

# Macroscopic Analysis of Adaptive Task Allocation in Robots

Kristina Lerman  
USC Information Sciences Institute  
4676 Admiralty Way  
Marina del Rey, CA 90292  
lerman@isi.edu

Aram Galstyan  
USC Information Sciences Institute  
4676 Admiralty Way  
Marina del Rey, CA 90292  
galstyan@isi.edu

## ABSTRACT

We describe a general mechanism for adaptation in multi-agent systems in which agents modify their behavior based on their memory of past events. These behavior changes can be elicited by environmental dynamics or arise as response to the actions of other agents. The agents use memory to estimate the global state of the system from individual observations and adjust their actions accordingly. We also present a mathematical model of the dynamics of collective behavior in such systems and apply it to study adaptive task allocation in mobile robots. In this application, the robots' task is to forage for Red or Green pucks (a robot can only forage for one puck type at a time). As it travels around the arena, a robot records observations of pucks and other robots, and uses these observations to compute the estimated density of each. If it finds there are not enough robots of a specific type, it may switch its foraging state to fill the gap. After a transient, we expect the number of robots in each foraging state to reflect the prevalence of each puck type in the environment. We modelled adaptive task allocation and studied the dynamics of the system for different transition rates between states. We find that for some rates lead to fast convergence times and a steady state solution.

## 1. INTRODUCTION

Biological control has been a dominant metaphor for the design of multi-robot systems (see [20, 2, 19, 6, 7] among others). Biological control is characterized by minimalist individual robot controller [3, 18] and local interactions, both direct (among robots) and indirect (through the environment). Biological metaphor offers several benefits for the design of agent-based systems, including (i) scalability: each agent has the same controller whether the group is composed of ten or 10,000 agents; (ii) robustness: group performance is robust to individual agent failure; (iii) flexibility: agents can be dynamically added or removed without significantly affecting the performance of the system; (iv) it is inherently

distributed: collective behavior arises via local interactions among autonomous agents. Biological metaphor also allows us to introduce learning and adaptation into the system that enable robots to operate in uncertain dynamic environments. Although learning has been an important research topic in Artificial Intelligence in general and robotics in particular, the issues of learning in the context of multi-robot systems have only recently begun to be investigated [11, 21, 10, 4, 17].

A final advantage of biologically-inspired systems is that they can be mathematically modelled and analyzed. Mathematical analysis is an alternative to experiment and simulation, the two principal tools used by computer scientists. While experiments allow researchers to directly study the systems' behavior in real environments, they are very costly and time-consuming to execute. Sensor-based computer simulations [22, 5] and probabilistic microscopic simulations [9], also do not scale well, and their results are not easily generalizable. As a result, many of the interesting questions about the behavior of agent-based systems, especially of large systems, remain unanswered. Using mathematical analysis, on the other hand, we can efficiently study dynamics of even very large agent groups, predict their long term behavior, gain insight into system design: *e.g.*, what parameters determine group behavior, optimize performance, prevent instabilities, *etc.* However, with few exceptions [24, 15, 13], little mathematical analysis of multi-robot systems exists.

In our earlier work we have developed an approach [14] that allows us to create a mathematical model of collective dynamics of multi-agent systems. The approach is valid for agents that obey the Markov property — *i.e.*, only the present state determines the agent's future state. Both reactive [9] and behavior-based [1] robotic systems satisfy this property. We have applied this approach to study foraging [13] and collaboration [15] in groups of robots and showed agreement between the predictions of mathematical analysis and the results experiment and simulations. At the same time, mathematical analysis has allowed us to draw conclusions about the behavior of the system that would have been difficult to obtain through experiment and simulation alone.

