A MODEL OF ADAPTATION IN COLLABORATIVE MULTI-AGENT SYSTEMS

Kristina Lerman

USC Information Sciences Institute, 4676 Admiralty Way, Marina del Rey, CA 90292, USA.

lerman@isi.edu

Phone:(310)448-8714

Fax:(310)822-0751

Abstract

Adaptation is an essential requirement for autonomous agent systems functioning in uncertain dynamic environments. Adaptation allows agents to change their behavior in order to improve the overall system performance. We describe a general mechanism for adaptation in multi-agent systems in which agents modify their behavior in response to changes in the environment or actions of other agents. The agents estimate the global state of the system from local observations and adjust their actions accordingly. We derive a mathematical model that describes the collective behavior of such adaptive systems. The model, consisting of coupled Rate Equations, governs how the collective behavior changes in time. We apply the model to study collaboration in a group of mobile robots. The system we study is an adaptive version of the collaborative stick pulling in a group of robots examined in detail in earlier works (Ijspeert, Martinoli, Billard & Gambardela, 2001; Lerman, Galstyan, Martinoli & Ijspeert, 2001). In adaptive stick pulling, robots estimate the number of robots and sticks in the system and adjust their individual behavior so as to improve collective performance. We solve the mathematical model and show that adaptation improves collective performance for all parameter values.

Keywords: robotics, mathematical models, adaptation

1 Introduction

Adaptation is an essential requirement for systems composed of autonomous agents functioning in dynamic environments that cannot be fully known or characterized in advance. Adaptation allows agents — be they robots, modules in an embedded system, nodes in a sensor network or software agents — to change their behavior in response to changes in the environment or actions of other agents, in order to improve the overall system performance. Biological systems continue to provide an inspiration for the design of adaptive agent systems: when individuals are flexible, the collective as a whole is much more efficient and resilient, even in face of near-catastrophic failures. In social insects such as ants, a large colony of relatively simple individuals can coordinate its actions, without apparent expressive communication or deliberation, to efficiently locate food sources, relocate and defend their nests, and maintain the well-being of the colony in hostile, dynamic environments.

Although learning and adaptation have long been a focus of the AI community, most of the work has been done in the context of one or few agents. The situation is much more complex in a multi-agent setting where the environment is inherently dynamic and stochastic due to the presence of many adaptive agents. Even though there has been a growing body of literature on multi-agent learning (Claus & Boutilier, 1998; Guestrin, Koller & Parr, 2001; Shoham, Grenager, & Powers, 2003), open questions remain. Existing approaches are mainly concerned with the equilibrium properties of the learning algorithms. In dynamically changing environment, on the other hand, these (*e.g.*, Nash) equilibria might not be well defined nor stable, hence, a more important question is how does the system react to changes in the environment. Unfortunately, this question cannot be answered as the tools for systematic study of collective behavior of adaptive multi-agent systems do not yet exist.

In this paper we present and study a simple general mechanism for adaptation in multi-agent systems. If each agent had instantaneous global knowledge of the environment and the state of other agents, it could dynamically change its behavior, allowing the system as a whole to adapt to changes. In most situations, such global knowledge is impractical or costly to collect. However, for sufficiently slow environmental dynamics, agents can correctly estimate the state of the environment through repeated local observations (Jones & Matarić, 2003). The agents then use this estimate to change their behavior in an appropriate way. We call this mechanism memory-based adaptation (Lerman & Galstyan, 2003a) because agents store local observations of the system in a rolling memory window.

In addition to describing an adaptation mechanism, we present a mathematical model of the collective behavior of adaptive agents using this mechanism. These agents are relatively simple: they only use memory of past observations to make decision about future actions, but do not rely on abstract representation, planning, or higher order reasoning functions. Such agents can be represented by a generalized stochastic Markov process. A differential equation, known as the generalized Stochastic Master Equation, governs the evolution of stochastic processes. The Master Equation is often too difficult to formulate and solve for real systems; therefore, we will work with the Rate Equation, which represents the mean field approximation to, or the first moment of, the Master Equation. The Rate Equation describes the dynamics of the average number of agents executing an action.

We illustrate the approach by applying it to study collaboration in groups of mobile robots. The illustration is based on the stick-pulling experiments in groups of robots carried out by Ijspeert, Martinoli, Billard & Gambardela (2001). In these experiments, the robots' task was to pull sticks out of their holes, and it could be successfully achieved only through collaboration between two robots. There was no explicit communication or coordination between the robots. Rather, when a robot found a stick, it lifted it partially out of the ground and held it for some period of time. If another robot found the first one during this time period, it grabbed the stick and lifted it out of the hole completely (successful collaboration); otherwise, the first robot released the stick (unsuccessful collaboration) and resumed the search. We show that a simplified model, in which rather than waiting a specified period of time, a robot has some probability of releasing the stick before the second robot has found it, produces qualitatively similar group behavior as the more complex model that explicitly includes the gripping time. More importantly, we show that in some range of the relevant parameter — the ratio of robots to sticks — there is an optimal stick release rate that maximizes group performance, the rate at which sticks are extracted. We derive an analytic expression for the optimal stick release rate.

