The Santa Fe Bar Problem Revisited:
Theoretical and Practical Implications

Amy Greenwald *†

Courant Institute of Mathematical Sciences, New York University

251 Mercer Street, New York, NY 10012

`amygreen@cs.nyu.edu`

Bud Mishra

Courant Institute of Mathematical Sciences, New York University

251 Mercer Street, New York, NY 10012

`mishra@cs.nyu.edu`

Rohit Parikh

Brooklyn College and CUNY Graduate Center

33 W. 42nd Street, New York, NY 10036

`parikh@csp.gc.cuny.edu`

December 1, 1998

Part I

Please do not cite or distribute without permission of authors

Keywords: repeated games, learning, rationality, computational agents, Russell's paradox, mixed strategy Nash equilibrium

*The authors would like to thank Ron Even and Dean Foster for insightful conversations regarding this material and related topics.

†The work of the first author was partially supported by the American Association of University Women. All three authors gratefully acknowledge the support of NSF Grant #EIA-9873469.

Abstract

This paper investigates the Santa Fe (*i.e.*, El Farol) bar problem from both a theoretical and a practical perspective. Theoretically, it is shown that belief-based learning (*e.g.*, Bayesian updating) yields unstable behavior in this repeated game. In particular, rationality and predictivity, two conditions sufficient for convergence to Nash equilibrium, are inherently incompatible. On the practical side, it is demonstrated via simulations that computational learning algorithms in which agents occasionally act irrationally do indeed give rise to near-equilibrium behavior.

1 Introduction

The *Santa Fe bar problem* (SFBP) was introduced by Brian Arthur [1], an economist at the Santa Fe Institute, in the study of bounded rationality and inductive reasoning. 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 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.* [1]

SFBP is as a non-cooperative, repeated game in which utility functions are determined by negative externalities. As such, SFBP can be viewed as an abstraction of the general problem of sharing resources of limited capacity, where the only interaction among agents occurs through the joint use of shared resources. This game-theoretic model is applicable to a number of real-world situations, ranging from fishermen fishing in common waters, to farmers polluting common water supplies, to various other versions of the tragedy of the commons [17]. Moreover, the set of applications is rapidly expanding, as game-like scenarios are emerging in the telecommunications infrastructure, where network links and queues serve as shared resources (see, for example, [19, 26]) as well as web sites and other shared databases (see, for example, [13]). In contrast with the recent trend towards proposing solutions to distributed resource allocation problems by pricing congestible resources [20, 21], Arthur suggests that bounded rationality together with learning can yield solutions to problems of resource allocation in decentralized environments, using SFBP as an example. Similarly, this paper advocates an approach to resource sharing which depends only on learning to play equilibrium strategies in repeated games. In particular, it is shown that strict rationality precludes learning, while ϵ-rationality, which may be interpreted as bounded rationality, facilitates learning.

[1] The problem was inspired by the El Farol bar in Santa Fe in which there is live Irish music on Thursday nights.

Consider the following analysis of SFBP under the standard economic assumption of rationality. We define an *undercrowded* bar as a bar in which attendance is less than or equal to c, and define an *overcrowded* bar as a bar in which attendance is strictly greater than c. For the sake of argument, let the utility of going to an undercrowded bar be $1/2$ and the utility of going to an overcrowded bar be $-1/2$; in addition, the utility of staying at home is 0, regardless of the state of the bar. [2] If a patron predicts that the bar will be undercrowded with probability p, then his best response is to go to the bar if $p > 1/2$ and to stay home if $p < 1/2$. [3] Now, if the patrons indeed learn to predict probability p accurately, then their predictions eventually come to match the actual probability that the bar is undercrowded, as it is determined by their (possibly mixed) strategic best responses. Herein lies a contradiction. If the patrons predict that the bar will be undercrowded with probability $p < 1/2$, then, in fact the bar will be empty with probability 1; on the other hand, if the patrons predict that the bar will be undercrowded with probability $p > 1/2$, then the bar will be full with probability 1. The conclusion is that rational patrons cannot learn via repeated play to make accurate predictions. In particular, rational behavior precludes learning.

This paradoxical outcome is arrived at via a diagonalization process not unlike Russell's paradox. [4] [25] Just as the truth of being in Russell's set depends on the fact of being in the set, the value of going to the bar depends on the act of going to the bar. Consider a bar of capacity 1 in a world of 2 potential patrons, say A and B. At equilibrium, at least one of A or B attends the bar. Assume that it is known that A is going to the bar; this ensures that attendance at the bar is at least 1. Given that the attendance at the bar is exactly 1, the bar is deemed undercrowded. Therefore, rationality dictates that B is to go to bar as well. If B does indeed go to the bar, however, the bar becomes overcrowded, at which point rationality dictates that B is to stay at home. In other words, rationality suggests that B should

[2]The definition of the utility of staying at home as 0, regardless of the state of the bar, can be replaced, without changing the argument, by: the utility of staying at home as $1/2$, whenever the bar is overcrowded, and the utility of staying at home as $-1/2$, whenever the bar is undercrowded.

[3]In the case where $p = 1/2$, the patrons are indifferent between attending the bar and staying home and may behave arbitrarily: *e.g.*, go to the bar with probability q. If $c \neq qN$, we show that this condition is incompatible with learning.

[4]Russell's set is defined as the set of all sets that are not elements of themselves: *i.e.*, $\mathcal{R} = \{X | X \notin X\}$. Note that $\mathcal{R} \in \mathcal{R}$ iff $\mathcal{R} \notin \mathcal{R}$.

go to the bar iff B should not go to the bar. Similarly, if it were known that B was surely to attend to the bar, by rationality, A should go to the bar iff A should not go to the bar.

The aforementioned paradox similarly arises in the two-player game of matching pennies, where player 1 aims to *match* player 2, while player 2 aims to *mismatch* player 1. In fact, matching pennies can be viewed as a special case of SFBP in which there are both positive and negative externalities: if player 1 prefers to go to the bar only when player 2 attends as well, while player 2 prefers to go to the bar only when player 1 stays at home, then player 1 is the matcher while player 2 is the mismatcher. In matching pennies, if player 1 plays *heads*, then player 2 prefers *tails*, but then player 1 prefers *tails*, at which point player 2 prefers *heads*. It follows that player 2 prefers *heads* iff player 2 prefers *tails*. Similarly, for player 1. The logical conflicts that arise in SFBP and in the game of matching pennies are a direct consequence of the fact that these games have no pure strategy Nash equilibria. [24] In order to resolve these paradoxes, game-theorists allow for the use of mixed strategies. In particular, the unique Nash equilibrium in matching pennies is for both players to play each of *heads* and *tails* with probability 1/2, while the unique Nash equilibrium in SFBP, is for all players to go to the bar with probability c/N and to stay at home with probability $1 - c/N$.

1.1 Overview

This paper is organized into two parts, the first of which concentrates on the theoretical implications of the Santa Fe bar problem, and the second on the practical implications. The first half presents a negative result on convergence to Nash equilibrium in SFBP which formalizes the diagonalization argument that is discussed in the introduction. In SFBP and other games that have a unique mixed strategy Nash equilibrium, two sufficient conditions for convergence to Nash equilibrium are *rationality* and *predictivity*. Rationality implies that players always choose (possibly mixed) strategic best responses to their beliefs. Predictivity is one way in which to capture the notion of learning: [5] a player is said to be predictive if that player's beliefs eventually coincide with the truth. If the players learn to predict (*i.e.*, if their beliefs do

[5] Learning can also be understood in terms of various merging properties, as described in Kalai and Lehrer [18].

indeed converge to actual probabilities) [6], then the (unique) best responses to their beliefs also converge, and moreover, these strategies constitute a Nash equilibrium. On the other hand, if beliefs do not converge, then strategies do not converge, and therefore play does not converge to a Nash equilibrium, unless the players irrationally settle on a mixed strategy Nash equilibrium, irrespective of their beliefs. In what follows, we argue that if the players employ predictive learning algorithms, then assuming that they are rational, their beliefs do not converge, and therefore play does not converge to Nash equilibrium in SFBP. In other words, *if play converges to Nash equilibrium, then play is either not rational or not learned.*

Our negative result complements the theorems of Foster and Young [9] and Nachbar [23]. Foster and Young [9] show that if players' beliefs are deterministic, then in two-player games with unique mixed strategy Nash equilibria, predictive learning algorithms cannot yield rational behavior; in other words, two conditions which together are sufficient for convergence to Nash equilibrium are deemed incompatible. Conversely, the main theoretical result in this paper states that rational behavior cannot yield predictivity in the SFBP, an N player game which has a unique mixed strategy Nash equilibrium. In addition, Nachbar [23] argues that unless players' initial beliefs magically coincide with a Nash equilibrium, repeated play of strategic form games among Bayesian rational players does not in general converge to a Nash equilibrium. Our negative result states that in SFBP, regardless of the initial conditions, no means of rational learning will ever converge in beliefs, and consequently, rational learning also does not converge in strategies to Nash equilibrium.