The simple models we investigated so far cannot describe adaptive robotic systems, where robots can use remembered or learned information about the past to guide future action. In this paper we describe an extension of the modeling

approach that allows us to incorporate history or memory of robots' actions into the mathematical model of a multi-robot system. This allows us to model robots that make decisions about future actions based on past experience, and thus study adaptation and learning in multi-agent systems. Although learning has been one of the most important topics in computer science, few mathematical descriptions of multi-agent systems composed of concurrent learners exist [23]. This is an important step towards creating a *science* of multi-agent systems that will allow quantitative understanding of the behavior of a collective of agents or robots in a given environment. Such a science of multi-agent systems will enable researchers design better and more robust robot controllers and understand the impact they will have on the collective behavior.

The rest of the paper is organized as follows. In Section 2 we present the model of the collective behavior of agents with memory. In Section 3 we illustrate our approach by studying an adaptive task allocation system in which robots use memory to change their actions and adjust to changes in the environment.

## 2. MODELING RETROSPECTIVE AGENTS

In earlier works [16, 15, 13, 14] we showed that the dynamics of collective behavior in a system of agents obeying the Markov property is captured by the Rate Equations that describe how the average number of agents in each state changes in time. *State* labels a set of related agent behaviors required to accomplish a task. As an example, consider a robot engaged in a foraging task, whose goal is to collect objects scattered around an arena and deliver them to a pre-specified home location. The foraging task consists of the following high-level behavioral requirements [1] or states: (i) *homing*, (ii) *puck pickup* and (iii) *searching*. Each of these states corresponds to a single action or behavior; however, it is often useful to coarse-grain the system by grouping related behaviors into a single state. Such coarse-graining not only helps in conceptualizing the system, it also keeps the mathematical model compact and tractable by reducing the number of states. For example, when a robot is said to be in the *searching* state, it is wandering around the arena, detecting objects and avoiding obstacles. In the course of accomplishing a task, the robot will transition from the *searching* to *pickup* and finally to *homing* states. It is clear that during a sufficiently short time interval each agent in a multi-agent system is in exactly one of a finite number of states.

If each agent had instantaneous global knowledge of the environment and the state of other agents, the system can dynamically adapt to any changes. In most situations, such global knowledge is impractical. However, for sufficiently slow dynamics, agents can correctly estimate the state of the environment through repeated observations (by storing them in memory). Here we present a theory of such a memory-based adaptation mechanism. Let  $p(n, t)$  be the probability an agent is in state  $n$  at time  $t$ . We note that for a homogenous system of independent and indistinguishable agents,  $p(n, t)$  describes the macroscopic state of the system, since it is simply the fraction of agents in the state  $n$ . Let us assume that the 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 evolution of the 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 the configuration of the system at time  $t$  (as in simple Markov systems), but also on configurations at times  $t - \Delta t$ ,  $t - 2\Delta t$ ,  $\dots$ ,  $t - (m - 1)\Delta t$ , which we refer to as history  $h$  of the system. In the derivation below we will employ the following identities:  $p(n, t + \Delta t|h) = \sum_{n'} p(n, t + \Delta t|n', t; h)p(n', t|h)$  and  $\sum_n p(n, t + \Delta t|n', t; h) = 1$ .

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 (if it continuous, one should use integration instead of summation for proper normalization). We can then write for the change in probability density  $\Delta p$  is:

$$\begin{aligned} \Delta p(n, t) &= p(n, t + \Delta t) - p(n, t) \\ &= \sum_h [p(n, t + \Delta t|h) - p(n, t|h)]p(h) \\ &= \sum_h \sum_{n'} p(n, t + \Delta t|n', t; h)p(n', t|h)p(h, t) \\ &\quad - \sum_h \sum_{n'} p(n', t + \Delta t|n, t; h)p(n, t|h)p(h, t) \end{aligned} \quad (1)$$

In the continuum limit, as  $\Delta t \rightarrow 0$ ,  $\Delta p/\Delta t$  can be written as

$$\begin{aligned} \frac{dp(n, t)}{dt} &= \sum_h \sum_{n'} W(n|n'; h)p(n', t|h)p(h, t) \\ &\quad - \sum_h \sum_{n'} W(n'|n; h)p(n, t|h)p(h, t), \end{aligned} \quad (2)$$

with transition rates

$$W(n|n'; h) = \lim_{\Delta t \rightarrow 0} \frac{p(n, t + \Delta t|n', t; h)}{\Delta t}. \quad (3)$$

Equation 2 is similar in form to the stochastic Master Equation widely studied in statistical physics and chemistry [12]. It 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  $n$ . In the most general form Eq.2 is analytically untractable due to strong correlations both in time and state-space. However, as we will show below, in some cases, when the history of the system can be represented as some aggregated quantity, interesting results can be obtained by using approximate techniques.