The result above suggests that if the number of robots and sticks is known in advance, the robots' stick release rate may be adjusted so as to maximize group performance. *The alternative is to build an adaptive version of the stick pulling system in which a robot can modify its own stick release rate based on its estimate of the number of sticks and other robots in the environment.* As it searches the arena, the robot records observations of sticks and other robots and uses these values to calculate the appropriate stick release rate. If the number of robots changes due to failure of robots or arrival of new ones, or the number of sticks changes as new ones are added, robots modify their individual behaviors to optimize group performance. We write down a model of adaptive stick pulling and analyze the collective behavior of the system in detail. Results show that adaptation improves collective performance of the system.

2 Collective Dynamics of Stochastic Processes

Even in a controlled laboratory setting, the actions of an individual agent, such as a robot, are stochastic and unpredictable: the robot is subject to forces that cannot be known in advance, including noise and fluctuations in the environment, interactions with other robots with complex, equally unpredictable trajectories, errors in its sensors and actuators, in addition to randomness that is often deliberately inserted into the robot controller by its designer, *e.g.*, in collision avoidance maneuvers, the robot often turns a random angle before proceeding. Although individual's behavior is stochastic and unpredictable, the collective behavior of many such individuals often has a simple probabilistic form. We claim that some types of robots can be represented as stochastic Markov processes. Of course, this does not apply to all robots, such as ones based on a hybrid architecture that use planning, reasoning or abstract representations; however, it is true of many simpler robots, including reactive, behavior-based and simple adaptive robots.

A reactive robot is one that makes a decision about what action to take based on its current state (*i.e.*, the action it is currently executing) and input from its sensors. A reactive robot can be considered an ordinary Markov process [1]; therefore, its actions can be represented by a (stochastic) finite state automaton. In fact, this representation has been used to describe robot controllers for more than two decades (Arbib, Kfoury & Moll, 1981; Arkin, 1999; Ijspeert *et al.*, 2001; Goldberg & Matarić, 2003). Each state of the automaton represents the action the agent is executing, with transitions coupling it to other states. Transitions are triggered by input from sensors. As an example, consider a robot engaged in the foraging task, whose goal is to collect objects scattered around an arena. This task consists of the following high-level behaviors: (i) *wandering* about the arena searching for pucks and (ii) *avoiding* obstacles, and (iii) puck *pickup*. Transition from *wandering* to *pickup* is triggered by a puck being sensed, from *wandering* to *avoiding* by an obstacle being sensed, and transition from *avoiding* to *wandering* is caused by the end of the *avoiding* behavior.

Agents can use an internal state to adapt to environmental changes. Consider, for example, a robot whose internal state holds m (local) observations of the environment. In this case, the robot's internal state is its *memory*, but internal state is a more general concept — it can hold the agent's beliefs about other agents or the utility of performing some actions, *etc.*An adaptive robot that makes decisions about future actions based on observations of the m past states of the system can be represented as a generalized Markov process of order m.

In earlier works (Lerman & Shehory, 2000; Lerman *et al.*2001; Lerman & Galstyan, 2002a, 2002b, 2003a, 2003b) we showed that dynamics of collective behavior of a homogeneous system of simple agents or robots is captured by a class of mathematical models known as the Rate Equations. The Rate Equation describe how the average number of robots executing a particular action changes in time and may be easily written down by analyzing individual robot controller. The Rate Equation approach has been used to model variety of dynamic processes in physics, chemistry, biology and ecology (Van Kampen, 1992; Barabasi & Stanley, 1995; Haberman, 1998); however with few exceptions (Huberman & Hogg, 1988; Sugawara & Sano, 1997; Agassounon, Martinoli & Easton, 2004) it has not found use in the robotics and AI communities.

The Rate Equations are usually phenomenological in nature, *i.e.*, they are not derived from microscopic theories. In most cases (*e.g.*, chemical processes, population dynamics, *etc.*), they can be easily written down by considering the important elements of the process. However, it is also possible to derive the Rate Equations from the Stochastic Master Equation (SME). Although SME exactly describes time evolution of the system, in most cases it is analytically intractable and approximate treatments are required. The Rate Equation represents the mean, or the first moment, of the SME.

We now derive the SME and the Rate Equation for the adaptive multi-agent system. In the treatment below, *state* represents the behavior or action an agent is executing in the process of completing its task. Let $p(n, t)$ be the probability an agent is in state n at time t. For a homogenous system of independent and indistinguishable agents, $p(n, t)$ also describes the macroscopic state of the system — the fraction of agents in state n. Let us assume that agents use a finite memory of length m of the past of the system in order to estimate the present state of the environment and make decisions about future actions. Then the agent (and therefore, the multi-agent system) can be represented as a generalized Markov processes of order m. This means that the state of an agent at time $t + \Delta t$ depends not only on its state at time t (as for ordinary Markov processes), but also on its observations at times $t - \Delta t$, $t - 2\Delta t$, ..., $t - (m - 1)\Delta t$, which we refer to collectively as its memory or history h. The following identities then hold:

$$p(n, t + \Delta t|h) = \sum_{n'} p(n, t + \Delta t|n', t; h)p(n', t|h) \quad (1)$$

$$1 = \sum_{n} p(n, t + \Delta t|n', t; h). \quad (2)$$

Let us introduce the probability distribution function over the histories (for a homogenous system this distribution is the same for all the agents): $p(h, t)$, $1 = \sum_{h \in H} p(h, t)$, where H is the set of all feasible histories. Evolution of the agent's state is given by:

$$\Delta p(n, t) = p(n, t + \Delta t) - p(n, t) = \sum_{h} \big[p(n, t + \Delta t|h) - p(n, t|h)\big]p(h).$$