The second half of this paper aims to resolve the paradoxes of the first half via simulation experiments in computational learning. In light of the suite of negative theoretical results, it appears necessary to relax one or more of the usual economic assumptions in order to approach equilibrium behavior. More specifically, we can relax either rationality or the implicit assumption that beliefs are deterministic. In a previous simulation paper [7], using a versatile Java-based system called CAFÉ [8], we considered the case of agents with non-deterministic belief systems. In this paper, we study learning among agents who are not perfectly rational; instead they are afflicted with a trembling hand, and therefore occasionally exhibit irrational behavior. Our simulation

[6] The existence of actual probabilities is guaranteed by the fact that players employ mixed strategies.

results counter the negative theoretical results of the first half of this paper by demonstrating that computational learning in SFBP – where agents are on occasion irrational – yields convergence to near-equilibrium behavior with attendance centering around the capacity of the bar. Moreover, we extend our study to a special case of the so-called New York City bar problem [7] in which there are two bars, and we observe similar convergence results.

2 Theoretical Implications

This section formalizes the Santa Fe bar problem in terms of the theory of learning in repeated games. It is proven that traditional models of belief-based learning (*e.g.*, Bayesian updating) among rational players do not in general give rise to equilibrium behavior in this game. In particular, it is shown no rational, belief-based learning algorithm yields both convergence and predictivity. Stated differently, if players are rational and if they utilize predictive learning algorithms, then beliefs do not converge, which implies that play does not converge to Nash equilibrium.

2.1 Formalization

The Santa Fe bar problem is a non-cooperative, repeated game. In this section, we formally define both the one-shot strategic form game, and the corresponding repeated game. The players 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* while 0 corresponds to *stay home*. Let Q_n denote the set of probability distributions over S_n, with mixed strategy $q_n \in Q_n$. The expected payoffs obtained by player n depend on the particular strategic choice taken by player n, the value to player n of attending the bar, and a negative externality, which are defined as follows.

Let s_n denote the realization of the mixed strategy q_n of player n; thus, $s = \sum_{n \in \mathcal{N}} s_n$ is the realized attendance at the bar. In addition, let $c < N$ denote the capacity of the bar. The externality f depends on s and c as follows: if the bar is undercrowded (*i.e.*, $s \leq c$), then $f(s) = 0$; on the other hand, if the bar is overcrowded (*i.e.*, $s > c$), then $f(s) = 1$. 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 for pure strategies $s_n \in S_n$ is given by $\pi_n(s_n, s) = \alpha_n - f(s)$, if $s_n = 1$, and $\pi_n(s_n, s) = 0$, otherwise. [7] As usual, the expected payoffs $\mathbb{E}_{q_n}[\pi_n(s_n, s)]$ obtained by player n via mixed strategy q_n are given by $\mathbb{E}_{q_n}[\pi_n(s_n, s)] = \sum_{s_n \in S_n} q_n(s_n)\pi_n(s_n, s)$. In this formulation, SFBP is a discretization of an ordered externality game in the sense of Friedman [11, 12].

The repeated SFBP over T periods is defined as: $(\mathcal{N}, (S_n, \pi_n)_{n\in\mathcal{N}})^T$. A history h^t of length $t \geq 0$ is a sequence of t outcomes drawn from the set $S = \{0, 1, 2, \ldots, N\}$. [8] In particular, the history $h^t = (s^1, \ldots, s^t)$ indicates the number of players who attended the bar during periods 1 through t. Let h^0 denote the null history, let H^t denote the set of all histories of length t, and finally, let $H = \bigcup_0^\infty H^t$. A belief-based learning algorithm is a function from the set of all possible histories to the set of possible beliefs. We assume that beliefs in the repeated SFBP take the form of aggregate statistics, with a belief as a subjective probability over the space of possible externality effects $\mathcal{E} = \{\textit{undercrowded}, \textit{overcrowded}\}$. [9] The event *undercrowded* obtains whenever $s^t \leq c$, and otherwise, the event *overcrowded* obtains. Let $\Delta(\mathcal{E})$ be the set of probability distributions over the set $\mathcal{E}$. Formally, a belief-based learning algorithm is a function $g : H \to \Delta(\mathcal{E})$. Since the event space $\mathcal{E}$ is of cardinality 2, the private sequence of probability distributions $\{(p_n^t, 1 - p_n^t)\}$ is denoted simply $\{p_n^t\}$, where p_n^t is the probability that player n attributes to the bar being undercrowded at time t.

The expected payoff for player n at time t is computed in terms of the beliefs that player n holds at time t:

$$\mathbb{E}_{p_n^t}[\pi_n(s_n, s)] = \begin{cases} p_n^t \alpha_n - (1 - p_n^t)(1 - \alpha_n) & \text{if } s_n = 1 \\ 0 & \text{otherwise} \end{cases}$$

Let $p_n^* = 1 - \alpha_n$. Now player n is indifferent between his two pure strategies whenever $p_n^t = p_n^*$, since this implies that $\mathbb{E}_{p_n^t}[\pi_n(1, s)] = \mathbb{E}_{p_n^t}[\pi_n(0, s)] = 0$. In order to simplify notation, in what follows we often write $\pi(q_n, s^t)$ for

[7] Our results also hold in the case where $\pi_n(s_n, s) = f(s) - \alpha_n$, if $s_n = 0$.

[8] Implicit in this notion of history is the assumption that players are anonymous: *i.e.*, they cannot distinguish other players, and moreover, they also cannot distinguish between themselves and others. This precludes any notion of correlated beliefs, in which players might attempt to correlate the behavior of an individual, such as oneself, with overall attendance at the bar.

[9] Implicit in this belief structure is the assumption that beliefs are *deterministic* in the sense of Foster and Young [9]. In particular, we do not consider beliefs of the form: with probability P_n^t, player n assigns probability p_n^t to the bar being undercrowded at time t, while with probability $P_n'^t$, player n assigns probability $p_n'^t$ instead.

$\mathbb{E}_{q_n}[\pi(s_n, s^t)]$ and $\pi(q_n, p_n^t)$ for $\mathbb{E}_{p_n^t}[\pi(q_n, s^t)]$. The actual probability that the bar is undercrowded at time t as determined by the players' mixed strategies is denoted p_0^t. The sequence of objective probabilities $\{p_0^t\}$ is unknown to any one player, however, since the players operate independently. [10]

Definition 2.1 SFBP *is* uniform *iff for all* $n, m \in \mathcal{N}$, $\alpha_n = \alpha_m \equiv \alpha$, *and consequently,* $p_n^* = p_m^* \equiv p^*$. [11]

2.2 Best-Reply Dynamics

This section formalizes an argument pertaining to the oscillatory behavior that is well-known to arise via Cournot best-reply dynamics [4] in congestion games that resemble the Santa Fe bar problem. We begin by reminding the reader of the assumptions implicit in best-reply dynamics.

Definition 2.2 *A strategy* $q_n^t \in Q_n$ *is said to be a* best-reply *for player* n *at time* t *iff*

$$q_n^t \in \arg\max_{q_n \in Q_n} \pi_n(q_n, p_n^t)$$

In other words, strategy q_n^t is a best-reply for player n at time t iff it is utility maximizing. Now player n utilizes best-reply dynamics iff for all times $t+1$, player n assumes that the outcome that is realized during round t will again be the outcome of round $t+1$, and consequently, she plays a best-reply to the outcome of round t.

Definition 2.3 *A given player* n *is said to employ* best-reply dynamics *in* SFBP *iff for all* t, *player* n *assumes that*

$$p_n^{t+1} = \begin{cases} 1 & \text{if } s^t \leq c \\ 0 & \text{otherwise} \end{cases}$$

and moreover, player n *plays only best-replies to these beliefs. In particular, if player* n *utilizes best-reply dynamics, then*

$$q_n^{t+1} \in \arg\max_{q_n \in Q_n} \pi_n(q_n, s^t)$$

[10] We often refer to the sequence of actual probabilities $\{p_0^t\}$ as objective probabilities, but technically, this is a misnomer, since these probabilities are not truly independent of the individual players' decisions. They might rather be termed intersubjective.

[11] In what follows, it is assumed that SFBP is uniform, although a weaker form of this assumption does suffice. Uniformity, however, serves to simplify the technical presentation.

Theorem 2.1 *In the uniform* SFBP*, learning via best-reply dynamics does not converge:* i.e.*, for all players n,*

$$\lim_{t,t' \to \infty} \mid p_n^t - p_n^{t'} \mid \neq 0$$

Proof 2.1 Assume that all players employ best-reply dynamics. If $s^t \leq c$, then $p_n^t = 1$ for all n, to which the best response at time $t+1$ is pure strategy $s_n^{t+1} = 1$. But then $s^{t+1} > c$, so that $p_n^{t+1} = 0$ for all n, to which the best response at time $t+2$ is pure strategy $s_n^{t+2} = 0$. Now, once again $s^t < c$. This pattern repeats itself indefinitely, generating oscillatory behavior that is far from the desired equilibrium. In particular, $\lim_{t,t' \to \infty} \mid s^t - s^{t'} \mid \neq 0$, from which it follows that beliefs do not converge.