At the end of this section we note that sometimes it suffices to study the evolution of the average quantities, such as average number of agents in a certain state rather than the whole probability density function itself. One can do this by deriving the so-called Rate Equation, that governs the evolution of the first moments with time. Let  $N_n$  the number of agents in state  $n$ . Assuming that during a sufficiently short time interval only one of the agents in state  $n'$  will make a transition to state  $n$ , we can derive the macroscopic equation for the rate of change of  $\langle N_n \rangle$ , the average number

of agents in state  $n$ :

$$\frac{d\langle N_n \rangle}{dt} = \sum_{h, n'} [W(n|n'; h)\langle N_{n'} \rangle - W(n'|n; h)\langle N_n \rangle] P(h, t)$$

Using the mean-field approximation  $\langle f(n) \rangle \approx f(\langle n \rangle)$  and the fact that  $N_n$  does not depend directly on  $h$ , we rewrite the above equation,

$$\frac{dN_n}{dt} = \sum_{n'} [\langle W(n|n') \rangle_h N_{n'} - \langle W(n'|n) \rangle_h N_n]. \quad (4)$$

Here for notational convenience  $\langle \dots \rangle_h$  denotes average over histories, and we have dropped angle brackets around  $N$ , although this variable still denotes an average quantity.

Equation 4 is very similar to the rate equation we used to study Markov-based agent systems [16, 15, 13], except that transition rates  $W(n|n')$  are now replaced by their history-averaged values. We will use the above equation to study how agents can use histories, or memories of past events, to improve the collective behavior of the system. We illustrate the approach by examining adaptive task allocation in robots.

### 3. ADAPTIVE TASK ALLOCATION

Chris Jones and Maja Mataric [10] have initiated a study of adaptive task allocation in multi-robot systems. This scenario is based on the foraging task. Consider an arena with some number of pucks scattered about it. The pucks can be of two distinct types, *Red* and *Green*. Each robot can be tasked to collect pucks of a specific type, say *Red*. When the robot's foraging state is set to *Red*, it is searching and collecting *Red* pucks. The robots can also recognize the foraging state of robots that are visible to it. The robots have no *a priori* information about the shape of the arena, the number of pucks left in it or the number of foraging robots. The goal of adaptive task allocation is to design a robot controller that will allow robots to dynamically adjust the division of labor, so that the number of robots searching for *Red* and *Green* pucks will, over time, correctly reflect their prevalence. To achieve this group behavior, each robot must be able to dynamically change its foraging type.

The solution proposed by Jones and Mataric is the following. As it wanders around the arena, each robot counts the number of pucks of each type in the environment as well as the number of robots in each foraging state. It does so by observing pucks and robots that are visible to it and adding these observations to history (memory). At some time interval, the robot uses the history array to estimate the fraction of pucks and robots of each type, and changes its foraging state according to a transition function. We will specify the transition function later. In general, it depends on the estimated density of robots and pucks of a particular type: specifically, the difference between these numbers. The robots should switch its state to *Red* if there are fewer than necessary robots in the *Red* state; likewise for *Green*.

#### 3.1 Dynamical Model of Adaptive Task Allocation

In this section we describe how the task allocation problem of the previous section can be analyzed using our formalism. The goal of this paper is to present a simplified model

that captures main features of the Jones and Mataric experiments, rather than to construct an exact model that quantitatively reproduces the details of these experiments. Here we are interested in establishing the conceptual framework for thinking about adaptive robot systems. In later works we will build upon this foundation by refining the model and quantitatively comparing its predictions to experimental results.