We expand Δp using identities Equation 1–2 and derive in the continuum limit the Stochastic Master

Equation for memory-based adaptive systems.

$$\begin{aligned}
\frac{dp(n,t)}{dt} &= \lim_{\Delta t \to 0} \frac{\Delta p(n,t)}{\Delta t} \\
&= \sum_h \sum_{n'} \left[W(n|n';h)p(n',t|h) - W(n'|n;h)p(n,t|h) \right] p(h)\,,
\end{aligned} \tag{3}$$

with transition rates

$$W(n|n';h) = \lim_{\Delta t \to 0} \frac{p(n,t+\Delta t|n',t;h)}{\Delta t}. \tag{4}$$

The generalized SME, Equation 3, describes the evolution of the probability density for an agent to be in state n at time t, or alternatively, the macroscopic probability density function for the agents in state n. It is similar to the stochastic Master Equation widely studied in statistical physics and chemistry (VanKampen, 1992). In its most general form this equation is often difficult to formulate and solve. Instead, we work with the Rate Equation, which represents the first moment, or the mean, of the SME. The Rate Equation describes how N_n, the average number of agents in state n, changes in time:

$$\frac{dN_n}{dt} = \sum_{n'} \left[\langle W(n|n') \rangle N_{n'} - \langle W(n'|n) \rangle N_n \right], \tag{5}$$

with history-averaged transition rates

$$\langle W(n|n') \rangle = \lim_{\Delta t \to 0} \frac{\sum_h p(n,t+\Delta t|n',t;h)p(h)}{\Delta t}. \tag{6}$$

Equation 5 also holds for systems composed reactive robots (Lerman & Galstyan, 2002b), which can be modeled as ordinary Markov processes, although the history term no longer appears in it. It is important to remember that the Rate Equations do not describe results of a specific experiment, rather, the behavior of quantities averaged over many experiments. We use the Rate Equation to study collective behavior of adaptive robot systems.

3 Collaboration in Robots

The stick-pulling experiments were carried out by Ijspeert *et al.* (2001) to investigate dynamics of collaboration among locally interacting reactive robots. Figure 1 is a snapshot of the physical set-up of the experiments. The robots' task was to locate sticks scattered around the arena and pull them out of their holes. A single robot cannot complete the task (pull the stick out) on its own — a collaboration between two robots is necessary for the task to be successfully completed. Each robot is governed by the same controller: each robot spends its time looking for sticks and avoiding obstacles. When a robot finds a stick, it lifts it partially out of its hole and waits for a period of time τ for a second robot to find it. If a second robot finds the first one, it will grip the stick and pull it out of the ground, successfully completing the task; otherwise, the first robot times out, releases the stick and returns to the searching state.

In Lerman *et al.*(2001) we have constructed a mathematical model of collective dynamics of this system and compared the model's predictions to experimental results. Here we examine a simplified scenario, where, instead of waiting a specified period of time, each robot releases the stick with some probability per unit time. As we show in Section 3.1, the behavior of such a simplified system is similar to that of the original system. Moreover, adaptive version of the simplified system is readily amenable to analysis. The adaptive version of the collaborative stick pulling task is described in Section 3.2.

3.1 Collective Behavior of Reactive Systems

On a macroscopic level, during a sufficiently short time interval, each robot will be in one of two states: *searching* or *gripping*. We assume that actions such as pulling the stick out or releasing it take place on a short enough time scale that they can be incorporated into the search state. Of course, in a model there can be a discrete state corresponding to every robot behavior or action in the controller. Martinoli & Easton (2003) have done this and found quantitative agreement between the model's prediction and simulations for systems of 16–24 robots. We have shown (Lerman *et al.*, 2001) that even a minimal model with only two states helps explain the main experimental findings.

In addition to states, we must also specify all possible transitions between states. When it finds a stick, the robot makes a transition from the search state to the gripping state. After both a successful and unsuccessful collaborations the robot releases the stick and makes a transition into the searching state, as shown in Figure 2. We will use the state diagram as the basis for writing down the rate equations for the dynamics of the system.

Each box in Figure 2 becomes a dynamic variable of the model: $N_s(t)$ and $N_g(t)$, the (average) number of robots in the searching and gripping states respectively, as well as $M(t)$, the number of uncollected sticks at time t. This is the environmental variable that couples the states by triggering transitions between them. The mathematical model of the stick-pulling system consists of a series of coupled rate equations, describing how the dynamic variables evolve in time:

$$\frac{dN_s}{dt} = -\alpha N_s(t)\Big(M(t) - N_g(t)\Big) + \tilde{\alpha} N_s(t) N_g(t) + \gamma N_g(t)\,, \tag{7}$$

$$\frac{dM}{dt} = -\tilde{\alpha} N_s(t) N_g(t) + \mu(t)\,, \tag{8}$$

where α, $\tilde{\alpha}$ are the rates at which a searching robot encounters a stick and a gripping robot respectively; γ is the rate at which robots release sticks ($1/\gamma$ is equivalent to the gripping time parameter τ in (Ijspeert *et al.*, 2001; Lerman *et al.*, 2001)); $\mu(t)$ is the rate at which new sticks are added by the experimenters. These parameters connect the model to the experiment: α and $\tilde{\alpha}$ are related to the size of the object, the robot's detection radius, or footprint, and the speed at which it explores the arena.