Recall that the SFBP falls into the class of ordered externality games. For such games, if best-reply dynamics converge, then it is known that the serially undominated set is a singleton [12, 22]. Consistent with this result, the serially undominated set obtained in the SFBP is not a singleton – on the contrary, it includes all the strategies of the game – and moreover, best-reply dynamics do not converge. Note, however, that stability results have been achieved regarding best-reply dynamics for special cases within the class of ordered externality games, some of which resemble non-uniform versions of SFBP [11].

2.3 Bayesian Learning

This section presents a generalization of the results obtained in the previous section. It is argued that if all players are rational and if they learn according to Bayes' rule, then play does not converge to equilibrium behavior. In fact, this result is not contingent on the assumption of Bayesian learning and is readily applicable to any effective belief-based learning mechanism in which players are rational.

Definition 2.4 *Given $\epsilon > 0$, a belief-based learning algorithm is said to be* ϵ-predictive *for player n iff it generates a sequence of probabilistic beliefs $\{p_n^t\}$* s.t.*:*

$$\lim_{t \to \infty} \mid p_n^t - p_0^t \mid < \epsilon$$

In other words, if player n utilizes a predictive learning algorithm, then player n's subjective beliefs p_n^t converge to the objective probabilities p_0^t. Notice that this definition does not require that the objective probabilities themselves converge, only that player n's subjective probabilities converge to an approximation of the objective probability.

Definition 2.5 *Given $\epsilon > 0$, a set of players $\mathcal{N}$ reaches* ϵ-consensus *iff for all $n, m \in \mathcal{N}$,*

$$\lim_{t \to \infty} \mid p_n^t - p_m^t \mid < \epsilon$$

Lemma 2.1 *If a set of players $\mathcal{N}$ is ϵ-predictive, then the players reach 2ϵ-consensus.*

Proof 2.2 Suppose not: *i.e.*, suppose that there exists some player m *s.t.* $p_m^t > p_n^t + 2\epsilon$, for all players $n \neq m$. Now it follows by the assumption of ϵ-predictivity that $p_m^t > p_0^t + \epsilon$ for players $n \neq m$. But this contradicts the assumption that player m is ϵ-predictive. On the other hand, suppose $p_m^t < p_n^t - 2\epsilon$, for all players $n \neq m$. Now it follows by the assumption of ϵ-predictivity for players $n \neq m$ that $p_m^t < p_0^t - \epsilon$. But this once again contradicts the assumption that player m is ϵ-predictive. Thus, the players do indeed reach 2ϵ-consensus.

Definition 2.6 *A player is* rational *iff she utilizes only best-replies.*

The following key lemma states that in the uniform SFBP, whenever the players exhibit both ϵ-predictive learning and rationality, if beliefs converge, then they converge in the range $(p^* - 2\epsilon,\ p^* + 2\epsilon)$.

Lemma 2.2 *In the uniform* SFBP*, if the players are rational and if they employ ϵ-predictive learning algorithms, then for all n, if beliefs converge,*

$$\lim_{t \to \infty} \mid p_n^t - p^* \mid < 2\epsilon$$

Proof 2.3 Suppose not: *i.e.*, suppose that there exists some player m *s.t.* $p_m^t > p^* + 2\epsilon$. By Lemma 2.1, we know that the players reach 2ϵ-consensus. It follows that for all n, $p_n^t > p^*$. Now by rationality, all players play best-replies, namely go to the bar: *i.e.*, $s_n^t = 1$, for all n. But this ensures that the bar will be overcrowded with probability 1, yielding $p_0^t = 0 < p_n^t - \epsilon$, for small values of ϵ. Contradiction.

On the other hand, if there exists some player m *s.t.* $p_m^t < p^* + 2\epsilon$, then by Lemma 2.1, namely the fact that the players reach 2ϵ-consensus, it follows that for all n, $p_n^t < p^*$. In this case rationality implies that all players stay at home: *i.e.*, $s_n^t = 0$, for all n. But this ensures that the bar will be undercrowded with probability 1, yielding $p_0^t = 1 > p_n^t + \epsilon$, for small values of ϵ. Contradiction.

The above lemma suggests that in the uniform SFBP, if players exhibit ϵ-predictive learning and rationality, then their beliefs converge to a range in which they are nearly indifferent between their pure strategies. A question arises as to what players might do in this range of indifference. Traditionally, game theorists assume that at the point of indifference play is arbitrary; after all, since players are indifferent, all mixed strategies are rational. We accept a slightly stronger view of the world, namely that true indifference implies nothing less than the flip of a fair coin, because if on the contrary a player were to favor one strategy or another, then that player would not be behaving as if he were indifferent. Our view is simultaneously weaker than traditional assumptions, since we allow for a small range of indifference, rather than a single point of indifference.

The following theorem presents the main theoretical result of this paper, namely that in SFBP, no mechanism of rational, ϵ-predictive, belief-based learning (including Bayesian updating) gives rise to objective probabilities, as determined by players' mixed strategies, that converge near the players' range of indifference. Assuming learning is predictive, it immediately follows that players subjective probabilities also do not converge near the range of indifference. Now since rational players only mix (*i.e.*, flip a fair coin) when their beliefs indeed lie within this range of indifference, and moreover, since the unique Nash equilibrium of SFBP is mixed, we conclude that rational learning does not converge to Nash equilibrium, unless the capacity of the bar fortuitously lies between $N/2 - k_1\sqrt{N}$ and $N/2 + k_2\sqrt{N}$, for $k_1, k_2 > 0$.

Theorem 2.2 *In the uniform* SFBP, *if the players are rational and if they employ ϵ-predictive belief-based learning algorithms, then*

$$\lim_{t\to\infty} \mid p_0^t - p^* \mid > \epsilon$$

provided that players flip a fair coin whenever $\mid p_n^t - p^* \mid < 2\epsilon$, [12] *and that the*

[12] Mathematically, this result holds for arbitrary probabilities p_n, p_m *s.t.* $\mid p_n - p_m \mid < \delta$,

capacity of the bar satisfies and $c \notin \{0, N\}$ *and*

$$c \leq N/2 - \sqrt{[1+\ln(1/(1-\alpha))]N} \quad \textit{and} \quad c \geq N/2 + \sqrt{[3+3\ln(1/\alpha)][N/2]}$$

Proof 2.4 Suppose not: *i.e.*, suppose that $\lim_{t\to\infty} \mid p_0^t - p^* \mid < \epsilon$. Since all players are ϵ-predictive learners, for all n, $\lim_{t\to\infty} \mid p_n^t - p_0^t \mid < \epsilon$. By Lemma 2.2, this implies that $\lim_{t\to\infty} \mid p_n^t - p^* \mid < 2\epsilon$. In this range, players are near-indifferent, and they consequently flip a fair coin. It follows that attendance at the bar is binomially distributed $\sim S(N, 1/2)$. Two distinct cases arise, depending on the capacity of the bar. Note that we exclude the degenerate cases in which $c \in \{0, N\}$. This proof utilizes the multiplicative variant of the Chernoff bound [3].

Case 2.3.9.1 Assume that

$$c \leq N/2 - \sqrt{[1+\ln(1/(1-\alpha))]N}$$

In this case,

$$\begin{aligned} p_0^t &= \Pr[S(N,1/2) < c] \\ &= \Pr[S(N,1/2) < \{1 - \sqrt{[4+4\ln(1/(1-\alpha))]/N}\}(N/2)] \\ &\leq e^{-([4+4\ln(1/(1-\alpha))]/2N)(N/2)} \\ &= (1-\alpha)/e \\ &< 1-\alpha-\epsilon \end{aligned}$$

for small values of ϵ. Therefore, $p_0^t < 1-\alpha-\epsilon = p^* - \epsilon$. Contradiction.

Case 2.3.9.2 Assume that

$$c \geq N/2 + \sqrt{[3+3\ln(1/\alpha)][N/2]}$$

In this case,

$$\begin{aligned} 1 - p_0^t &= \Pr[S(N,1/2) > c] \\ &= \Pr[S(N,1/2) > \{1 + \sqrt{[6+6\ln(1/\alpha)]/N}\}(N/2)] \\ &\leq e^{-([6+6\ln(1/\alpha)]/3N)(N/2)} \\ &= \alpha/e \\ &< \alpha - \epsilon \end{aligned}$$

for small values of $\delta > 0$, where p_n and p_m denote the probabilities that players n and m, respectively, go to the bar. We do not present this case, however, since this assumption is more difficult to justify.

for small values of ϵ. Therefore, $p_0^t > 1 - \alpha + \epsilon = p^* + \epsilon$. Contradiction.

