Multiagent Systems: A Survey from a Machine Learning Perspective

Peter Stone*
AT&T Labs — Research
180 Park Ave., room A273
Florham Park, NJ 07932
pstone@research.att.com
www.research.att.com/~pstone

Manuela Veloso
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213
veloso@cs.cmu.edu
www.cs.cmu.edu/~mmv

In *Autonomous Robotics* volume 8, number 3. July, 2000.

Abstract

Distributed Artificial Intelligence (DAI) has existed as a subfield of AI for less than two decades. DAI is concerned with systems that consist of multiple independent entities that interact in a domain. Traditionally, DAI has been divided into two sub-disciplines: Distributed Problem Solving (DPS) focuses on the information management aspects of systems with several components working together towards a common goal; Multiagent Systems (MAS) deals with behavior management in collections of several independent entities, or agents. This survey of MAS is intended to serve as an introduction to the field and as an organizational framework. A series of general multiagent scenarios are presented. For each scenario, the issues that arise are described along with a sampling of the techniques that exist to deal with them. The presented techniques are not exhaustive, but they highlight how multiagent systems can be and have been used to build complex systems. When options exist, the techniques presented are biased towards machine learning approaches. Additional opportunities for applying machine learning to MAS are highlighted and robotic soccer is presented as an appropriate test bed for MAS. This survey does not focus exclusively on robotic systems. However, we believe that much of the prior research in non-robotic MAS is relevant to robotic MAS, and we explicitly discuss several robotic MAS, including all of those presented in this issue.

1 Introduction

Extending the realm of the social world to include autonomous computer systems has always been an awesome, if not frightening, prospect. However it is now becoming both possible and necessary through advances in the field of Artificial Intelligence (AI). In the past several years, AI techniques have become more and more robust and complex. To mention just one of the many exciting successes, a car steered itself more than 95% of the way across the United States using the ALVINN system [Pormerleau, 1993]. By meeting this and other such daunting challenges, AI researchers have earned the right to start examining the implications of multiple autonomous "agents" interacting in the real world. In fact, they have rendered this

*This work was carried out while a member of the Computer Science Department at Carnegie Mellon University.

examination indispensable. If there is one self-steering car, there will surely be more. And although each may be able to drive individually, if several autonomous vehicles meet on the highway, we must know how their behaviors interact.

Multiagent Systems (MAS) is the subfield of AI that aims to provide both principles for construction of complex systems involving multiple agents and mechanisms for coordination of independent agents' behaviors. While there is no generally accepted definition of "agent" in AI [Russell and Norvig, 1995], for the purposes of this article, we consider an agent to be an entity, such as a robot, with goals, actions, and domain knowledge, situated in an environment. The way it acts is called its "behavior." (This is not intended as a general theory of agency.) Although the ability to consider coordinating behaviors of autonomous agents is a new one, the field is advancing quickly by building upon pre-existing work in the field of Distributed Artificial Intelligence (DAI).

DAI has existed as a subfield of AI for less than two decades. Traditionally, DAI is broken into two sub-disciplines: Distributed Problem Solving (DPS) and MAS [Bond and Gasser, 1988]. The main topics considered in DPS are information management issues such as task decomposition and solution synthesis. For example, a constraint satisfaction problem can often be decomposed into several not entirely independent subproblems that can be solved on different processors. Then these solutions can be synthesized into a solution of the original problem.

MAS allows the subproblems of a constraint satisfaction problem to be subcontracted to different problem solving *agents* with their own interests and goals. Furthermore, domains with multiple agents of any type, including autonomous vehicles and even some human agents, are beginning to be studied.

This survey of MAS is intended as an introduction to the field. The reader should come away with an appreciation for the types of systems that are possible to build using MAS as well as a conceptual framework with which to organize the different types of possible systems.

The article is organized as a series of general multiagent scenarios. For each scenario, the issues that arise are described along with a sampling of the techniques that exist to deal with them. The techniques presented are not exhaustive, but they highlight how multiagent systems can be and have been used to build complex systems.

Because of the inherent complexity of MAS, there is much interest in using machine learning techniques to help deal with this complexity [Weiß and Sen, 1996; Sen, 1996]. When several different systems exist that could illustrate the same or similar MAS techniques, the systems presented here are biased towards those that use machine learning (ML) approaches. Furthermore, every effort is made to highlight additional opportunities for applying ML to MAS. This survey does not focus exclusively on robotic systems. However, we believe that much of the prior research in non-robotic MAS is relevant to robotic MAS, and we explicitly discuss several robotic MAS (referred to as *multi-robot systems*), including all of those presented in this issue.

Although there are many possible ways to divide MAS, the survey is organized along two main dimensions: agent heterogeneity and amount of communication among agents. Beginning with the simplest multiagent scenario, homogeneous non-communicating agents, the full range of possible multiagent systems, through highly heterogeneous communicating agents, is considered.

For each multiagent scenario presented, a single example domain is presented in an appropriate instantiation for the purpose of illustration. In this extensively-studied domain, the Predator/Prey or "Pursuit" domain [Benda *et al.*, 1986], many MAS issues arise. Nevertheless, it is a "toy" domain. At the end of the article, a much more complex domain—robotic soccer—is presented in order to illustrate the full power of MAS.

The article is organized as follows. Section 2 introduces the field of MAS, listing several of its strong points and presenting a taxonomy. The body of the article, Sections 3 – 7, presents the various multiagent scenarios, illustrates them using the pursuit domain, and describes existing work in the field. A domain that facilitates the study of most multiagent issues, robotic soccer, is advocated as a test bed in Section 8. Section 9 concludes.

2 Multiagent Systems

Two obvious questions about any type of technology are:

- What advantages does it offer over the alternatives?
- In what circumstances is it useful?

It would be foolish to claim that MAS should be used when designing all complex systems. Like any useful approach, there are some situations for which it is particularly appropriate, and others for which it is not. The goal of this section is to underscore the need for and usefulness of MAS while giving characteristics of typical domains that can benefit from it. For a more extensive discussion, see [Bond and Gasser, 1988].

Some domains require MAS. In particular, if there are different people or organizations with different (possibly conflicting) goals and proprietary information, then a multiagent system is needed to handle their interactions. Even if each organization wants to model its internal affairs with a single system, the organizations will not give authority to any single person to build a system that represents them all: the different organizations will need their own systems that reflect their capabilities and priorities.

For example, consider a manufacturing scenario in which company X produces tires, but subcontracts the production of lug-nuts to company Y. In order to build a single system to automate (certain aspects of) the production process, the internals of both companies X and Y must be modeled. However, neither company is likely to want to relinquish information and/or control to a system designer representing the other company. Perhaps with just two companies involved, an agreement could be reached, but with several companies involved, MAS is necessary. The only feasible solution is to allow the various companies to create their own agents that accurately represent their goals and interests. They must then be combined into a multiagent system with the aid of some of the techniques described in this article.

Another example of a domain that requires MAS is hospital scheduling as presented in [Decker, 1996c]. This domain from an actual case study requires different agents to represent the interests of different people within the hospital. Hospital employees have different interests, from nurses who may want to minimize the patient's time in the hospital, to x-ray operators who may want to maximize the throughput on their machines. Since different people evaluate candidate schedules with different criteria, they must be represented by separate agents if their interests are to be justly considered.

Even in domains that could conceivably use systems that are not distributed, there are several possible reasons to use MAS. Having multiple agents could speed up a system's operation by providing a method for parallel computation. For instance, a domain that is easily broken into components—several independent tasks that can be handled by separate agents—could benefit from MAS. Furthermore, the parallelism of MAS can help deal with limitations imposed by time-bounded or space-bounded reasoning requirements.

While parallelism is achieved by assigning different tasks or abilities to different agents, robustness is a benefit of multiagent systems that have redundant agents. If control and responsibilities are sufficiently shared among different agents, the system can tolerate failures by one or more of the agents. Domains that must degrade gracefully are in particular need of this feature of MAS: if a single entity—processor or agent—controls everything, then the entire system could crash if there is a single failure. Although a multiagent system need not be implemented on multiple processors, to provide full robustness against failure, its agents should be distributed across several machines.

Another benefit of multiagent systems is their scalability. Since they are inherently modular, it should be easier to add new agents to a multiagent system than it is to add new capabilities to a monolithic system. Systems whose capabilities and parameters are likely to need to change over time or across agents can also benefit from this advantage of MAS.

From a programmer's perspective the modularity of multiagent systems can lead to simpler programming. Rather than tackling the whole task with a centralized agent, programmers can identify subtasks and assign control of those subtasks to different agents. The difficult problem of splitting a single agent's time among different parts of a task solves itself. Thus, when the choice is between using a multiagent system or a single-agent system, MAS may be the simpler option. Of course there are some domains that are more naturally approached from an omniscient perspective—because a global view is given—or with centralized control—because no parallel actions are possible and there is no action uncertainty [Decker, 1996b]. Single-agent systems should be used in such cases.

Multiagent systems can also be useful for their illucidation of fundamental problems in the social sciences and life sciences [Cao *et al.*, 1997], including intelligence itself [Decker, 1987],. As Weiß put it: "Intelligence is deeply and inevitably coupled with interaction" [Weiß, 1996]. In fact, it has been proposed that the best way to develop intelligent machines at all might be to start by creating "social" machines [Dautenhahn, 1995]. This theory is based on the socio-biological theory that primate intelligence first evolved because of the need to deal with social interactions [Minsky, 1988].

While all of the above reasons to use MAS apply generally, there are also some arguments in favor of multi-*robot* systems in particular. In tasks that require robots to be in particular places, such as robot scouting, a team of robots has an advantage over a single robot in that it can take advantage of geographic distribution. While a single robot could only sense the world from a single vantage point, a multi-robot system can observe and act from several locations simultaneously.

Finally, as argued in [Jung and Zelinsky, 2000], multi-robot systems can exhibit benefits over single-robot systems in terms of the "performance/cost ratio." By using heterogeneous robots each with a subset of the capabilities necessary to accomplish a given task, one can use simpler robots that are presumably less expensive to engineer than a single monolithic robot with all of the capabilities bundled together. Reasons presented above to use MAS are summarized in Table 1.

Table 1: Reasons to use Multiagent Systems

• Some domains require it	• Simpler programming
• Parallelism	• To study intelligence
• Robustness	• Geographic distribution
• Scalability	• Cost effectiveness

2.1 Taxonomy

Several taxonomies have been presented previously for the related field of Distributed Artificial Intelligence (DAI). For example, Decker presents four dimensions of DAI [Decker, 1987]:

1. Agent granularity (coarse vs. fine);
2. Heterogeneity of agent knowledge (redundant vs. specialized);
3. Methods of distributing control (benevolent vs. competitive, team vs. hierarchical, static vs. shifting roles);
4. and Communication possibilities (blackboard vs. messages, low-level vs. high-level, content).

Along dimensions 1 and 4, multiagent systems have coarse agent granularity and high-level communication. Along the other dimensions, they can vary across the whole ranges. In fact, the remaining dimensions are very prominent in this article: degree of heterogeneity is a major MAS dimension and all the methods of distributing control appear here as major issues.

More recently, Parunak [1996] has presented a taxonomy of MAS from an application perspective. From this perspective, the important characteristics of MAS are:

- System function;
- Agent architecture (degree of heterogeneity, reactive vs. deliberative);
- System architecture (communication, protocols, human involvement).

A useful contribution is that the dimensions are divided into agent and system characteristics. Other overviews of DAI and/or MAS include [Lesser, 1995; Durfee, 1992; Durfee *et al.*, 1989; Bond and Gasser, 1988].

There are also some existing surveys that are specific to multi-*robot* systems. Dudek et al. [1996] presented a detailed taxonomy of multiagent robotics along seven dimensions, including robot size, various communication parameters, reconfigurability, and unit processing. Cao et al. [1997] presented a "taxonomy based on problems and solutions," using the following five axes: group architecture, resource conflicts, origins of cooperation, learning, and geometric problems. It specifically does not consider competitive multi-robot scenarios. This article contributes a taxonomy that encompasses MAS along with a detailed chronicle of existing systems as they fit in to this taxonomy.

The taxonomy presented in this article is organized along what we believe to be the most important aspects of agents (as opposed to domains): degree of heterogeneity and degree of communication. Communication is presented as an agent aspect because it is the degree to which the agents communicate (or *whether* they communicate), not the communication protocols that are available to them, that is considered. Other aspects of agents in MAS are touched upon within the heterogeneity/communication framework. For example, the degree to which different agents play different roles is certainly an important MAS issue, but here it is framed within the scenario of heterogeneous non-communicating agents (it arises in the other three scenarios as well).

All four combinations of heterogeneity and communication—homogeneous non-communicating agents; heterogeneous non-communicating agents; homogeneous communicating agents; and heterogeneous communicating agents—are considered in this article. Our approach throughout the article is to categorize the issues as they are reflected in the literature. Many of the issues could apply in earlier scenarios, but do not in the articles that we have come across. On the other hand, many of the issues that arise in the earlier scenarios also apply in the later scenarios. Nevertheless, they are only mentioned again in the later scenarios to the degree that they differ or become more complex.

The primary purpose of this taxonomy is as a framework for considering and analyzing the challenges that arise in MAS. This survey is designed to be useful to researchers as a way of separating out the issues that arise as a result of their decisions to use homogeneous versus heterogeneous agents and communicating versus non-communicating agents.

The multiagent scenarios along with the issues that arise therein and the techniques that currently exist to address these issues are described in detail in Sections 4 – 7. Table 2 gives a preview of these scenarios and associated issues as presented in this article.

2.2 Single-Agent vs. Multiagent Systems

Before studying and categorizing MAS, we must first consider their most obvious alternative: centralized, single-agent systems. Centralized systems have a single agent which makes all the decisions, while the others act as remote slaves. For the purposes of this survey, a "single-agent system" should be thought of as a centralized system in a domain which also allows for a multiagent approach.

A single-agent system might still have multiple entities — several actuators, or even several physically

Homogeneous Non-communicating Agents	Heterogeneous Non-communicating Agents
• Reactive vs. deliberative agents • Local or global perspective • Modeling of other agents' states • How to affect others	• Benevolence vs. competitiveness • Fixed vs. learning agents (arms race, credit-assignment) • Modeling of others' goals, actions, and knowledge • Resource management (interdependent actions) • Social conventions • Roles
Homogeneous Communicating Agents	**Heterogeneous Communicating Agents**
• Distributed sensing • Communication content	• Understanding each other • Planning communicative acts • Benevolence vs. competitiveness • Negotiation • Resource management (schedule coordination) • Commitment/decommitment • Collaborative localization • Changing shape and size

Table 2: Issues arising in the various scenarios as reflected in the literature.

separated components. However, if each entity sends its perceptions to and receives its actions from a single central process, then there is only a single agent: the central process. The central agent models all of the entities as a single "self." This section compares the single-agent and multiagent approaches.

2.2.1 Single-Agent Systems

In general, the agent in a single-agent system models itself, the environment, and their interactions. Of course the agent is itself part of the environment, but for the purposes of this article, agents are considered to have extra-environmental components as well. They are independent entities with their own goals, actions, and knowledge. In a single-agent system, no other such entities are recognized by the agent. Thus, even if there are indeed other agents in the world, they are not modeled as having goals, etc.: they are just considered part of the environment. The point being emphasized is that although agents are *also* a part of the environment, they are explicitly modeled as having their own goals, actions, and domain knowledge (see Figure 1).

2.2.2 Multiagent Systems

Multiagent systems differ from single-agent systems in that several agents exist which model each other's goals and actions. In the fully general multiagent scenario, there may be direct interaction among agents (communication). Although this interaction could be viewed as environmental stimuli, we present inter-agent communication as being separate from the environment.

From an individual agent's perspective, multiagent systems differ from single-agent systems most significantly in that the environment's dynamics can be affected by other agents. In addition to the uncertainty

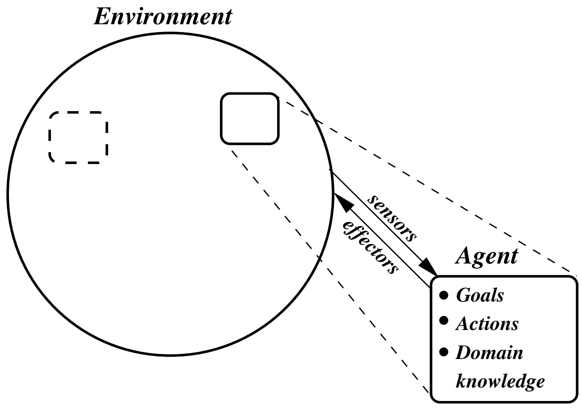

Figure 1: A general single-agent framework. The agent models itself, the environment, and their interactions. If other agents exist, they are considered part of the environment.

that may be inherent in the domain, other agents intentionally affect the environment in unpredictable ways. Thus, all multiagent systems can be viewed as having dynamic environments.

Figure 2 illustrates the view that each agent is both part of the environment and modeled as a separate entity. There may be any number of agents, with different degrees of heterogeneity and with or without the ability to communicate directly. From the fully general case depicted here, we begin by eliminating both the communication and the heterogeneity to present homogeneous non-communicating MAS (Section 4). Then, the possibilities of agent heterogeneity and inter-agent communication are considered one at a time (Sections 5 and 6). Finally, in Section 7, we arrive back at the fully general case by considering agents that can interact directly.

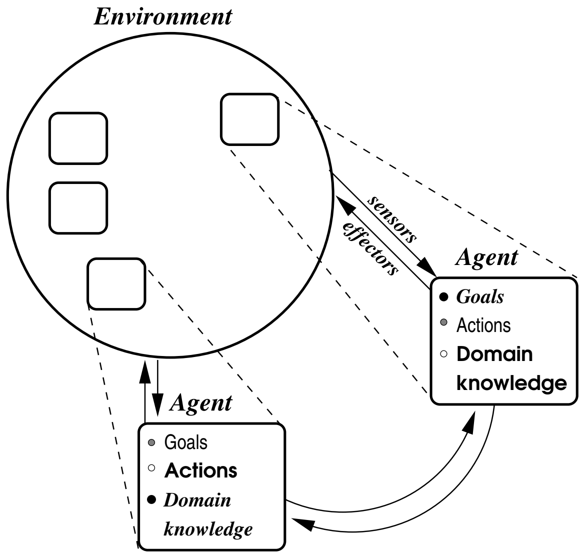

Figure 2: The fully general multiagent scenario. Agents model each other's goals, actions, and domain knowledge, which may differ as indicated by the different fonts. They may also interact directly (communicate) as indicated by the arrows between the agents.