The first term in Equation 7 accounts for the decrease in the number of searching robots as robots find and grip sticks; the second term describes successful collaborations between two robots (sticks are pulled out), and the third term accounts for the failed collaborations (when a robot releases a stick without a second robot present), both of which lead to an increase the number of searching robots. We do not need a separate equation for N_g, since this quantity may be calculated from the conservation of robots condition, $N = N_s + N_g$. The last equation, Equation 8, states that the number of sticks, $M(t)$, decreases in time at the rate of successful collaborations. The equations are subject to the initial conditions that at $t = 0$ the number of searching robots in N and the number of sticks is M.

We introduce the following transformations on variables in order to rewrite equations in dimensionless form: $n(t) = N_s(t)/N$ and $m(t) = M(t)/M$ are fractions of searching robots and uncollected sticks at time t; $\beta = N/M$, ratio of total number of robots to the total number of sticks; $R_G = \tilde{\alpha}/\alpha$ and $\tilde{\beta} = R_G\beta$. The fraction of gripping robots is simply $1 - n(t)$. Dimensionless versions of Equation 7–8 are:

$$\frac{dn}{dt} = -n(t)[m(t) + \beta n(t) - \beta] + \tilde{\beta} n(t)[1 - n(t)] + \gamma[1 - n(t)] \tag{9}$$

$$\frac{dm}{dt} = -\beta\tilde{\beta} n(t)[1 - n(t)] + \mu' \tag{10}$$

Note that only two parameters, β and γ, appear in the equations and, thus determine the behavior of solutions. The third parameter $\tilde{\beta} = R_G\beta$ is fixed experimentally and is not independent. Note that we do not need to specify α and $\tilde{\alpha}$ — they enter the model only through R_G (throughout this paper we will use $R_G = 0.35$). [2]

We assume that the number of sticks does not change with time ($m(t) = m(0) = 1$) because new sticks are added (*e.g.*, by the experimenter) at the rate the robots pull them out. A steady-state solution, if it exists, describes the long term time-independent behavior of the system. To find it, we set the left hand side of Equation 9 to zero:

$$-n[1 + \beta n - \beta] + \tilde{\beta} n[1 - n] + \gamma[1 - n] = 0. \tag{11}$$

This quadratic equation can be solved to obtain steady state values of $n(\beta, \gamma)$.

Collaboration rate is the rate at which robots pull sticks out of their holes. The steady-state collaboration rate is

$$R(\gamma, \beta) = \beta\tilde{\beta} n(\gamma, \beta)[1 - n(\gamma, \beta)]\,, \tag{12}$$

where $n(\gamma, \beta)$ is the steady-state number of searching robots for a particular value of γ and β. Figure

3(a) depicts the collaboration rate as a function of $1/\gamma$. Note, that there exists a critical value of β, so that for $\beta > \beta_c$, collaboration rate remains finite as $1/\gamma \to \infty$, while for $\beta < \beta_c$, it vanishes. The intuitive reason for this was presented in (Ijspeert *et al.*, 2001): when there are fewer robots than sticks, and each robot holds the stick indefinitely (vanishing release probability), after a while every robot is holding a stick, and no robots are available to help pull sticks out. Also, for $\beta < \beta_c$ there is an optimal value of γ which maximizes the collaboration rate and can be computed from the condition $dR(\gamma, \beta)/d\gamma = \beta\tilde{\beta}d(n - n^2)/d\gamma = 0$, with n given by roots of Equation 11. Another way to compute the optimal release rate is by noting that for a given value of β below some critical value, the collaboration rate is greatest when half of the robots are gripping and the other half are searching. Substituting $n = 1/2$ into Equation 11, leads to

$$\gamma_{opt} = 1 - (\beta + \tilde{\beta})/2 \ \texttt{for} \ \beta < \beta_c = 2/(1 + R_G). \tag{13}$$

No optimal release rate exists when β exceeds its critical value β_c.

Figure 3(b) shows results of experiments and simulation for groups of two to six robots (Ijspeert *et al.*, 2001). The three curves in Figure 3(a) are qualitatively similar to those in Figure 3(b) for 2 robots ($\beta = 0.5$), 4 robots ($\beta = 1.0$) and 6 robots ($\beta = 1.5$). Even the grossly simplified model reproduces the main conclusions of the experimental work: existence of β_c, the critical value of the ratio of robots to sticks, and the optimal release rate (or conversely, the gripping time) that maximizes the collaboration rate for $\beta < \beta_c$. In addition, analysis gives analytic form for important parameters, such as β_c and γ_{opt} — values we will exploit in constructing adaptive version of collaborative stick pulling.