The following corollary states that no mechanism of rational, learning gives rise to subjective probabilities that converge to objective probabilities in SFBP, if the players' beliefs converge. The contrapositive of this corollary reads: if the players exhibit predictive learning, then their beliefs do not converge. By rationality, it follows that play does not converge to Nash equilibrium.

Corollary 2.1 *In the uniform* SFBP*, if the provisos of Theorem 2.2 hold, and if the players are rational and their beliefs converge, then no belief-based learning algorithm exhibits $\epsilon/3$-predictivity.*

Proof 2.5 Suppose not: *i.e.*, assume that the players employ ϵ-predictive learning algorithms. By definition, for all n, $\lim_{t\to\infty} \mid p_0^t - p_n^t \mid < \epsilon/3$. By Lemma 2.2, if beliefs converge, then $\lim_{t\to\infty} \mid p_n^t - p^* \mid < 2/3\epsilon$. Now, we see that $\lim_{t\to\infty} \mid p_0^t - p^* \mid < \epsilon$, which contradicts the proof of Theorem 2.2. Therefore, no belief-based learning algorithm exhibits predictivity in SFBP.

This section demonstrated that it is inconsistent to conclude that belief-based learning among rational players in the Santa Fe bar problem exhibits properties of learning. In light of these negative results, it appears necessary to relax one or more of the usual economic assumptions in order to approach equilibrium behavior.

2.4 Generalizations

The above theorem relies on the assumption that rather than a single point of indifference, there is a small band. Without this assumption, we construct the following counterexample to Theorem 2.2.

- Let G players hold beliefs $p^* + \frac{1}{t}$. Notice that these players beliefs converge to p^*, but by rationality, these players *always* go to the bar.
- Let H players hold beliefs $p^* - \frac{1}{t}$. Notice that these players beliefs converge to p^*, but by rationality, these players *never* go to the bar.
- Let I players hold beliefs p^* at all times t. These players are indifferent and flip a fair coin every period.

Assume the capacity of the bar $c = G + d$, in order to accommodate the G players who go to the bar in every period, with some additional capacity d to accommodate the indifferent players who go to the bar some of the time. Suppose indifferent players go to the bar whenever their coin flip shows heads. Now the probability that the bar is undercrowded at time t is the probability that with I flips of a coin, at most d heads appear. In particular,

$$p_0^t = \sum_{j=1}^{d} \frac{\binom{I}{j}}{2^I}$$

In this scenario, if $p^* = p_0^t$, then players are both rational and predictive, and moreover, their beliefs converge to the true probabilities. Note that this can also be viewed as a constraint on α, since $\alpha = 1 - p^*$; in particular, if

$$\alpha = \sum_{j=d+1}^{I} \frac{\binom{I}{j}}{2^I}$$

then once again, players are both rational and predictive, and moreover, their beliefs converge to the true probabilities. This leads us to the following conjecture: In the uniform SFBP, if the players are all rational, then belief-based learning algorithms are unlikely to exhibit ϵ-predictivity, provided that players flip a fair coin only when they are truly indifferent, the capacity of the bar $c \neq G + I/2$ (roughly), and $\alpha \neq \sum_{j=d+1}^{I} \binom{I}{j}/2^I$.

3 Practical Implications

In this section, we study learning among computational agents who are not perfectly rational; instead, at times, they are irrational, as they implement a trembling hand. The results of simulations counter the negative theoretical results of the first half of this paper by demonstrating that computational learning in the Santa Fe bar game – where agents are on occasion irrational – yields convergence to near-equilibrium behavior with attendance centering around the capacity of the bar. The results obtained in this section are of particular interest because they demonstrate that the theory of learning in repeated games affords solutions to problems of resource allocation in decentralized environments.

The Santa Fe bar paradigm is applicable in a wide range of practical areas, such as network control and optimization and financial management. For example, the Santa Fe bar problem is analogous to a network flow control problem which a set of software agents might face in deciding whether or not to transmit information over a communication link. In this scenario, if the network informs the agents that delays due to congestion are minor, then all the agents might decide to transmit, causing congestion to in fact be major; but now, if the congestion delays are perceived to be substantial, then all agents might decide *not* to transmit, causing congestion to once again be minimal; and so on. Similarly, this problem also parallels an investment scenario which a set of computational agents might face in deciding whether or not to buy a certain security. If the market price is low, then all the agents might decide to *buy*, but this increase in demand in turn causes an increase in market price; now if the market price is high, then all the agents might decide to *sell*, but this increase in supply causes the market price to fall once again. This pattern repeats itself indefinitely in this naive implementation of computational investors.

An interesting extension of this problem, dubbed the *New York City bar problem* [6, 15], considers this problem in a city with many bars. In this case, the networking analog is a routing problem which concerns the choice of a route (or vector of routes) by which to transmit a fixed amount of data so as to minimize congestion. In financial terms, this problem corresponds to the management of an investment portfolio. This section describes simulations of both the Santa Fe bar problem in its original form, and the New York city bar problem in the case of two bars. In our experiments, oscillatory behavior is avoided and instead convergence to near-equilibrium is achieved.

3.1 Learning Algorithms

The simulation experiments discussed in this section were conducted using non-Bayesian learning algorithms because it is not possible to use Bayesian updating in any realistic setting due to the complexity of prior probabilities. According to these algorithms, players do not maintain prior probabilities, or beliefs, regarding the space of their opponents' strategies or their payoff structures. Instead, they *explore* their own strategy space by initially playing all strategies with non-zero probability, and they *exploit* successful strategies by increasing the probability of employing those strategies that generate high payoffs.

The learning algorithms considered in this paper include an extension [13] of the additive updating procedure studied by Roth and Erev [5], and two related multiplicative updating procedures due to Freund and Schapire [10] and Auer, Cesa-Bianchi, Freund, and Schapire [2]. These learning algorithms are simulated in two distinct settings. In the first, so-called *informed* settings, complete information is available regarding payoffs, including the surmised payoffs of strategies which are not in fact played; in the second, so-called *naive* settings, the only information pertaining to payoff functions that is available at a given time is the payoff of the strategy that is in fact employed at that time. [14]

Notation

The learning algorithms are described from the point of view of a single player engaging in a game against nature, such that the role of the externality is assumed to be played by nature. [15] The player chooses strategic actions from strategy set S with $i, j \in S$. The payoff awarded at time t by employing strategy i is denoted π_i^t. The learning algorithm that determines strategic choices is described in terms of a time-dependent vector of weights, namely $w^t = (w_i^t)_{i \in S}$, where w_i^t denotes the weight, or probability, of playing strategy i at time t.

[13] This algorithm was developed by the authors in collaboration with Dean Foster.

[14] This model is applicable for example in studying the multi-armed bandit problem [2].

[15] Note that this point of view gives rise to notation which deviates slightly from that of the first half of the paper.

Additive Updating

Additive updating is based on a cumulative sum of weighted payoffs achieved by all strategies. This weighted sum is computed as if randomized strategies, or interior points, were attainable. In particular, define σ_i^t as follows:

$$\sigma_i^t \quad = \quad \sum_{x=0}^{t} w_i^x \pi_i^x \tag{1}$$

Now the weight of strategy i at time $t+1$ is the ratio of the weighted payoffs achieved by strategy i to the sum of the weighted payoffs achieved by all strategies:

$$w_i^{t+1} \quad = \quad \frac{\sigma_i^t}{\sum_{j \in S} \sigma_j^t} \tag{2}$$

In naive settings, the only information pertaining to payoff functions that is available at a given time is the payoff of the strategy that is in fact employed at that time. Thus, the naive updating rule must utilize an estimate of weighted payoffs which depends only on this limited information. Such an estimate is given by $\hat{\sigma}_i^t$, which is defined as follows:

$$\hat{\sigma}_i^t = \sum_{x=0}^{t} \mathbf{I}_i^x r_i^x \tag{3}$$

where $\mathbf{I}_i^t$ is the indicator function [16] Note that the expected value of this estimator is the actual value of cumulative weighted payoffs: *i.e.*, $\mathbb{E}[\hat{\sigma}_i^t] = \sigma_i^t$.

In naive settings, the updating rule given in Equation 2 is modified such that $\hat{\sigma}_i^t$ is used in place of σ_i^t. This update procedure yields a set of weights which must then adjusted in order to ensure that the space of possible payoffs be adequately explored. This is achieved by imposing an artificial lower bound on the probability with which strategies are played. In particular, for $\epsilon > 0$,

$$\hat{w}_i^t = (1-\epsilon) w_i^t + \frac{\epsilon}{|S|} \tag{4}$$

Roth and Erev [5] study a related procedure for updating strategic weights which is also applicable in naive settings but depends only on finite history.

[16] The indicator function is defined as follows:

$$\mathbf{I}_i^t = \begin{cases} 1 & \text{if strategy } i \text{ is played at time } t \\ 0 & \text{otherwise} \end{cases}$$

Multiplicative Updating

Freund and Schapire [10] introduce a multiplicative procedure for updating strategic weights. The multiplicative update rule depends on the cumulative payoffs achieved by all strategies including the surmised payoffs of those strategies which are not played (notation ρ_i^t):

$$\rho_i^t = \sum_{x=0}^{t} \pi_i^x \tag{5}$$

The weight assigned to strategy i at time $t+1$ is given by:

$$w_i^{t+1} = \frac{(1+\beta)^{\rho_i^t}}{\sum_{j \in S}(1+\beta)^{\rho_j^t}} \tag{6}$$

where $\beta > 0$ is a parameter which controls the tradeoff between learning rapidly (for large values of β) and accuracy (for small values of β).

Auer, Cesa-Bianchi, Freund, and Schapire [2] study a related procedure for updating strategic weights which is applicable in naive settings. First of all, as in the case of additive updating, it is necessary to assign minimal probabilities to all strategies in order that the space of payoff functions be adequately explored (see Equation 4). Now let

$$\hat{\pi}_i^t = \mathbf{I}_i^t \frac{\pi_i^t}{\hat{w}_i^t} \tag{7}$$

In other words, $\hat{\pi}_i^t$ is equal to 0 if strategy s_i is not employed at time t; otherwise, $\hat{\pi}_i^t$ is the payoff achieved by strategy s_i at time t scaled by the likelihood of playing strategy s_i. Estimated cumulative payoffs (notation $\hat{\rho}_i^t$) are given by:

$$\hat{\rho}_i^t = \sum_{x=0}^{t} \hat{\pi}_i^x \tag{8}$$

Note that the expected value of the above estimator is the actual value of cumulative payoffs: *i.e.*, $\mathbb{E}[\hat{\rho}_i^t] = \rho_i^t$. The naive variant of multiplicative updating uses $\hat{\rho}_i^t$ in place of ρ_i^t in Equation 6.

In what follows, we describe a series of experiments conducted using these learning algorithms in simulations of the Santa Fe bar problem (one-bar problem) and the New York City bar problem (two-bar problem) in informed and naive settings.

3.2 One-bar Problem

The one-bar problem was simulated assuming 100 computational agents and a single bar of capacity 60. Figure 1 depicts the results of simulations of both the additive and the multiplicative updating schemes by plotting attendance at the bar over time. (For simulation purposes, $\alpha = .5$ and $\beta = .01$.) Note that attendance centers around 60, which is in fact the capacity of the bar. In particular, these learning algorithms, which do not necessitate perfectly rational behavior, yield near-equilibrium outcomes. These results are robust in that they hold regardless of the capacity of the bar. More specifically, Figure 2 depicts attendance at the bar over time when the capacity of the bar begins at 60, but at time 1000 it changes to 40, and at time 2000 it changes to 80. In this scenario, attendance at the bar fluctuates accordingly. [17]

Careful inspection of the mixed strategies which the agents employ reveals that these algorithms converge to δ-Nash equilibrium (see, for example, [14]) in the one-bar problem. In particular, when the capacity of the bar is fixed at 60, the additive updating algorithm yields mixed strategies ranging from $(.397, .603)$ to $(.403, .597)$ (*i.e.*, $\delta = .003$), while the multiplicative updating algorithm yields mixed strategies ranging from $(.371, .629)$ to $(.423, .577)$ (*i.e.*, $\delta = .029$). [18] Consequently, these algorithms produce a fair solution to the one-bar problem in which all agents attend the bar an equivalent number of times, namely approximately 60% of the time. It remains to prove mathematically that additive and multiplicative updating converge to δ-Nash equilibrium in the one-bar problem.

In addition to simulating the uniform Santa Fe bar game, with $\alpha = .5$ for all agents, the non-uniform problem was also simulated with the values of α_n uniformly distributed in the range $[0, 1]$. The results of these simulations are depicted in Figure 3. In this scenario attendance once again centers around the capacity of the bar; moreover, we observe substantially more stability. This is a result of the fact that those agents with small values of α learn not to attend the bar at all, while those agents with large values α learn to always attend the bar, and only those on the cusp vary between sometimes going

[17] In fact, this result was obtained for additive updating using a responsive variant of the algorithm (see [16]) in the spirit of Roth and Erev [5]. In contrast, in the case of multiplicative updating, the agents readily learn to adjust weights within the space of serially undominated strategies.

[18] Note that in the multiplicative updating algorithm, the values of these extrema are directly correlated to the choice of β. For example, for $\beta = .001$, the mixed strategies range from $(.392, .608)$ to $(.404, .596)$ (*i.e.*, $\delta = .008$).

to the bar and sometimes staying at home. In fact, theoretical results on stability have been obtained for best-reply dynamics in related non-atomic, non-uniform games [11]. Note, however, that while stable, that this solution is not fair.

The last set of simulations discussed in this section considered the one-bar problem in a naive setting. The agents learned via the naive versions of the algorithms (with $\epsilon = .02$.), and they were provided with incomplete information: if they attended a bar, they were informed of its attendance, otherwise, they were not. The naive algorithms generally approximate the informed algorithms, such that we observe convergence to δ-Nash equilibrium in informed settings if and only if convergence to $(\delta + \epsilon)$-Nash equilibrium occurs in the corresponding naive case. Figure 4 (a) depicts the results of learning via naive multiplicative updating for $\beta = .01$ for which attendance once again centers around the capacity of the bar.

Finally, it is interesting to note that in naive settings, for alternative choices of β in the multiplicative updating algorithm, it is possible to decrease variance and thereby increase stability. In particular, if the agents' rates of learning are increased, learning becomes less accurate, and some number of (say m) agents stop attending the bar while (say l) others settle on always attending the bar. The remaining $N - m - l$ agents consequently adjust their behavior as if the total population were $N - m$ and the capacity of the bar were $c - l$, which yields lower variance in the attendance (see Figure 4 (b) where $\beta = .1$). After a while, however, since the agents are continually experimenting, agents who were once fixed on a certain strategy learn that their payoffs could be improved by altering their strategy, and they do so. Thus, in the long-run, all agents attend the bar 60% of the time.

3.3 Two-bar Problem

In this section, we discuss simulations of the two-bar problem. We show that in spite of the additional complexity of this problem, the near-equilibrium behavior observed in the one-bar problem extends to the two-bar problem because of the robust nature of the computational learning algorithms. The first scenario which we consider is analogous to the one-bar problem; in particular, there is excess demand. Figure 5 depicts the attendance at two bars, say bar A and bar B, each of capacity 40, with a population of 100 agents who learn according to the informed version of the multiplicative updating algorithm. [19] Note that attendance in each of the bars centers around 40, with about 20 agents choosing to stay at home each round. This is once again a fair solution in that each agent attends each bar 40% of the time and stays at home 20% of the time.

In the two-bar problem, in addition to the study of excess demand, it is also of interest to consider the case of excess supply: *e.g.*, two bars, each of capacity 60, and a population of 100 agents. Figure 6 depicts the results of simulation of this scenario where agents learn according to the informed version of the additive updating algorithm. In this case, agents learn to play mixed strategies of approximately $(1/2, 1/2)$, and consequently each agent attends each bar approximately 50% of the time. The space of equilibria, however, in this instantiation of the two-bar problem with excess supply, ranges from 40 agents at bar A and 60 agents at bar B to 60 agents at bar A and 40 agents at bar B.

In naive settings, equilibria other than the symmetric outcome which is achieved in informed settings can persist. Figure 7 depicts the results of simulations of naive additive learners who experiment until they arrive at a situation in which they appear satisfied, namely 40 regular attendees at bar A, and 60 regular attendees at bar B, and there they seem to remain. In fact, the population at bar B is slightly declining over time, while the attendance at bar A is correspondingly increasing ever so slightly, because of the downward pressure which exists when attendance at bar B exceeds capacity. It is likely that attendance at each bar would ultimately settle close to 50, but this warrants further investigation.

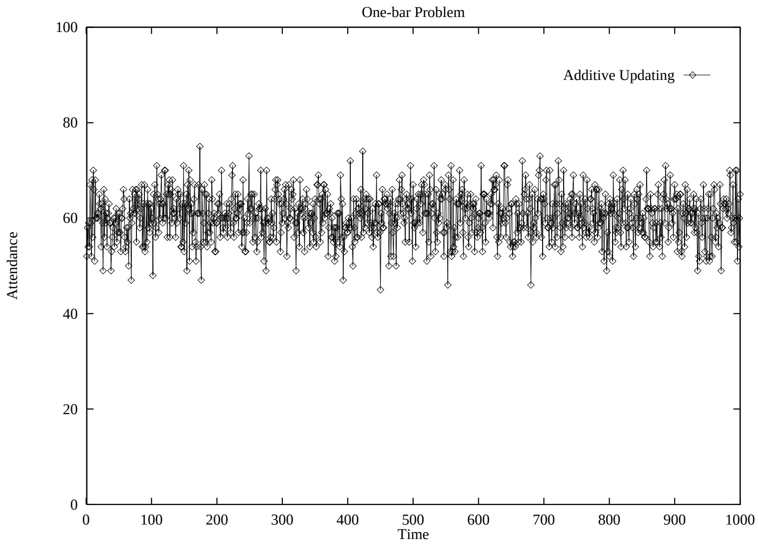
Finally, in the case of excess supply, naive multiplicative learners who learn quickly ($\beta = .1$) arrive at a stable solution in which the variance is quite low (see Figure 8). In contrast with the observed behavior that results via additive updating, naive multiplicative learners experiment until they arrive at a situation in which they are satisfied, namely 40 regular attendees at bar A, and 60 regular attendees at bar B, and there they remain. In this case, it does not appear that further iterations will ultimately lead to equal attendance in both bars, since the lower variance elmininates the downward pressure which was apparent in Figure 7.

[19] No qualitative differences between additive and multiplicative updating were observed in simulations of the two-bar problem. Consequently, we present results in this section which depict attendance at two bars, rather than attendance resulting from two algorithms.

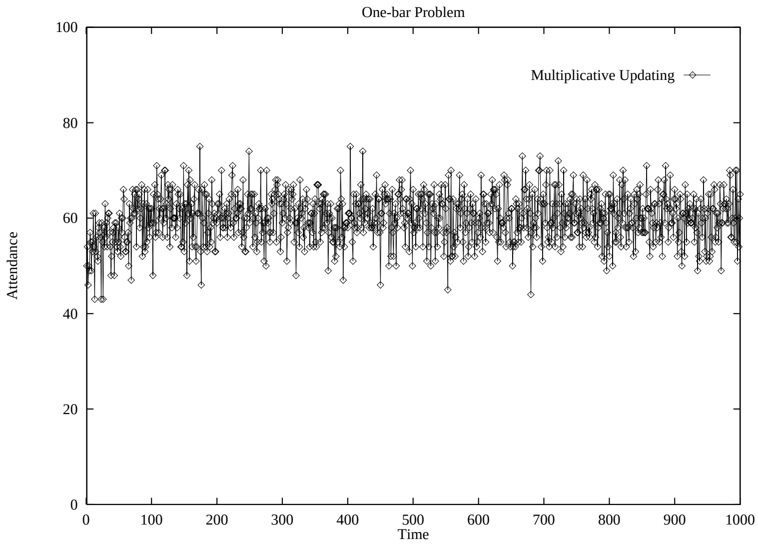
4 Conclusion

This paper investigated the Santa Fe (*i.e.*, El Farol) bar problem from both a theoretical and a practical perspective. Theoretically, it was shown that belief-based learning (*e.g.*, Bayesian updating) yields unstable behavior in this repeated game. In particular, rationality and predictivity, two conditions sufficient for convergence to Nash equilibrium, are inherently incompatible. On the practical side, it is demonstrated via simulations that computational learning algorithms in which agents occasionally act irrationally do indeed give rise to near-equilibrium behavior. It remains to prove theoretically that this is the necessary outcome.

The results presented in this paper are of particular interest in the design of computational agents for dissemination on the Internet. For example, we envision automated information retrieval agents learning our preferences and retrieving relevant articles on game theory. We have learned, however, that straightforward implementations of such agents can give rise to outcomes far from the desired equilibrium. Consequently, it is essential that we come to thoroughly understand the dynamics of the interactions among heterogeneous computational agents before engaging them to work on our behalf.

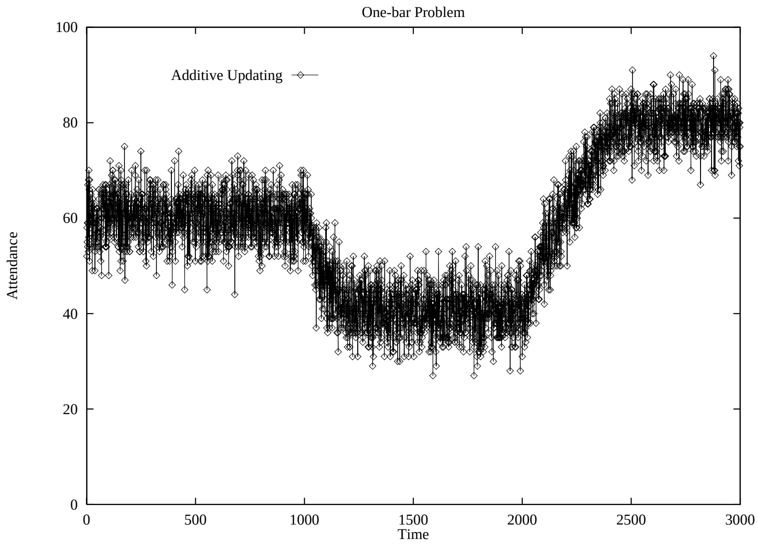


(a) Additive Updating

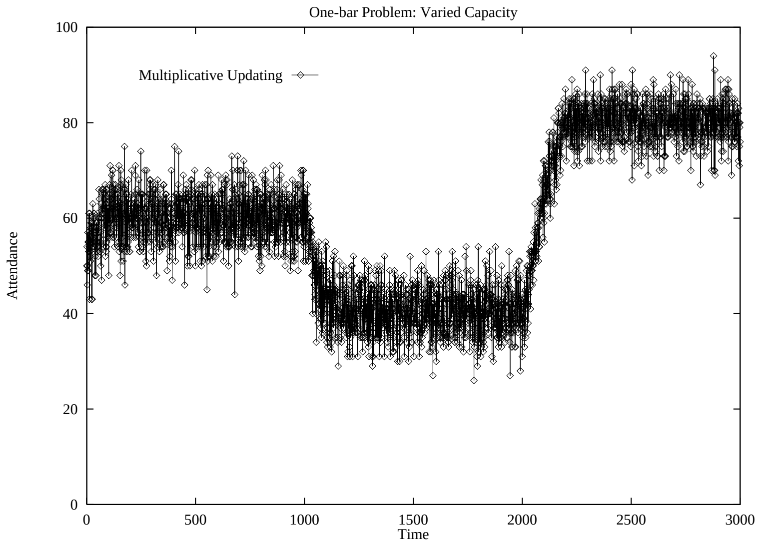


(b) Multiplicative Updating

Figure 1: Attendance vs. Time in the One-bar Problem

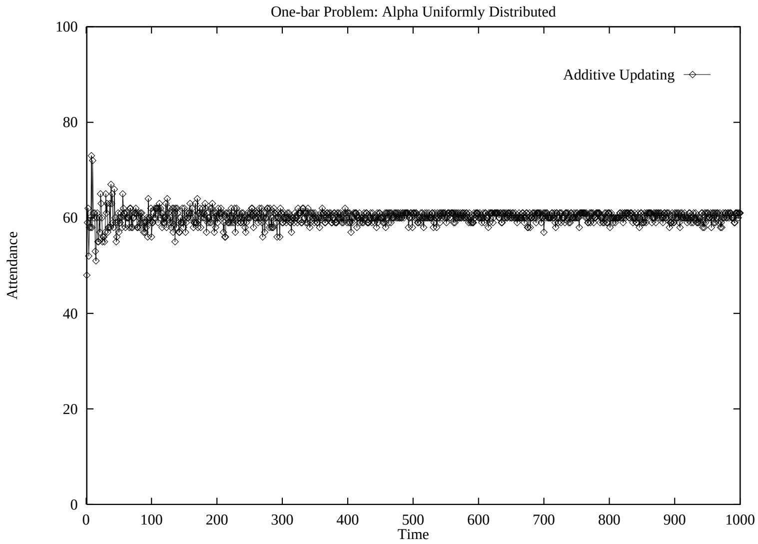


(a) Additive Updating

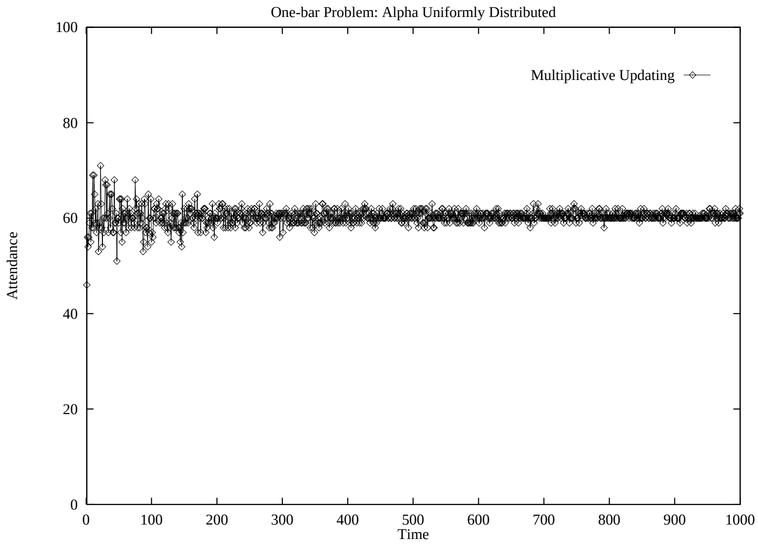


(b) Multiplicative Updating

Figure 2: Attendance vs. Time in the One-bar Problem with Varied Capacity

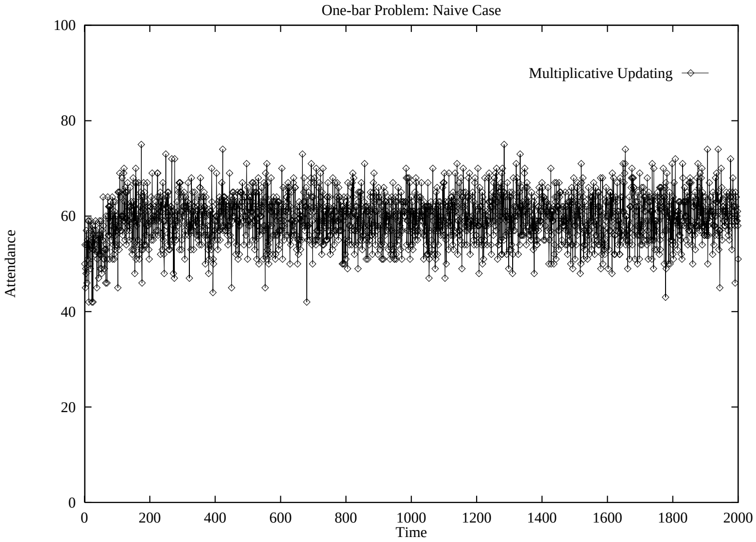


(a) Additive Updating

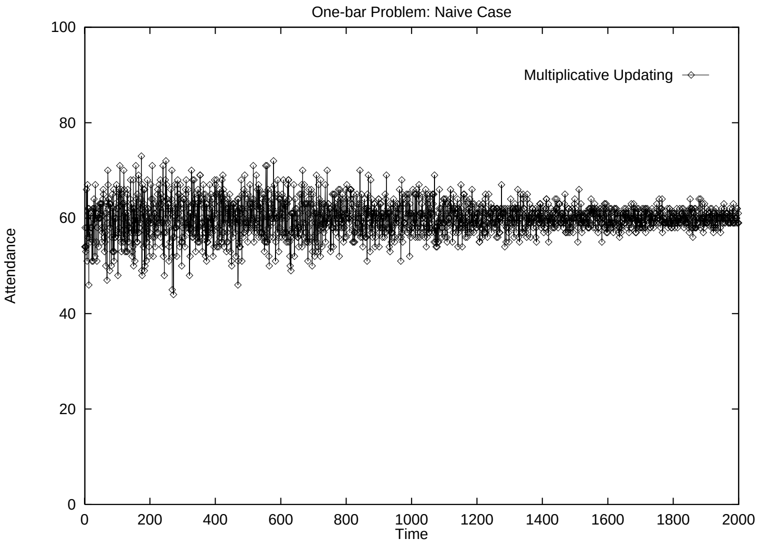


(b) Multiplicative Updating

Figure 3: Attendance vs. Time in the One-bar Problem with α_n Uniformly Distributed

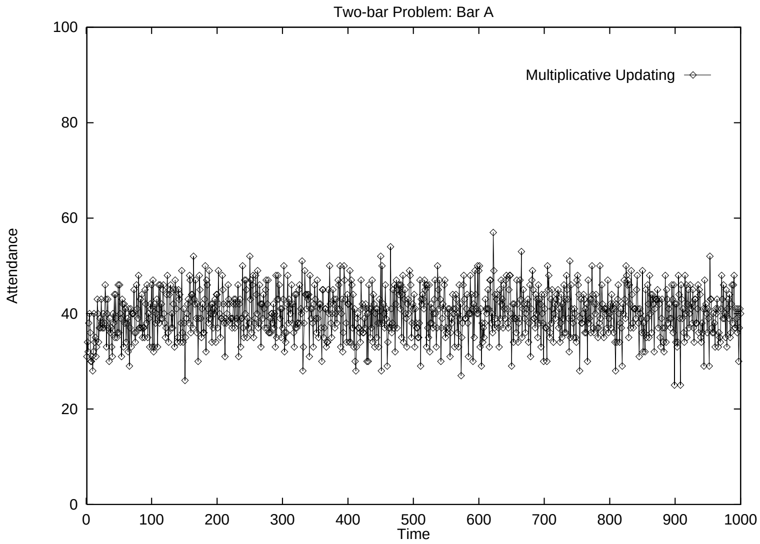


(a) Multiplicative Updating ($\beta = .01$)

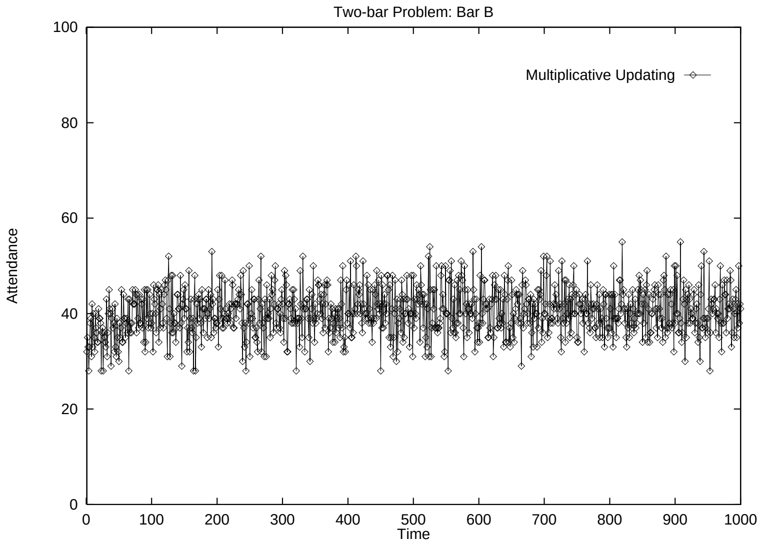


(b) Multiplicative Updating ($\beta = .1$)

Figure 4: Attendance vs. Time in the One-bar Problem: Naive Case

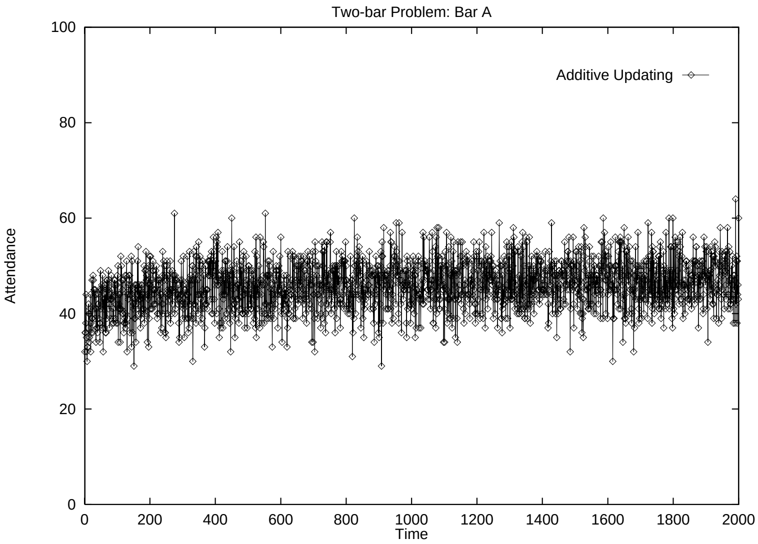


(a) Multiplicative Updating: Bar A

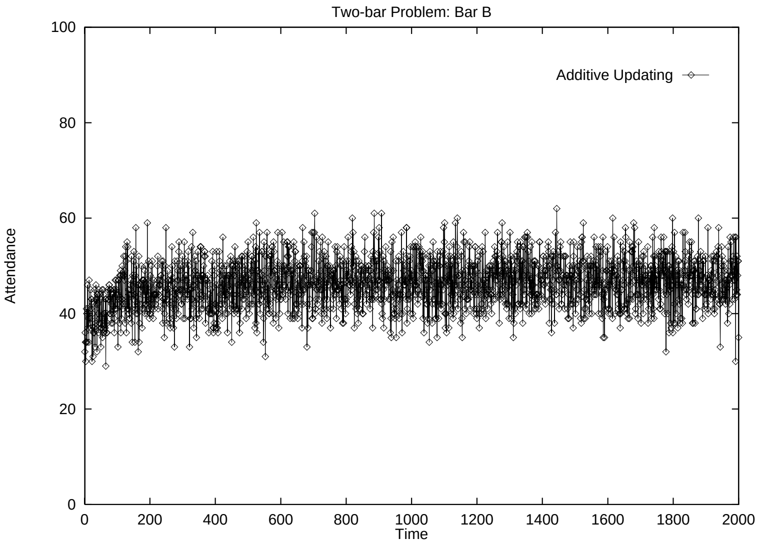


(b) Multiplicative Updating: Bar B

Figure 5: Attendance vs. Time in the Two-bar Problem: Excess Demand

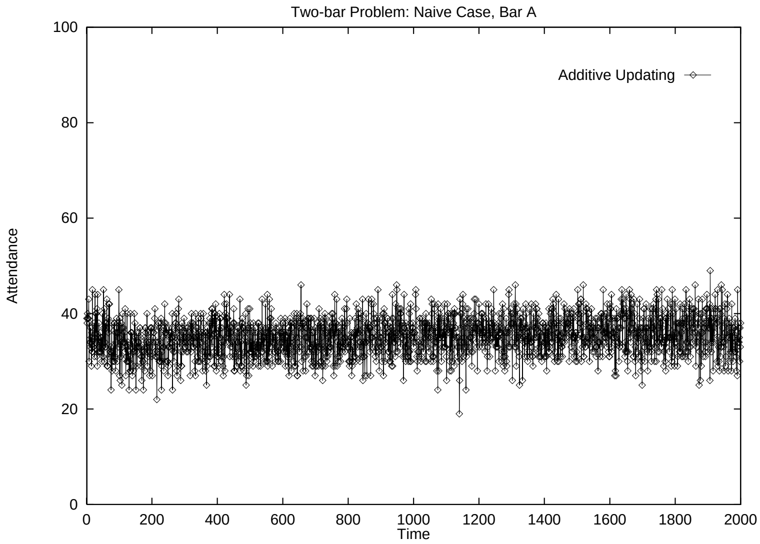


(a) Additive Updating: Bar A

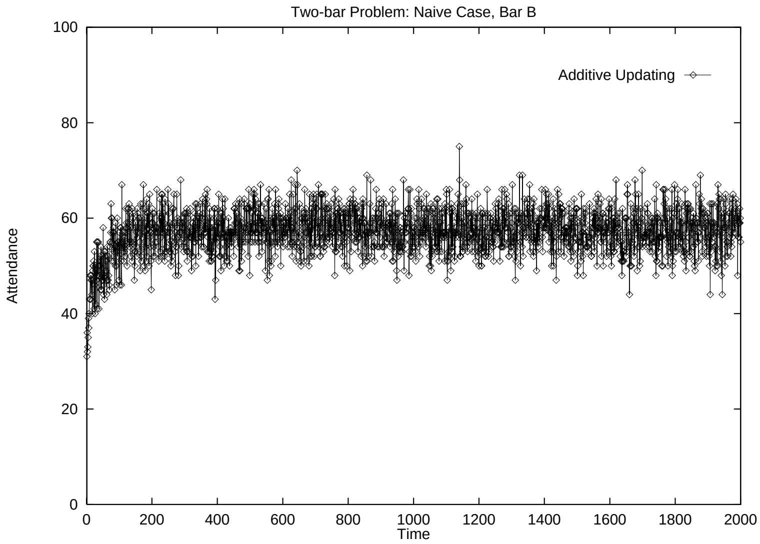


(b) Additive Updating: Bar B

Figure 6: Attendance vs. Time in the Two-bar Problem: Excess Supply

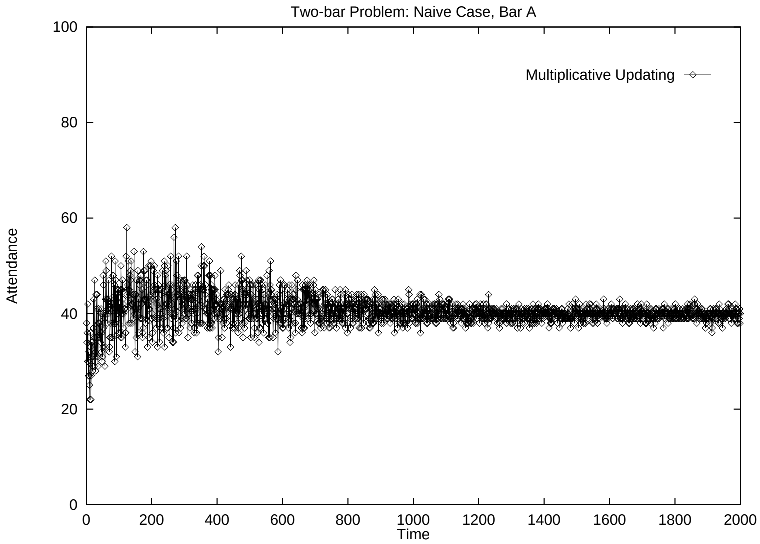


(a) Additive Updating: Bar A

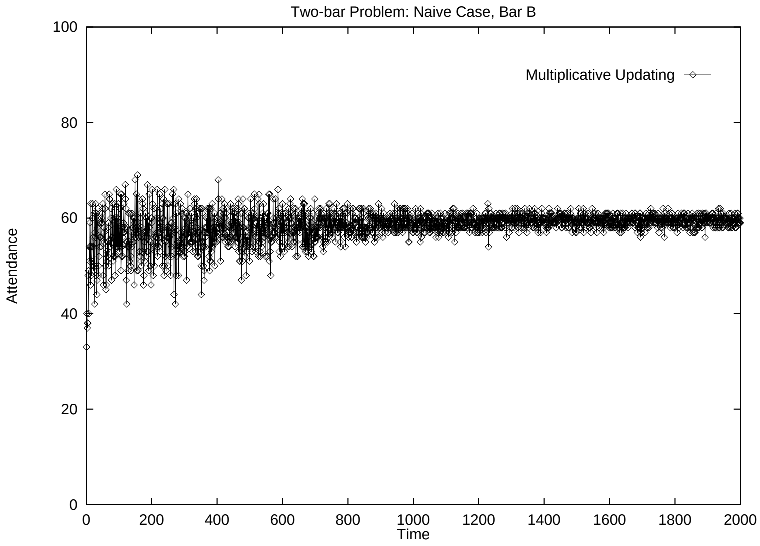


(b) Additive Updating: Bar B

Figure 7: Attendance vs. Time in the Two-bar Problem: Naive Case



(a) Multiplicative Updating ($\beta = .1$): Bar A



(b) Multiplicative Updating ($\beta = .1$): Bar B

Figure 8: Attendance vs. Time in the Two-bar Problem: Naive Case

References

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

[2] P. Auer, N. Cesa-Bianchi, Y. Freund, and R. Schapire. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *Proceedings of the 36th Annual Symposium on Foundations of Computer Science*. ACM Press, November 1995.

[3] H. Chernoff. A measure of the asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Annals of Mathematical Statistics*, 23:493–509, 1952.

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

[5] I. Erev and A.E. Roth. On the need for low rationality, cognitive game theory: Reinforcement learning in experimental games with unique mixed strategy equilibria. *Draft*, May 1996.

[6] R. Even. *Bounded Rationality and Computationally Complex Adaptive Systems*. Ph.D. Dissertation, New York University, New York, Expected December, 1998.

[7] R. Even, A. Greenwald, and B. Mishra. CAFÉ society: Network simulations using market models. *Submitted for Publication*, March 1998.

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

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

[10] Y. Freund and R. Schapire. Game theory, on-line prediction, and boosting. In *Proceedings of the 9th Annual Conference on Computational Learning Theory*. ACM Press, May 1996.

[11] E. Friedman. Dynamics and rationality in ordered externality games. *Games and Economic Behavior*, 15:65–76, 1996.

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

[13] E. Friedman and S. Shenker. Learning and implementation on the Internet. *Mimeo*, 1997.

[14] D. Fudenberg and J. Tirole. *Game Theory.* MIT Press, Cambridge, 1991.

[15] A. Greenwald. *Learning to Play Network Games.* Ph.D. Dissertation, New York University, New York, Expected May, 1999.

[16] A. Greenwald, E. Friedman, and S. Shenker. Learning in network contexts: Results from experimental simulations. *Presented at Fourth Informs Telecommunications Conference*, March 1998.

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

[18] E. Kalai and E. Lehrer. Weak and strong merging of opinions. *Journal of Mathematical Economics*, 23:73–86, 1994.

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

[20] J. Mackie-Mason and H. Varian. Pricing congestible resources. *IEEE Journal of Selected Areas in Communications*, 13(7):1141–1149, September 1995.

[21] J. Mackie-Mason and H. Varian. Pricing the Internet. In *Public Access to the Internet*, pages 269–314. MIT Press, Cambridge, 1995.

[22] P. Milgrom and J. Roberts. Adaptive and sophisticated learning in normal form games. *Games and Economic Behavior*, 3:82–100, 1991.

[23] J. Nachbar. Prediction, optimization, and learning in repeated games. *Econometrica*, 65:275–309, 1997.

[24] R. Parikh. A connection between nash equilibria and propositional logic. *Mimeo*, November 1998.

[25] B. Russell. Letter from russell to frege. In *From Frege to Gödel*, pages 124–125, Cambridge, 1967. Harvard University Press.

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