3 Organization of Existing Work

The following sections present many different MAS techniques that have been previously published. They present an extensive, but not exhaustive, list of work in the field. Space does not permit exhaustive coverage. Instead, the work mentioned is intended to illustrate the techniques that exist to deal with the issues that arise in the various multiagent scenarios. When possible, ML approaches are emphasized.

All four multiagent scenarios are considered in the following order: homogeneous non-communicating agents, heterogeneous non-communicating agents, homogeneous communicating agents, and heterogeneous communicating agents. For each of these scenarios, the research issues that arise, the techniques that deal with them, and additional ML opportunities are presented. The issues may appear across scenarios, but they are presented and discussed in the first scenario to which they apply.

In addition to the existing learning approaches described in the sections entitled “Issues and Techniques”, there are several previously unexplored learning opportunities that apply in each of the multiagent scenarios. For each scenario, a few promising opportunities for ML researchers are presented.

Many existing ML techniques can be directly applied in multiagent scenarios by delimiting a part of the domain that only involves a single agent. However *multiagent learning* is more concerned with learning issues that arise because of the multiagent aspect of a given domain. As described by Weiß, multiagent learning is “learning that is done by several agents and that becomes possible only because several agents are present” [Weiß, 1995]. This type of learning is emphasized in the sections entitled “Further Learning Opportunities.”

For the purpose of illustration, each scenario is accompanied by a suitable instantiation of the Predator/Prey or “Pursuit” domain.

3.1 The Predator/Prey (“Pursuit”) Domain

The Predator/Prey, or “Pursuit” domain (hereafter referred to as the “pursuit domain”), is an appropriate one for illustration of MAS because it has been studied using a wide variety of approaches and because it has many different instantiations that can be used to illustrate different multiagent scenarios. Since it involves agents moving around in a world, it is particularly appropriate as an abstraction of robotic MAS. The pursuit domain is not presented as a complex real-world domain, but rather as a toy domain that helps concretize many concepts. For discussion of a domain that has the full range of complexities characteristic of more real-world domains, see Section 8.

The pursuit domain was introduced by Benda et al. [1986]. Over the years, researchers have studied several variations of its original formulation. In this section, a single instantiation of the domain is presented. However, care is taken to point out the parameters that can be varied.

The pursuit domain is usually studied with four *predators* and one *prey*. Traditionally, the predators are blue and the prey is red (black and grey respectively in Figure 3). The domain can be varied by using different numbers of predators and prey.

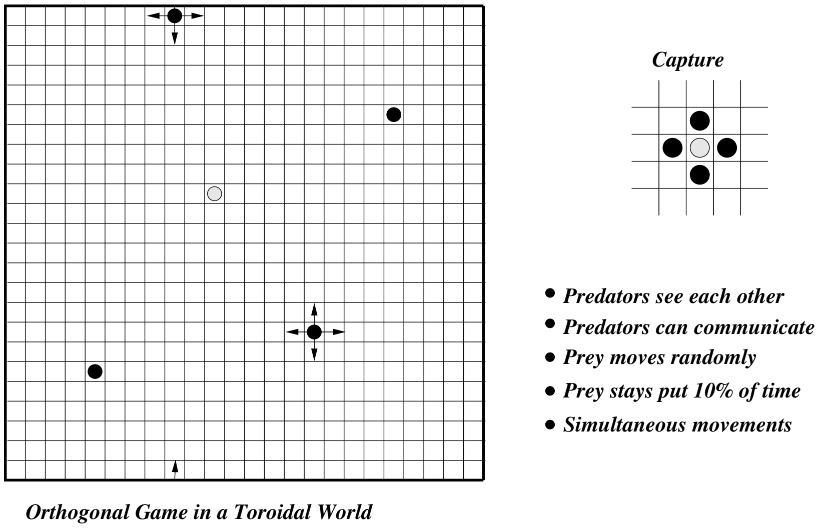

Figure 3: A particular instantiation of the pursuit domain. Predators are black and the prey is grey. The arrows on top of two of the predators indicate possible moves.

The goal of the predators is to "capture" the prey, or surround it so that it cannot move to an unoccupied position. A capture position is shown in Figure 3. If the world has boundaries, fewer than four predators can capture the prey by trapping it against an edge or in a corner. Another possible criterion for capture is that a predator occupies the same position as the prey. Typically, however, no two players are allowed to occupy the same position.

As depicted in Figure 3, the predators and prey move around in a discrete, grid-like world with square spaces. They can move to any adjacent square on a given turn. Possible variations include grids with other shapes as spaces (for instance hexagons) or continuous worlds. Within the square game, players may be allowed to move diagonally instead of just horizontally and vertically. The size of the world may also vary from an infinite plane to a small, finite board with edges. The world pictured in Figure 3 is a toroidal world: the predators and prey can move off one end of the board and come back on the other end. Other parameters of the game that must be specified are whether the players move simultaneously or in turns; how much of the world the predators can see; and whether and how the predators can communicate.

Finally, in the original formulation of the domain, and in most subsequent studies, the prey moves randomly: on each turn it moves in a random direction, staying still with a certain probability in order to simulate being slower than the predators. However, it is also possible to allow the prey to actively try to escape capture. As is discussed in Section 5, there has been some research done to this effect, but there is still much room for improvement. The parameters that can be varied in the pursuit domain are summarized in Table 3.

The pursuit domain is a good one for the purposes of illustration because it is simple to understand and because it is flexible enough to illustrate a variety of scenarios. The possible actions of the predators and

Table 3: Variable parameters in the pursuit domain

- Definition of capture
- Size and shape of the world
- Predators' legal moves
- Simultaneous or sequential movement
- Visible objects and range
- Predator communication
- Prey movement

prey are limited and the goal is well-defined. In terms of the reasons to use MAS as presented in Table 1, the pursuit domain does not necessarily require MAS. But in certain instantiations it can make use of the parallelism, robustness, and simpler programming offered by MAS.

In the pursuit domain, a single-agent approach is possible: the agent can observe the positions of all four predators and decide how each of them should move. Since the prey moves randomly rather than intentionally, it is not associated with any agent. Instead it is considered part of the environment as shown in Figure 4. It is also possible to consider DPS approaches to the pursuit domain by breaking the task into subproblems to be solved by each predator. However, most of the approaches described here model the predators as independent agents with a common goal. Thus, they comprise a multiagent system.

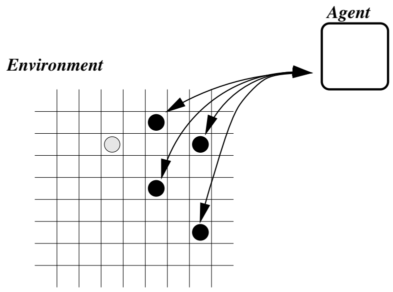

Figure 4: The pursuit domain with just a single agent. One agent controls all predators and the prey is considered part of the environment.

For each of the multiagent scenarios presented below, a new instantiation of the pursuit domain is defined. Their purpose is to illustrate the different scenarios within a concrete framework.

3.2 Domain Issues

Throughout this survey, the focus is upon agent capabilities. However, from the point of view of the system designer, the characteristics of the domain are at least as important. Before moving on to the agent-based categorization of the field in Sections 4 – 7, a range of domain characteristics is considered.

Relevant domain characteristics include: the number of agents; the amount of time pressure for generating actions (is it a real-time domain?); whether or not new goals arrive dynamically; the cost of communication; the cost of failure; user involvement; and environmental uncertainty. The first four of these characteristics are self-explanatory and do not need further mention.

With respect to cost of failure, an example of a domain with high cost of failure is air-traffic control [Rao and Georgeff, 1995]. On the other hand, the directed improvisation domain considered by Hayes-Roth et al. [1995] has a very low cost of failure. In this domain, entertainment agents accept all improvisation suggestions from each other. The idea is that the agents should not be afraid to make mistakes, but rather should "just let the words flow" [Hayes-Roth *et al.*, 1995].

Several multiagent systems include humans as one or more of the agents. In this case, the issue of communication between the human and computer agents must be considered [Sanchez *et al.*, 1995]. Another example of user involvement is user feedback in an information filtering domain [Ferguson and Karakoulas, 1996].

Decker [1995] distinguishes three different sources of uncertainty in a domain. The state transitions in the domain itself might be non-deterministic; agents might not know the actions of other agents; and agents might not know the outcomes of their own actions. This and the other domain characteristics are summarized in Table 4.

Table 4: Domain characteristics that are important when designing MAS

- Number of agents
- Amount of time pressure (real time?)
- Dynamically arriving goals?
- Cost of communication
- Cost of failure
- User involvement
- Environmental uncertainty: [Decker, 1995]
  - a priori in the domain
  - in the actions of other agents
  - in outcomes of an agent's own actions

4 Homogeneous Non-Communicating Multiagent Systems

In homogeneous, non-communicating multiagent systems, all of the agents have the same internal structure including goals, domain knowledge, and *possible* actions. They also have the same procedure for selecting among their actions. The only differences among agents are their sensory inputs and the actual actions they take: they are situated differently in the world.

4.1 Homogeneous Non-Communicating Multiagent Pursuit

In the homogeneous non-communicating version of the pursuit domain, rather than having one agent controlling all four predators, there is one *identical* agent per predator. Although the agents have identical capabilities and decision procedures, they have limited information about each other's internal state and sensory inputs. Thus they are not be able to predict each other's actions. The pursuit domain with homogeneous agents is illustrated in Figure 5.

Within this framework, Stephens and Merx [1990] propose a simple heuristic behavior for each agent that is based on local information. They define *capture positions* as the four positions adjacent to the prey. They then propose a "local" strategy whereby each predator agent determines the capture position to which

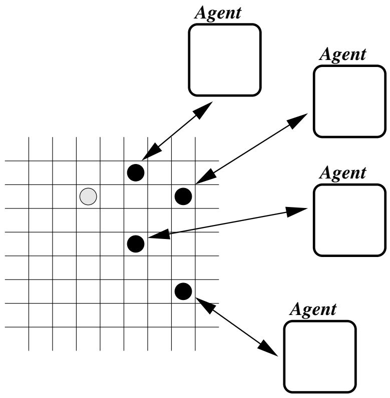

Figure 5: The pursuit domain with homogeneous agents. There is one identical agent per predator. Agents may have (the same amount of) limited information about other agents' internal states.

it is closest and moves towards that position. The predators cannot see each other, so they cannot aim at different capture positions. Of course a problem with this heuristic is that two or more predators may move towards the same capture position, blocking each other as they approach. This strategy is not very successful, but it serves as a basis for comparison with two other control strategies—"distributed" and "central"—that are discussed in Sections 6 and 7.

Since the predators are identical, they can easily predict each other's actions given knowledge of each other's sensory input. Such prediction can be useful when the agents move simultaneously and would like to base their actions on where the other predators will be at the next time step. Vidal and Durfee [1995] analyze such a situation using the Recursive Modeling Method (RMM). RMM is discussed in more detail below, but the basic idea is that predator A bases its move on the predicted move of predator B and vice versa. Since the resulting reasoning can recurse indefinitely, it is important for the agents to bound the amount of reasoning they use either in terms of time or in terms of levels of recursion. Vidal and Durfee's [1995] Limited Rationality RMM algorithm is designed to take such considerations into account.

Levy and Rosenschein [1992] use a game theoretical approach to the pursuit domain. They use a payoff function that allows selfish agents to cooperate. A requirement for their model is that each predator has full information about the location of other predators. Their game model mixes game-theoretical cooperative and non-cooperative games.

Korf [1992] also takes the approach that each agent should try to greedily maximize its own local utility. He introduces a policy for each predator based on an attractive force to the prey and a repulsive force from the other predators. Thus the predators tend to approach the prey from different sides. This policy is very successful, especially in the diagonal (agents can move diagonally as well as orthogonally) and hexagonal (hexagonal grid) games. Korf draws the conclusion that explicit cooperation is rarely necessary or useful, at least in the pursuit domain and perhaps more broadly:

> We view this work as additional support for the theory that much coordination and cooperation in both natural and man-made systems can be viewed as an emergent property of the interaction

of greedy agents maximizing their particular utility functions in the presence of environmental constraints. [Korf, 1992]

However, whether or not altruism occurs in nature, there is certainly some use for benevolent agents in MAS, as shown below. More pressingly, if Korf's claim that the pursuit domain is easily solved with local greedy heuristics were true, there would be no point in studying the pursuit domain any further. Fortunately, Haynes and Sen [1998] show that Korf's heuristics do not work for certain instantiations of the domain (see Section 5).

4.2 General Homogeneous Non-Communicating MAS

In the general multiagent scenario with homogeneous agents, there are several different agents with identical structure (sensors, effectors, domain knowledge, and decision functions), but they have different sensor input and effector output. That is to say, they are situated differently in the environment and they make their own decisions regarding which actions to take. Having different effector output is a necessary condition for MAS: if the agents all act as a unit, then they are essentially a single agent. In order to realize this difference in output, homogeneous agents must have different sensor input as well. Otherwise they will act identically. For this scenario, in which we consider non-communicating agents, assume that the agents cannot communicate directly. Figure 6 illustrates the homogeneous, non-communicating multiagent scenario, indicating that the agents' goals, actions, and domain knowledge are the same by representing them with identical fonts.

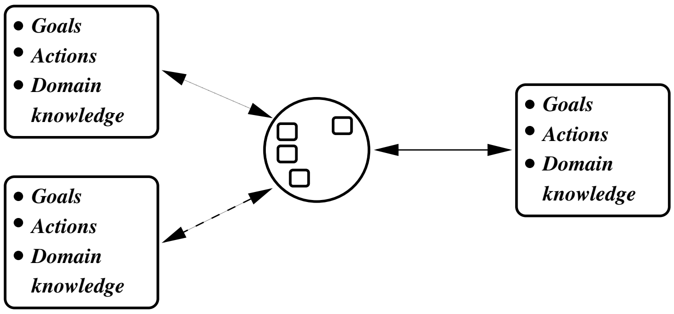

Figure 6: MAS with homogeneous agents. Only the sensor input and effector output of agents differ, as represented by the different arrow styles. The agents' goals, actions, and/or domain knowledge are all identical as indicated by the identical fonts.

4.3 Issues and Techniques

Even in this most restrictive of multiagent scenarios, there are several issues with which to deal. The techniques provided here are representative examples of ways to deal with the presented issues. The issues and techniques, as well as the learning opportunities discussed later, are summarized in Table 5.

Homogeneous Non-Communicating

Issues

- Reactive vs. deliberative agents
- Local or global perspective
- Modeling of other agents' states
- How to affect others

Learning opportunities

- Enable others' actions
- Sensor data $\rightarrow$ Other agent's sensor data

Techniques

- Reactive behaviors for formation maintenance. [Balch and Arkin, 1995]
- Deliberative behaviors for pursuit. [Levy and Rosenschein, 1992]
- Mixed reactive and deliberative behaviors. [Sahota, 1994; Rao and Georgeff, 1995]
- Local knowledge sometimes better. [Sen *et al.*, 1998]
- (limited) Recursive Modeling Method (RMM). [Durfee, 1995]
- Don't model others–just pay attention to reward. [Schmidhuber, 1996]
- Stigmergy. [Goldman and Rosenschein, 1994; Holland, 1996]
- Q-learning for behaviors like foraging, homing, etc. [Mataric, 1994a]

Table 5: The issues, techniques, and learning opportunities for homogeneous MAS as reflected in the literature.

4.3.1 Reactive vs. Deliberative agents

When designing any agent-based system, it is important to determine how sophisticated the agents' reasoning will be. Reactive agents simply retrieve pre-set behaviors similar to reflexes without maintaining any internal state. On the other hand, deliberative agents behave more like they are thinking, by searching through a space of behaviors, maintaining internal state, and predicting the effects of actions. Although the line between reactive and deliberative agents can be somewhat blurry, an agent with no internal state is certainly reactive, and one which bases its actions on the predicted actions of other agents is deliberative. Here we describe one system at each extreme as well as two others that mix reactive and deliberative reasoning.

Balch and Arkin [1995] use homogeneous, reactive, non-communicating agents to study formation maintenance in autonomous robots. The robots' goal is to move together in a military formation such as a diamond, column, or wedge. They periodically come across obstacles which prevent one or more of the robots from moving in a straight line. After passing the obstacle, all robots must adjust in order to regain their formation. The agents reactively convert their sensory data (which includes the positions of the other robots) to motion vectors for avoiding obstacles, avoiding robots, moving to a goal location, and formation maintenance. The actual robot motion is a simple weighted sum of these vectors.

At the deliberative end of the spectrum is the pursuit domain work by Levy and Rosenschein [1992] that is mentioned above. Their agents assume that each will act in service of its own goals. They use game theoretic techniques to find equilibrium points and thus to decide how to act. These agents are clearly deliberative, considering that they search for actions rather than simply retrieving them.

There are also several existing systems and techniques that mix reactive and deliberative behaviors. One example is the OASIS system which reasons about when to be reactive and when to follow goal-directed plans [Rao and Georgeff, 1995]. Another example is *reactive deliberation* [Sahota, 1994]. As the name implies, it mixes reactive and deliberative behavior: an agent reasons about which reactive behavior to follow under the constraint that it must choose actions at a rate of 60 Hz. Reactive deliberation was developed on the first robotic soccer platform [Barman *et al.*, 1993]. Reactive deliberation was not explicitly designed for MAS, but because it was designed for real-time control in dynamic environments, it is likely to be extendible to multiagent scenarios.

4.3.2 Local or global perspective

Another issue to consider when building a multiagent system is how much sensor information should be available to the agents. Even if it is feasible within the domain to give the agents a global perspectives of the world, it may be more effective to limit them to local views.

Sen et al. consider a case of multiple agents sharing a set of identical resources in which they have to learn (adapt) their resource usage policies [Sen *et al.*, 1998]. Since the agents are identical and do not communicate, if they all have a global view of the current resource usage, they will all move simultaneously to the most under-used resource. However, if they each see a partial picture of the world, then different agents gravitate towards different resources: a preferable effect. Better performance by agents with less knowledge is occasionally summarized by the cliche "Ignorance is Bliss."

4.3.3 Modeling of other agents' states

Durfee [1995] provides another example of "Blissful Ignorance," mentioning it explicitly in the title of his paper: "Blissful Ignorance: Knowing Just Enough to Coordinate Well." Now rather than referring to resource usage, the saying applies to the limited recursive modeling method (RMM). When using RMM, agents explicitly model the belief states of other agents, including what they know about each other's beliefs. If agents have too much knowledge, RMM could recurse indefinitely. Even if further information can be obtained by reasoning about what agent A thinks agent B thinks agent A thinks . . . , endless reasoning can lead to inaction. Durfee contends that for coordination to be possible, some potential knowledge must be ignored. As well as illustrating this concept in the pursuit domain [Vidal and Durfee, 1995], Durfee goes into more detail and offers more generally applicable methodology in [Durfee, 1995].

The point of the RMM is to model the internal state of another agent in order to predict its actions. Even though the agents know each other's goals and structure (they are homogeneous), they may not know each other's future actions. The missing pieces of information are the internal states (for deliberative agents) and sensory inputs of the other agents. How and whether to model other agents is a ubiquitous issue in MAS. In the more complex multiagent scenarios presented in the next sections, agents may have to model not only the internal states of other agents, but also their goals, actions, and abilities.

Although it may be useful to build models of other agents in the environment, agent modeling is not done universally. A form of multiagent RL is defined in which agents do not model each other as agents [Schmidhuber, 1996]. Instead they consider each other as parts of the environment and affect each other's policies only as sensed objects. The agents pay attention to the reward they receive using a given policy and checkpoint their policies so they can return to successful ones. Schmidhuber shows that the agents can learn to cooperate without modeling each other.

4.3.4 How to affect others

When no communication is possible, agents cannot interact with each other directly. However, since they exist in the same environment, the agents can affect each other indirectly in several ways. They can be sensed by other agents, or they may be able to change the state of another agent by, for example, pushing it. Agents can also affect each other by one of two types of *stigmergy* [Beckers *et al.*, 1994]. First, *active stigmergy* occurs when an agent alters the environment so as to affect the sensory input of another agent. For example, a robotic agent might leave a marker behind it for other agents to observe. Goldman and Rosenschein [1994] demonstrate an effective form of active stigmergy in which agents heuristically alter the environment in order to facilitate future unknown plans of other agents. Second, *passive stigmergy* involves altering the environment so that the effects of another agent's actions change. For example, if one agent turns off the main water valve to a building, the effect of another agent subsequently turning on the kitchen faucet is altered.

Holland [1996] illustrates the concept of passive stigmergy with a robotic system designed to model the behavior of an ant colony confronted with many dead ants around its nest. An ant from such a colony tends to periodically pick up a dead ant, carry it for a short distance, and then drop it. Although the behavior appears to be random, after several hours, the dead ants are clustered in a small number of heaps. Over time, there are fewer and fewer large piles until all the dead ants end up in one pile. Although the ants behave homogeneously and, at least in this case, we have no evidence that they communicate explicitly, the ants manage to cooperate in achieving a task.

Holland [1996] models this situation with a number of identical robots in a small area in which many pucks are scattered around. The robots are programmed reactively to move straight (turning at walls) until they are pushing three or more pucks. At that point, the robots back up and turn away, leaving the three pucks in a cluster. Although the robots do not communicate at all, they are able to collect the pucks into a single pile over time. This effect occurs because when a robot approaches an existing pile directly, it adds the pucks it was already carrying to the pile and turns away. A robot approaching an existing pile obliquely might take a puck away from the pile, but over time the desired result is accomplished. Like the ants, the robots use passive stigmergy to affect each other's behavior.

A similar scenario with more deliberative robots is explored by Mataric. In this case, the robots use Q-learning to learn behaviors including foraging for pucks as well as homing and following [Mataric, 1994a]. The robots learn independent policies, dealing with the high-dimensional state space with the aid of *progress* *estimators* that give intermediate rewards, and with the aid of boolean value predicates that condense many states into one. Mataric's robots actively affect each other through observation: a robot learning to follow another robot can base its action on the relative location of the other robot.

4.4 Further Learning Opportunities

In addition to the existing learning approaches described above, there are several unexplored learning opportunities that apply to homogeneous non-communicating systems (see Table 5).

One unexplored learning opportunity that could apply in domains with homogeneous non-communicating agents is learning to enable others' actions. Inspired by the concept of stigmergy, an agent may try to learn to take actions that will not directly help it in its current situation, but that may allow other similar agents to be more effective in the future. Typical RL situations with delayed reward encourage agents to learn to achieve their goals directly by propagating local reinforcement back to past states and actions [Kaelbling *et al.*, 1996]. However if an action leads to a reward by another agent, the acting agent may have no way of reinforcing that action. Techniques to deal with such a problem would be useful for building multiagent systems.

In terms of modeling other agents, there is much room for improvement in the situation that a given agent does not know the internal state or sensory inputs of another agent (recall that in this homogeneous scenario, everything else about the other agent is known, being identical across agents). When such information is known, RMM can be used to determine future actions of agents. However, if the information is not directly available, it would be useful for an agent to learn it. Thus, the function mapping agent X's sensor data (which might include a restricted view of agent Y) to agent Y's current sensor data is a useful function to learn. If effectively learned, agent X can then use (limited) RMM to predict agent Y's future actions.

5 Heterogeneous Non-Communicating Multiagent Systems

To this point, we have only considered agents that are homogeneous. Adding the possibility of heterogeneous agents in a multiagent domain adds a great deal of potential power at the price of added complexity. Agents might be heterogeneous in any of a number of ways, from having different goals to having different domain models and actions. An important subdimension of heterogeneous agent systems is whether agents are benevolent or competitive. Even if they have different goals, they may be friendly to each other's goals or they may actively try to inhibit each other. The degree of heterogeneity within a MAS can be measured in an information-theoretic way using Balch's *social entropy* [2000].

5.1 Heterogeneous Non-Communicating Multiagent Pursuit

Before exploring the general multiagent scenario involving heterogeneous non-communicating agents, consider how this scenario can be instantiated in the pursuit domain. As in the previous scenario, the predators are controlled by separate agents. But they are no longer necessarily identical agents: their goals, actions and domain knowledge may differ. In addition, the prey, which inherently has goals different from those of the predators, can now be modeled as an agent. The pursuit domain with heterogeneous agents is illustrated in Figure 7.

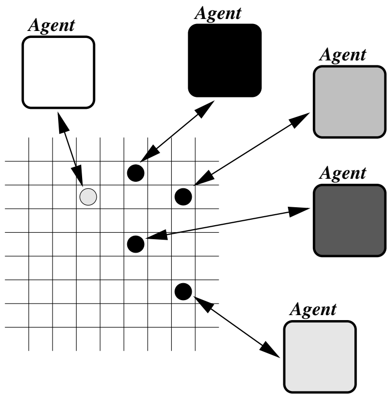

Figure 7: The pursuit domain with heterogeneous agents. Goals and actions may differ among agents. Now the prey may also be modeled as an agent.

Haynes and colleagues have done various studies with heterogeneous agents in the pursuit domain [1995; 1998; 1996]. They have evolved teams of predators, equipped predators with case bases, and competitively evolved the predators and the prey.

First, Haynes et al. use genetic programming (GP) to evolve teams of four predators [1995]. Rather than evolving predator agents in a single evolutionary pool and then combining them into teams to test performance, each individual in the population is actually a team of four agents already specifically assigned to different predators. Thus the predators can evolve to cooperate. This co-evolution of teammates is one possible way around the absence of communication in a domain. In place of communicating planned actions to each other, the predators can evolve to *know*, or at least act as if knowing, each other's future actions.

In a separate study, Haynes and Send use case-based reasoning to allow predators to learn to cooperate [1998]. They begin with identical agents controlling each of the predators. The predators move simultaneously to their closest capture positions. But because predators that try to occupy the same position all remain stationary, cases of deadlock arise. When deadlock occurs, the agents store the negative case so as to avoid it in the future, and they try different actions. Keeping track of which agents act in which way for given deadlock situations, the predators build up different case bases and thus become heterogeneous agents. Over time, the predators learn to stay out of each other's way while approaching the prey.

Finally, Haynes and Sen [1996] explore the possibility of evolving both the predators and the prey so that they all try to improve their behaviors. Working in a toroidal world and starting with predator behaviors such as Korf's greedy heuristic [Korf, 1992] and their own evolved GP predators, they then evolve the prey to behave more effectively than randomly. Although one might think that continuing this process would lead to repeated improvement of the predator and prey behaviors with no convergence, a prey behavior emerges that always succeeds: the prey simply moves in a constant straight line. Even when allowed to re-adjust to the "linear" prey behavior, the predators are unable to reliably capture the prey. Haynes and Sen conclude that Korf's greedy solution to the pursuit domain relies on random prey movement which guarantees locality of movement. Although there may yet be greedy solutions that can deal with different types of prey behavior, they have not yet been discovered. Thus the predator domain retains value for researchers in MAS.

Although Haynes and Sen [1996] show that the pursuit domain is still worth studying, the co-evolutionary results are less than satisfying. As mentioned above, one would intuitively expect the predators to be able to adapt to the linearly moving prey. For example, since they operate in a toroidal world, a single predator could place itself in the prey's line of movement and remain still. Then the remaining predators could surround the prey at their leisure. The fact that the predators are unable to re-evolve to find such a solution suggests that either the predator evolution is not performed optimally, or slightly more "capable" agents (e.g. agents able to reason more about past world states) would lead to a more interesting study. Nevertheless, the study of competitive co-evolution in the pursuit domain started by Haynes and Sen is an intriguing open issue.

5.2 General Heterogeneous Non-Communicating MAS

The general multiagent scenario with heterogeneous non-communicating agents is depicted in Figure 8. As in the homogeneous case (see Figure 6), the agents are situated differently in the environment which causes them to have different sensory inputs and necessitates their taking different actions. However in this scenario, the agents have much more significant differences. They may have different goals, actions, and/or domain knowledge as indicated by the different fonts in Figure 8. In order to focus on the benefits (and complexity) of heterogeneity, the assumption of no communication is retained for this section.

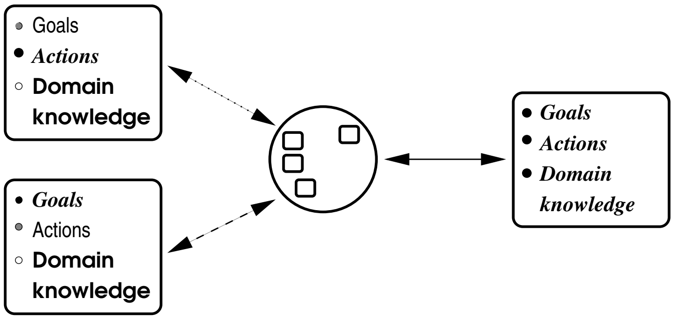

Figure 8: The general heterogeneous MAS scenario. Now agents' goals, actions, and/or domain knowledge may differ as indicated by the different fonts. The assumption of no direct interaction remains.

5.3 Issues and Techniques

Even without communication, numerous issues that were not present in the homogeneous agent scenario (Section 4) arise in this scenario. Some have already been touched upon above in the context of the pursuit domain. These issues and existing techniques to deal with them, along with further learning opportunities, are described below and summarized in Table 6.

Heterogeneous Non-Communicating

Issues

- Benevolence vs. competitiveness
- Fixed vs. learning agents (arms race, credit-assignment)
- Modeling of others' goals, actions, and knowledge
- Resource management (interdependent actions)
- Social conventions
- Roles

Learning opportunities

- Credit-assignment in competitive scenarios
- Behaviors that blend well with team
- Prediction of others' actions
- Dynamic role assumption

Techniques

- Game theory, iterative play. [Mor and Rosenschein, 1995; Sandholm and Crites, 1996]
- Minimax-Q. [Littman, 1994]
- Competitive co-evolution. [Rosin and Belew, 1995; Haynes and Sen, 1996; Grefenstette and Daley, 1996; Stone, 2000]
- Deduce intentions, abilities through observation. [Huber and Durfee, 1995; Wang, 1996]
- Autoepistemic reasoning (ignorance). [Permpoontanalarp, 1995]
- Model as a team (individual $\rightarrow$ role). [Tambe, 1995, 1996]
- Social reasoning: depend on others for goal ($\neq$ game theory). [Sichman and Demazeau, 1995]
- GAs to deal with Braes' paradox (more resource $\rightarrow$ worse). [Glance and Hogg, 1995; Arora and Sen, 1996]
- Multiagent RL for adaptive load balancing. [Schaerf *et al.*, 1995]
- Focal points/emergent conventions. [Fenster *et al.*, 1995; Walker and Wooldridge, 1995]
- Locker-room agreements. [Stone and Veloso, 1999]
- Agents filling different roles. [Prasad *et al.*, 1998; Tambe, 1997; Balch, 1998]

Table 6: The issues, techniques, and learning opportunities for heterogeneous MAS as reflected in the literature.

5.3.1 Benevolence vs. competitiveness

One of the most important issues to consider in a multiagent system is whether the different agents are benevolent or competitive. Even if they have different goals, the agents can be benevolent if they are willing to help each other achieve their respective goals [Goldman and Rosenschein, 1994]. On the other hand, the agents may be selfish and only consider their own goals when acting. In the extreme, the agents may be involved in a zero-sum situation so that they must actively oppose other agents' goals in order to achieve their own. This last case describes competitive agents.

Some people only consider using selfish agents, claiming that they are both more effective when building real systems and more biologically plausible. Of course if agents have the same goals, they will help each other, but people rarely consider agents that help each other achieve different goals for no apparent reason: when agents cooperate, they usually do so because it is in their own best interest. As we have already seen in the pursuit domain, Korf [1992] advocates using greedy agents that minimize their own distance to the prey, and similarly, Levy and Rosenschein [1992] use Game Theory to study how the predators can cooperate despite maximizing their own utilities. Some advocates of selfish agents point to nature for their justification, claiming that animals are not altruistic, but rather act always in their own self-interest [Korf, 1992]. On the other hand, Ridley [1997] provides a detailed chronicle and explanation of apparent altruism in nature (usually explainable as kin selection) and cooperation in human societies.

Whether or not altruism exists, in some situations it may be in an animal's (or agent's) interest to cooperate with other agents. Mor and Rosenschein [1995] illustrate this possibility in the context of the prisoner's dilemma, a standard problem in game theory. In the prisoner's dilemma, two agents try to act so as to maximize their own individual rewards. They are not actively out to thwart each other since it is not a zero-sum game, yet they place no inherent value on the other receiving reward. The prisoner's dilemma is constructed so that each agent is given two choices: defect or cooperate. No matter what the other agent does, a given agent receives a greater reward if it defects. Yet if both agents cooperate, they receive a greater reward than if they both defect. In any given play, an agent is better off defecting. Nevertheless, Mor and Rosenschein show that if the same agents come up against each other repeatedly (iterated prisoner's dilemma), cooperative behavior can emerge. In effect, an agent can serve its own self-interest by establishing a *reputation* for being cooperative. Then when coming up against another cooperative agent, the two can benefit from a sense of trust for each other: they both cooperate rather than both defecting. Only with repeated play can cooperation emerge among the selfish agents in the prisoner's dilemma.

In the prisoner's dilemma, the agents are selfish but not inherently competitive: in specific circumstances, they are willing to act benevolently. However, when the agents are actually competitive (such as in zero-sum games), cooperation is no longer sensible. For instance, Littman considers a zero-sum game in which two players try to reach opposite ends of a small discrete world. The players can block each other by trying to move to the same space. Littman [1994] introduces a variant of Q-learning called Minimax-Q which is designed to work on Markov games as opposed to Markov Decision Processes. The competitive agents learn probabilistic policies since any deterministic policy can be completely counteracted by the opponent.

The issue of benevolence (willingness to cooperate) vs. competitiveness comes up repeatedly in the systems described below. Were a third dimension to be added to the categorization of MAS (in addition to degrees of heterogeneity and communication), this issue would be it.

5.3.2 Fixed vs. learning agents

Another important characteristic of multiagent systems is whether the agents' behaviors are fixed or whether they can learn from experience. Learning agents can be desirable in dynamic environments. But particularly when using competitive agents, allowing them to alter their behavior can lead to complications. We refer to such agents as *competitive learning* agents. Systems benevolent learning agents are said to use *cooperative learning*. The evolution of both predator and prey agents by Haynes and Sen [1996] qualifies as competitive learning.

The robotic soccer domain also presents an opportunity for both cooperative and competitive learning. Team-Partitioned, Opaque-Transition Reinforcement Learning (TPOT-RL) [Stone, 2000] implements cooperative learning. Individual soccer-playing agents learn ball-passing policies simultaneously, eventually creating a compatible set of policies. Using a learned feature space that aggressively compacts a very large state space, agents are able to learn to alter their behaviors based on the opponents' perceived strategy.

Grefenstette and Daley [1996] conduct a preliminary study of competitive and cooperative learning in a domain that is loosely related to the pursuit domain. Their domain has two robots that can move continuously and one morsel of (stationary) food that appears randomly in the world. In the cooperative task, both robots must be at the food in order to "capture" it. In a competitive task in the same domain, agents try to be the first to reach the food [Grefenstette and Daley, 1996].

One problem to contend with in competitive rather than cooperative learning is the possibility of an escalating "arms race" with no end. Competing agents might continually adapt to each other in more and more specialized ways, never stabilizing at a good behavior. Of course in a dynamic environment, it may not be feasible or even desirable to learn a stable behavior. Applying RL to the iterated prisoner's dilemma, Sandholm and Crites [1996] find that a learning agent is able to perform optimally against a fixed opponent. But when both agents are learning, there is no stable solution.

Another issue in competitive learning is the credit-assignment problem. When performance of an agent improves, it is not necessarily clear whether the improvement is due to an improvement in that agent's behavior or a negative change in the opponent's behavior. Similarly, if an agent's performance gets worse, the blame or credit could belong to that agent or to the opponent.

One way to deal with the credit-assignment problem is to fix one agent while allowing the other to learn and then to switch. This method encourages the arms race more than ever. Nevertheless, Rosin and Belew [1995] use this technique, along with an interesting method for maintaining diversity in genetic populations, to evolve agents that can play TicTacToe, Nim, and a simple version of Go. When it is a given agent's turn to evolve, it executes a standard genetic algorithm (GA) generation. Individuals are tested against individuals from the competing population, but a technique called "competitive fitness sharing" is used to maintain diversity. When using this technique, individuals from agent X's population are given more credit for beating opponents (individuals from agent Y's population) that are not beaten by other individuals from agent X's population. More specifically, the reward to an individual for beating individual y is divided by the number of other individuals in agent X's population that also beat individual y. Competitive fitness sharing shows much promise for competitive learning.