3.2 Collective Behavior of Adaptive Systems

Figure 3(a) suggests that if the number of sticks and robots is known in advance, the robot's release rate can be set to a value that maximizes the group collaboration rate. If the number of sticks or the number of robots is not known or changing in time (due to robot failure, for example), the robots can still tune their individual parameters to maximize group performance. They accomplish this through the memory-based adaptation mechanism. As they search the arena, robots record observed numbers of sticks and other robots, estimate the density of each from these values, and compute the appropriate stick release rate according to the following rules:

$$\gamma = 1 - \frac{\beta_{obs}(1+R_G)}{2} \;\; \texttt{for} \;\; \beta_{obs} < 2/(1+R_G) \quad (14)$$

$$\gamma = 0 \;\; \texttt{for} \;\; \beta_{obs} \geq 2/(1+R_G), \quad (15)$$

where $\beta_{obs} = N_{obs}/M_{obs}$, the ratio of the observed numbers of robots and sticks. Suppose each robot has a memory window of size h. As it makes observations, robot adds them to memory, replacing older observations with more recent ones. For a particular robot, the values in most recent memory slot are N^0_{obs} and M^0_{obs}, the observed number of robots and sticks at time t; in the next latest slot, the values are N^1_{obs} and M^1_{obs}, the observed numbers at time $t - \Delta$, and so on. Robot computes γ_{opt} from $N_{obs} = \sum_{j=0}^{h-1} N^j_{obs}$ and $M_{obs} = \sum_{j=0}^{h-1} M^j_{obs}$.

Dynamics of the adaptive system are specified by Eqs. 9-10, where γ is now the history-averaged stick release rate, the aggregate of individual decisions made according to rules in Equation 14–15. It is computed in the following way. When observations of all robots are taken into account, the mean of the observed number of robots in the first memory slot is $\frac{1}{N}\sum_{i=1}^{N} N^0_{i,obs} \approx N(t)$, where $N(t)$ is the average number of robots at time t. Likewise, the mean value observed value in memory slot j is $\frac{1}{N}\sum_{i=1}^{N} N^j_{i,obs} \approx N(t - j\Delta)$, the average number of robots at time $t - j\Delta$. In general, the actual value will fluctuate because of measurement errors; however, on average, it will be the average number of robots (or sticks) in the system at that time. This system is trivial — the average number of robots and sticks does not change in time. In other systems, however, parameters may depend on variables that change in time, for example, the number of searching robots (Lerman & Galstyan, 2003). The rate equations for such systems will be time delay equations, since parameters will depend on the delayed values of the dynamic variables.

Figure 4(a) shows how the solution, the fraction of searching robots, relaxes in both adaptive and reactive systems. In all cases, solutions reach a steady-state. Note that in reactive systems, the steady-state value of n_s depends on β, while in adaptive systems, by design $n_s = 0.5$.

Figure 4(b) shows the difference between collaboration rate in adaptive and reactive systems for different values of γ (the value of collaboration depends on γ only in reactive systems). The difference is always positive, meaning that adaptation always improves collaboration rate, by as much as 15% in this range of β. The two sets of curves are for two values of R_G, an experimental parameter that measures how easy it is for the second robot to grip the stick. When the first robot is gripping the stick, it constrains the angle at which the second robot can approach and grip the stick. The angle of approach is parametrized by R_G. In experiments R_G was measured to be 0.35 (Ijspeert *et al.*, 2001), and this is the value we used in this paper. Essentially, R_G gives the angle the second robot can approach the first one and still be able to grip the stick. As we can see from the figure, this experimental parameter influences collaboration rate. If robots are redesigned, so that one robot can approach a gripping robot from a wider angle (bigger value of R_G), the benefit of adaptation in such a system will be even greater.

4 Prior Work

Mathematical analysis of the behavior of MAS is a relatively new field with approaches and methodologies borrowed from other fields, such as mathematics, physics and biology.

Analysis of Robot Systems In recent years, a number of studies appeared that attempted to mathematically model and analyze collective behavior of distributed robot systems. These include analysis of the effect of collaboration in foraging (Sugawara & Sano, 1997) and stick-pulling (Lerman *et al.*, 2001; Martinoli & Easton, 2003) experiments, the effect of interference in robot foraging (Lerman & Galstyan, 2002a), and robot aggregation task (Agassounon *et al.*, 2004). This type of analysis has been limited to simple reactive or behavior-based robots in which perception and action are tightly coupled. Such robots take input from sensors or behaviors and send output to actuators or other behaviors. They make no use of memory or internal state, nor do they change their behavior in response to environmental changes.

Closest to ours is the work of Huberman and Hogg (1988), who studied collective behavior of a system of adaptive agents using game dynamics as a mechanism for adaptation. In game dynamical systems, winning strategies are rewarded, and agents use the best performing strategies to decide their next move. They constructed a mathematical model of the dynamics of such systems and studied them under variety of conditions, including imperfect knowledge and delayed information. Although the mechanism for adaptation is different, their approach, which they termed "computational ecology" is similar in spirit to ours, as it is based on the foundations of stochastic processes and models of average behavior. Their work, however, does not explicate any general principles or a framework for analysis that would apply to other systems.

Another example of the stochastic approach is the probabilistic microscopic model (Martinoli, Ijspeert & Gambardella, 1999; Ijspeert *et al.*, 2001) developed to study collective behavior of a group of robots. Rather than compute the exact trajectories and sensory information of individual robots, each robot's interactions with other robots and the environment is modeled as a series of stochastic events, with probabilities determined by simple geometric considerations. Running several series of stochastic events in parallel, one for each robot, allows one to study the group behavior of the multi-robot system.