During a sufficiently short time interval, each robot can be considered to belong to a *Green* foraging state or the *Red* state. This is a very high level, coarse-grained description. In reality, each state is composed of several robot actions and behaviors, such as wandering the arena, detecting pucks, avoiding obstacles, *etc.* However, since we want the model to capture how the fraction of robots in each foraging state evolves in time, it is a sufficient level of abstraction to consider only these states. If we find that additional levels of detail are required to explain robot behaviors, we can elaborate the model by breaking each of the high level states into its underlying components.

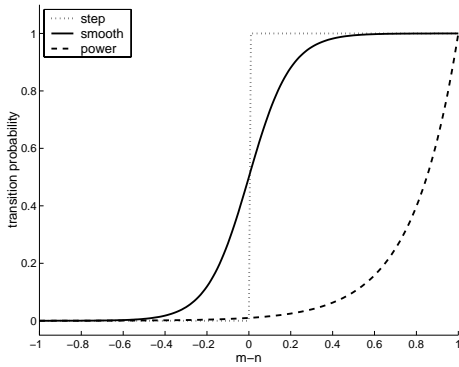
A robot uses information in its history to make a transition between states. A robot makes a transition to *Red* foraging state according to a transition function that depends on the difference between the estimated fraction of *Red* robots and *Red* pucks; otherwise it makes a transition to the *Green* foraging state.

Let  $N_R(t)$  and  $N_G(t)$  be the number of robots in *Red* and *Green* foraging states respectively at time  $t$ . Also, let  $M_R(t)$  and  $M_G(t)$  be the number of uncollected *Red* and *Green* pucks in the arena. It is important to remember that the dynamic variables correspond to quantities that have been averaged over many experiments or simulations. The following set of differential equations govern how the average numbers of robots and pucks evolve in time.<sup>1</sup>

$$\begin{aligned} \frac{dN_R}{dt} &= \alpha_R(t)N_G(t) - \alpha_G(t)N_R(t) \\ \frac{dN_G}{dt} &= \alpha_G(t)N_R(t) - \alpha_R(t)N_G(t) \\ \frac{dM_R}{dt} &= \beta_R M_R(t)N_R(t) + \mu_R \\ \frac{dM_G}{dt} &= \beta_G M_G(t)N_G(t) + \mu_G \end{aligned}$$

Quantities  $\alpha_R$  and  $\alpha_G$  govern the rate at which robots switch to *Red* and *Green* states respectively. In an adaptive system, these are time-dependent. Parameters  $\beta_R$  and  $\beta_G$  are the rate at which robots encounter *Red* and *Green* pucks, while  $\mu_R$  and  $\mu_G$  are the rates at which new *Red* and *Green* pucks are deposited in the arena. For simplicity, in the model described here, these rates are such that the total

<sup>1</sup>The differential equations describing the evolution of a dynamical system are usually derived as a continuous limit of discrete time difference equation, for example:  $N_R(t+1) = N_R(t) - \alpha_G \Delta t N_R(t) + \alpha_R \Delta t N_G(t)$ . The problem with this approach is that it models a *synchronous* system, where all robots make decision at the same time. Although feasible, such a model is not realistic; moreover, most choices of transition rates  $\alpha_R$  and  $\alpha_G$  lead to severe oscillations in the dynamic variables. The differential equations model we are working with is derived from the stochastic master equation, and is applicable to *asynchronous* systems.



**Figure 1: Transition probability vs difference between puck and robot densities for three different transition functions**

number of pucks of each type remains constant. Experimentally, this is realized by the replacing a puck in a new random location after a robot picks it up.

