Automated Learning in Network Games

Principal Investigator: Bud Mishra

In collaboration with:

Rohit Parikh

and

Amy Greenwald

Courant Institute of Mathematical Sciences
New York University
251 Mercer Street, New York, NY 10012

May 13, 1998

Contents

1 Introduction **1**
1.1 Our Theoretical Model 2
1.2 Statement of Goals 3

2 The Failure of Rationality **4**
2.1 The Santa Fe Bar Problem 5
2.2 Best-Reply Dynamics 6

3 Tools **8**
3.1 Logical Tools 8
3.2 Experimental Tools 9
3.3 The Computational Model 9

4 Success of Bounded Rationality **10**
4.1 Simulation Results 10

5 Related Work **14**

6 Research Methodologies **15**

7 Results from Prior Support **17**
7.1 Bud Mishra 17
7.2 Rohit Parikh 19

C. Project Description

Objective

We propose to study the design and implementation of automated agents suitable for controlling and optimizing resource allocation in large-scale networks. We begin with the standard assumptions of economics and game theory, which we generalize and enhance with a framework for logical reasoning. In this way, we create models applicable in more general economic settings, as well as in network contexts, which we intend to use to analyze the behavior of intelligent agents that abide by adaptive learning algorithms.

1 Introduction

Recently, there has been a dramatic expansion of the telecommunications infrastructure, in terms of the degree of connectivity, the bandwidth of communication links, and the intelligence embedded in the switches. This growth has been paralleled by the potential of creating a new range of network capabilities, such as software agents that collaboratively mine and warehouse information on the World Wide Web, multi-media data transfer and display using shared bandwidth and buffer space, and a financial infrastructure supporting E-commerce. The challenges associated with these developments, however, are many – primarily because any realistic resource allocation scheme at this massive a scale and with such a diverse set of applications cannot rely on complete and common knowledge of network properties. As a result, the vast body of literature on distributed computing, where it is often assumed that all but a few "malicious" agents cooperate according to some commonly known and centrally mandated protocol, is frustratingly ineffective at network control.

In contrast with computer science, theories on the interaction of complex agents in dynamic and distributed environments form an integral part of mathematical economics. In particular, there exist elegant economic models that describe optimal techniques for allocating resources and coordinating behavior in multi-agent systems, where decisions are based on local, delayed, and often conflicting information. However, the underlying mathematical theories are tractable only as long as one assumes infinitely many rational and homogeneous agents [2, 5]. Furthermore, it has been observed that these economic analyses lead to efficient-market models that widely deviate from the short-term behavior patterns observed in real markets [27, 38]. While it remains an open problem to develop realistic yet mathematically rigorous market models, it is possible to build computational economies in which to study more general economic assumptions, such as lack of common knowledge, deductive and inductive rationality, and heterogeneous agents [4]. Moreover, based on the ongoing developments in the logic community – specifically non-monotonic logics that incorporate models of belief-revision [1] and the logic of games [44] – it is also possible to formalize the strategic reasoning of intelligent agents, something which standard economic theories fail to consider.

The proposed research is inherently interdisciplinary, as it lies at the intersection of recent advances in telecommunication networks, computational economics, and logic. We advocate the design and implementation of network protocols which utilize intelligent and adaptive computational agents that learn to manage resources efficiently in the face of limited information. Initially, we rely on heuristics developed by economists to model the behavior of our agents; for example, our work builds on the model of bounded rationality and inductive learning discussed in Arthur [3]. Ultimately, however, we are interested in providing a rigorous treatment of the reasoning behavior of automated agents by using belief-revision logics to model belief-based learning (*e.g.*, Bayesian updating). Our agents interact in simulated network economies that are built within a powerful software infrastructure called CAFÉ (Complex Adaptive Financial Environment), which provides a reliable tool for the evaluation of a wide class of agents. The long-term goal of the proposed research is the dissemination of automated agents (provably) suitable for control and optimization in large-scale, geographically-distributed networks, without the benefit of common knowledge of the state of the system.

1.1 Our Theoretical Model

Network control and optimization problems such as flow control and routing are resource allocation problems. Consequently, at first glance, it appears that economic theory, and in particular, the theory of games, can be directly applied to networking. Independently, others have also made this observation – in economics, see, for example, Varian [55] and Shenker [53]; in game theory, see, for example, Lazar *et. al.* [31, 36]. Upon closer scrutiny, however, it is revealed that this relationship is *not* so straightforward. In particular, fundamental assumptions on which economic theories are based are not valid in network environments – specifically, common knowledge, rationality, and homogeneity of agents. [1] One of the primary goals of the proposed research is to investigate precisely to what extent equilibrium behavior depends on these assumptions. Towards this end, we introduce our model of decision-theoretic problems (*i.e.*, games), which enhances standard economic models in two ways: (i) provides a logical framework with which to analyze the behavior of intelligent learning agents, and (ii) generalizes traditional assumptions in order to render economics applicable in network contexts.

Our model of games, in its most general form, is described as follows. We assume a finite set of (possibly heterogeneous) agents $\mathcal{N}$, where each $n \in \mathcal{N}$ is equipped with its own internal model of the decision theoretic problem it is facing. Typically, this model consists of a set of possible states of the world; or this model may be a probability

[1] In fact, it is debatable whether or not these assumptions hold true in economic environments; thus, our investigation has the potential to influence economic theory as well.

distribution over such a set. Each state is a pair consisting of (i) a picture of the agent's beliefs about the world, and (ii) a picture of how the agent believes that other agents see the world. [2] Let M be the set of all such possible models. Associated with each agent n is a set S_n of possible actions (or pure strategies) and a payoff function π_n on $S_n \times M$ such that given a model m and action s_n, the payoff is $\pi_n(s_n, m)$. Initially, each agent has an internal model m_n^0 and a belief-revision function ρ_n such that given internal model m_n^t at time t, and new information μ^{t+1} at time $t+1$, the agent's revised model after taking this information into account is $\rho_n(m_n^t, \mu^{t+1})$.

This model can be restricted for a particular domain to account for certain domain-specific knowledge. For instance, in the domain of telecommunications, an appropriately restricted model is one in which software agents are viewed as playing *network games*, where agents make decisions pertaining to the management of network resources. In such games, the underlying payoff structure is inherently unknown; in particular, agents cannot precisely describe their preferences in terms of the trade-off between throughput and congestion delays, given the increasingly dynamic nature of large-scale computer networks. Thus, learning is essential, since it provides a robust and efficient means of adapting to unexpected environmental changes. In our framework, the internal model which a network agent maintains describes what that agent believes its payoff function to be at any time, and via belief-revision techniques, the agent learns over time, thereby improving the estimate of its payoff function.

Our model is also suitable for applications in which the agents are divided into two or more classes, such as producers and consumers in an economy. In particular, the CAFÉ system incorporates our proposed framework as well as standard economics, with computational agents (*e.g.*, consumers, producers, and speculators), resources, and prices as the essential components of a CAFÉ economy. All CAFÉ agents maintain private models describing their beliefs about the future prices of the various resources. The role of an agent is to decide on some quantity of resources to demand or supply, where this decision is motivated by the intent of maximizing individual utility with respect to a private belief system. This belief system is updated regularly in the face of newly observed information. Using CAFÉ, our goal is to study the design of mechanisms by which we can influence the behavior of agents such that the (selfishly-motivated) decisions taken by individuals jointly yield globally desired properties.