Learning in Multi-agent Systems Although learning has been one of the most important topics in computer science, few mathematical descriptions of the collective behavior of MAS composed of large numbers of concurrent learners exist (Wolpert & Tumer, 1999; Sato & Crutchfield, 2003). These are microscopic models, which only allow one to study collective behavior of relatively small systems. We are interested in approaches that will enable us to analyze even very large systems.

Application-level studies of learning in the context of multi-robot systems have recently been carried out (Kaelbling, 1991; Matarić, 1997; Riedmiller & Merke, 2001; Stone, 2001; Li, Martinoli & Abu-Mostafa, 2002; Jones & Matarić, 2003). Specifically, Li *et al.* (2002) introduced learning into collaborative stick pulling robots and showed in simulation that learning does improve system performance by allowing robots to specialize. No analysis of the collective behavior or performance of the system have been attempted in any of these studies.

5 Conclusion

We have presented a general mechanism for adaptation in multi-agent systems in which the agents can modify their behavior in response to environmental dynamics or actions of other agents. The agents estimate the global state of the system from individual observations stored in memory and adjust their behaviors accordingly. We have also derived a system of equations that describes the dynamics of collective behavior of such adaptive systems. We have applied the mathematical model to study adaptive collaboration in robots, where robots compute internal parameters based on the observations stored in memory. We explicitly took finite memory size into account, although in the aggregate approach considered here, the size of the memory window does not impact the behavior of the system. We showed that adaptation improves performance (*i.e.*, collaboration rate) of the robot system with respect to the reactive system.

Although no experimental studies of adaptive stick-pulling have been carried out, the memory-based adaptation mechanism described here has been applied to the problem of dynamic task allocation in robots, where it was studied both in simulation (Jones & Matarić, 2003) and theoretically (Lerman & Galstyan, 2003). In this scenario, robots decide on the optimal division of labor based on the observed numbers of tasks and robots engaged in those tasks. Embodied simulations show that the memory-based adaptation mechanism does lead to the desired division of labor. Moreover, theoretical results closely reproduce experimental observations. These works give us confidence that the adaptation mechanism described in this paper will indeed work for collaborative stick-pulling.

There are many issues that remain to be addressed by analysis. One of the more important ones is the effect of noisy observations on collective behavior. The observed numbers of robots and sticks will vary from robot to robot, because robots are sampling different areas of the system. We have recently studied the exact stochastic model of the simplified dynamic task allocation scenario which has allowed us to directly study the effect of noise-induced fluctuations in robot's performance (Galstyan & Lerman, 2004). In future research we will expand on these results to further characterize noise-induced variations and their effect on the performance of the system.

We believe that mathematical analysis is a powerful tool for studying multi-robot (and multi-agent) systems and will play an increasingly important role in the design of these systems. Although past applications of mathematical analysis have targeted relatively simple (reactive) robots, we have shown that analysis can also describe more complex robots. Integrating analysis in the design cycle of robot controllers will allow researchers to efficiently test controllers before they are deployed on physical systems. Analysis will enable researchers not only to confirm that the controllers do indeed produce desired collective behavior, but also to quickly find parameters that improve collective performance of the multi-robot system.

End Notes

1. An ordinary Markov process's future state depends only on its present state and none of the past states. For a semi-Markov process, the transition also depends on how long the process has been in the current state. A generalized Markov process's future state depends on the past m states.

2. The parameter α can be easily calculated from experimental values quoted in (Ijspeert *et al.*, 2001). As a robot travels through the arena, it sweeps out some area during time dt and will detect objects that fall in that area. This detection area is $V_R W_R dt$, where $V_R = 8.0\ cm/s$ is robot's speed, and $W_R = 14.0\ cm$ is robot's detection width. If the arena radius is $R = 40.0\ cm$, a robot will detect sticks at the rate $\alpha = V_R W_R / \pi R^2 = 0.02\ s^{-1}$. According to (Ijspeert *et al.*, 2001), a robot's probability to grab a stick already being held by another robot is 35% of the probability of grabbing a free stick. Therefore, $R_G = \tilde{\alpha}/\alpha = 0.35$. R_G is an experimental value obtained with systematic experiments with two real robots, one holding the stick and the other one approaching the stick from different angles.

Ackowledgements

The research reported here was supported in part by the Defense Advanced Research Projects Agency (DARPA) under contract number F30602-00-2-0573. The author would like to thank Aram Galstyan and Tad Hogg for many helpful discussions.

References

Agassounon, W., Martinoli, A. and Easton, K. 2004 Macroscopic Modeling of Aggregation Experiments using Embodied Agents in Teams of Constant and Time-Varying Sizes. Special issue on Swarm Robotics, Dorigo, M. and Sahin, E. editors, *Autonomous Robots*, **17**(2-3):163–191.

Arbib, M. A., Kfoury, A. J. and Moll, R. N. 1981. *A Basis for Theoretical Computer Science.* Springer Verlag, New York, NY.

Arkin, R. C. 1999. *Behavior-Based Robotics.* The MIT Press, Cambridge, MA, USA.

Barabasi, A.-L. and Stanley, H. 1995. *Fractal Concepts in Surface Growth.* Cambridge University Press, Cambridge, England.

Claus, C. and Boutilier, C. 1998. The dynamics of reinforcement learning in cooperative multiagent systems. pp. 746–752. In *Proc. of the Fifteenth National Conf. on Artificial Intelligence (AAAI-98).*

Galstyan, A. and Lerman, K. 2004. *Proc. of Engineering of Self-Organizing Systems workshop, International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS'04)*, New York, New York.