5.3.3 Modeling of others' goals, actions, and knowledge

For the case of homogeneous agents, it was useful for agents to model the internal states of other agents in order to predict their actions. With heterogeneous agents, the problem of modeling others is much more complex. Now the goals, actions, and domain knowledge of the other agents may also be unknown and thus need modeling.

Without communication, agents are forced to model each other strictly through observation. Huber and Durfee [1995] consider a case of coordinated motion control among multiple mobile robots under the assumption that communication is prohibitively expensive. Thus the agents try to deduce each other's plans by observing their actions. In particular, each robot tries to figure out the destinations of the other robots by watching how they move. Plan recognition of this type is also useful in competitive domains, since knowing an opponent's goals or intentions can make it significantly easier to defeat.

In addition to modeling agents' goals through observation, it is also possible to learn their actions. The OBSERVER system [Wang, 1996] allows an agent to incrementally learn the preconditions and effects of planning actions by observing domain experts. After observing for a time, the agent can then experimentally refine its model by practicing the actions itself.

When modeling other agents, it may be useful to reason not only about what is true and what is false, but also about what is not known. Such reasoning about ignorances is called *autoepistemic reasoning* [Permpoontanalarp, 1995].

Just as RMM is useful for modeling the states of homogeneous agents, it can be used in the heterogeneous scenario as well. Tambe [1995] takes it one step further, studying how agents can *learn* models of teams of agents. In an air combat domain, agents can use RMM to try to deduce an opponents' plan based on its observable actions. For example, a fired missile may not be visible, but the observation of a preparatory maneuver commonly used before firing could indicate that a missile has been launched.

When teams of agents are involved, the situation becomes more complicated. In this case, an opponent's actions may not make sense except in the context of a team maneuver. Then the agent's *role* within the team must be modeled [Tambe, 1996].

One reason that modeling other agents might be useful is that agents sometimes depend on each other for achieving their goals. Unlike in the prisoner's dilemma where agents can cooperate or not depending on their utility estimation, there may be actions that *require* cooperation for successful execution. For example, two robots may be needed to successfully push a box, or, as in the pursuit domain, several agents may be needed to capture an opponent. Sichman and Demazeau [1995] analyze how the case of conflicting mutual models of different co-dependent agents can arise and be dealt with.

5.3.4 Resource management

Heterogeneous agents may have interdependent actions due to limited resources needed by several of the agents. Example domains include network traffic problems in which several different agents must send information through the same network; and load balancing in which several computer processes or users have a limited amount of computing power to share among them.

One interesting network traffic problem called Braess' paradox has been studied from a multiagent perspective using GAs [Glance and Hogg, 1995]. Braess' paradox is the phenomenon of adding more resources to a network but getting worse performance. When using a particular GA representation to represent different parts of a sample network that has usage-dependent resource costs, agents that are sharing the network and reasoning separately about which path of the network to use cannot achieve global optimal performance [Glance and Hogg, 1995]. When the GA representation is improved, the system is able to find the globally optimal traffic flow [Arora and Sen, 1996]. TPOT-RL, mentioned above as having been applied in a robotic soccer domain, has also been applied in a network traffic flow scenario [Stone, 2000].

Adaptive load balancing has been studied as a multiagent problem by allowing different agents to decide which processor to use at a given time. Using RL, heterogeneous agents can achieve reasonable load balance without any central control and without communication among agents [Schaerf *et al.*, 1995]. The agents keep track of how long a job takes when it is scheduled on a given resource, and they are given some incentive to explore untried processors or processors that did poorly in the past.

5.3.5 Social conventions

Although the current multiagent scenario does not allow for communication, there has been some very interesting work done on how heterogeneous agents can nonetheless reach "agreements," or make coinciding choices, if necessary. Humans are able to reach tacit agreements as illustrated by the following scenario:

> Imagine that you and a friend need to meet today. You both arrived in Paris yesterday, but you were unable to get in touch to set a time and place. Nevertheless, it is essential that you meet today. Where will you go, and when?

Vohra posed this question to an audience of roughly 40 people at the AAAI-95 Fall Symposium on Active Learning: roughly 75% of the people wrote down (with no prior communication) that they would go to the Eifel tower at noon. Thus even without communicating, people are sometimes able to coordinate actions. Apparently features that have been seen or used often present themselves as obvious choices.

In the context of MAS, Fenster et al. [1995] define the Focal Point method. They discuss the phenomenon of cultural (or programmed) preferences allowing agents to "meet" without communicating. They propose that, all else being equal, agents who need to meet should choose rare or extreme options.

Rather than coming from pre-analysis of the options as in the focal point method, conventions can emerge over time if agents are biased towards options that have been chosen, for example, most recently or most frequently in the past [Walker and Wooldridge, 1995].

Locker-room agreements [Stone and Veloso, 1999] are another convention-based mechanism for agents to coordinate without communication. Protocols and task decompositions are pre-defined. Agents rely on each other to follow the protocols so that they can anticipate each other's actions and resolve conflicts without communication. Locker-room agreements have been used in the robotic soccer domain, both in simulation [Stone *et al.*, 1999] and on real robots [Veloso *et al.*, 1998b].

5.3.6 Roles

When agents have similar goals, they can be organized into a team. Each agent then plays a separate *role* within the team. With such a benevolent team of agents, one must provide some method for assigning different agents to different roles. This assignment might be obvious if the agents are very specific and can each only do one thing. However in some domains, the agents are flexible enough to interchange roles.

Prasad et al. [1998] study design agents that can either initiate or extend a design of a steam pump. In different situations, different agents are more effective at initiation and at extension. Thus a supervised learning technique is used to help agents learn what roles they should fill in different situations.

STEAM [Tambe, 1997] allows a team of agents to fill and switch roles dynamically. Particularly if a critical agent fails, another agent is able to replace it in its role so that the team can carry on with its mission. Similarly, Stone and Veloso's [1999] flexible teamwork structure allows agents to seamlessly switch roles based upon globally sensed trigger conditions.

If allowed to learn independently, a group of agents might end up filling different roles in the domain or all end up with the same behavior. Balch [1998] investigates methods of encouraging behavioral diversity in a team of agents.

5.4 Further Learning Opportunities

Throughout the above investigation of issues and techniques in the heterogeneous non-communicating multiagent scenario, many learning approaches are described. A few of the other most obvious future ML applications to this scenario are described here and summarized in Table 6.

One challenge when using learning agents is dealing with the credit-assignment problem. When several different agents are learning at the same time, changes in an agent's fitness could be due to its own behavior or due to the behavior of others. Yet if agents are to learn effectively, they must have a reasonable idea of whether a given change in behavior is beneficial or detrimental. Methods of objective fitness measurement are also needed for testing various evolution techniques. In competitive (especially zero-sum) situations, it is difficult to provide adequate performance measurements over time. Even if all agents improve drastically (as measured by testing against common fixed agents), if they all improve the same amount, the actual results when interacting with each other could remain the same. One possible way around this problem is to test agents against past agents in order to measure improvement. However this solution is not ideal: the current agent may have adapted to the current opponent rather than past opponents. A reliable measurement method would be a valuable contribution to ML in MAS.

In cooperative situations, agents ideally learn to behave in such a way that they can help each other. Unfortunately, most existing ML techniques focus on exploring behaviors that are likely to help an agent with its own "personal" deficiencies. An interesting contribution would be a method for introducing into the learning space a bias towards behaviors that are likely to interact favorably with the behaviors of other agents.

Many of the techniques described for heterogeneous non-communicating agents pertained to modeling other agents. However the true end is not just knowledge of another agent's current situation, but rather the ability to predict its future actions. For example, the reason it is useful to deduce another mobile robot's goal location is that its path to the goal may then be predicted and collision avoided. There is still much room for improvement of existing techniques and for new techniques that allow agents to predict each other's future actions.

In the context of teams of agents, it has been mentioned that agents might be suited to different roles in different situations. In a dynamic environment, these flexible agents are more effective if they can switch roles dynamically. For example, if an agent finds itself in a position to easily perform a useful action that is not usually considered a part of its current role, it may switch roles and leave its old role available for another agent. A challenging possible approach to this problem is to enable the agents to learn which roles they should assume in what situations. Dynamic role assumption is a particularly good opportunity for ML researchers in MAS.

6 Homogeneous Communicating Multiagent Systems

To this point, we have not considered MAS in which agents can communicate with each other directly. Admittedly, communication could be viewed as simply part of an agent's interaction with the environment. However just as agents are considered special parts of the environment for the purposes of this survey, so is communication among agents considered extra-environmental. With the aid of communication, agents can coordinate much more effectively than they have been able to up to this point. However, communication also introduces several challenges. In this scenario, we consider homogeneous agents that can communicate with one another.

6.1 Homogeneous Communicating Multiagent Pursuit

In the pursuit domain, communication creates new possibilities for predator behavior. Since the prey acts on its own in the pursuit domain, it has no other agents with which to communicate. However the predators can freely exchange information in order to help them capture the prey more effectively. The current situation is illustrated in Figure 9.

Recall the "local" strategy defined by Stephens and Merx in which each predator simply moved to its closest "capture position" (see Sections 4). In their instantiation of the domain, the predators can see the prey, but not each other. With communication possible, they define another possible strategy for the predators [Stephens and Merx, 1990]. When using a "distributed" strategy, the agents are still homogeneous, but they communicate to insure that each moves toward a different capture position. In particular, the predator farthest from the prey chooses the capture position closest to it, and announces that it will approach that position. Then the next farthest predator chooses the closest capture position from the remaining three, and so

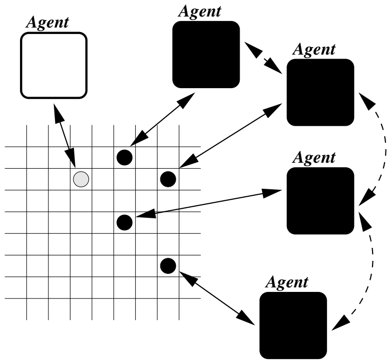

Figure 9: The pursuit domain with homogeneous communicating agents. Now the predators can communicate with one another.

on. This simple protocol encourages the predators to close in on the prey from different sides. A distributed strategy, it is much more effective than the local policy and does not require very much communication. However there are situations in which it does not succeed.

6.2 General Homogeneous Communicating MAS

The multiagent scenario with homogeneous, communicating agents is depicted in Figure 10. As in the homogeneous, non-communicating case (Figure 6), the agents are identical except that they are situated differently in the environment. However in this scenario, the agents can communicate directly as indicated by the arrows connecting the agents in Figure 10. From a practical point of view, the communication might be broadcast or posted on a "blackboard" for all to interpret, or it might be targeted point-to-point from an agent to another specific agent.

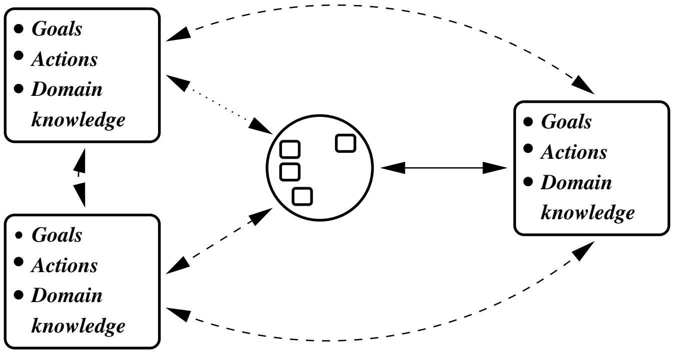

Figure 10: MAS with homogeneous, communicating agents. Only the sensor input and effector output of agents differ. Information can be transmitted directly among agents as indicated by the arrows between agents. Communication can either be broadcast or transmitted point-to-point.

6.3 Issues and Techniques

Communication raises several issues to be addressed in multiagent systems. However, in most cases, the issues are addressed in the literature with heterogeneous, communicating agents. In this section, we consider a limited number of issues which are addressed with homogeneous, communicating agents as indicated in Table 7. Many more communication-related issues are addressed in Section 7, which is devoted to *heterogeneous*, communicating multiagent systems.

Homogeneous, Communicating MAS

Issues

- Distributed sensing
- Communication content

Learning opportunities

- What and when to communicate

Techniques

- Active sensing [Matsuyama, 1997]
- Query propagation for distributed traffic mapping [Moukas and Maes, 1997]
- State vs. goal communication [Balch and Arkin, 1994; Stone and Veloso, 1999]

Table 7: The issues and techniques for homogeneous, communicating multiagent systems as reflected in the literature.

6.3.1 Distributed Sensing

The cooperative distributed vision project [Matsuyama, 1997] aims to construct and monitor a broad visual scene for dynamic three dimensional scene understanding by using multiple cameras, either stationary or on mobile robots. For example, consider the problem of tracking an individual car using cameras mounted at urban intersections. When the car leaves one camera's range and enters another's, there needs to be a way of identifying the two images as representing the same car, even though it probably looks different in the two cases (i.e. it is driving away from one camera and towards the other). The project combines active sensing—the ability to shift attention towards an area of higher uncertainty or interest—and communication among multiple sensing agents.

Another distributed sensing project is the trafficopter system [Moukas and Maes, 1997]. In trafficopter, cars themselves collect and propagate traffic information to help each other decide on the best route to a given location. For example, a car driving in one direction might query an oncoming vehicle about traffic conditions up the road. By propagating such queries among vehicles, the original car can build a map of traffic conditions along different routes to its goal.

6.3.2 Communication Content

One important issue for communicating agents is *what* they should communicate. In the distributed sensing applications mentioned above, agents communicated with each other regarding their sensed states of the world. However, it is also possible for agents to share information regarding their individual goals.

In three multi-robot applications, Balch and Arkin [1994] study the effects of allowing agents to communicate their states and their goals with one another. They found that agents that communicated goal information performed slightly better than those that communicated state information. Both conditions exhibited far superior behavior when compared with non-communicating agents.

It has also been observed that allowing agents to communicate internal state information can be very effective in the robotic soccer domain [Stone and Veloso, 1999]. In this application, opponent and ball locations are communicated among agents that can therefore increase their knowledge via communication.

6.4 Further Learning Opportunities

While it has been demonstrated that communicating state information can be advantageous in MAS, in many domains, bandwidth considerations do not allow for constant, complete exchange of such information. In addition, if communications are delayed, as opposed to being instantaneous, they may become obsolete before arriving at their intended destinations. In such cases, it may be possible for agents to learn what and when to communicate with other agents based on observed affects of utterances on group performance.

7 Heterogeneous Communicating Multiagent Systems

The scenario examined in Section 5 included agents that differ in any number of ways, including their sensory data, their goals, their actions, and their domain knowledge. Such heterogeneous multiagent systems can be very complex and powerful. However the full power of MAS can be realized when adding the ability for heterogeneous agents to communicate with one another. In fact, combining communication and heterogeneity introduces the possibility of having a multiagent system operate similarly to a single-agent system in terms of control. By sending their sensor inputs to and receiving their commands from one agent, all the other agents can surrender control to that single agent. In this case, control is no longer distributed.

7.1 Heterogeneous Communicating Multiagent Pursuit

Allowing for both heterogeneity and communication in the pursuit domain opens up new control possibilities. The current situation is illustrated in Figure 11.

Tan [1993] uses communicating agents in the pursuit domain to conduct some interesting multiagent Q-learning experiments. In his instantiation of the domain, there are several prey agents and the predators have limited vision so that they may not always know where the prey are. Thus the predators can help each other by informing each other of their sensory input. Tan shows that they might also help each other by exchanging reinforcement episodes and/or control policies.

Stephens and Merx [1990] present one more strategy that always succeeds but requires much more communication than the distributed approach presented in Section 6.1: the “central” strategy. The central

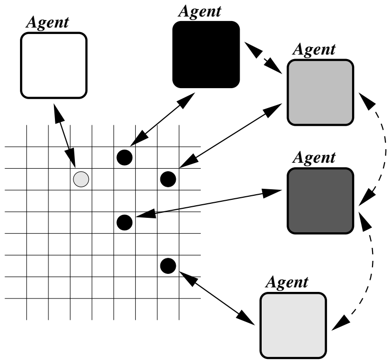

Figure 11: The pursuit domain with heterogeneous communicating agents. Agents can be fully heterogeneous as well as being able to communicate with one another.

strategy is effectively a single control agent with distributed actuators. Three predators transmit all of their sensory inputs to one central agent which then decides where all the predators should move and transmits its decision back to them. In this case, there is really only one intelligent controlling agent and three executors.

Benda et al. [1986], in the original presentation of the pursuit domain, also consider the full range of communication possibilities. They consider the possible organizations of the four predators when any pair can either exchange data, exchange data and goals, or have one make the decisions of the other. The tradeoff between lower communication costs and better decisions is described. Communication costs might come in the form of limited bandwidth or consumption of reasoning time.

Another way to frame this tradeoff is as one between cost and freedom: as communication cost (time) increases, freedom decreases. Osawa suggests that the predators should move through four phases. In increasing order of cost (decreasing freedom), they are: autonomy, communication, negotiation, and control [Osawa, 1995]. When the predators stop making sufficient progress toward the prey using one strategy, they should move to the next most expensive strategy. Thus they can close in on the prey efficiently and effectively.

7.2 General Heterogeneous Communicating MAS

The fully general multiagent scenario appears in Figure 12. In this scenario, we allow the agents to be heterogeneous to any degree from homogeneity to full heterogeneity.

7.3 Issues and Techniques

Since heterogeneous communicating agents can choose not to communicate, and in some cases can also choose to be homogeneous or at least to minimize their heterogeneity, most of the issues discussed in the previous three scenarios apply in this one as well. Two of the most studied issues are communication protocols and theories of commitment. Already discussed in the context of the heterogeneous, non-communicating MAS scenario, the issue of benevolence vs. competitiveness becomes more complicated in the current con-

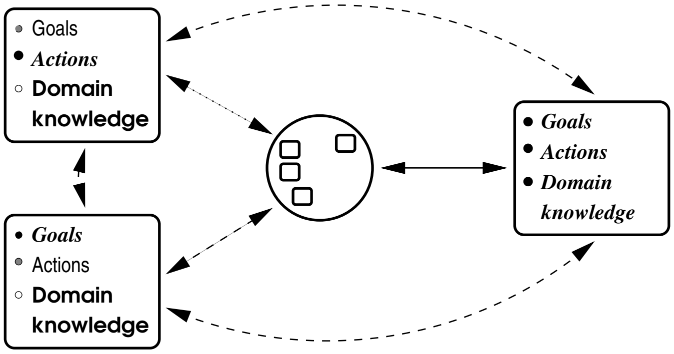

Figure 12: The general communicating MAS scenario. Agents can be heterogeneous to any degree. Information can be transmitted directly among agents as indicated by the arrows between agents. Communication can either be broadcast or transmitted point-to-point.

text. These issues and others along with some of the existing techniques to deal with them are described below and summarized in Table 8.

7.3.1 Understanding each other

In all communicating multiagent systems, and particularly in domains with agents built by different designers, there must be some set language and protocol for the agents to use when interacting. Independent aspects of protocols are information content, message format, and coordination conventions. Among others, existing protocols for these three levels are: KIF for content [Genesereth and Fikes, 1992], KQML for message format [Finin *et al.*, 1994], and COOL for coordination [Barbuceanu and Fox, 1995]. There has been a lot of research done on refining these and other communication protocols.

One challenge that arises in using symbolic communication among agents is making sure that the symbols are grounded similarly in the internal representations of the different agents. In an approach related to the social conventions discussed in Section 5, it is possible to use shared past experiences to ground a symbolic representation. This technique has been used in a heterogeneous multi-robot vacuum cleaning task [Jung and Zelinsky, 2000].

One of the first industrial multiagent systems, ARCHON [Jennings and Wittig, 1992] successfully integrated several legacy systems. Applied in five different industrial settings, ARCHON successfully allows independently developed, heterogeneous computer systems to communicate in order to create collaborative, process control systems.

Creatures [Grand and Cliff, 1998] is a multiagent computer game based on sophisticated biological models, such as genetic codes, hormones, and energy systems. Agents have the ability to grow and learn, including a simple verb-object language, based on interactions with a human user or other agents in the environment.

Heterogeneous Communicating

Issues

- Understanding each other
- Planning communicative acts
- Benevolence vs. competitiveness
- Negotiation
- Resource management (schedule coordination)
- Commitment/decommitment
- Collaborative localization
- Changing shape and size

Learning opportunities

- Evolving language
- Effects of speech acts on global dynamics
- Communication utility and truthfulness
- Commitment utility

Techniques

- Language protocols: KIF [Genesereth and Fikes, 1992], KQML [Finin *et al.*, 1994], COOL. [Barbuceanu and Fox, 1995]
- Grounding meaning via shared experience. [Jung and Zelinsky, 2000]
- Legacy systems integration. [Jennings and Wittig, 1992]
- Language learning. [Grand and Cliff, 1998]
- Speech acts. [Cohen and Levesque, 1995; Lux and Steiner, 1995]
- Learning social behaviors. [Mataric, 1994b]
- Reasoning about truthfulness. [Rosenschein and Zlotkin, 1994; Sandholm and Lesser, 1996]
- Multiagent Q-learning. [Tan, 1993; Weiß, 1995]
- Training other agents' Q-functions (track driving). [Clouse, 1996]
- Minimize the need for training. [Potter *et al.*, 1995]
- Cooperative co-evolution. [Bull *et al.*, 1995]
- Contract nets for electronic commerce. [Sandholm and Lesser, 1995b]
- Market-based systems. [Huberman and Clearwater, 1995]
- Bayesian learning in negotiation: model others. [Zeng and Sycara, 1998]
- Market-based methods for distributed constraints. [Parunak *et al.*, 1998]
- Generalized partial global planning (GPGP). [Decker and Lesser, 1995; Lesser, 1998]
- Learning to choose among coordination methods. [Sugawara and Lesser, 1995]
- Query response in information networks. [Sycara *et al.*, 1996]
- Division of independent tasks. [Parker, 1994, 2000]
- Internal, social, and collective (role) commitments. [Castelfranchi, 1995]
- Commitment states (potential, pre, and actual) as planning states. [Haddadi, 1995]
- Belief/desire/intention (BDI) model: OASIS. [Rao and Georgeff, 1995]
- BDI commitments only over intentions. [Rao and Georgeff, 1995]
- Coalitions. [Zlotkin and Rosenschein, 1994; Shehory and Kraus, 1995; Sandholm and Lesser, 1995a]
- Fusing uncertain sensor data. [Fox *et al.*, 2000; Grabowski *et al.*, 2000]
- Inter-component communication in metamorphic robots. [no *et al.*, 2000]

Table 8: The issues, techniques, and learning opportunities for heterogeneous, communicating multiagent systems as reflected in the literature.

7.3.2 Planning communicative acts

When an agent transmits information to another agent, it has an effect just like any other action would have. Thus within a planning framework, one can define preconditions and effects for communicative acts. When combined with a model of other agents, the effect of a communication act might be to alter an agent's belief about the state of another agent or agents. The theory of communication as action is called *speech* *acts* [Cohen and Levesque, 1995; Lux and Steiner, 1995].

Mataric adds a learning dimension to the idea of speech acts. Starting with the foraging behavior mentioned above [Mataric, 1994a], the agents can then learn to choose from among a set of social behaviors that includes broadcasting and listening [Mataric, 1994b]. Q-learning is extended so that reinforcement can be received for direct rewards obtained by the agent itself or for rewards obtained by other agents.

When using communication as a planning action, the possibility arises of communicating misinformation in order to satisfy a particular goal. For instance, an agent may want another agent to believe that something is true. Rather than actually making it true, the agent might just *say* that it is true. For example, Sandholm and Lesser [1996] analyze a framework in which agents are allowed to "decommit" from agreements with other agents by paying a penalty to these other agents. They consider the case in which an agent might not be truthful in its decommitment, hoping that the other agent will decommit first. In such situations, agents must also consider what communications to believe [Rosenschein and Zlotkin, 1994].

7.3.3 Benevolence vs. competitiveness

Some studies involving competitive agents were described in the heterogeneous non-communicating scenario (see Section 5). In the current scenario, there are many more examples of competitive agents.

Similar to Tan's work on multiagent RL in the pursuit domain [Tan, 1993] is Weiß's work with competing Q-learners. The agents compete with each other to earn the right to control a single system [Weiß, 1995]. The highest bidder pays a certain amount to be allowed to act, then receives any reward that results from the action.

Another Q-learning approach, this time with benevolent agents, has been to explore the interesting idea of having one agent teach another agent through communication [Clouse, 1996]. Starting with a *trainer* that has moderate expertise in a task, a learner can be rewarded for mimicking the trainer. Furthermore, the trainer can recommend to the learner what action to take in a given situation so as to direct the learner towards a reward state. Eventually, the learner is able to perform the task without any guidance.

While training is a useful concept, some research is driven by the goal of reducing the role of the human trainer. As opposed to the process of *shaping*, in which the system designer develops simple behaviors and slowly builds them into more complex ones, populations appropriately seeded for competitive co-evolution can reduce the amount of designer effort. Potter and Grefenstette [1995] illustrate this effect in their domain described above in which two robots compete for a stationary pellet of food. Subpopulations of rules used by GAs are seeded to be more effective in different situations. Thus specialized subpopulations of rules corresponding to shaped behaviors tend to emerge.

GAs have also been used to evolve separate communicating agents to control different legs of a quadrapedal robot using cooperative co-evolution [Bull *et al.*, 1995].

7.3.4 Negotiation

Drawing inspiration from competition in human societies, several researchers have designed negotiating multiagent systems based on the law of supply and demand. In the *contract nets* framework [Smith, 1980], agents all have their own goals, are self-interested, and have limited reasoning resources. They bid to accept tasks from other agents and then can either perform the tasks (if they have the proper resources) or subcontract them to other agents. Agents must pay to contract their tasks out and thus shop around for the lowest bidder. Many multiagent issues arise when using contract nets [Sandholm and Lesser, 1995b].

In a similar spirit is an implemented multiagent system that controls air temperature in different rooms of a building [Huberman and Clearwater, 1995]. A person can set one's thermostat to any temperature. Then depending on the actual air temperature, the agent for that room tries to "buy" either hot or cold air from another room that has an excess. At the same time, the agent can sell the excess air at the current temperature to other rooms. Modeling the loss of heat in the transfer from one room to another, the agents try to buy and sell at the best possible prices. The market regulates itself to provide equitable usage of a shared resource.

Zeng and Sycara [1998] study a competitive negotiation scenario in which agents use Bayesian Learning techniques to update models of each other based on bids and counter bids in a negotiation process.

The MarCon system [Parunak *et al.*, 1998] uses market-based methods for distributed constraint problems. Designers at different points along a supply chain negotiate the characteristics of the overall design by buying and selling characteristics and propagating the resulting constraints.

7.3.5 Resource management

MarCon is an example of multiagent resource management: the design characteristics desired by one agent may consume the resources of another.

Similarly, generalized partial global planning (GPGP) allows several heterogeneous agents to post constraints, or commitments to do a task by some time, to each other's local schedulers and thus coordinate without the aid of any centralized agent [Decker and Lesser, 1995]. A proposed general multiagent architecture based on GPGP contains five components: "local agent scheduling, multiagent coordination, organizational design, detection, and diagnosis [Lesser, 1998]."

In a heterogeneous, communicating multiagent system applied to diagnosis of a local area network, agents learn to choose among different coordination strategies based on the current situation [Sugawara and Lesser, 1993, 1995]. Less sophisticated coordination methods require fewer network and time resources, but may lead to tasks failing to be executed or to redundant actions by multiple agents.

RETSINA [Sycara *et al.*, 1996] uses three classes of heterogeneous, communicating agents to deliver information in response to specific user queries in information networks. RETSINA is able to satisfy the information requests of multiple users by searching multiple information sources, while considering network constraints and resource limitations of information agents. RETSINA has been used to implement several distributed network applications including a financial portfolio manager, a personal information manager and meeting scheduler, and a satellite visibility forecaster.

ALLIANCE and its learning variant L-ALLIANCE [Parker, 1994, 2000] use communication among heterogeneous robots to help divide independent tasks among the robots. With an emphasis on fault tolerance, agents only broadcast the task that they are currently working on. If the communication fails, multiple robots might temporarily try to do the same task, but they will eventually realize the conflict by observation and one will move on to a different task. In L-ALLIANCE, robots learn to evaluate each other's abilities with respect to specific tasks in order to more efficiently divide their tasks among the team.

7.3.6 Commitment/decommitment

When agents communicate, they may decide to cooperate on a given task or for a given amount of time. In so doing, they make *commitments* to each other. Committing to another agent involves agreeing to pursue a given goal, possibly in a given manner, regardless of how much it serves one's own interests. Commitments can make systems run much more smoothly by providing a way for agents to "trust" each other, yet it is not obvious how to get self-interested agents to commit to others in a reasonable way. The theory of commitment and decommitment (when the commitment terminates) has consequently drawn considerable attention.

Castelfranchi [1995] defines three types of commitment: internal commitment—an agent binds itself to do something; social commitment—an agent commits to another agent; and collective commitment—an agent agrees to fill a certain role. Setting an alarm clock is an example of internal commitment to wake up at a certain time.

Commitment states have been used as planning states: potential cooperation, pre-commitment, and commitment [Haddadi, 1995]. Agents can then use means-ends analysis to plan for goals in terms of commitment opportunities. This work is conducted within a model called belief/desire/intention, or BDI.

BDI is a popular technique for modeling other agents. Other agents' domain knowledge (beliefs) and goals (desires) are modeled as well as their "intentions," or goals they are currently trying to achieve and the methods by which they are trying to achieve them. The BDI model is used to build a system for air-traffic control, OASIS [Rao and Georgeff, 1995], which has been implemented for testing (in parallel with human operators who retain full control) at the airport in Sydney, Australia. Each aircraft is represented by a controlling agent which deals with a global sequencing agent. OASIS mixes reactive and deliberative actions in the agents: they can break out of planned sequences when coming across situations that demand immediate reaction. Since agents cannot control their beliefs or desires, they can only make commitments to each other regarding their intentions.

Finally, groups of agents may decide to commit to each other. Rather than the more usual two-agent or all-agent commitment scenarios, there are certain situations in which agents may want to form coalitions [Zlotkin and Rosenschein, 1994]. Since this work is conducted in a game theory framework, agents consider the utility of joining a coalition in which they are bound to try to advance the utility of other members in exchange for reciprocal consideration. Shehory and Kraus [1995] present a distributed algorithm for task allocation when coalitions are either needed to perform tasks or more efficient that single agents. Sandholm and Lesser [1995a] use a vehicle routing domain to illustrate a method by which agents can form valuable coalitions when it is intractable to discover the optimal coalitions.

7.3.7 Collaborative localization

Localization is a common challenge for autonomous robots. For most robotic tasks, a robot must know where it is situated in the world before it can act effectively. One common approach to localization is Markov localization, in which a robot maintains a probabilistic belief over its current position based on its observations and a map of the environment. Recent work has extended this approach to multiple robots [Fox *et al.*, 2000]. When a robot R_1 detects another robot R_2, it can use R_2's current belief about R_2's position along with the detected relative position of R_2 to increase the data available for R_1's own effort to localize. This approach was successfully implemented both on homogeneous robots and on robots with different sensors.

Millibots [Grabowski *et al.*, 2000] are the smallest-scale components of another heterogeneous, communicating, multi-robot system that is able to perform collaborative localization and mapping. Each millibot (a robot with dimensions roughly $6cm^3$) is specialized with a subset of sensors that can collect data from the environment. In order to maintain localization, three millibots from the group stay still so that they can be used as reference points for the other robots. Periodically, one of the three can move to a new location (or be replaced by another robot) so that the group as a whole can move. Meanwhile, the sensing robots broadcast their sensory data to a larger robot, which acts as a team leader. The team leader can then fuse the data from the exploring robots and send back tasks for them to accomplish.

7.3.8 Changing shape and size

CONRO, a "deployable robot with inter-robot metamorphic capabilities," [no *et al.*, 2000] is a particularly ambitious project involving heterogeneous, communicating robots. The goal is to create a robot that can change shape and size by reconfiguring its components, splitting into parts, or joining back together again. While the project has thus far focussed on the considerable challenge of creating the necessary hardware components, Castaño et al. discuss the need for wireless inter-component communication to support docking and remote sensing.

7.4 Further Learning Opportunities

Once again, there are many possible ways in the current scenario to enhance MAS with ML techniques. Within this heterogeneous communicating multiagent scenario there is a clear need to pre-define a language and communication protocol for use by the agents. However, an interesting alternative would be to allow the agents to learn for themselves what to communicate and how to interpret it. For example, an agent might be given a small language of utterances and a small set of meanings, but no mapping between the two. Agents would then have to learn both what to say and how to interpret what they hear. A possible result would be more efficient communications: they would need to be understandable only by the agents rather than by both agents and humans.

When considering communications as speech acts, agents could be allowed to learn the effects of speech on the global dynamics of the system. In domains with low bandwidth or large time delays associated with communication, the utility of communicating at a given moment might be learned. In addition, if allowed to learn to communicate, agents are more likely to avoid being reliably conned by untruthfulness in communication: when another agent says something that turns out not to be true, it will not be believed so readily in the future.

Finally, commitment—the act of taking on another agent's goals—has both benefits and disadvantages. System builders may want to allow their agents to learn when to commit to others. The learning opportunities in this scenario are summarized in Table 8.

8 Robotic Soccer

Several multiagent domains have been mentioned throughout the course of this survey, including design, planning, entertainment, games, air-traffic control, air combat, personal assistants, load-balancing, network routing, and robotic leg control. In this section a single domain which embodies most multiagent issues is presented.

Robotic soccer is a particularly good domain for studying MAS. Originated by Mackworth [1993], it has been gaining popularity in recent years, with several international competitions taking place [Kitano, 1996; Kim, 1996; Kitano, 1998; Asada and Kitano, 1999; Veloso *et al.*, 2000]. It is also the subject of an official IJCAI-97 Challenge [Kitano *et al.*, 1997b]. It can be used to evaluate different MAS techniques in a direct manner: teams implemented with different techniques can play against each other.

Although the pursuit domain serves us well for purposes of illustration, robotic soccer is much more complex and interesting as a general test bed for MAS. Even with many predators and several prey, the pursuit domain is not complex enough to simulate the real world. Although robotic soccer is a game, most real-world complexities are retained. A key aspect of soccer's complexity is the need for agents not only to control themselves, but also to control the ball which is a passive part of the environment.

8.1 Overview

Robotic soccer can be played either with real robots or in a simulator. The first robotic soccer system was the Dynamo system [Barman *et al.*, 1993]. Barman et al. built a 1 vs. 1 version of the game.

Some robotic issues can only be studied in the real-world instantiation, but there are also many issues that can be studied in simulation. A particularly good simulator for this purpose is the "soccer server" developed by Noda [1998] and pictured in Figure 13. This simulator is realistic in many ways: the players'

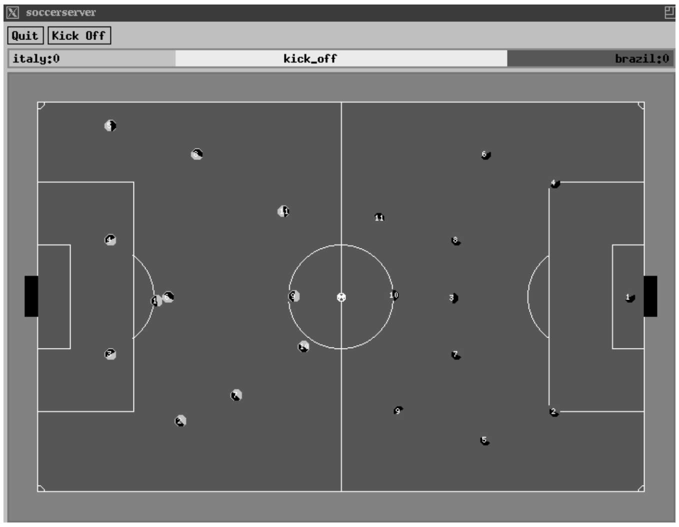

Figure 13: The soccer server system

vision is limited; the players can communicate by posting to a blackboard that is visible to all players; all players are controlled by separate processes; each team has 11 players; each player has limited stamina; actions and sensors are noisy; and play occurs in real time. The simulator provides a domain and supports users who wish to build their own agents. Furthermore, teams of agents can be evaluated by playing against each other, or perhaps against standard teams. The simulator was first used for a competition among twenty-nine teams from around the world in 1997 [Kitano *et al.*, 1997a] and continues to be used for this purpose currently. Thus robotic soccer satisfies Decker's criteria for DAI test beds [Decker, 1996a]. The advantages of robotic soccer as a test bed for MAS are summarized in Table 9.

Table 9: Advantages of (simulated) robotic soccer as a MAS test bed

- Complex enough to be realistic
- Easily accessible
- Embodies most MAS issues
- Direct comparisons possible
- Good multiagent ML opportunities

8.2 MAS in Robotic Soccer

The main goal of any test bed is to facilitate the trial and evaluation of ideas that have promise in the real world. A wide variety of MAS issues can be studied in simulated robotic soccer. In fact, most of the MAS issues listed in Table 2 can be feasibly studied in the soccer simulator.

Homogeneous, non-communicating MAS can be studied in robotic soccer by fixing the behavior of the opposing team and populating the team being studied with identical, mute players. To keep within the homogeneous agent scenario, the opponents must not be modeled as agents.

- In this context, the players can be *reactive or deliberative* to any degree. The extremely reactive agent might simply look for the ball and move straight at it, shooting whenever possible. At this extreme, the players may or may not have any knowledge that they are part of a team.
- On the other hand, players might *model each other*, thus enabling deliberative reasoning about whether to approach the ball or whether to move to a different part of the field in order to defend or to receive a pass.
- With players modeling each other, they may also reason about how to *affect each other's behaviors* in this inherently dynamic environment.
- It is possible to study the relative merits of *local and global perspectives* on the world. Robots can be given global views with the help of an overhead camera, and the soccer server comes equipped with an omniscient mode that permits global views. Simulated robotic soccer is usually approached as a problem requiring local sensing.

Heterogeneous, non-communicating MAS can also be studied in the robotic soccer domain.

- Since each player has several teammates with the same global goal and several opponents with the opposite goal, each player is both *benevolent and competitive* at the same time. This possibility for combination of collaborative and adversarial reasoning is a major feature of the domain.
- If the teams are learning during the course of a single game or over several games, all the issues of *learning agents*, including the "arms race" and the credit-assignment problem, arise.
- In the soccer server, stamina is a resource assigned to each individual agent. At the team level, stamina is important for *resource management*: if too many agents are tired, the team as a whole will be ineffective. Therefore, it is to the team's advantage to distribute the running among the different agents.
- When trying to collaborate, players' actions are usually interdependent: to execute a successful pass, both the passer and the receiver must execute the appropriate actions. Thus *modeling each other* for the purpose of coordination is helpful. In addition, if opponents' actions can be predicted, then proactive measures might be taken to render them ineffective.
- *Social conventions*, such as programmed notions of when a given agent will pass or which agents should play defense, can also help coordination. The locker-room agreement is an example of social conventions within a team.
- Since communication is still not allowed, the players must have a reliable method for filling the different team *roles* needed on a soccer team (e.g. defender, forward, goaltender) which can be achieved via a flexible teamwork structure.

Homogeneous, communicating MAS can be studied by again fixing the behavior of the opposing team and allowing teammates to communicate.

- *Distributed sensing* can be studied in this context due to the large amount of hidden state inherent in the soccer server. At any given moment, a particular agent can see only a small portion of the world. Only by communicating with teammates can it get a more complete picture of the world.
- It is particularly important in the soccer server to choose *communication content* carefully. Since players have only a limited hearing frequency, a frivolous utterance can cause subsequent more important information to go unheard.

Heterogeneous, communicating MAS is perhaps the most appropriate scenario to study within the context of robotic soccer. Since the agents indeed are heterogeneous and can communicate, the full potential of the domain is realized in this scenario.

- With players sending messages to each other, they must have a language in order to *understand each other*.
- Especially in the single-channel, low-bandwidth communication environment modeled by the soccer server, agents must *plan their communicative acts*. If the opponents can understand the same language, a planned utterance can affect the knowledge of both teammates and opponents. The utility of communication must be carefully considered and the possibility of lying in order to fool the opponent arises. In addition, the low-bandwidth creates the condition that sending a message may prevent other messages from getting through.
- Like in the heterogeneous, non-communicating scenario, since agents have both teammates and adversaries, they must reason about being both *benevolent and competitive*.
- *Negotiation* protocols may be useful in the robotic soccer domain if different agents, based on their different sensory perspectives, have different opinions about what course of action would be best for the team.
- In a real-time environment, timing is very important for any team play, including a simple pass. Thus, *resource management* in terms of timing, or action coordination, is crucial.
- Protocols are also needed for *commitment* to team plays: the passer and receiver in a pass play must both agree to execute the pass. For more complex team plays, such as our set-plays, several players may need to commit to participate. But then the issue arises of how single-mindedly they must adhere to the committed play: when may they react to more pressing situations and ignore the commitment?
- When an agent is unsure of its position in the environment, it can take cues from other agents, via either observation or communication, thus exhibiting *collaborative localization*.

In terms of the reasons to use MAS presented in Table 1, robotic soccer systems usually require separate agents for controlling the separate players, and they can benefit from the parallelism, robustness, and simpler programming of MAS. Systems whose players have onboard sensors are necessarily multiagent, since no single agent has access to all of the players' sensory inputs. Some competitions also stipulate in their rules that the robots must be controlled by separate agents. At the very least, the two teams must be controlled by separate agents. Even teams that could theoretically be controlled by a single agent stand to gain by using MAS. By processing the sensory inputs of the different players separately, multiple agents can control their players in parallel, perhaps contending with different tasks on the field. One player might be in position to defend its goal, while another is preparing an offensive attack. These players need not be controlled by the same agent: they can go about their tasks in parallel. Furthermore, if any of the agents fails for some reason (as often happens in real robotic systems), the other agents can attempt to compensate and continue playing. Finally, it is empirically easier to program a single agent per player than it is to control an entire team centrally.

As demonstrated above, most of the MAS issues summarized in Table 2 can be studied in robotic soccer. We now review the research that has been conducted in this domain. First, we describe research conducted in the "early years", before organized robotic soccer workshops, that served as the foundations for the recent popularity of the domain. Second, we review some of the research presented at dedicated robotic soccer workshops held in conjunction with the international competitions, as well as other contemporary robotic soccer-related research.

8.3 Foundations

Producing natural language commentary from real-time input, the SOCCER system [Andre *et al.*, 1988] was the first AI research related to soccer. SOCCER analyzed *human* soccer games. By looking for triggers and terminations of events such as a player running or the ball being passed, SOCCER aims to announce important events without redundancy.

Robotic soccer was introduced as an interesting and promising domain for AI research at the Vision Interface conference in June, 1992 [Mackworth, 1993]. The first working robotic soccer systems [Barman *et al.*, 1993; Sahota *et al.*, 1995] were also described at that time. A ground-breaking system for robotic soccer, and the one that served as the inspiration and basis for the authors' own research in the robotic soccer domain, the Dynamite test bed was designed to be capable of supporting several robots per team, but most work has been done in a 1 vs. 1 scenario. It uses an overhead camera and color-based detection to provide global sensory information to the robots. Dynamite was used to introduce a decision making strategy called reactive deliberation which was used to choose from among seven hard-wired behaviors [Sahota, 1994]. Subsequently, an RL approach based on high-level sensory predicates was used to choose from among the same hard-wired behaviors [Ford *et al.*, 1994].

Asada et al. [1994a] developed the first robots equipped with on-board sensing capabilities. These robots use learning from easy missions, an RL training technique, to learn to hit a stationary ball into the goal. One contribution of this work is the construction of state and action spaces that reduce the complexity of the learning task [Asada *et al.*, 1996]. As opposed to the action-dependent features used by TPOT-RL which create an abstract feature space prior to learning, states are clustered during learning based on the best action to take from each state. Another contribution is the combination of low-level behaviors, such as shooting and avoiding an opponent, that are learned using RL [Asada *et al.*, 1994b; Uchibe *et al.*, 1996]. Rather than building the learned behaviors at different behavior levels as in layered learning, two previously learned control strategies are used to produce a new one, which then replaces the original two.

Minimax-Q learning for Markov games was first applied in an abstract simulated soccer game [Littman, 1994]. This version of the domain is much simpler than the soccer server, having 800 states, 5 actions, and no hidden information. One player on each team moves in a grid world and the ball is always possessed by one of the players. Using minimax-Q, players learn optimal probabilistic policies for maneuvering past each other with the ball.

The authors conducted machine learning experiments in a simulator based closely on the Dynasim simulator [Sahota, 1996] which simulates the Dynamite robots mentioned above. First, we used memory-based learning to allow a player to learn when to shoot and when to pass the ball [Stone and Veloso, 1996a]. We then used neural networks to teach a player to shoot a moving ball into particular parts of the goal [Stone and Veloso, 1998]. Based on training in a small region of the field, our agent was able to learn to successfully time its approach to a moving ball such that it could score from all areas of the field. These experiments served as the basis for our initial learning experiments in the soccer server [Stone and Veloso, 1996b].

In another early learning experiment in the soccer server, a player learned when to shoot and when to pass [Matsubara *et al.*, 1996]. The agent bases its decision on the positions of the ball, the goaltender, and one teammate.

8.4 The Competition Years

The research reported in Section 8.3 confirmed the potential of robotic soccer as an AI research domain and justified the value of having large-scale competitions from a research perspective. Starting with the first competitions held in 1996 (Pre-RoboCup-96 and MiroSot-96) and continuing since then, there has been a great deal of robotic soccer-related research. It has been presented both at dedicated robotic soccer workshops held in conjunction with the competitions and in other scientific forums. In this subsection we review some of this recent robotic soccer research.

8.4.1 Robot Hardware

Much of the research inspired by competitions has been devoted to building robot hardware that is suitable for this challenging environment, e.g. [Achim *et al.*, 1996; Hong *et al.*, 1996; Hsia and Soderstrand, 1996; Kim *et al.*, 1996; Shim *et al.*, 1996; Veloso *et al.*, 1998a]. The emphasis in hardware approaches varies greatly. Some research focuses on fast and robust visual perception of the environment [Sargent *et al.*, 1997; Cheng and Zelinsky, 1998; Han and Veloso, 1998]. And some research focuses on automatic calibration of vision parameters [Shen *et al.*, 1998; Veloso and Uther, 1999] in response to the need for vision systems that work under various lighting conditions (conditions at competitions are never the same as in the lab). Instead of vision, one alternative approach is to use a laser range-finder for localization in the environment [Gutmann *et al.*, 1998].

Other research focuses on robot path planning in crowded, dynamic, environments [Han *et al.*, 1996; Kim and Chung, 1996; Bowling and Veloso, 1999]. Path planning is particularly challenging with non-holonomic robots because they can only move straight in the direction that they are facing or in curved paths starting from their current location and direction. Omnidirectional robots can simplify path planning considerably: they do not have to consider the direction they are facing as a constraint [Price *et al.*, 1998; Yokota *et al.*, 1998].

In addition to robots developed specifically for the competitions, there have been robots created to exhibit special soccer-related skills. Shoobot [Mizuno *et al.*, 1996, 1998] is a nomad-based robot that can dribble and shoot a soccer ball as it moves smoothly through an open space. The Sony legged robots [Fujita and Kageyama, 1997] walk on four legs. They have been used as the basis of an exclusively legged-robot soccer competition [Veloso *et al.*, 1998c]. And the Honda humanoid robots [Hirai, 1997] have been demonstrated kicking a real soccer ball and performing a penalty shot with a shooting and a goaltending robot. This demonstration indicates the feasibility of RoboCup's long-term goal of having a humanoid robot soccer competition on a real soccer field [Kitano *et al.*, 1998].

8.4.2 Soccer Server Accessories

In addition to soccer-playing agent development, the soccer server has been used as a substrate for 3-dimensional visualization, real-time natural language commentary, and education research.

Figure 13 shows the 2-dimensional visualization tool that is included in the soccer server software. SPACE [Shinjoh, 1998] converts the 2-dimensional image into a 3-dimensional image, changing camera angle and rendering images in real time.

Another research challenge being addressed within the soccer server is producing natural language commentary of games as they proceed. Researchers aim to provide both low-level descriptions of the action, for example announcing which team is in possession of the ball, and high-level analysis of the play, for example commenting on the team strategies being used by the different teams. Commentator systems for the soccer server include ROCCO [Andre *et al.*, 1998], MIKE [Matsubara *et al.*, 1999], and Byrne [Binsted, 1999].

8.4.3 Multiagent Control and Robotic Soccer Strategy

The robotic soccer domain has inspired many different approaches to building and organizing teams of agents.

Some research is based on applying existing programming methodologies to the robotic soccer domain. Team GAMMA [Noda, 1998] is built using Gaea [Nakashima *et al.*, 1995], a logic programming language that is essentially a multi-threaded, multi-environment version of prolog. Gaea implements a dynamic subsumption architecture, allowing agents to override behaviors in different ways based on the current environment, or behavior context. Team ROGI [de la Rosa *et al.*, 1997] is built using another programming methodology, namely agent-oriented programming [Shoham, 1990].

Other research, introduces new multiagent control methodologies and applies them to robotic soccer. For example, the MICROB robotic soccer team is an implementation of the Cassiopeia programming method [Drogoul and Collinot, 1998]. Cassiopeia focuses on the organizational issues of multiagent tasks, analyzing the interdependencies of low-level skills and facilitating the formation of groups based on these inter-dependencies. Temporary organizations are formed based on the contract net framework [Smith, 1980]. For example, the player with the ball might contract with another player to place itself in a particular location to receive a pass. This approach differs from that of the CMUnited-98 small-robot team [Veloso *et al.*, 1999] which uses strategic positioning using attraction and repulsion (SPAR). There, the agents position themselves autonomously, and the agent with the ball decides autonomously where to pass: no negotiation is involved, enabling the players to act as quickly as possible.

Scerri [1998] presents another multi-layered approach to robotic soccer. However, unlike our own hierarchical approach, it does not involve the learning of any behaviors. In this approach, the different abstraction layers deal with different granularities of sensory input. For example, a low-level move-to-ball behavior is given the ball's precise location, while a high-level defend behavior—which might call go-to-ball—knows only that the ball is in the defensive half of the field. The Samba control architecture [Riekki and Roening, 1998] uses two behavior layers: the reactive layer which defines action maps from sensory input to actuator output; and the task layer which selects from among the action maps.

ISIS [Tambe *et al.*, 1998] is a role-based approach to robotic soccer based on STEAM [Tambe, 1997]. STEAM defines team behaviors that can be invoked dynamically. There has also been another formation-based approach to positioning agents on the soccer field [Matsumoto and Nagai, 1998]. However, unlike in our dynamic formations with flexible positions, the player positions are static and the team formation cannot change dynamically. Several other researchers recognize the importance of decomposing the soccer task into different roles, e.g. [Coradeschi and Karlsson, 1998; Ch'ng and Padgham, 1998].

One approach with dynamically changing roles is developed in a soccer simulator other than the soccer server [Balch, 1998]. Balch uses his behavioral diversity measure to encourage role learning in an RL framework, finding that providing a uniform reinforcement to the entire team is more effective than providing local reinforcements to individual players.

Often, definitions of robotic soccer positions involve fixed locations at which an agent should locate itself by default, e.g. [Gutmann *et al.*, 1998; Matsumoto and Nagai, 1998]. In contrast, the within a locker-room agreement as described above, flexible positions allow players to adjust their locations within their roles [Stone and Veloso, 1999]. The ranges of flexibility are defined a priori as a part of the locker-room agreement. Observational reinforcement learning [Andou, 1998] allows agents to *learn* their positions dynamically based on the distribution of past ball locations in a game. A similar approach is also described in [Inoue and Wilkin, 1997].

In another learning approach, teammate and opponent capabilities are learned through repeated trials of specific actions [Nadella and Sen, 1997]. This research is conducted in a soccer simulator in which the ball is always in possession of a player, eliminating the necessity for fine ball control. Each player has an assigned efficiency in the range $[0, 1]$ for the execution of actions such as passing, tackling, and dribbling corresponding to the probability that the action will succeed. Agents do not know the abilities of themselves, their teammates, or the opponents. Instead, they learn to estimate them based on repeated trials. The agents can then base action decisions on the learned parameters.

Layered learning [Stone, 2000] has been implemented in the simulated robotic soccer domain. Layered learning is a general-purpose machine learning paradigm for complex domains in which learning a mapping directly from agents' sensors to their actuators is intractable. Given a hierarchical task decomposition, layered learning allows for learning at each level of the hierarchy, with learning at each level directly affecting learning at the next higher level. TPOT-RL [Stone, 2000] (mentioned above) is used for one of learned layers in a layered learning implementation.

All of the learning approaches described above are used to learn portions of an agent's behavior. Other aspects are created manually. In contrast, a few entirely learned soccer behaviors have been created.

Hexcer [Uther and Veloso, 1997] is an extension of the grid world soccer game described above [Littman, 1994]. Rather than square grid locations, the world is defined as a lattice of hexagons. Thus the action space is increased and the geometric constraints are altered. The added complexity necessitates the development of generalized U-trees to allow agents to learn successful policies [Uther and Veloso, 1997]. In Hexcer, it is possible for agents to learn straight from sensors to actuators because, like Littman's simulation, Hexcer has a much smaller state space than the soccer server and the agents have no hidden state.

The RoboCup-97 and RoboCup-98 competitions each included one team created using genetic programming [Koza, 1992]. In both cases, the goal was to learn entirely from agent sensors to actuators in the soccer server. The first attempt [Luke *et al.*, 1998] was eventually scaled down, although a successful team was created based on some manually created low-level skills. The following year, Darwin United [Andre and Teller, 1999] entered an entirely learned team.

9 Conclusion

This survey is presented as a description of the field of MAS. It is designed to serve both as an introduction for people unfamiliar with the field and as an organizational framework for system designers. This framework is presented as a series of four increasingly complex and powerful scenarios. The simplest systems are those with homogeneous non-communicating agents. The second scenario involves heterogeneous non-communicating agents. The third deals with homogeneous, communicating agents. Finally, the general

MAS scenario involves communicating agents with any degree of heterogeneity.

Each multiagent scenario introduces new issues and complications. In the MAS literature, several techniques and systems already address these issues. After summarizing a wide range of such existing work, useful future directions are presented. Throughout the survey, Machine Learning approaches are emphasized.

Although each domain requires a different approach, from a research perspective the ideal domain embodies as many issues as possible. Robotic soccer is presented here as a useful domain for the study of MAS. Systems with a wide variety of agent heterogeneity and communication abilities can be studied. In addition, collaborative and adversarial issues can be combined in a real-time situation. With the aid of research in such complex domains, the field of MAS should continue to advance and to spread in popularity among designers of real systems.

MAS is an active field with many open issues. Continuing research is presented at dedicated conferences and workshops such as the International Conference on Multiagent Systems [Weiß and Sen, 1996; Sen, 1996; AAA, 1995]. MAS work also appears in many of the DAI conferences and workshops [Distributed, 1990; Weiß, 1996]. This survey provides a framework within which the reader can situate both existing and future work.

Acknowledgements

We would like to thank Keith Decker, Astro Teller, and the anonymous reviewers for their helpful comments and suggestions.

References

AAAI. *Proceedings of the First International Conference on Multi-Agent Systems*, Menlo Park,CA, June 1995. AAAI Press. Victor Lessor–General Chair.

Sorin Achim, Peter Stone, and Manuela Veloso. Building a dedicated robotic soccer system. In *Proceedings of the IROS-96 Workshop on RoboCup*, pages 41–48, Osaka, Japan, November 1996.

Tomohito Andou. Refinement of soccer agents' positions using reinforcement learning. In Hiroaki Kitano, editor, *RoboCup-97: Robot Soccer World Cup I*, pages 373–388. Springer Verlag, Berlin, 1998.

David Andre and Astro Teller. Evolving team Darwin United. In Minoru Asada and Hiroaki Kitano, editors, *RoboCup-98: Robot Soccer World Cup II*. Springer Verlag, Berlin, 1999.

Elisabeth Andre, Gerd Herzog, and Thomas Rist. On the simultaneous interpretation of real world image sequences and their natural language description: The system soccer. In *Proc. of the 8th ECAI*, pages 449–454, Munich, 1988.

Elisabeth Andre, Gerd Herzog, and Thomas Rist. Generating multimedia presentations for RoboCup soccer games. In Hiroaki Kitano, editor, *RoboCup-97: Robot Soccer World Cup I*, pages 200–215. Springer Verlag, Berlin, 1998.

Neeraj Arora and Sandip Sen. Resolving social dilemmas using genetic algorithms. In *Adaptation, Coevolution and Learning in Multiagent Systems: Papers from the 1996 AAAI Spring Symposium*, pages 1–5, Menlo Park,CA, March 1996. AAAI Press. AAAI Technical Report SS-96-01.

Minoru Asada and Hiroaki Kitano, editors. *RoboCup-98: Robot Soccer World Cup II*. Lecture Notes in Artificial Intelligence 1604. Springer Verlag, Berlin, 1999.

Minoru Asada, Shoichi Noda, Sukoya Tawaratsumida, and Koh Hosoda. Purposive behavior acquisition on a real robot by vision-based reinforcement learning. In *Proc. of MLC-COLT (Machine Learning Confernce and Computer Learning Theory) Workshop on Robot Learning*, pages 1–9, 1994.

Minoru Asada, Eiji Uchibe, Shoichi Noda, Sukoya Tawaratsumida, and Koh Hosoda. Coordination of multiple behaviors acquired by vision-based reinforcement learning. In *Proc. of IEEE/RSJ/GI International Conference on Intelligent Robots and Systems 1994*, pages 917–924, 1994.

Minoru Asada, Shoichi Noda, and Koh Hosoda. Action-based sensor space categorization for robot learning. In *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems 1996 (IROS '96)*, pages 1502–1509, 1996.

Tucker Balch and Ron C. Arkin. Communication in reactive multiagent robotic systems. *Autonomous Robots*, 1(1):1–25, 1994.

Tucker Balch and Ronald C. Arkin. Motor schema-based formation control for multiagent robot teams. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 10–16, Menlo Park, California, June 1995. AAAI Press.

Tucker Balch. *Behavioral Diversity in Learning Robot Teams*. PhD thesis, College of Computing, Georgia Institute of Technology, 1998.

Tucker Balch. Social entropy: An information theoretic measure of robot team diversity. *Autonomous Robots*, 2000. In this issue.

Mihai Barbuceanu and Mark S. Fox. COOL: A language for describing coordination in multi agent systems. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 17–24, Menlo Park, California, June 1995. AAAI Press.

R. A. Barman, S. J. Kingdon, J. J. Little, A. K. Mackworth, D. K. Pai, M. Sahota, H. Wilkinson, and Y. Zhang. DYNAMO: Real-time experiments with multiple mobile robots. In *Intelligent Vehicles Symposium*, pages 261–266, Tokyo, July 1993.

R. Beckers, E. Holland, and J. L. Deneubourg. From local actions to global tasks: Stigmergy and collective robotics. In R. Brooks and P. Maes, editors, *Artifical Life IV: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*, pages 181–189. MIT Press, 1994.

M. Benda, V. Jagannathan, and R. Dodhiawala. On optimal cooperation of knowledge sources - an empirical investigation. Technical Report BCS–G2010–28, Boeing Advanced Technology Center, Boeing Computing Services, Seattle, Washington, July 1986.

Kim Binsted. Character design for soccer commentary. In Minoru Asada and Hiroaki Kitano, editors, *RoboCup-98: Robot Soccer World Cup II*. Springer Verlag, Berlin, 1999.

Alan H. Bond and Les Gasser. An analysis of problems and research in DAI. In Alan H. Bond and Les Gasser, editors, *Readings in Distributed Artificial Intelligence*, pages 3–35. Morgan Kaufmann Publishers, San Mateo, CA, 1988.

Michael Bowling and Manuela Veloso. Motion control in dynamic multi-robot environments. In *Proceedings of The 1999 IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA'99)*, monterrey, november 1999.

L. Bull, T.C. Fogarty, and M. Snaith. Evolution in multi-agent systems: Evolving communicating classifier systems for gait in a quadrupedal robot. In Stephanie Forrest, editor, *Proceedings of the Sixth International Conference on Genetic Algorithms*, pages 382–388, San Mateo,CA, July 1995. Morgan Kaufman.

Y. Uny Cao, Alex S. Fukunaga, and Andrew B. Kahng. Cooperative mobile robotics: Antecedents and directions. *Autonomous Robots*, 4:7–27, 1997.

Cristiano Castelfranchi. Commitments: From individual intentions to groups and organizations. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 41–48, Menlo Park, California, June 1995. AAAI Press.

Gordon Cheng and Alexander Zelinsky. Real-time vision processing for a soccer playing mobile robot. In Hiroaki Kitano, editor, *RoboCup-97: Robot Soccer World Cup I*, pages 144–155. Springer Verlag, Berlin, 1998.

Simon Ch'ng and Lin Padgham. From roles to teamwork: a framework and architecture. *Applied Artificial Intelligence*, 12:211–232, 1998.

Jeffery A. Clouse. Learning from an automated training agent. In Gerhard Weiß and Sandip Sen, editors, *Adaptation and Learning in Multiagent Systems*. Springer Verlag, Berlin, 1996.

Philip R. Cohen and Hector J. Levesque. Communicative actions for artificial agents. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 65–72, Menlo Park, California, June 1995. AAAI Press.

Silvia Coradeschi and Lars Karlsson. A role-based decision-mechanism for teams of reactive and coordinating agents. In Hiroaki Kitano, editor, *RoboCup-97: Robot Soccer World Cup I*, pages 112–122. Springer Verlag, Berlin, 1998.

Kerstin Dautenhahn. Getting to know each other—artificial social intelligence for autonomous robots. *Robotics and Autonomous Systems*, 16:333–356, 1995.

J. Ll. de la Rosa, A. Oller, J. Vehi, and J. Puyol. Soccer team based on agent-oriented programming. *Robotics and Autonomous Systems*, 21(2):161–176, October 1997.

Keith S. Decker and Victor R. Lesser. Designing a family of coordination algorithms. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 73–80, Menlo Park, California, June 1995. AAAI Press.

Keith S. Decker. Distributed problem solving: A survey. *IEEE Transactions on Systems, Man, and Cybernetics*, 17(5):729–740, September 1987.

Keith S. Decker. *Environment Centered Analysis and Design of Coordination Mechanisms*. PhD thesis, University of Massachusetts, 1995.

Keith S. Decker. Distributed artificial intelligence testbeds. In G. M. P. O'Hare and N. R. Jennings, editors, *Foundations of Distributed Artificial Intelligence*, pages 119–138. Wiley Interscience, 1996.

Keith S. Decker. Personal correspondence, May 1996.

Keith S. Decker. Task environment centered simulation. In M. Prietula, K. Carley, and L. Gasser, editors, *Simulating Organizations: Computational Models of Institutions and Groups*. AAAI Press/MIT Press, 1996.

Proceedings of the 10th International Workshop on Distributed Artificial Intelligence, Bandera, Texas, October 1990.

Alexis Drogoul and Anne Collinot. Applying an agent-oriented methodology to the design of artificial organizations: A case study in robotic soccer. *Autonomous Agents and Multi-Agent Systems*, 1(1):113–129, 1998.

Gregory Dudek, Michael R. M. Jenkin, Evangelos Milios, and David Wilkes. A taxonomy for multi-agent robotics. *Autonomous Robots*, 3(4):375–397, 1996.

Edmund H. Durfee, Victor R. Lesser, and Daniel D. Corkill. Trends in cooperative distributed problem solving. *IEEE Transactions on Knowledge and Data Engineering*, 1(1):63–83, March 1989.

Edmund H. Durfee. What your computer really needs to know, you learned in kindergarten. In *Proceedings of the Tenth National Conference on Artificial Intelligence*, Philadelphia, PA, 1992. Morgan Kaufman.

Edmund H. Durfee. Blissful ignorance: Knowing just enough to coordinate well. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 406–413, Menlo Park, California, June 1995. AAAI Press.

Maier Fenster, Sarit Kraus, and Jeffrey S. Rosenschein. Coordination without communication: Experimental validation of focal point techniques. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 102–108, Menlo Park, California, June 1995. AAAI Press.

Innes A. Ferguson and Grigoris J. Karakoulas. Multiagent learning and adaptation in an information filtering market. In *Adaptation, Coevolution and Learning in Multiagent Systems: Papers from the 1996 AAAI Spring Symposium*, pages 28–32, Menlo Park,CA, March 1996. AAAI Press. AAAI Technical Report SS-96-01.

Tim Finin, Don McKay, Rich Fritzson, and Robin McEntire. KQML: An information and knowledge exchange protocol. In Kazuhiro Fuchi and Toshio Yokoi, editors, *Knowledge Building and Knowledge Sharing*. Ohmsha and IOS Press, 1994.

Roger Ford, Craig Boutilier, and Keiji Kanazawa. Exploiting natural structure in reinforcement learning: Experience in robot soccer-playing, 1994. Unpublished Manuscript.

Dieter Fox, Wolfram Burgard, Hannes Kruppa, and Sebastian Thrun. A probabilistic approach to collaborative multi-robot localization. *Autonomous Robots*, 2000. In this issue.

Masahiro Fujita and Koji Kageyama. An open architecture for robot entertainment. In *Proceedings of the First International Conference on Autonomous Agents*, pages 435–442, Marina del Rey, CA, February 1997.

M. R. Genesereth and R. E. Fikes. Knowledge interchange format, version 3.0 reference manual. Technical Report Logic-92-1, Computer Science Department, Stanford University, 1992.

Natalie S. Glance and Tad Hogg. Dilemmas in computational societies. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 117–124, Menlo Park, California, June 1995. AAAI Press.

Claudia Goldman and Jeffrey Rosenschein. Emergent coordination through the use of cooperative state-changing rules. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, pages 408–413, Philadelphia, PA, 1994. Morgan Kaufman.

Robert Grabowski, Luis E. Navarro-Serment, Christiaan J. J. Paredis, and Pradeep K. Khosla. Heterogeneous teams of modular robots for mapping and exploration. *Autonomous Robots*, 2000. In this issue.

Stephen Grand and Dave Cliff. Creatures: Entertainment software agents with artificial life. *Autonomous Agents and Multi-Agent Systems*, 1(1):39–58, 1998.

John Grefenstette and Robert Daley. Methods for competitive and cooperative co-evolution. In *Adaptation, Coevolution and Learning in Multiagent Systems: Papers from the 1996 AAAI Spring Symposium*, pages 45–50, Menlo Park,CA, March 1996. AAAI Press. AAAI Technical Report SS-96-01.

J-S. Gutmann, W. Hattzack, I. Herrmann, B. Nebel, F. Rittinger, A. Topor, T. Weigel, and B. Welsch. The CS Freiburg team. In Minoru Asada, editor, *Proceedings of the second RoboCup Workshop*, pages 451–458, Paris, France, July 1998.

Afsaneh Haddadi. Towards a pragmatic theory of interactions. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 133–139, Menlo Park, California, June 1995. AAAI Press.

Kwun Han and Manuela Veloso. Reactive visual control of multiple non-holonomic robotic agents. In *Proceedings of the International Conference on Robotics and Automation*, Leuven,Belgium, May 1998.

Wong-Gie Han, Seung-Min Baek, Tae-Yong Kuc, and Sung Kyun Kwan. Path planning of visual-served multiple mobile robots using the genetic algorithms. In *Proceedings of the Micro-Robot World Cup Soccer Tournament*, pages 57–63, Taejon, Korea, November 1996. IEEE Robotics and Automation Society.

Barbara Hayes-Roth, Lee Brownston, and Robert van Gent. Multiagent collaboration in directed improvisation. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 148–154, Menlo Park, California, June 1995. AAAI Press.

Thomas Haynes and Sandip Sen. Evolving behavioral strategies in predators and prey. In Gerhard Weiß and Sandip Sen, editors, *Adaptation and Learning in Multiagent Systems*, pages 113–126. Springer Verlag, Berlin, 1996.

Thomas Haynes and Sandip Sen. Learning cases to resolve conflicts and improve group behavior. volume 48, pages 31–49, January 1998.

Thomas Haynes, Roger Wainwright, Sandip Sen, and Dale Schoenefeld. Strongly typed genetic programming in evolving cooperation strategies. In Stephanie Forrest, editor, *Proceedings of the Sixth International Conference on Genetic Algorithms*, pages 271–278, San Mateo, CA, July 1995. Morgan Kaufman.

Kazuo Hirai. Development of the Honda humanoid robot. Presentation at the CMU robotics institute seminar, November 1997. At URL http://www.honda.co.jp/home/hpr/e_news/robot/.

O.E. Holland. Multiagent systems: Lessons from social insects and collective robotics. In *Adaptation, Coevolution and Learning in Multiagent Systems: Papers from the 1996 AAAI Spring Symposium*, pages 57–62, Menlo Park,CA, March 1996. AAAI Press. AAAI Technical Report SS-96-01.

Sun-Gi Hong, Tae-Dok Eom, Il-Kwon Jeong, Changkyu Choi, Jin-Ho Shin, and Ju-Jang Lee. Development of soccer-playing robots: Control, cooperation, and strategy. In *Proceedings of the Micro-Robot World Cup Soccer Tournament*, pages 134–141, Taejon, Korea, November 1996. IEEE Robotics and Automation Society.

Tien C. Hsia and Michael Soderstrand. Development of a micro robot system for playing soccer games. In *Proceedings of the Micro-Robot World Cup Soccer Tournament*, pages 149–152, Taejon, Korea, November 1996. IEEE Robotics and Automation Society.

Marcus J. Huber and Edmund H. Durfee. Deciding when to commit to action during observation-based coordination. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 163–170, Menlo Park, California, June 1995. AAAI Press.

Bernardo Huberman and Scott H. Clearwater. A multi-agent system for controlling building environments. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 171–176, Menlo Park, California, June 1995. AAAI Press.

Naomichi Inoue and Samuel Wilkin. Learning system for the formation and the pass-play. Presented at the poster session of the IROS-96 workshop on RoboCup, November 1997.

N. R. Jennings and T. Wittig. Archon: Theory and practice. In N. M. Avouris and L. Gasser, editors, *Distributed Artificial Intelligence: Theory and Praxis*, pages 179–195. Kluwer Academic Press, 1992.

David Jung and Alexander Zelinsky. Grounded symbolic communication between heterogeneous cooperating robots. *Autonomous Robots*, 2000. In this issue.

Leslie Pack Kaelbling, Michael L. Littman, and Andrew W. Moore. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285, May 1996.

Do-Yoon Kim and Myung Jin Chung. Path planning for multi-mobile robots in the dynamic environment. In *Proceedings of the Micro-Robot World Cup Soccer Tournament*, pages 127–132, Taejon, Korea, November 1996. IEEE Robotics and Automation Society.

Kyun-Hoon Kim, Kuk-Won Ko, Joo-Gon Kim, Su-Ho Lee, and Hyung Suck Cho. Multiple micro robots playing robot soccer game. In *Proceedings of the Micro-Robot World Cup Soccer Tournament*, pages 38–43, Taejon, Korea, November 1996. IEEE Robotics and Automation Society.

Jong-Hwan Kim, editor. *Proceedings of the Micro-Robot World Cup Soccer Tournament*, Taejon, Korea, November 1996.

Hiroaki Kitano, Yasuo Kuniyoshi, Itsuki Noda, Minoru Asada, Hitoshi Matsubara, and Eiichi Osawa. RoboCup: A challenge problem for AI. *AI Magazine*, 18(1):73–85, Spring 1997.

Hiroaki Kitano, Milind Tambe, Peter Stone, Manuela Veloso, Silvia Coradeschi, Eiichi Osawa, Hitoshi Matsubara, Itsuki Noda, and Minoru Asada. The RoboCup synthetic agent challenge 97. In *Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence*, pages 24–29, San Francisco, CA, 1997. Morgan Kaufmann.

Hiroaki Kitano, Minoru Asada, Itsuki Noda, and Hitoshi Matsubara. RoboCup: Robot world cup. *Crossroads*, 4.3, February 1998.

Hiroaki Kitano, editor. *Proceedings of the IROS-96 Workshop on RoboCup*, Osaka, Japan, November 1996.

Hiroaki Kitano, editor. *RoboCup-97: Robot Soccer World Cup I*. Springer Verlag, Berlin, 1998.

Richard E. Korf. A simple solution to pursuit games. In *Working Papers of the 11th International Workshop on Distributed Artificial Intelligence*, pages 183–194, February 1992.

John R. Koza. *Genetic Programming*. MIT Press, 1992.

Victor R. Lesser. Multiagent systems: An emerging subdiscipline of AI. *ACM Computing Surveys*, 27(3):340–342, September 1995.

Victor Lesser. Reflections on the nature of multi-agent coordination and its implications for an agent architecture. *Autonomous Agents and Multi-Agent Systems*, 1(1):89–112, 1998.

Ran Levy and Jeffrey S. Rosenschein. A game theoretic approach to the pursuit problem. In *Working Papers of the 11th International Workshop on Distributed Artificial Intelligence*, pages 195–213, February 1992.

Michael L. Littman. Markov games as a framework for multi-agent reinforcement learning. In *Proceedings of the Eleventh International Conference on Machine Learning*, pages 157–163, San Mateo, CA, 1994. Morgan Kaufman.

Sean Luke, Charles Hohn, Jonathan Farris, Gary Jackson, and James Hendler. Co-evolving soccer softbot team coordination with genetic programming. In Hiroaki Kitano, editor, *RoboCup-97: Robot Soccer World Cup I*, pages 398–411, Berlin, 1998. Springer Verlag.

Andreas Lux and Donald Steiner. Understanding cooperation: an agent's perspective. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 261–268, Menlo Park, California, June 1995. AAAI Press.

A. K. Mackworth. On seeing robots. In A. Basu and X. Li, editors, *Computer Vision: Systems, Theory, and Applications*, pages 1–13. World Scientific Press, Singapore, 1993.

Maja J. Mataric. Interaction and intelligent behavior. MIT EECS PhD Thesis AITR–1495, MIT AI Lab, August 1994.

Maja J. Mataric. Learning to behave socially. In *Third International Conference on Simulation of Adaptive Behavior*, 1994.

Hitoshi Matsubara, Itsuki Noda, and Kazuo Hiraki. Learning of cooperative actions in multi-agent systems: a case study of pass play in soccer. In *Adaptation, Coevolution and Learning in Multiagent Systems: Papers from the 1996 AAAI Spring Symposium*, pages 63–67, Menlo Park,CA, March 1996. AAAI Press. AAAI Technical Report SS-96-01.

Hitoshi Matsubara, Ian Frank, Kumiko Tanaka-Ishii, Itsuki Noda, Hideyuki Nakashima, and Koiti Hasida. Automatic soccer commentary and RoboCup. In Minoru Asada and Hiroaki Kitano, editors, *RoboCup-98: Robot Soccer World Cup II*. Springer Verlag, Berlin, 1999.

Akihiro Matsumoto and Hiromasa Nagai. Decision making by the characteristics and the interaction in multi-agent robotics soccer. In Hiroaki Kitano, editor, *RoboCup-97: Robot Soccer World Cup I*, pages 132–143. Springer Verlag, Berlin, 1998.

Takashi Matsuyama. Cooperative distributed vision. In *Proceedings of the First International workshop on Cooperative Distributed Vision*, Kyoto, Japan, October 1997.

Marvin L. Minsky. *The Society of Mind*. Simon & Schuster, 1988.

Hironori Mizuno, Masakatu Kourogi, and Yoichi Muraoka. Building shoobot possible to dribble and shoot. In *Proceedings of the IROS-96 Workshop on RoboCup*, November 1996.

Hironori Mizuno, Masakatu Kourogi, Yukihide Kawamoto, and Yoichi Muraoka. A method applied for soccer's behaviors using proper feedback and feedforward control. In Hiroaki Kitano, editor, *RoboCup-97: Robot Soccer World Cup I*, pages 156–167. Springer Verlag, Berlin, 1998.

Yishay Mor and Jeffrey Rosenschein. Time and the prisoner's dilemma. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 276–282, Menlo Park, California, June 1995. AAAI Press.

Alexandros Moukas and Pattie Maes. Trafficopter: A distributed collection system for traffic information. At URL http://trafficopter.www.media.mit.edu/projects/trafficopter/, 1997.

Rajani Nadella and Sandip Sen. Correlating internal parameters and external performance: learning soccer agents. In Gerhard Weiß, editor, *Distributed Artificial Intelligence Meets Machine Learning*, pages 137–150. Springer-Verlag, 1997.

Hideyuki Nakashima, Itsuki Noda, and Ichiro Ohsawa. Organic programming for multi-agents. In Victor Lesser, editor, *Proceedings of the First International Conference on Multi-Agent Systems*, page 459, San Francisco, CA, 1995. MIT Press.

Andrés Casta no, Wei-Min Shen, and Peter Will. CONRO: Towards deployable robots with inter-robot metamorphic capabilities. *Autonomous Robots*, 2000. In this issue.

Itsuki Noda, Hitoshi Matsubara, Kazuo Hiraki, and Ian Frank. Soccer server: A tool for research on multi-agent systems. *Applied Artificial Intelligence*, 12:233–250, 1998.

Itsuki Noda. Team GAMMA: Agent programming on gaea. In Hiroaki Kitano, editor, *RoboCup-97: Robot Soccer World Cup I*, pages 500–507. Springer Verlag, Berlin, 1998.

Ei-Ichi Osawa. A metalevel coordination strategy for reactive cooperative planning. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 297–303, Menlo Park, California, June 1995. AAAI Press.

Lynne E. Parker. *Heterogeneous Multi-Robot Cooperation*. PhD thesis, Massachusetts Institute of Technology, 1994.

Lynne E. Parker. Life-long adaptation in heterogeneous multi-robot teams: response to continual variation in individual robot performance. *Autonomous Robots*, 2000. In this issue.

H. Van Dyke Parunak, Allen Ward, and John Sauter. A systematic market approach to distributed constraint problems. In *Proceedings of the Third International Conference on Multi-Agent Systems*, pages 455–456, 1998.

H. Van Dyke Parunak. Applications of distributed artificial intelligence in industry. In G. M. P. O'Hare and N. R. Jennings, editors, *Foundations of Distributed Artificial Intelligence*, pages 139–164. Wiley Interscience, 1996.

Yongyuth Permpoontanalarp. Generalised proof-theory for multi-agent autoepistemic reasoning. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 304–311, Menlo Park, California, June 1995. AAAI Press.

Dean A. Pormerleau. *Neural Network Perception for Mobile Robot Guidance*. Kluwer Academic Publishers, 1993.

Mitchell A. Potter, Kenneth A. De Jong, and John J. Grefenstette. A coevolutionary approach to learning sequential decision rules. In Stephanie Forrest, editor, *Proceedings of the Sixth International Conference on Genetic Algorithms*, pages 366–372, San Mateo,CA, July 1995. Morgan Kaufman.

M. V. Nagendra Prasad, Susan E. Lander, and Victor R. Lesser. Learning organizational roles for negotiated search in a multiagent system. volume 48, pages 51–67, January 1998.

Andrew Price, Andrew Jennings, and John Kneen. RoboCup97: An omnidirectional perspective. In Hiroaki Kitano, editor, *RoboCup-97: Robot Soccer World Cup I*, pages 320–332. Springer Verlag, Berlin, 1998.

Anand S. Rao and Michael P. Georgeff. BDI agents: From theory to practice. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 312–319, Menlo Park, California, June 1995. AAAI Press.

Matt Ridley. *The Origins of Virtue : Human Instincts and the Evolution of Cooperation*. Viking Press, April 1997.

Jukka Riekki and Juha Roening. Playing soccer by modifying and combining primitive reactions. In Hiroaki Kitano, editor, *RoboCup-97: Robot Soccer World Cup I*, pages 74–87. Springer Verlag, Berlin, 1998.

Jeffrey S Rosenschein and Gilad Zlotkin. *Rules of Encounter*. MIT Press, 1994.

Christopher D. Rosin and Richard K. Belew. Methods for competitive co-evolution: Finding opponents worth beating. In Stephanie Forrest, editor, *Proceedings of the Sixth International Conference on Genetic Algorithms*, pages 373–380, San Mateo,CA, July 1995. Morgan Kaufman.

Stuart J. Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, Englewood Cliffs, NJ, 1995.

Michael K. Sahota, Alan K. Mackworth, Rod A. Barman, and Stewart J. Kingdon. Real-time control of soccer-playing robots using off-board vision: the Dynamite testbed. In *IEEE International Conference on Systems, Man, and Cybernetics*, pages 3690–3663, 1995.

Michael K. Sahota. Reactive deliberation: An architecture for real-time intelligent control in dynamic environments. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, pages 1303–1308, 1994.

Michael K. Sahota. Dynasim user guide. At URL http://www.cs.ubc.ca/nest/lci/soccer, January 1996.

J. Alfredo Sanchez, Flavio S. Azevedo, and John J. Leggett. Paragente: Exploring the issues in agent-based user interfaces. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 320–327, Menlo Park, California, June 1995. AAAI Press.

Tuomas W. Sandholm and Robert H. Crites. On multiagent Q-learning in a semi-competitive domain. In Gerhard Weiß and Sandip Sen, editors, *Adaptation and Learning in Multiagent Systems*. Springer Verlag, Berlin, 1996.

Tuomas Sandholm and Victor Lesser. Coalition formation among bounded rational agents. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, pages 662–669, Los Angeles, CA, 1995. Morgan Kaufman.

Tuomas Sandholm and Victor Lesser. Issues in automated negotiation and electronic commerce: Extending the contract net framework. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 328–335, Menlo Park, California, June 1995. AAAI Press.

Tuomas Sandholm and Victor Lesser. Advantages of a leveled commitment contracting protocol. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, pages 126–133, Menlo Park,California, 1996. AAAI Press.

Randy Sargent, Bill Bailey, Carl Witty, and Anne Wright. Dynamic object capture using fast vision tracking. *AI Magazine*, 18(1):65–72, Spring 1997.

Paul Scerri. A multi-layered behavior based system for controlling RoboCup agents. In Hiroaki Kitano, editor, *RoboCup-97: Robot Soccer World Cup I*, pages 467–474. Springer Verlag, Berlin, 1998.

Andrea Schaerf, Yoav Shoham, and Moshe Tennenholtz. Adaptive load balancing: A study in multi-agent learning. *Journal of Artificial Intelligence Research*, 2:475–500, 1995.

Jurgen Schmidhuber. A general method for multi-agent reinforcement learning in unrestricted environments. In *Adaptation, Coevolution and Learning in Multiagent Systems: Papers from the 1996 AAAI Spring Symposium*, pages 84–87, Menlo Park,CA, March 1996. AAAI Press. AAAI Technical Report SS-96-01.

Sandip Sen, Neeraj Arora, and Shounak Roychowdhury. Using limited information to enhance group stability. volume 48, pages 69–82, January 1998.

Sandip Sen, editor. *Adaptation, Coevolution and Learning in Multiagent Systems: Papers from the 1996 AAAI Spring Symposium*, Menlo Park,CA, March 1996. AAAI, AAAI Press. AAAI Technical Report SS-96-01.

Onn Shehory and Sarit Kraus. Task allocation via coalition formation among autonomous agents. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, pages 655–661, Los Angeles, CA, 1995. Morgan Kaufman.

Wei-Min Shen, Jafar Adibi, Rogelio Adobbati, Bonghan Cho, Ali Erdem, Hadi Moradi, Behnam Salemi, and Sheila Tejada. Building integrated mobile robots for soccer competition. In *Proceedings of the International Conference on Robotics and Automation*, 1998.

Hyun-Sik Shim, Myung-Jin Jung, Heung-Soo Kim, In-Hwan Choi, Woong-Suk Han, and Jong-Hwan Kim. Designing distributed control architecture for cooperative multiagent systems. In *Proceedings of the Micro-Robot World Cup Soccer Tournament*, pages 19–25, Taejon, Korea, November 1996. IEEE Robotics and Automation Society.

Atsuchi Shinjoh. RoboCup-3D: The construction of intelligent navigation system. In Hiroaki Kitano, editor, *RoboCup-97: Robot Soccer World Cup I*, pages 188–199. Springer Verlag, Berlin, 1998.

Yoav Shoham. Agent-oriented programming. Technical Report CS-1335-90, Computer Science Dept., Stanford University, 1990.

Jaime Simao Sichman and Yves Demazeau. Exploiting social reasoning to deal with agency level inconistency. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 352–359, Menlo Park, California, June 1995. AAAI Press.

Reid G. Smith. The contract net protocol: High-level communication and control in a distributed problem solver. *IEEE Transactions on Computers*, C-29(12):1104–1113, December 1980.

Larry M. Stephens and Matthias B. Merx. The effect of agent control strategy on the performance of a dai pursuit problem. In *Proceedings of the 10th International Workshop on Distributed Artificial Intelligence*, Bandera, Texas, October 1990.

Peter Stone and Manuela Veloso. Beating a defender in robotic soccer: Memory-based learning of a continuous function. In David S. Touretzky, Michael C. Mozer, and Michael E. Hasselmo, editors, *Advances in Neural Information Processing Systems 8*, pages 896–902, Cambridge, MA, 1996. MIT Press.

Peter Stone and Manuela Veloso. Using machine learning in the soccer server. In *Proceedings of the IROS-96 Workshop on RoboCup*, pages 19–27, Osaka, Japan, November 1996.

Peter Stone and Manuela Veloso. Towards collaborative and adversarial learning: A case study in robotic soccer. *International Journal of Human-Computer Studies*, 48(1):83–104, January 1998.

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

Peter Stone, Manuela Veloso, and Patrick Riley. The CMUnited-98 champion simulator team. In Minoru Asada and Hiroaki Kitano, editors, *RoboCup-98: Robot Soccer World Cup II*. Springer Verlag, Berlin, 1999.

Peter Stone. *Layered Learning in Multiagent Systems: A Winning Approach to Robotic Soccer*. MIT Press, 2000.

Toshiharu Sugawara and Victor Lesser. On-line learning of coordination plans. COINS technical report 93-27, University of Massachussetts Computer Science Department, 1993.

Toshiharu Sugawara and Victor Lesser. Learning coordination plans in distributed OS environments. In Victor Lesser, editor, *Proceedings of the First International Conference on Multi-Agent Systems*, page 462, San Francisco, CA, 1995. MIT Press.

K. Sycara, K. Decker, A. Pannu, M. Williamson, and D. Zeng. Distributed intelligent agents. *IEEE Expert*, 11(6), December 1996.

Milind Tambe, Jafar Adibi, Yaser Al-Onaizan, Ali Erdem andGal A. Kaminka, Stacy C. Marsela, Ion Muslea, and Marcelo Tallis. Using an explicit model of teamwork in RoboCup-97. In Hiroaki Kitano, editor, *RoboCup-97: Robot Soccer World Cup I*, pages 123–131. Springer Verlag, Berlin, 1998.

Milind Tambe. Recursive agent and agent-group tracking in a real-time , dynamic environment. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 368–375, Menlo Park, California, June 1995. AAAI Press.

Milind Tambe. Tracking dynamic team activity. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, Menlo Park, California, 1996. AAAI Press.

Milind Tambe. Towards flexible teamwork. *Journal of Artificial Intelligence Research*, 7:81–124, 1997.

Ming Tan. Multi-agent reinforcement learning: Independent vs. cooperative agents. In *Proceedings of the Tenth International Conference on Machine Learning*, pages 330–337, 1993.

Eiji Uchibe, Minoru Asada, and Koh Hosoda. Behavior coordination for a mobile robot using modular reinforcement learning. In *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems 1996 (IROS '96)*, pages 1329–1336, 1996.

William T. B. Uther and Manuela M. Veloso. Generalizing adversarial reinforcement learning. In *Proceedings of the AAAI Fall Symposium on Model Directed Autonomous Systems*, 1997.

Manuela Veloso and William Uther. The CMTrio-98 Sony legged robot team. In Minoru Asada and Hiroaki Kitano, editors, *RoboCup-98: Robot Soccer World Cup II*. Springer Verlag, Berlin, 1999.

Manuela Veloso, Peter Stone, and Kwun Han. CMUnited-97: RoboCup-97 small-robot world champion team. *AI Magazine*, 19(3):61–69, 1998.

Manuela Veloso, Peter Stone, Kwun Han, and Sorin Achim. The CMUnited-97 small-robot team. In Hiroaki Kitano, editor, *RoboCup-97: Robot Soccer World Cup I*, pages 242–256. Springer Verlag, Berlin, 1998.

Manuela Veloso, William Uther, Masahiro Fujita, Minoru Asada, and Hiroaki Kitano. Playing soccer with legged robots. In *Proceedings of IROS-98, Intelligent Robots and Systems Conference*, Victoria, Canada, October 1998.

Manuela Veloso, Michael Bowling, Sorin Achim, Kwun Han, and Peter Stone. The CMUnited-98 champion small robot team. In Minoru Asada and Hiroaki Kitano, editors, *RoboCup-98: Robot Soccer World Cup II*. Springer Verlag, Berlin, 1999.

Manuela Veloso, Enrico Pagello, and Hiroaki Kitano, editors. *RoboCup-99: Robot Soccer World Cup III*. Springer Verlag, Berlin, 2000. To appear.

Jose M. Vidal and Edmund H. Durfee. Recursive agent modeling using limited rationality. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 376–383, Menlo Park, California, June 1995. AAAI Press.

Adam Walker and Michael Wooldridge. Understanding the emergence of conventions in multi-agent systems. In *Proceedings of the First International Conference on Multi-Agent Systems*, pages 384–389, Menlo Park, California, June 1995. AAAI Press.

Xuemei Wang. Planning while learning operators. In *Proceedings of the Third International Conference on AI Planning Systems*, May 1996.

Gerhard Weiß and Sandip Sen, editors. *Adaptation and Learning in Multiagent Systems*. Springer Verlag, Berlin, 1996.

Gerhard Weiß. Distributed reinforcement learning. *Robotics and Autonomous Systems*, 15:135–142, 1995.

Gerhard Weiß. ECAI-96 workshop on learning in distributed artificial intelligence. Call For Papers, 1996.

K. Yokota, K. Ozaki, A. Matsumoto, K. Kawabata, H. Kaetsu, and H. Asama. Omni-directional autonomous robots cooperating for team play. In Hiroaki Kitano, editor, *RoboCup-97: Robot Soccer World Cup I*, pages 333–347. Springer Verlag, Berlin, 1998.

Dajun Zeng and Katia Sycara. Bayesian learning in negotiation. volume 48, pages 125–141, January 1998.

Gilad Zlotkin and Jeffrey S. Rosenschein. Coalition, cryptography, and stability: Mechanisms for coalition formation in task oriented domains. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, pages 432–437, Menlo Park, California, August 1994. AAAI Press.