1.2 Statement of Goals

Our model provides a unified framework in which to study the long-term behavior of a population of individuals, and moreover, to determine whether or not certain globally desired properties are satisfied, such as fairness, stability, and convergence to equilibrium. More specifically, the goals of our research program are as follows:

[2] Item (ii) is relevant for issues like common knowledge of rationality, but for rationality *simpliciter*, item (i) is adequate.

1. **Reasoning about Games**: Given a mathematical specification of a game (or class of games), and rules for strategic decision making for each agent specified in a game logic, to provide efficient decision procedures that determine if the asymptotic play of the game satisfies the desired properties. If it does not satisfy one or more of these properties, the procedure should provide a counterexample in terms of a sequence of plays that results in violation of the property.

2. **Learning Algorithms**: Given a mathematical specification of a game (or class of games), to provide learning algorithms based on a partially-observable history of past plays that allow agents to efficiently make optimal decisions. These algorithms will be analyzed in terms of their computational-efficiency, the complexity of the space of necessary observables, and their competitiveness (with respect to some clairvoyant agent) and/or some measure of regret (external or internal). Certain games may not yield any such learning algorithms under the usual assumptions of game theory (*e.g.*, DRIP, D = Deterministic, R = Rational, I = Information independence and P=Predictive); we shall provide an exact characterization of such games.

3. **Mechanism Design**: Given a set of desired global properties which describe the behavior of a population of agents operating under some global dynamics (expressed in terms of a set of constraints), to devise mechanisms (*i.e.*, utility functions and strategy spaces) such that the long-term behavior of agents. We shall also formulate metrics to classify these mechanisms in terms of their efficiency, convergence rate, and dependence on partial information and/or common knowledge.

These goals are successively more complex and form a natural progression; thus, we will start from the first goal and enrich our research methodology as we encounter each step.

2 The Failure of Rationality

Throughout this proposal, we use the Santa Fe bar problem as a motivating example. In its original formulation, this problem is an abstraction of the problem of designing efficient network congestion control algorithms, and moreover, in an extended form can be viewed as an abstraction of the problem of routing network packets over a system of parallel links. Simple analysis of this basic resource allocation problem reveals the short-comings of current economic theory; in particular, this section focuses on the assumption of rationality. Later, we propose solutions based on forms of bounded rationality and inductive reasoning which we validate using CAFÉ tools. Our research may also be of interest in economics since this problem arises in a number of a real-world situations, ranging from farmers polluting common water supplies, to fisherman fishing in common waters, to other versions of the tragedy of the commons [30].

2.1 The Santa Fe Bar Problem

The *Santa Fe bar problem* (SFBP) was introduced by Brian Arthur [3], an economist at the Santa Fe Institute. Here is the scenario:

> *N [say, 100] people decide independently each week whether to go to a bar that offers entertainment on a certain night ... Space is limited, and the evening is enjoyable if things are not too crowded – especially, if fewer than* 60 *[or, some fixed but perhaps unknown capacity c] percent of the the possible* 100 *are present ... a person or agent goes (deems it worth going) if he expects fewer than* 60 *to show up or stays home if he expects more than* 60 *to go. Choices are unaffected by previous visits; there is no collusion or prior communication among the agents; and the only information available is the number who came in past weeks.*[3]

Arthur first analyzed the Santa Fe bar problem assuming only that the inhabitants of Santa Fe are both rational and homogeneous. He noted the following. Let the utility of going to an uncrowded bar be equal to 1, while the utility of going to a crowded bar is equal to -1, and finally, the utility of staying home is 0, regardless of the state of the bar. Now, if an agent believes that the bar will be crowded with a probability p, then his best-reply is to go to the bar if $p < 1/2$ and to stay home if $p > 1/2$. However, since the agents are homogeneous, all their beliefs and best-reply are identical. Herein lies a paradox. If all the agents believe that the bar will be undercrowded with probability $p < 1/2$, then, in fact the bar will be empty with probability 1; in contrast, if all the agents believe that the bar will be undercrowded with probability $p > 1/2$, then the bar will be full with probability 1. [4] The conclusion is that there is no common set of best-replies and beliefs that the agents can learn over time which maximizes utility.

As mentioned earlier, SFBP is analogous to a network flow control problem which a software agent might face in deciding whether or not to transmit data at a given time. Characteristic of both the flow control problem and SFBP is the fact that since the decision of any one agent does not have significant impact on the state of the world obtained, the utility obtained by an individual agent can be viewed as arising via the effect of an *externality*.[5] An interesting extension of SFBP is choosing the precise amount of data to transmit at a given time, rather than merely deciding whether or not to transmit. Again the total flow imposes a cost on all agents which is modeled as an externality, but in this case the utility obtained is proportional to both the amount of data transmitted and the cost incurred. Economically, these network flow control problems address the management of resources in situations in the face of excess demand.

[3] The problem was inspired by the El Farol bar in Santa Fe which offers live music on Thursday nights.

[4] In the case where $p = 1/2$, agents attend the bar and stay at home with equal probability. It is straightforward to show, however, that this condition is not sustainable.

[5] An *externality* is a standard economics term used to describe third-party effects, such as pollution.

A further extension of SFBP, which we dubbed the *New York City bar problem* [15, 25], considers this problem in a city with many bars. In this case, the networking analog is a routing problem which is concerned with the choice of a route (or vector of routes) by which to transmit a fixed amount of data so as to minimize overall congestion. In contrast to the flow control problems discussed above, the network routing problem corresponds to a situation in which globally, there is excess supply. In this case, the challenge is to build automated agents that independently learn the optimal distribution channels, without the benefit of central management.

2.2 Best-Reply Dynamics

The Santa Fe bar problem is a non-cooperative game. As such, it can be expressed formally as a repeated strategic form game. The players in this game are the inhabitants of Santa Fe; notation $\mathcal{N} = \{1, \ldots, N\}$, with $n \in \mathcal{N}$. For player n, the strategy set $S_n = \{0, 1\}$, where 1 corresponds to *go to the bar* and 0 corresponds to *stay home*. The payoffs obtained by a given player depend on the particular strategic choice taken by that player and an externality. In particular, in this formulation, the Santa Fe bar game is a discretization of a simple finite externality game in the sense of Friedman [20].

Let s_n^t be the strategic choice of player n at time t and let $s_{\mathcal{N}}^t = \sum_{n \in \mathcal{N}} s_n^t$. In addition, let $c < N$ denote the capacity of the bar. The externality f depends on $s_{\mathcal{N}}^t$ and c as follows: if the bar is undercrowded (*i.e.*, $s_{\mathcal{N}}^t \leq c$), then $f(s_{\mathcal{N}}^t) = 0$; on the other hand, if the bar is overcrowded (*i.e.*, $s_{\mathcal{N}}^t > c$), then $f(s_{\mathcal{N}}^t) = 1$. Finally, let $0 < \alpha_n < 1$ denote the value to player n of attending the bar, and without loss of generality assume $\alpha_n \leq \alpha_{n+1}$. Now the payoff function for player n is given by $\pi_n(s_n^t, s_{\mathcal{N}}^t) = \alpha_n - f(s_{\mathcal{N}}^t)$, if $s_n^t = 1$, and $\pi_n(s_n^t, s_{\mathcal{N}}^t) = 0$, otherwise. The expected payoff for player n at time t is computed in terms of the true probability $p_{\mathcal{N}}^t$ that the bar is undercrowded at time t:

$$\mathcal{E}_{p_{\mathcal{N}}^t}[\pi_n(s_n^t, s_{\mathcal{N}}^t)] = \begin{cases} p_{\mathcal{N}}^t \alpha_n - (1 - p_{\mathcal{N}}^t)(1 - \alpha_n) & \text{if } s_n^t = 1 \\ 0 & \text{otherwise} \end{cases}$$

Let $p_n^* = 1 - \alpha_n$. Note that a given player n is indifferent between the two strategies whenever $p_{\mathcal{N}}^t = p_n^*$, since $\mathcal{E}[\pi_n(1, s_{\mathcal{N}}^t)] = \mathcal{E}[\pi_n(0, s_{\mathcal{N}}^t)] = 0$. The sequence of probabilities $\{p_{\mathcal{N}}^t\}$ is unknown to any one player, however, since the players operate independently. Instead, associated with each player n is a private sequence $\{p_n^t\}$ of probabilities, or beliefs, that the bar will be undercrowded at time t.

Definition 2.1 *The Santa Fe bar game is* uniform *iff for all $n \neq m \in \mathcal{N}$, $\alpha_n = \alpha_m$.* [6]

In what follows, we formalize the intuitive argument in Arthur [3] pertaining to the oscillatory behavior that arises via best-reply dynamics in the Santa Fe Bar game. We begin by explicitly defining the necessary assumptions.

[6] In [26], we consider the non-uniform case; we analyze stability results obtained via simulations.

Definition 2.2 *A given player n is said to be* rational *at time t iff*

$$s_n^t \in \arg\max_{s_n \in S_n} \pi_n(s_n, m_n^t)$$

In other words, player n is rational at time t, if the strategic action taken by player n is a best-reply to his beliefs about the state of the world (notation m_n^t). Now player n utilizes Cournot best-reply dynamics [11] iff for all times $t+1$, player n assumes that the outcome obtained during round t will be the outcome of round $t+1$, and consequently, he plays a best-reply to the outcome of round t. Note that this corresponds to the belief-revision function specified in our general model.

Definition 2.3 *A given player n is said to employ* best-reply dynamics *iff for all t, player n assumes that $m_n^{t+1} = s_{\mathcal{N}}^t$, and moreover, player n is rational. In particular, if player n utilizes best-reply dynamics, then*

$$s_n^{t+1} \in \arg\max_{s_n \in S_n} \pi_n(s_n, s_{\mathcal{N}}^t)$$

Theorem 2.1 *In the uniform Santa Fe bar game, learning via best-reply dynamics does not converge.*

Proof 2.1 (Sketch) Assume that all players employ best-reply dynamics. If $s_{\mathcal{N}}^t < c$, then the best response at time $t+1$ for all n is $s_n^{t+1} = 1$. But then $s_{\mathcal{N}}^{t+1} > c$, and now the best response at time $t+2$ for all n is $s_n^{t+2} = 0$. Now, once again $s_{\mathcal{N}}^t < c$. This patterns repeats itself indefinitely, generating oscillatory behavior that is far from the desired equilibrium. Finally, note that if ever $s_{\mathcal{N}}^t = c$, this situation cannot persist since it gives rise to an unstable equilibrium. In particular, the best response to $s_{\mathcal{N}}^t = c$ is mixed strategy $(1/2, 1/2)$, but $\Pr[s_{\mathcal{N}}^{t+1} = \ldots = s_{\mathcal{N}}^{t+k} = c] \to 0$ as $k \to \infty$.

In this section, we demonstrated that it is inconsistent to conclude that rational and homogeneous agents in the Santa Fe bar problem are capable of even very weak forms of belief-based learning. Moreover, if all players are rational, even if they learn via Bayesian updating, play still does not converge to equilibrium behavior [26]. These negative results are part of a more general phenomenon also noted by Nachbar [42] and Foster and Young [19] who state that repeated play of normal form games among rational players does not converge to a Nash equilibrium, unless players' initial beliefs coincide with an equilibrium. Our negative result states that in the Santa Fe bar game, assuming the stated conditions, no learning algorithm will ever converge to Nash equilibrium, even if players' initial beliefs coincide with an equilibrium. This theorem can be further generalized to show that the weaker property of calibrated beliefs [13] can also never be achieved. Since calibrated learning gives rise to correlated equilibrium [17], no learning algorithm will ever converge to the more general concept of correlated equilibrium either.

The formalization presented in this section revealed that Arthur's original intuition was founded on the interplay of homogeneity and rationality. However, this analysis failed to incorporate any notion of knowledge or reasoning on the part of the agents. In the next section, we discuss the logical and experimental tools which we intend to use to resolve the paradox inherent in the bar problems.

3 Tools

We intend to use two types of tools in the pursuit of our research program. In particular, we draw from a set of logical tools, with which we plan to investigate informational requirements, and a set of experimental, or computational, tools with which we study boundedly rational, heterogeneous agents.

3.1 Logical Tools

There are well-studied logics of knowledge applicable in distributed contexts which can be used to calculate what knowledge an agent has at a given time and how an agent should use new information to revise its current model. These logics enter into our model in two important ways.

In [44, 43], the cake cutting algorithm is used as a paradigm for a game-theoretic situation in which every player has a winning strategy. In this problem, a set of agents must divide a cake among themselves in a manner which is clearly seen by all the agents to be fair. The algorithm ensures that no agent may succeed in satisfying its greed at the expense of fairness for the other agents. This is possible because although it seems that the greed of some agents may prevent other agents from obtaining their fair share, in fact, the goals of the different agents are compatible. A logic of games is used to prove correctness of this algorithm. This logic is closely related to dynamic logic, but it differs in that each action is associated with an agent. There is a complete axiomatization and decidability result for the propositional case (see [44]).

Logical considerations also enter our model at the stage of belief revision. In a purely logical framework, the Alchourron, Gärdenfors, and Makinson (AGM) axioms describe boundary conditions on revision functions which various agents may describe [1]. Their model is suitable for investigating belief revision when newly received information is incompatible with current beliefs. The AGM axioms have been studied extensively, and various completeness results have been proven (see, for example, Grove [28]). The revision process is necessarily non-monotonic, and Grove uses an ordering among possible models to achieve this non-monotonicity. In [47], it is shown how the revision process can be made more efficient by revising only that portion of an agent's beliefs which are in the same domain of information as the new information; *e.g.*, new information obtained from the dentist about the health of one's teeth should not affect beliefs about Alan Greenspan.

In a purely arithmetical situation, the results contained in [50] and [46] describe how new (but consistent) information results in the revision process, leading ultimately to stability and consensus. These results come as a sort of culmination of a strain of thought which began with the fundamental paper by Aumann [6], with further results by Bacharach [7], Cave [9], and Geanakoplos and Polemarchakis [24]. Typically, an agent works with knowledge rather than belief so that the actual state of the world is among the agent's possible models. As the agent receives more information, the set of models is culled and the information becomes more refined. This theme is also followed in [46]. Moreover, in [48], some of these techniques are extended to fuzzy situations.

3.2 Experimental Tools

For experimentation purposes, we intend to use a Java-based system called CAFÉ [16] (Complex Adaptive Financial Environment), designed for the simulation of complex adaptive systems. Examples of such complex interactions occur in market economies, biological systems, and potentially among Internet software agents. The prototype of the CAFÉ system is described in detail in Even and Mishra [16]. Inherent in this system is an object-oriented design that provides an easily extensible framework in which to simulate the behavior of agents that interact in complex ways. The top-level structure is the `Agent` class, of which the sub-classes `Patron` and `Bar` are applicable in simulations of the bar problems. In the study of the bar problems, two types of agents are derived, namely patrons and bars, both of which are modeled as boundedly rational economic agents that act in a way so as to maximize utility.

The mathematical description of CAFÉ closely resembles classical economic models. Consequently, standard economic equilibrium analysis, assuming for example continuous and differentiable utility functions, can be applied to CAFÉ. The computational model of CAFÉ, however, is far richer in that it does not depend on these assumptions, and moreover, it does not assume that economic agents are either homogeneous or rational. Since modeling of networks games using traditional economic assumptions may give rise to control protocols that result in undesirable behavior, we utilize a computational model in the CAFÉ system which allows for heterogeneous and boundedly rational agents, and is suitable for network control.

3.3 The Computational Model

The computational model of CAFÉ is based on standard economic models; however, this model incorporates an important extension, namely the introduction of a private belief system for each agent. Moreover, it is assumed that agents act in a way so as to maximize utility with respect to their beliefs. The belief systems of the agents are represented by sets of predictor functions, which gives rise to agents that exhibit bounded rationality in the sense discussed in Arthur [3].

This section describes the particular choices of predictor functions that comprise the computational model of CAFÉ in terms of the bar problems, where the producers in the economy are the bars, and the consumers are the patrons. The agents employ bounded rationality and inductive learning to predict the attendance at the bars. In CAFÉ, the bounded rationality of the agents is modeled by a pool of simple functions which utilize historical data to predict attendance at the bar. For example, some predictor functions are trend-based, while others depend on adaptive expectations, while still others make constant predictions. The predictor functions correspond to the beliefs of agents in belief-based learning models.

Let $G = \{g_1, \ldots, g_K\}$ denote the set of predictor functions. Initially, the agents (randomly) select a fixed number of predictor functions from the pool; say $\{g_{i1}, \ldots, g_{ik}\}$ is the selection for patron i, with $k << K$. Throughout the simulation, the agents monitor the accuracy of their predictor functions. Any predictor functions that are consistently inaccurate are discarded and replaced with alternatives from the original pool. At time t, the attendance predicted at the bars by patron i is the output of the currently most accurate predictor function, say g_i^{*t}. Let h_j^t denote the attendance history at bar j through time t. If $g_i^{*t}(h_j^t) \geq c$, then the optimal predictor function for patron i at time t predicts excess demand at bar j. Otherwise, If $g_i^{*t}(h_j^t) < c$, then the optimal predictor function for patron i at time t predicts excess capacity at bar j.

4 Success of Bounded Rationality

In light of the paradox inherent in SFBP, Arthur's insight was to propose the study of boundedly rational agents who use inductive learning to build expectational models. By simulating the behavior of such agents, Arthur obtained an efficient solution to SFBP in which the overall attendance at the bar stabilized near capacity. In contrast to standard solutions to distributed resource allocation problems which involve pricing congestible resources [55], this approach depends only on learning to play equilibrium strategies in repeated games. The CAFÉ system, however, is suited for the study of externality effects as well as direct pricing mechanisms in repeated games. In this section, we present the results of preliminary simulations of the bar problems which demonstrate the potential of our model to afford solutions to general problems of resource allocation in decentralized environments.

4.1 Simulation Results

This section presents the results of preliminary simulations of the bar problems. In the original SFBP, there is one bar of capacity 60, and there are 100 patrons. Figure 1 plots the attendance over time of agents that employ bounded rationality. Note that attendance stabilizes near the capacity of the bar.

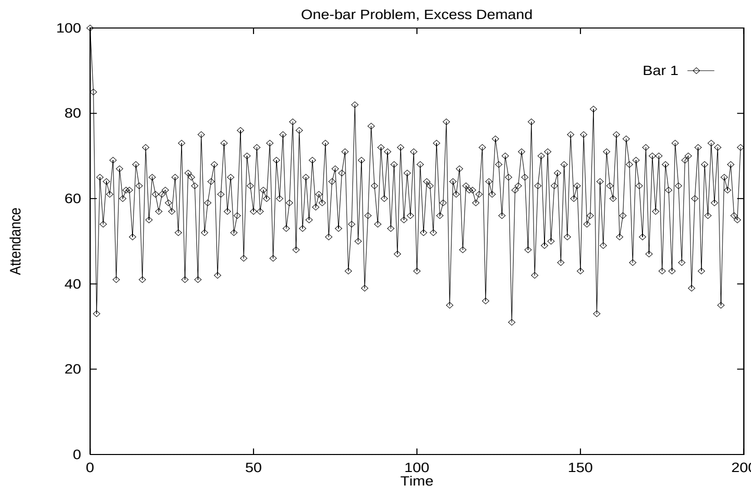

Figure 1: Attendance vs. time at one bar assuming excess demand

We now consider the two bar problem. Initially, this problem is studied in two scenarios, the first assuming excess supply (*i.e.*, more seats in the bars than patrons), and the second assuming excess demand (*i.e.*, more patrons than seats). Assuming excess supply, we find that boundedly rational learning is sufficient for obtaining an efficient solution to the two-bar problem. This scenario corresponds to a network routing problem in which the capacity of the network exceeds the (fixed) demand of the users. The results of simulations of this problem assuming 100 patrons and two bars, each of capacity 60, are depicted in Figure 2. Note that the system reaches equilibrium with each bar attracting a population in the neighborhood of 50 patrons.

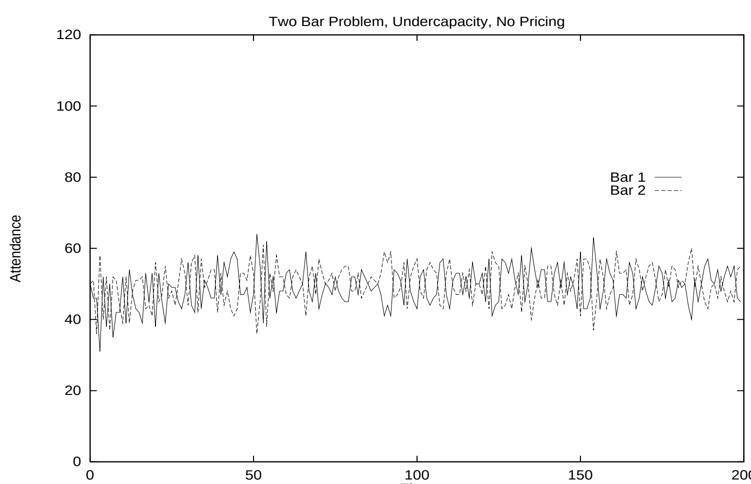

Figure 2: Attendance vs. time at two bars assuming excess capacity

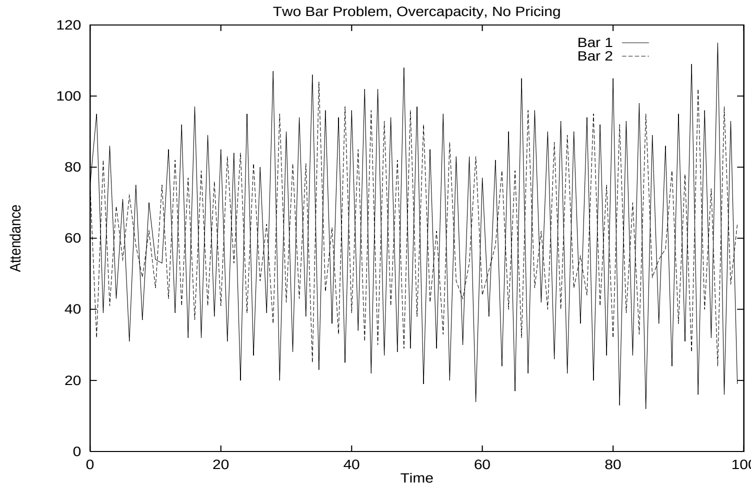

Figure 3: Attendance vs. time at two bars assuming excess demand

The ever-growing interest in the Internet, however, shows that it may not be realistic to assume an excess supply of network resources. Since our research is geared towards network applications, the second scenario that is considered assumes excess demand, but in this case the results are not so promising. In particular, bounded rationality is *not* sufficient for obtaining an efficient solution to the two-bar problem assuming excess demand. For simulation purposes, the population is set at 150 and the capacity of each bar is equal to 60. The results are presented in Figure 3. While this mechanism seems to be adequate for either population control at one bar or population routing between two bars, it is insufficient for achieving both simultaneously.

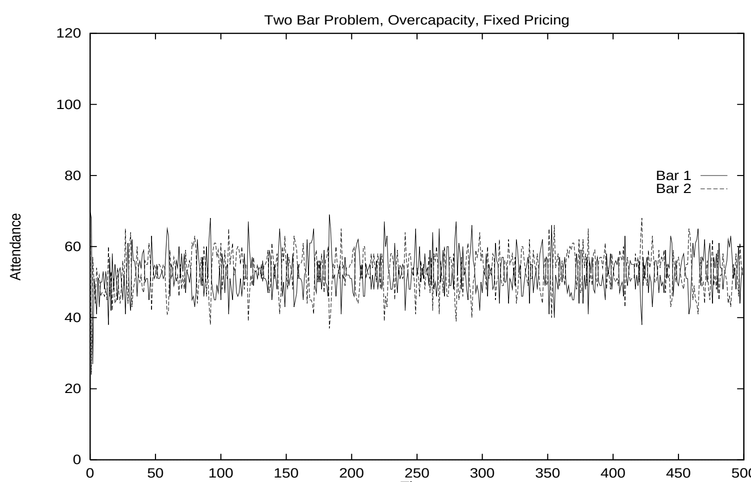

Figure 4: Attendance vs. time at two bars with fixed pricing and excess demand

We address this problem by introducing pricing into our simulations in order to curb demand. We simulate the two-bar problem assuming two distinct pricing regimes: fixed and competitive pricing. The results demonstrate the ability of pricing to control demand as well as the potential of pricing as a tool for guaranteeing different qualities of service in multi-media networks.

Two scenarios are considered. In the first case, the price of both bars is set to 1 (see Figure 4). This scenario demonstrates that the introduction of pricing in fact successfully controls demand. Enough of the patrons are "priced out" of the market so that when the remaining patrons choose bars using bounded rationality the system reaches equilibrium. In the second scenario, the price of one bar is set to 1, while the price of the second bar is set to 2 (see Figure 5). Note that the more expensive bar draws approximately half the population of the cheaper bar. These simulation results demonstrate the potential of pricing as a tool for guaranteeing different qualities of service in multi-media networks, since higher priced services are requested less often, thereby making it possible to meet maximal delay requirements.

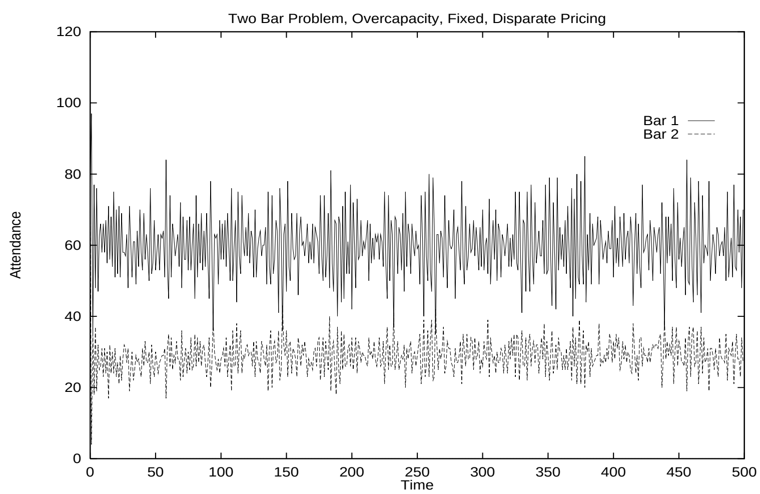

Figure 5: Attendance vs. time at two bars with varied pricing and excess demand

The final simulations of the two-bar problem consider competitive pricing, a situation in which each bar independently varies its price. Each bar maintains its price as long as its population is within a pre-defined neighborhood of its capacity. If it is below this range it lowers its price, and if it is above, it raises its price. Figure 6 shows that the bars reach equilibrium with a population of approximately 50. Figure 7 show the prices of the bars as they vary with time. Note that when the bars reach equilibrium, their prices remain fixed.

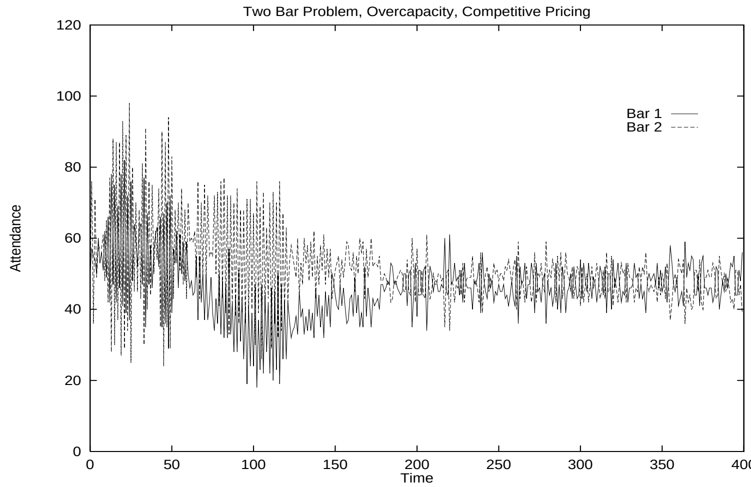

Figure 6: Attendance vs. time at two bars with competitive pricing and excess demand

The simulations presented in this section exemplify the relationship between the many aspects of our research program. In particular, we are both interested in the design of automated intelligent agents which exhibit optimal performance in network environments, as well as the design of mechanisms which motivate agents to behave efficiently. These problems are inherently intertwined; we propose to attack them by incorporating aspects of networking, learning in repeated games, and logic.

5 Related Work

Presently, there is an expanding body of literature on theory and systems which apply economic ideas to control and optimization problems in networking. For example, Varian and MacKie-Mason [55] describe simple pricing mechanisms that induce the efficient use of network resources, from the point of view of network service providers. In addition, Shenker [54] advocates the use of implementation theory (see, for example, Myerson [41]), in an effort to induce socially desirable network operating points in the presence of non-cooperative and self-interested users. Moreover, Korilis *et. al.* [35] study pricing mechanisms as a means of achieving Pareto optimality. Regarding systems, SPAWN is a computational economy that was designed at Xerox PARC to manage and coordinate distributed tasks on multiple processors [56]. Similarly, WALRAS is an asynchronous distributed system developed jointly at Michigan and MIT which operates via a market pricing mechanism [57]. In contrast, we propose to conduct conduct extensive studies of learning algorithms and mechanism design using CAFÉ, and moreover, we later intend to disseminate automated agents which abide by these algorithms on a large-scale network, such as the Internet.

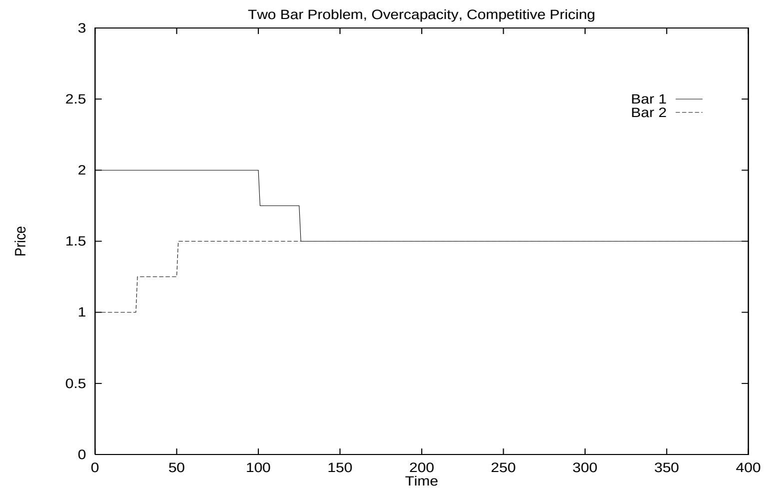

Figure 7: Price vs. time at two bars with competitive pricing and excess demand

There is also a vast literature on learning through repeated play of games, and we make no attempt here to provide a detailed review; see the review by Fudenberg and Levine [23] for a comprehensive discussion. The work on learning falls roughly into two camps. The "high-rationality" approach involves learning algorithms which aim to predict the strategies of their opponents, and myopically optimize with respect to those predictions. The prediction methods can be Bayesian (as in Kalai and Lehrer [33]), calibrated (as in Foster and Vohra [18]), or consistent (as in Fudenberg and Levine [22, 21]). Typically the asymptotic play of such algorithms are either correlated or Nash equilibria. Since these algorithms depend on knowledge of the underlying structure of the game, they are not applicable in the network contexts which we are considering here. In contrast, the "low-rationality" approaches to learning are concerned with situations similar to that which we consider here; in particular, agents have no information other than the payoffs which they receive for the strategies they employ. For examples of such work, see Roth and Erev [52], Erev and Roth [14], Borgers and Sarin [8], Mookerji and Sopher [40], and Van Huyck *et. al.* [32]. The focus of these papers is typically on matching the results of human experiments. We focus instead on the nature of asymptotic play.

6 Research Methodologies

The Santa Fe bar problem forms the bare bones of an analogy between economics and network control and optimization. However, in the absence of a well-understood solution to this problem, this relationship cannot be exploited. Thus, we are interested in gaining a thorough understanding of the paradox inherent in SFBP, and further applying this understanding to solving the generalized class of bar problems and other more general problems which are characterized by our model.

The main outcomes of our research are expected to be (i) a thorough understanding of the foundations for verification, learning schemes and mechanism design for network games, and (ii) enhancement of experimental tools, namely CAFÉ, and implementation of the resulting algorithms on large scale networks. The research on the foundational work will be disseminated through publications in journal and conference proceedings. The work on implementation will be carried out and evaluated in collaboration with an industrial partner. We have set for ourselves the following research milestones:

Year 1: Foundational work on the logic of games. Design and implementation of verification algorithms. Test software on the current version of CAFÉ.

Year 2: Foundational work on learning schemes. Implementation of selected subset of these algorithms. Perform experiments using these algorithms on enhanced version of CAFÉ.

Year 3: Foundational work on mechanism design. Design efficient mechanisms for network games. Empirical verification with CAFÉ. Finally, transfer the technology to a large-scale network such as the Internet.

7 Results from Prior Support

7.1 Bud Mishra

1. *Reactive Algorithms in Robotics*: B. Mishra. 1995–98, *National Science Foundation* (IRIS), IRI-9414862, $ 228,993.
2. *CISE Research Instrumentation*: P. Dasgupta, Z. Kedem, B. Mishra, K. Palem and D. Shasha. 1995-96, *National Science Foundation* (CISE), CDA-9421935, $ 75,429.

Ph.D. Theses

[**1**] This dissertation developed under the supervision of B. Mishra studies a number of problems arising in the context of grasping and fixturing, various measures of goodness of the grasp and a new algorithmic approach ("reactive robotics") to build specialized grippers.

[**2**] This dissertation developed under the supervision of B. Mishra focuses on the automatic development of hybrid (combining discrete as well as continuous plant models).

The resulting "control compiler," called CONTROL-D, has been successfully used to build a controller for a food-manufacturing system and a gait-controller for a walking machine.

The following additional Ph.D. theses have been developed under the supervision of B. Mishra and deal with real-time system applications for robotics (C. Frenandes, G. Koren, N. Silver, L. Salkind and D. Clark), algorithmic algebra (G. Gallo and P. Pedersen), complexity of computational logic (L. Ericson), theory of learning (P. Caianiello) and parallel debugging systems (A. Dinning).

Other Results

[**5, 6, 7, 8, 9**] These series of papers develop the theory underlying the CONTROL-D system, used for automatic development of hybrid controllers (combining discrete as well as continuous plant models).

[**10, 11, 12, 13**] These series of papers develop the theory and applications of the "reactive robotics" scheme that can be successfully used to build grippers and hands, where sensor values are used to determine immediate actuation in a simple table driven manner. Several grippers based on these ideas have been successfully implemented.

[**14, 15, 16**] These and some related sequence of papers study the problem of analyzing and synthesizing efficient and optimal grasps with multi-fingered robot hands and clamps (the fixturing problem). A related problem studied deals with the problem of continuously reorienting a grasped object. Recent work also deals with the so-called "reactive" algorithms with special applications to manufacturing.

References

[1] M. Teichmann. *Grasping and Fixturing: a Geometric Study and an Implementation*, Ph.D. Thesis, Courant Institute of Mathematical Sciences, New York University, New York, 1995.

[2] M. Antoniotti. *Synthesis and Verification of Controllers for Robotics and Manufacturing Devices with Temporal Logic and the Control-D System*, Ph.D. Thesis, Courant Institute of Mathematical Sciences, New York University, New York, 1995.

[3] B. Mishra. "Computational Real Algebraic Geometry," **CRC Handbook of Discrete and Computational Geometry**, (Edited by J.E. Goodman and J. O'Rourke), CRC Series, Discrete and Combinatorial Mathematics, CRC Press, 1997.

[4] B. Mishra. "Grasp Metrics: Optimality and Complexity," **Algorithmic Foundations of Robotics**, (Edited by K. Goldberg, D. Halperin, J.-C. Latombe and R. Wilson), pp. 137–166, A.K. Peters, Wellesley, Massachusetts, 1995.

[5] B. Mishra. "Hybrid Controllers for Robotics." *Robotics and Manufacturing: Recent Trends in Research and Applications*, **Vol. 6**, (Eds: M. Jamshidi, F. Pin and P. Daucher), *The Proceedings of International Symposium on Robotics and Manufacturing, ISRAM*, World Automation Congress (Montpellier, France) pp. 435–440, ASME Press, New York, 1996.

[6] M. Antoniotti and B. Mishra. "$\mathcal{NP}$-completeness of the Supervisor Synthesis Problem for Unrestricted CTL Specifications." *Workshop on Discrete Event Systems*: WODES '96, Edinburgh, Scotland, U.K. August 19–21, 1996.

[7] M. Antoniotti, B. Mishra and M. Jafari. "Applying Temporal Logic Verification and Synthesis to Manufacturing Systems." IEEE System Man and Cybernetics Conference, Vancouver, Canada. 1995.

[8] M. Antoniotti and B. Mishra. "Discrete Event Models + Temporal Logic = Supervisory Controller: Automatic Synthesis of Locomotion Controllers," *1995 IEEE International Conference on Robotics and Automation*: ICRA'95, Nagoya, Japan, May 21-27, 1995.

[9] M. Antoniotti and B. Mishra. "Automatic Synthesis Algorithms for Supervisory Controllers," *Proceedings of the Fourth International Conference on Computer Integrated Manufacturing and Automation Technology*, pp. 151–156, Troy, NY, October 10–12, 1994.

[10] F. Hansen, M. Adeshnik B. Mishra and M. Teichmann, "Implementing A Reactive Robotic Gripper," *1998 NSF Design and Manufacturing Grantees Conference*, Monterrey, Mexico, January 5–8, 1998.

[11] M. Teichmann and B. Mishra. "Reactive Algorithms for 2 and 3 Finger Grasping," *Proceedings of the 1994 International Workshop on Intelligent Robots and Systems*: IRS'94, Grenoble, (France), July 11–14, 1994.

[12] B. Mishra and M. Teichmann. "Three Finger Optimal Planar Grasp," *Proceedings of the 1994 International Workshop on Intelligent Robots and Systems*: IRS'94, Grenoble, (France), July 11–14, 1994.

[13] M. Teichmann and B. Mishra. "Reactive Algorithms for Grasping Using a Modified Parallel Jaw Gripper," *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*: ICRA'94, San Diego, California, May 8–13, 1994.

[14] M. Teichmann and B. Mishra. "Probabilistic Algorithms for Efficient Grasping and Fixturing," Special Issue: Algorithms in Robotics, (Guest editor: R. Motwani and P. Raghavan) Accepted *Algorithmica*, Springer-International, 1997.

[15] M. Teichmann and B. Mishra. "The Power of Friction: Quantifying the "Goodness" of Frictional Grasps." **Algorithms for Robotic Motion and Manipulation**, (Edited by J.P. Laumond and M. Overmars), pp. 311–320, A.K. Peters, Wellesley, Massachusetts, 1997.

[16] D. Kirkpatrick, B. Mishra and C.-K. Yap. "Quantitative Steinitz's Theorem with Applications to Multifingered Grasping," *Discrete & Computational Geometry*, Springer-Verlag, New York, pp. 295–318,Volume **7**, Number 3, 1992.

7.2 Rohit Parikh

1. "A Logical Study of Distributed Transition Systems", with Lodaya, Ramanujam and Thiagarajan. *Information and Computation* **119** May 1995, 91-119.

We study a number of logics which formalize transition systems and prove decidability undecidability and completeness results.

2. "Vagueness and Utility: the Semantics of Common Nouns" in *Linguistics and Philosophy* **17** 1994, 521-35.

We point out that to date there do not exist satisfactory logics or semantics for vague predicates. We show that these predicates are person dependent, i.e. the way they are applied varies from person to person and also from occasion to occasion. Hence a theory is needed of why they are useful in communicatio and do not lead to misunderstandings. We show how there are settings where despite some differences in application by the various individuals involved, communication is useful. These are the settings in which we do in fact use these predicates, avoiding them in other areas where such sturdiness does not obtain.

3. "Logical omniscience", in *Logic and Computational Complexity* Ed. Leivant, Springer Lecture Notes in Computer Science no. 960, (1995) 22-29.

Current logics of knowledge have the property that under their definition of what it means for i to know some formula A, i knows all valid formulas and also the consequences of anything that i knows. This is implausible and to find more plausible definitions of knowledge is the problem of logical omniscience. We make some algorithim based suggestions.

4. "Language as social software" (abstract) International Congress on Logic, Methodology and Philosophy of Science (1995), page 417. To appear in *Future Pasts*, Ed. Floyd and Shieh, Harvard U. Press, 1998.

One can view language as playing the role of a system of signals to facilitate social behaviour. It turns out that this view is very flexible and can explain various philosophical puzzles like Searle's Chinese room puzzle or Quine's indetermincy of translation thesis.

5. "Knowledge based computation (Extended abstract)" in *Proceedings of AMAST-95* Montreal, July 1995, Edited by Alagar and Nivat, LNCS no. 936, 127-42.

A short survey of work in this area done to date.

6. "Topological Reasoning and The Logic of Knowledge" (with Dabrowski and Moss) *Annals of Pure and Applied Logic* **78** (1996) 73-110

While it is true that one's knowledge depends on one's evidence, traditional definitions of knowledge leave out the fact that one can *gather* or improve one's knowledge. E.g. a measurement of some quantity can be made more accurate by using better instruments. This observation allows us to develop a logic with two modalities, one for knowledge and the other for effort. Some topological notions like closed or perfect can be defined in this logic. We prove axiomatizations and provide completeness results.

7. "How far can we formalize language games?" in *The Foundational Debate* edited by DePauli-Scimanovich, Köhler and Stadler, Kluwer Academic (1995) pp. 89-100.

Wittgenstein's views in the Philosophy of Mathematics are examined and shown to be very modern in spirit. We raise the question how far one can provide formal versions of language games as a way of making certain problems more explicit.

8. "Vague predicates and language games", *Theoria* Spain, vol XI, no. 27, Sep 1996, pp. 97-107.

Further research along the lines of #2, above.

9. "Belief revision and language splitting" to appear in *Proc. ITALLC*, CSLI 1998.

The celebrated AGM axioms for belief revision allow the trivial revision under which all old information is lost. We show how we can incorporate a formal notion of relevance which allows one's information to be split uniquely into a number of disjoint subject areas. Revising information only in those areas where new information is received blocks the trivial revision.

References

[1] C. Alchourron, P. Gärdenfors, and D. Makinson. On the logic of theory change: partial meet contraction and revision functions. *Symoblic Logic*, 50:510–530, 1985.

[2] K. Arrow and F. Hahn. *General Competitive Analysis*. North-Holland, Amsterdam, 1971.

[3] W.B. Arthur. Inductive reasoning and bounded rationality. *Complexity in Economic Theory*, 84(2):406–411, May 1994.

[4] W.B. Arthur, J.H. Holland, B. LeBaron, R. Palmer, and P. Taylor. Asset pricing under endogenous expectations in an artificial stock market. *Santa Fe Institute Technical Report*, 96–12–093, 1996.

[5] R. Aumann. Markets with a continuum of traders. *Econometrica*, 32:39–50, 1964.

[6] R. Aumann. Agreeing to disagree. *Annals of Statistics*, 4:1236–1239, 1976.

[7] M. Bacharach. Some extensions of a claim of aumann in an axiomatic model of knowledge. *Journal of Economic Theory*, 37:167–190, 1985.

[8] T. Borgers and R. Sarin. Learning through reinforcement and replicator dynamics. Mimeo, 1995.

[9] J.A.K. Cave. Learning to agree. *Economic Letters*, 12:147–152, 1983.

[10] M. Chandy and J. Misra. How processes learn. In *4th ACM Conference on Principles of Distributed Computing*, pages 204–214, 1985.

[11] A. Cournot. *Recherches sur les Principes Mathematics de la Theorie de la Richesse*. Hachette, 1838.

[12] A. Darwiche and J. Pearl. On the logic of iterated theory revision. In R. Fagin, editor, *Theoretical Aspects of Reasoning about Knowledge*, pages 5–22. Morgan Kaufmann, 1994.

[13] A.P. Dawid. The well-calibrated Bayesian. *Journal of the American Statistical Association*, 77(379):605–613, 1982.

[14] I. Erev and A. Roth. On the need for low rationality cognitive game theory: reinforcement learning in experimental games with unique mixed strategy equilibria. Mimeo, 1996.

[15] R. Even. *Market Models and Multi-Agent Systems*. Ph.D. Dissertation, New York University, New York, Expected August, 1998.

[16] R. Even and B. Mishra. CAFÉ: A complex adaptive financial environment. In *Conference of Computational Intelligence for Financial Engineering*, pages 20–25, March 1996.

[17] D. Foster and R. Vohra. Calibrated learning and correlated equilibrium. *Preprint*, 1995.

[18] D. Foster and R. Vohra. Regret in the on-line decision problem. *Preprint*, 1997.

[19] D. Foster and P. Young. Learning with hazy beliefs. *Preprint*, 1996.

[20] E. Friedman. Learnability in a class of non-atomic games arising on the internet. *Mimeo*, 1998.

[21] D. Fudenberg and D. Levine. Conditional universal consistency. Mimeo, 1995.

[22] D. Fudenberg and D. Levine. Consistency and cautious fictitious play. *Journal of Economic Dynamics and Control*, 19:1065–1089, 1995.

[23] D. Fudenberg and D. Levine. *Theory of Learning in Games*. Mimeo, 1996.

[24] J. Geanakoplos and H. Polemarchakis. We can't disgraee forever. *Journal of Economic Theory*, 28:192–200, 1982.

[25] A. Greenwald. Learning to play network games. Technical Report 958, New York University, May 1997.

[26] A. Greenwald, B. Mishra, and R. Parikh. The Santa Fe bar problem revisited: Theoretical and practical implications. *Unpublished Manuscript*, April 1998.

[27] S. Grossman. On the efficiency of competitive stock markets where traders have diverse information. *Journal of Finance*, 31:573–585, 1976.

[28] A. Grove. Two modellings for theory change. *Philosophical Logic*, 17:157–170, 1988.

[29] J. Halpern and Y. Moses. Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3):549–587, July 1990.

[30] G. Hardin. The tragedy of the commons. *Science*, 162:1243–1248, 1968.

[31] M.T. Hsiao and A. Lazar. A game theoretic approach to decentralized flow control of markovian queueing networks. In Courtois and Latouche, editors, *Performance'87*, pages 55–73. North-Holland, 1988.

[32] J. Van Huyck, R. Battalio, and F. Rankin. Selection dynamics and adaptive behavior without much information. Mimeo, 1996.

[33] E. Kalai and E. Lehrer. Rational learning leads to Nash equilibrium. *Econometrica*, 61:1019–1045, 1993.

[34] H. Katsuno and A. Mendelzon. Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52:263–294, 1991.

[35] Y. A. Korilis, T. A. Varvarigou, and S. R. Ahuja. Pricing non-cooperative networks. *Submitted to the IEEE Transactions on Networking*, May 1997.

[36] Y.A. Korilis, A. Lazar, and A. Orda. The designer's perspective to noncooperative networks. In *Infocom*, Boston, April 1995.

[37] P. Krasucki, R. Parikh, and G. Ndjatou. Probabilistic knowledge and probabilistic common knowledge. *ISMIS*, pages 1–8, 1990.

[38] M. Kurz. On the structure and diversity of rational beliefs. *Journal of Economic Theory*, 1994.

[39] D. Lehmann. Belief revision, revised. In *Proceedings of 14th IJCAI*, pages 1534–1541, 1995.

[40] D. Mookherjee and B. Sopher. Learning behavior in an experiemental matching pennies game. *Games and Economic Behavior*, 7:62–91, 1994.

[41] R.B. Myerson. *Game Theory: Analysis of Conflict*. Harvard University Press, Cambridge, 1991.

[42] J. Nachbar. Prediction, optimization, and learning in games. *Preprint*, 1996.

[43] R. Parikh. Propositional logics of programs: New directions. *FCT-83 LNCS*, #158:347–359.

[44] R. Parikh. The logic of games and its applications. *Annals of Discrete Mathematics*, 24:111–140, 1985.

[45] R. Parikh. Levels of knowledge in distributed computing. In *IEEE Symposium on Logic in Computer Science*, pages 314–321, June 1986.

[46] R. Parikh. Finite and infinite dialogues. In Moschovakis, editor, *Workshop on Logic from Computer Science*, pages 481–498. MSRI Publications, 1991.

[47] R. Parikh. Logical omniscience. In D. Leivant, editor, *Logic and Computational Complexity*, pages 22–29. Springer Lecture Notes in Computer Science LNCS #960, 1995.

[48] R. Parikh. Vague predicates and language games. *Theoria*, XI(27):97–107, September 1996.

[49] R. Parikh. Belief revision and language splitting. In *To Appear in ITALLC*, 1998.

[50] R. Parikh and P. Krasucki. Communication, consensus and knowledge. *Journal of Economic Theory*, 52:178–189, 1990.

[51] R. Parikh and R. Ramanujam. Distributed processes and the logic of knowledge. In R. Parikh, editor, *Logics of Programs*, volume 193, pages 256–268. Springer Lecture Notes in Computer Science, 1985.

[52] A. Roth and I. Erev. Learning in extensive form games: experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, 8:164–212, 1995.

[53] S. Shenker. Efficient network allocations with selfish users. In P.J.B. King, I. Mitrani, and R.J. Pooley, editors, *Performance '90*, pages 279–285. North Holland, New York, 1995.

[54] S. Shenker. Making greed work in networks: A game-theoretic analysis of switch service disciplines. *IEEE/ACM Transactions on Networking*, 3:819–831, 1995.

[55] H. Varian. Pricing the Internet. In *Public Access to the Internet*, JFK School of Government, 1993.

[56] C.A. Waldsburger, T. Hogg, B.A. Huberman, J.O. Kephart, and W.S. Stornetta. SPAWN: A distributed computational economy. IEEE Transactions on Software Engineering, 18:103–117, 1992.

[57] M. Wellman and J. Cheng. The WALRAS algorithm: A convergent distributed implementation of general equilibrium outcomes. *Submitted for Publication*, 1996.