Goldberg, D. and Matarić, M. J. 2003. Maximizing reward in a non-stationary mobile robot environment. *Autonomous Agents and Multi-Agent Systems*, **6**(3):281–316.

Guestrin, C., Koller, D. and Parr, R. 2001. Multiagent Planning with Factored MDPs. In *Advances in Neural Information Processing Systems (NIPS)*, Vancouver, British Columbia, Canada.

Haberman, R. 1998. *Mathematical Models: Mechanical Vibrations, Population Dynamics, and Traffic Flow.* Society of Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Huberman, B. A. and Hogg, T. 1988. The behavior of computational ecologies. pp. 77–115: In B. A. Huberman, editor, *The Ecology of Computation.* Elsevier (North-Holland), Amsterdam.

Ijspeert, A. J., Martinoli, A., Billard, A. and Gambardella L. M. 2001. Collaboration through the Exploitation of Local Interactions in Autonomous Collective Robotics: The Stick Pulling Experiment. *Autonomous Robots* **11**(2):149–171.

Jones, C. V. and Matarić, M. J. 2003. Adaptive task allocation in large-scale multi-robot systems. In *Proc. of the 2003 (ICRA'03)*, Las Vegas, NV.

Kaelbling, L. P. 1991. *Learning in Embedded Systems.* MIT Press, Cambridge, MA, USA.

Lerman, K. and Galstyan, A. 2002a. Mathematical model of foraging in a group of robots: Effect of interference. *Autonomous Robots*, **13**(2):127–141.

Lerman, K. and Galstyan, A. 2002b. Two paradigms for the design of artificial collectives. In *Proc. of the First Annual workshop on Collectives and Design of Complex Systems*, NASA-Ames, CA.

Lerman, K. and Galstyan, A. 2003. Macroscopic Analysis of Adaptive Task Allocation in Robots. In *Proc. of the Int. Conf. on Intelligent Robots and Systems (IROS-2003)*, Las Vegas, NV.

Lerman, K., Galstyan, A., Martinoli, A. and Ijspeert, A. 2001. A macroscopic analytical model of collaboration in distributed robotic systems. *Artificial Life Journal*, **7**(4):375–393.

Lerman, K. and Shehory, O. 2000. Coalition Formation for Large-Scale Electronic Markets. pp. 167–174. In *Proc. of the Int. Conf. on Multi-Agent Systems (ICMAS'2000)*, Boston, MA.

Li, L., Martinoli, A., and Abu-Mostafa, Y. 2002. Emergent Specialization in Swarm Systems. pp. 261–266. In *Lecture Notes in Computer Science: 2412*, Springer Verlag, New York, NY.

Martinoli, A. and Easton, K. 2003. Modeling swarm robotic systems. pp. 297–306. In B. Siciliano and P. Dario, editors, *Proc. of the Eight Int. Symp. on Experimental Robotics (ISER-02)*, Springer Verlag, New York, NY.

Martinoli, A., Ijspeert, A. J., and Gambardella, L. M. 1999. A probabilistic model for understanding and comparing collective aggregation mechanisms. pp. 575–584. In D. Floreano, J.-D. Nicoud, and F. Mondada, editors, *LNAI:1674*, Springer, New York, NY.

Matarić, M. J. 1997. Reinforcement learning in the multi-robot domain. *Autonomous Robots*, **4**(1):73–83.

Riedmiller, M. and Merke, A. 2001. Karlsruhe brainstormers - a reinforcement learning approach to robotic soccer ii. In *RoboCup-01: Robot Soccer World Cup V*, LNCS. Springer.

Sato, Y. and Crutchfield, J. P. 2003. Coupled replicator equations for the dynamics of learning in multiagent systems. *Physical Review*, **E67**, 015206.

Shoham, Y., Grenager, T., and Powers, R. 2003. Multi-agent reinforcement learning: A critical survey. unpublished manuscript http://robotics.stanford.edu/ shoham/YoavPublications.htm.

Sugawara, K. and Sano, M. 1997. Cooperative acceleration of task performance: Foraging behavior of interacting multi-robots system. *Physica* **D100**:343–354.

Van Kampen, N. G. 1992. *Stochastic Processes in Physics and Chemistry.* Elsevier Science, Amsterdam.

Wolpert, D. and Tumer, K. 1999. An introduction to collective intelligence. Technical Report NASA-ARC-IC-99-63, NASA Ames Research Center.

Figure 1 Physical set-up of the stick-pulling experiment (courtesy of A. Martinoli).

Figure 2 Macroscopic state diagram of the multi-robot system. The arrow marked 's' corresponds to the transition from the gripping to the searching state after a successful collaboration, while the arrow marked 'u' corresponds to the transition after an unsuccessful collaboration, *i.e.*, when the robots releases the stick without a successful collaboration taking place.

Figure 3 (a) Collaboration rate per robot vs inverse stick release rate $1/\gamma$ for $\beta = 0.5$, $\beta = 1.0$, $\beta = 1.5$. These values of β correspond, respectively, to two, four, and six robots in the experiments with four sticks. (b) Collaboration rate vs. the gripping time parameter for groups of two to six robots and four sticks (from (Ijspeert *et al*, 2001)). Heavy symbols represent experimental results, while lines represent results of two different types of simulations.