It is more convenient to work with the average densities rather than numbers of robots. Also, because the number of pucks does not change and does not enter the equations describing time evolution of the number of robots, we may safely ignore these equations. Dividing both sides of the first two equations above by  $N$ , the total number of robots, reduces them to:

$$\frac{dn_R}{dt} = \alpha_R n_G(t) - \alpha_G n_R(t) \quad (5)$$

$$\frac{dn_G}{dt} = \alpha_G n_R(t) - \alpha_R n_G(t), \quad (6)$$

where  $n_R(t)$  and  $n_G(t)$  are the densities of robots in each foraging state. We can also divide the number of pucks of each color by the total number of pucks, resulting in  $m_R$  and  $m_G$ , the densities of *Red* and *Green* pucks.

### 3.2 Transition Rates

Each of the Equations 5–6 is a special case of the Rate Eq.4 describing an adaptive system, with  $\alpha_R$  and  $\alpha_G$  representing the history averaged transition rates  $\langle W \rangle_h$ . At regular time intervals, the robot looks at the history of observations and estimates the density of *Red* pucks and robots in *Red* state. In general, the transition probability should be a function of  $\hat{m}_R - \hat{n}_R$ , the difference between the estimated fractions of pucks and robots in a particular state (degenerate to choice of *R* or *G*).

There are a number of options for the microscopic (per robot) transition rates. The probability of a single robot making a transition to *Red* state may be one of the following:

**Step function** robot switches to *Red* state if  $\hat{m}_R - \hat{n}_R \geq 0$ ; otherwise, it switches to *Green* state:  $f_{step} = \Theta(\hat{m}_R - \hat{n}_R)$ .

**Smooth approximation** to the step function, *e.g.*, robot switches to *Red* state with probability  $f_{smooth} = (1 +$

$\tanh(c(\hat{m}_R - \hat{n}_R)))/2$ . The function can be made arbitrarily close to the step function by increasing the value of  $c$ . In this work, we use  $c = 5$ .

**Power function** transition probability is

$f_{pow} = 100^{(\hat{m}_R - \hat{n}_R)}/100$ . This is the value used by Jones and Mataric in simulations.

In all transition probabilities above,  $\hat{m}_R$  and  $\hat{n}_R$  are the estimated densities of pucks and robots. One can think of each value as being drawn from a distribution centered on the true mean of density, with variance that depends on the length of the history window. Figure 1 shows the three possible microscopic transition probabilities.

At the collective level of Equations 5–6, the macroscopic transition rates  $\alpha_R$  and  $\alpha_G$  are in fact simply averaged microscopic transition probabilities:

$$\begin{aligned} \alpha_G &= \alpha \langle f(\hat{n}_R - \hat{m}_R) \rangle_{P(\hat{n}_R, \hat{m}_R)} \\ \alpha_R &= \alpha \langle f(\hat{n}_G - \hat{m}_G) \rangle_{P(\hat{n}_G, \hat{m}_G)} \end{aligned} \quad (7)$$

where  $\alpha$  is to assure the proper time scale,  $\langle \dots \rangle_P$  stands for averaging over the distribution  $P$ , and  $P(\hat{n}, \hat{m})$  is the joint probability that a robot has observed the fraction of robots and pucks of a corresponding color to be  $\hat{n}$  and  $\hat{m}$  respectively. Although calculating  $P(\hat{n}, \hat{m})$  goes beyond the scope of this paper we note that for sufficiently large history lengths it can be approximated by a sharply peaked distribution around its mean ( $\langle n \rangle, \langle m \rangle$ ). This suggests that if the microscopic transition functions are smooth enough (as  $f_{pow}$ , or  $f_{smooth}$  with small enough  $c$ ), then the effect of averaging is to replace the estimated values of densities with their mean values (in the case of the step function, the effect of averaging is to smear out the discontinuity).

A steady state is one in which the densities of robots in *Red* or *Green* states no longer change. Existence of the steady state is of prime interest to the robot designer, because if a system has a stable steady state, one can reliably predict its long term behavior. In the adaptive task allocation problem, the desired steady state is one in which the distribution of robots is equal to the distribution of pucks, namely,  $n_{R,ss} = m_R$  and  $n_{G,ss} = m_G$ .

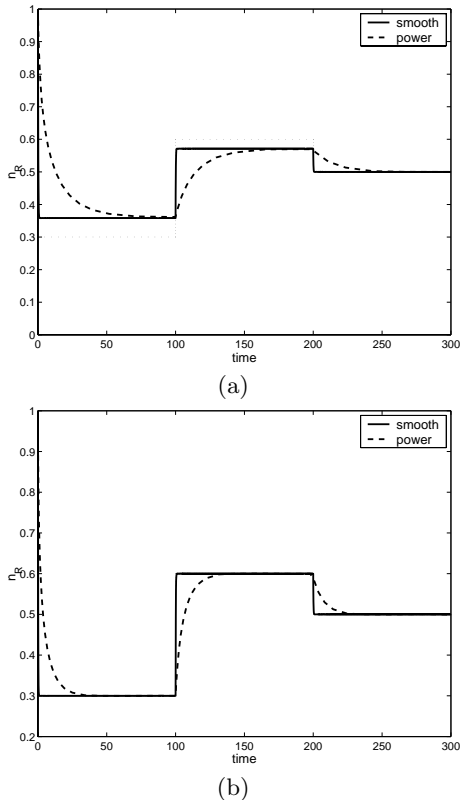
The steady state is achieved when the rhs of Eqs. 5 are each equal to zero. The total number of robots is conserved; therefore, we need to examine the behavior of one dynamic variable only, and compute the value of the other from the conservation condition:  $n_G = 1 - n_R$ . The steady state exists when  $\alpha_R n_G = \alpha_G n_R$  which can be rewritten in terms of the transition functions as

$$n_R f(n_R - m_R) = (1 - n_R) f(-n_R + m_R) \quad (8)$$

It is easy to check that Eq. 8 will have the desired solution  $n_R = m_R$  when  $f(0) = 0$ . If  $f(0) \neq 0$ , on the other hand, the steady state solution will be different from the desired one and generally speaking will depend on the details of the transition function. This can be overcome by slightly adjusting the transition rates in the following way:

$$f(n - m) \rightarrow f(n - m)/m \quad (9)$$

with the only requirement that  $f(x)$  be continuous at  $x = 0$ . With the transition rate of this form, Eq. 8 will be non-trivially satisfied when  $n_R = m_R$ .



**Figure 2: Evolution of the density of robots in the Red foraging state for different transition rates: (a) smooth and power functions shown in Fig.1, and (b) modified rates given by Eq. 8, where functions  $f$  are the same as in (a). The density of Red pucks changes from  $m_R = 0.3$  to  $m_R = 0.6$  at  $t = 100$  and again to  $m_R = 0.5$  at  $t = 200$ .**

### 3.3 Results

We studied the behavior of solutions to Eq. 5–6 for different forms of transition rates. Figure 2(a) shows the evolution of the density of robots in the Red foraging state starting from  $n_R(t = 0) = 1.0$ , for two different transition rates:  $\alpha_R = (1 + \tanh(5(m_R - n_R)))/2$  (smooth) and  $\alpha_R = 100^{(m_R - n_R)}/100$  (power function).<sup>2</sup> The density of Red pucks was initially  $m_R = 0.3$ . At time  $t = 100$  the distribution abruptly changes to  $m_R = 0.6$ , and at  $t = 200$  to  $m_R = 0.5$ . In all cases,  $n_R$  converges to a steady state where its value remains the same until the next change in the puck distribution; however, the steady state value depends on the transition rate. In fact, as we make the  $\tanh$  function steeper (by increasing  $c$ ), the steady state solution shifts and approaches the desired value  $m_R$ . Although we have made several simplifying assumptions, the model reproduces several important features of the simulations data [10], including convergence to a steady state as puck distribution

<sup>2</sup>Here and in the rest of the discussion, by  $m_R$  and  $n_R$  we mean mean values of the Red puck and Red robot densities respectively.

changes, and a shift in the steady state value from the puck distribution.<sup>3</sup>

Next, we solved equations under identical conditions, but this time choosing transition rates given by Eq. 8 (with the functional part the same as in Fig.2(a)). Results are shown in Fig. 2(b). In this case the solutions do converge to correct values, even as the distribution of pucks changes. However, convergence time still depends strongly on the transition rate — there are two orders of magnitude difference between power and smooth functions. The reason power function leads to slow convergence rate is apparent from Fig.1: transition probability has a non-negligible value only at the very beginning, while  $m_R - n_R$  is large, and is very small the rest of the time.

## 4. DISCUSSION

We have developed a general mathematical model to describe adaptation in multi-agent systems. In these systems the agents can modify their behavior in response to environmental dynamics or actions of other agents. The agents use history to estimate the global state of the system from individual observations and adjust their behaviors accordingly. We have derived a model that describes the dynamics of collective behavior in such adaptive systems. We have applied the formalism to study adaptive task allocation in mobile robots. In this application, the robots’ task is for forage for Red or Green pucks (a robot can only forage for one puck type at a time). As it travels around the arena, a robot records observations of pucks and other robots, and uses these observations to compute an estimated density of each. If it finds there are not enough robots of a specific type, it may switch its foraging state to fill the gap. After some transient, we expect the number of robots in each foraging state reflect the prevalence of those types of pucks in the environment. We modelled adaptive task allocation and studied the dynamics of the system for different transition rates. We find that an intelligent choice of transition rates will lead to fast convergence times and the desired steady state.

We made important simplifying assumptions in the model. First, we assumed very long history. Effectively, such a long window smoothes out fluctuations observed in the simulations data. In a steady state, long histories are formally equivalent to robots having full view of the arena, and thus, knowing the average densities of relevant quantities. We have extended this equivalence beyond the steady state to the transient regime. Although not fully justified, this assumption is probably approximately correct.

We plan to address the issues above by developing an exact microscopic model of the system. Such a model will enable us to directly study transients, the magnitude of fluctuations in the steady state robot density and its relationship to experimental parameters, such as history length. Another

<sup>3</sup>Jones & Mataric give results for two different microscopic transition functions: step and power functions. Because macroscopic averaging smoothes out transition rates, the microscopic step and power functions correspond to smooth and power transitions functions in the macroscopic model, although possibly with different coefficients than used in this work.

simplifying assumption we made was that of uniform puck distribution. If there are local clusters of pucks of a given type, the robot density will reflect local distribution rather than the global one. Hopefully, we will be able to better model this with a microscopic model. Obviously, much work remains to be done on relating modeling results to those of simulations. This is our next challenge.

## 5. ACKNOWLEDGEMENTS

The research reported here was supported in part by the Defense Advanced Research Projects Agency (DARPA) under contracts number F30602-00-2-0573, in part by the National Science Foundation under Grant No. 0074790. We are grateful to Chris Jones for describing the problem and sharing simulations results.

## 6. REFERENCES

- [1] R. C. Arkin. *Behavior-Based Robotics*. The MIT Press, Cambridge, MA, 1999.
- [2] R. Beckers, O. E. Holland, and J. L. Deneubourg. From local actions to global tasks: Stigmergy and collective robotics. In R. A. Brooks and P. Maes, editors, *Artificial Life IV*, pp. 181–189, Cambridge, MA, USA, July 1994. MIT Press.
- [3] K. Boehringer, R. Brown, B. Donald, J. Jennings, and D. Rus. Distributed robotic manipulation: Experiments in minimalism. In O. Khatib and J. K. Salisbury, editors, *Proc. of the Fourth Int. Symp. on Experimental Robotics, Stanford*, pp. 11–25. Springer Verlag, 1995.
- [4] T. S. Dahl, M. J Matarić, and G. S. Sukhatme. Multi-robot task-allocation through vacancy chains. In *Proc. of the 2003 IEEE Int. Conf. on Robotics and Automation (ICRA'03), Taipei, Taiwan, May 12 - 17, 2003*, 2003.
- [5] B. P. Gerkey, R. T. Vaughan, K. Sty, A. Howard, G. S. Sukhatme, and M. J Matarić. Most valuable player: A robot device server for distributed control. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS-01), Wailea, Hawaii, 2001*. <http://www-robotics.usc.edu/player/>.
- [6] D. Goldberg and M. J Matarić. Robust behavior-based control for distributed multi-robot collection tasks. Technical Report IRIS-00-387, USC Institute for Robotics and Intelligent Systems, 2000.
- [7] O. Holland and C. Melhuish. Stigmergy, self-organization and sorting in collective robotics. *Artificial Life*, 5(2), 2000.
- [8] B. A. Huberman and T. Hogg. The behavior of computational ecologies. In B. A. Huberman, editor, *The Ecology of Computation*, pp. 77–115, Amsterdam, 1988. Elsevier (North-Holland).
- [9] A. J. Ijspeert, A. Martinoli, A. Billard, and L. M. Gambardella. Collaboration through the exploitation of local interactions in autonomous collective robotics: The stick pulling experiment. *Autonomous Robots*, 11(2):149–171, 2001.
- [10] C. Jones and M. J. Matarić. Adaptive task allocation in large-scale multi-robot systems. Technical Report CRES-03-008, USC Center for Robotics and Embedded Systems, February 2003.
- [11] L. P. Kaelbling. *Learning in Embedded Systems*. PhD thesis, Stanford University, 1990.
- [12] N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry*. Elsevier Science, Amsterdam, 1992.
- [13] K. Lerman and A. Galstyan. Mathematical model of foraging in a group of robots: Effect of interference. *Autonomous Robots*, 13(2):127–141, 2002.
- [14] K. Lerman and A. Galstyan. A methodology for mathematical analysis of multi-agent systems. *submitted to Journal of Agents and Multi-Agent Systems*, 2002.
- [15] K. Lerman, A. Galstyan, Alcherio Martinoli, and Auke Ijspeert. A macroscopic analytical model of collaboration in distributed robotic systems. *Artificial Life Journal*, 7(4):375–393, 2001.
- [16] K. Lerman and O. Shehory. Coalition Formation for Large-Scale Electronic Markets. In *Proc. of the Int. Conf. on Multi-Agent Systems (ICMAS'2000), Boston, MA, 2000.*, pp. 167–174, 2000.
- [17] L. Li, A. Martinoli., and Y. Abu-Mostafa. *Emergent Specialization in Swarm Systems, Lecture Notes in Computer Science 2412*, pp. 261–266. Springer Verlag, New York, NY, 2002.
- [18] A. Martinoli. *Swarm Intelligence in Autonomous Collective Robotics: From Tools to the Analysis and Synthesis of Distributed Control Strategies*. PhD thesis, No 2069, EPFL, 1999.
- [19] A. Martinoli, A. J. Ijspeert, and L. M. Gambardella. A probabilistic model for understanding and comparing collective aggregation mechanisms. In D. Floreano, J. D. Nicoud, and F. Mondada, editors, *LNAI 1674*, pp. 575–584, Berlin, September 13–17 1999. Springer.
- [20] M. Matarić. Minimizing complexity in controlling a mobile robot population. In *Proc. of the 1992 IEEE Int. Conf. on Robotics and Automation*, pp. 830–835, Nice, France, 1992.
- [21] M. J Matarić. Reinforcement learning in the multi-robot domain. *Autonomous Robots*, 4(1):73–83, 1997.
- [22] O. Michel. Webots: Symbiosis between virtual and real mobile robots. In J.-C. Heudin, editor, *Proc. of the First Int. Conf. on Virtual Worlds, Paris, France.*, pp. 254–263. Springer Verlag, 1998.
- [23] Y. Sato and J. P. Crutchfield. Coupled replicator equations for the dynamics of learning in multiagent systems. *Physical Review*, E, 2002.
- [24] K. Sugawara and M. Sano. Cooperative acceleration of task performance: Foraging behavior of interacting multi-robots system. *Physica*, D100:343–354, 1997.