Figure 4 (a) Time evolution of the fraction of searching robots for adaptive and reactive systems. (b) Difference between collaboration rates for adaptive and reactive systems for different values of experimental parameters R_G and γ.

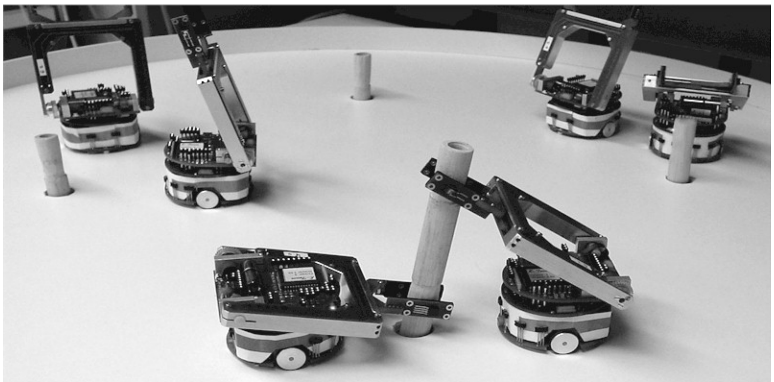


Figure 1:

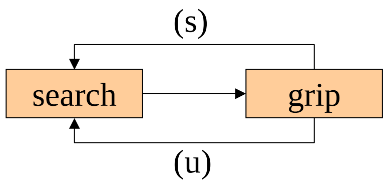
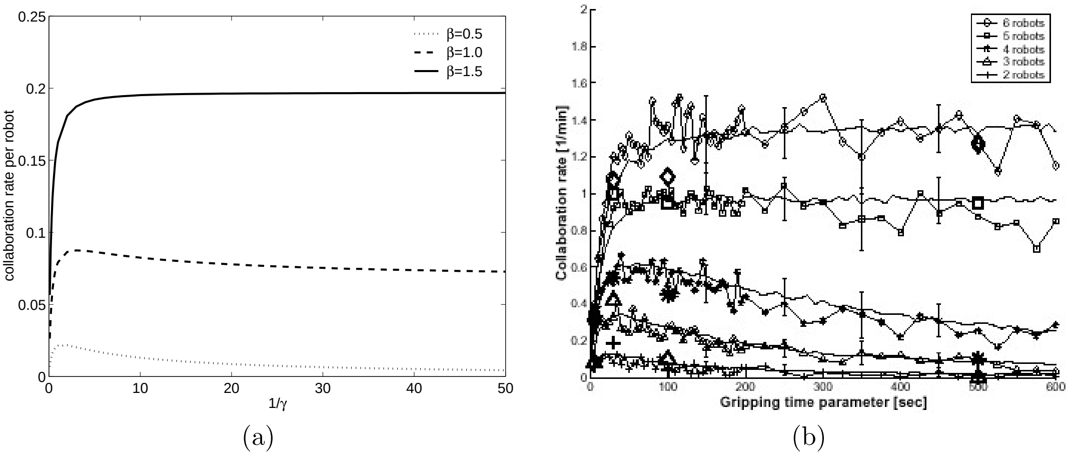
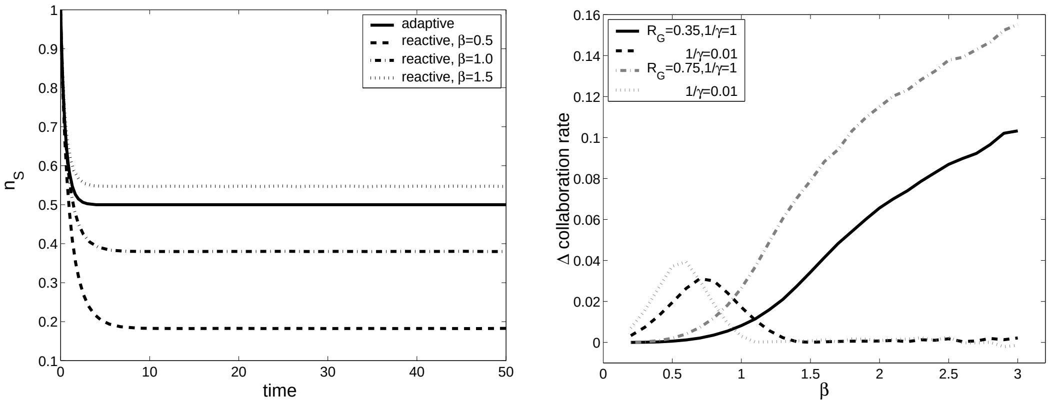


Figure 2:



0.25
0.2
0.15
0.1
0.05
0
collaboration rate per robot
β=0.5
β=1.0
β=1.5
0
10
20
30
40
50
1/γ
(a)
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Collaboration rate [1/min]
6 robots
5 robots
4 robots
3 robots
2 robots
0
100
200
300
400
500
600
Gripping time parameter [sec]
(b)

Figure 3:



adaptive
reactive, β=0.5
reactive, β=1.0
reactive, β=1.5
n_S
time
R_G=0.35,1/γ=1
1/γ=0.01
R_G=0.75,1/γ=1
1/γ=0.01
Δ collaboration rate
β

(a) (b)

Figure 4: