On the Impossibility of Rational Learning in the Santa Fe Bar Problem

Amy Greenwald
Department of Computer Science, Brown University
Box 1910, Providence, RI 02906
amygreen@cs.brown.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
365 Fifth Avenue, New York, NY 10016
ripbc@cunyvm.cuny.edu

August 31, 2000

Keywords: rationality, learning, Nash equilibrium, Russell's paradox, repeated games

Abstract

This paper investigates the Santa Fe (El Farol) bar problem from the point of view of rational learning. It is argued that rationality together with belief-based learning (*e.g.*, Bayesian updating) yields unstable behavior in this game. More specifically, two conditions sufficient for convergence to Nash equilibrium, namely rationality and predictivity, are shown to be incompatible.

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 learning. 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]

We motivate this paper with the following intuitive analysis of SFBP under the standard economic assumption of rationality. Define an *uncrowded* bar as one in which attendance is less than or equal to c, and define a *crowded* bar as one in which attendance is strictly greater than c. Let the utility of going to an uncrowded bar be $1/2$ and the utility of going to a crowded bar be $-1/2$; in addition, the utility of staying at home is 0, regardless of the state of the bar.[2] If a agent predicts that the bar will be uncrowded with probability p, then his rational (*i.e.*, best) reply is to go to the bar if $p > 1/2$ and to stay home if $p < 1/2$. Now, if the agents indeed learn to predict probability p accurately, then their predictions eventually come to match the actual probability that the bar is uncrowded, as it is determined by their (possibly randomized) strategic best-replies. Herein lies a contradiction. If the agents learn to

[1] This problem was inspired by the El Farol bar in Santa Fe which offers live music on Thursdays.

[2] The assumption that 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 is $1/2$, whenever the bar is crowded, and the utility of staying at home is $-1/2$, whenever the bar is uncrowded.

predict that the bar will be uncrowded with probability $p < 1/2$, then, in fact the bar will be empty with probability 1; on the other hand, if the agents learn to predict that the bar will be uncrowded with probability $p > 1/2$, then the bar will be empty with probability 0.[3][4] We conclude that rational agents cannot learn to make accurate predictions. Rationality precludes learning.

1.1 Logical Implications

This paradoxical outcome in SFBP is arrived at via a diagonalization process in the spirit of Russell's paradox [15]. Russell's set is the set of all sets that are not elements of themselves: *i.e.*, $\mathcal{R} = \{X | X \notin X\}$. Thus, $\mathcal{R} \in \mathcal{R}$ iff $\mathcal{R} \notin \mathcal{R}$. Just as the truth of being in Russell's set depends on the fact of (not) being in the set, the value of going to the bar depends on the act of going (or not going) to the bar. For the sake of argument, consider a bar of capacity 1/2 in a world of a single agent.[5] If the agent does not go to the bar, then the bar is uncrowded, in which case her best-reply is to go to the bar. But now the bar is crowded, and so her best-reply is to stay at home. Thus rationality dictates that this agent should go to the bar if and only if she 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 variant of SFBP in which there are two players and both positive and negative externalities:[6] 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

[3] In the case where $p = 1/2$, the agents are indifferent between attending the bar and staying home and may behave arbitrarily. This arbitrary behavior is unlikely to coincide with the outcome "the bar is empty with probability p."

[4] Schelling [16] refers to phenomena of this kind as self-negating prophecies.

[5] Similarly, one could consider a bar of capacity 1 and a married couple who act in unison.

[6] See Section 2 in order to understand SFBP in terms of solely negative externalities.

the mismatcher. In matching pennies, if player 1 prefers *heads*, then player 2 prefers *tails*, but then player 1 prefers *tails*, at which point player 2 actually prefers *heads*, and finally, player 1 prefers *heads* once again. It follows that player 1 prefers *heads* iff player 1 prefers *tails*. Similarly, for player 2.

The logical conflict that arises in the game of matching pennies is closely related to the fact that the game has no pure strategy Nash equilibria [14]; similarly, SFBP has no symmetric pure strategy Nash equilibria, except in degenerate cases. In order to resolve these paradoxes, game-theorists introduce mixed strategies. The unique Nash equilibrium in matching pennies is for both players to play each of *heads* and *tails* with probability $1/2$; a mixed strategy Nash equilibrium in SFBP is for all agents to go to the bar with probability $p \approx c/N$ and to stay at home with probability $1-p$.[7]

1.2 Game-Theoretic Implications

This paper presents negative results on convergence to Nash equilibrium in SFBP which formalizes the above diagonalization argument. Two sufficient conditions for convergence to Nash equilibrium are *rationality* and *predictivity*. By rationality, we mean that players play best-replies to their beliefs. Predictivity is one way in which to capture the notion of learning:[8] a player is said to be predictive if that player's beliefs eventually coincide with (or approach) the truth of what he is predicting. If players learn to predict (*i.e.*, if beliefs indeed converge to other players' actual strategies), then best-replies to beliefs constitute a Nash equilibrium. In what follows, we argue that if the players employ predictive learning algorithms, assuming rationality, play does not converge to Nash equilibrium in SFBP. Equivalently, *if play converges to Nash equilibrium, then either play is not rational or play is not learned.*

[7]Technically, this symmetric Nash equilibrium is the solution p to the following equation: $\sum_{x=0}^{c} \binom{N}{x} p^x (1-p)^{N-x} = \sum_{x=c+1}^{N} \binom{N}{x} p^x (1-p)^{N-x}$, which is approximately c/N.

[8]Learning can also be understood in terms of various merging properties; see Kalai and Lehrer [9].

In a seminal work by Kalai and Lehrer [8], sufficient conditions are presented for predictivity—specifically, an absolute continuity assumption—which suggests that convergence to Nash equilibrium is at least possible. Our negative results complement the work of Nachbar [13] and Foster and Young [4], who argue that the conditions sufficient for prediction are unlikely to ever hold. Nachbar shows that unless players' initial beliefs somehow magically coincide with Nash equilibrium, repeated play of strategic form games among Bayesian rational players does not converge to Nash equilibrium. Similarly, Foster and Young prove that in two-player games of incomplete information with unique mixed strategy Nash equilibria, rationality is not compatible with predictivity. Our theorems argue in a similar vein that unless certain strict regularity conditions are satisfied, no means of rational learning converges to Nash equilibrium in SFBP.

1.3 Computer Science Implications

SFBP and its natural extensions (*e.g.*multiple bars) serve as abstractions of various congestion control problems that arise in networking. Many authors (see, for example, [10, 17]) who capitalize on the potential for the theory of repeated games as a model of networking environments do so because of the difficulty to enforce cooperation in large-scale networks; instead, it is more realistic and more general to assume non-cooperative networks. This generality is modeled in repeated games by assuming that agents are rational. Those same authors who study networking games assuming rationality often also assume that the network operating point is a Nash equilibrium. One might hope to justify this assumption on the grounds that Nash equilibrium is the outcome of rational learning. It is the conclusion of this study, however, that Nash equilibrium is *not* the outcome of rational learning in games that model networking environments.

1.4 Overview

This next section formalizes SFBP in terms of the theory of learning in repeated games. In Section 3, it is shown that best-reply dynamics yields oscillatory behavior in SFBP; in particular, play does not converge to Nash equilibrium or otherwise. The section that follows presents a seemingly positive result, namely that beliefs can in fact converge to Nash equilibrium. But then, we counter this result by observing that the conditions which ensure convergence are highly improbable. Our first negative result relies on the usual definition of rationality, while our second negative result holds assuming even weaker conditions. In summary, rational learning (*e.g.*, Bayesian updating) does not in general give rise to equilibrium behavior in SFBP.

2 Formalization

The Santa Fe bar problem is a repeated game of negative externalities.[9] We now 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 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 $c \in \{0, \ldots, N\}$ denote the capacity of the bar. The externality E depends on s and c as follows: if the bar is

[9] An externality is a third-party effect. An example of a negative externality is pollution; an example of a positive externality is standardization. Although externalities are so-called because they are external to the game, it is natural to consider payoffs in terms of externalities when there are large numbers of players.

uncrowded (*i.e.*, $s \leq c$), then $E(s) = 0$; on the other hand, if the bar is crowded (*i.e.*, $s > c$), then $E(s) = 1$. Let $0 \leq \alpha_n \leq 1$ denote the value to player n of attending the bar, and without loss of generality assume $\alpha_n \leq \alpha_{n+1}$.[10] Now the payoff function[11] for player n is given by:

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

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 [5, 6].

The one-shot strategic form SFBP is described by the tuple $\Gamma = (\mathcal{N}, (S_n, \pi_n)_{n\in\mathcal{N}}, c)$, and the infinitely repeated SFBP is given by Γ^∞. Following Foster and Young[12] [4], a history h^t of length $t \in \mathbb{N}$ is defined to be a sequence of t outcomes drawn from the set $S = \{0, 1, \ldots, N\}$; 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 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 a subjective probability over the space of possible externality effects $\mathcal{E} = \{uncrowded, crowded\}$.[13] Recall that the event *uncrowded* obtains at time t whenever $s^t \leq c$; otherwise, the event *crowded* obtains. Let $\Delta(\mathcal{E})$ be the set of

[10] The results in this paper are restricted to the *uniform* SFBP in which $\alpha_n = \alpha_m$ for all $n, m \in \mathbb{N}$.

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

[12] 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.

[13] Implicit in this belief structure is the assumption that beliefs are first-order, or *deterministic* in the sense of Foster and Young [4]. For example, we do not consider beliefs of the form: with probability P, player n assigns probability p_n^t to the bar being uncrowded at time t, while with probability $1 - P$, player n assigns probability $p_n'^t$ to the bar being uncrowded at time t instead.

probability distributions over the set $\mathcal{E}$. Formally, a belief-based learning algorithm for player n is a function $f_n : H \to \Delta(\mathcal{E})$.[14] Since the event space $\mathcal{E}$ is of cardinality 2, the sequence of beliefs $\{f_n^t(h^t)\} = \{(p_n^{t+1}, 1 - p_n^{t+1})\}$ is denoted simply $\{p_n^{t+1}\}$, where p_n^{t+1} is the probability that player n attributes to the bar being uncrowded at time $t+1$.

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

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

Let $p_n^* \equiv 1 - \alpha_n$. Player n is indifferent between his two pure strategies whenever $p_n^t = p_n^*$, since this implies $\mathbb{E}_n^t[\pi_n(1, s)] = \mathbb{E}_n^t[\pi_n(0, s)] = 0$. The actual/objective[15] probability that the bar is uncrowded at time t is denoted by p_0^t. The existence of such probabilities is implied by the assumption that in general players employ mixed strategies: *i.e.*, p_0^t can be computed directly from the players' strategies q_n^t.

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

3 An Example: Best-Reply Dynamics

We now show that best-reply dynamics, a learning algorithm for which Cournot proved convergence to pure strategy Nash equilibrium in models of duopoly [3], yield

[14] A belief-based learning algorithm is *not* a function $f_n : H \to Q_{-n}$, where $Q_{-n} \equiv \prod_{m \neq n} Q_m$. If it were, this would violate the assumption that players are anonymous, as players could distinguish their own play from that of the aggregate. The given definition precludes any notion of correlated beliefs, in which players might attempt to correlate the behavior of an individual, such as oneself, with attendance at the bar.

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

oscillatory behavior in SFBP. Note that since best-reply dynamics is one method of belief-based learning the result presented in this section follows as an immediate corollary of the more general results derived in later sections. We begin by reminding the reader of the assumptions implicit in best-reply dynamics.

Definition 3.1 A strategy $q_n^t \in Q_n$ is said to be a *best-reply* for player n at time t iff $q_n^t \in \arg\max_{q_n \in Q_n} \pi_n(q_n, p_n^t)$: *i.e.*, $\pi_n(q_n^t, p_n^t) \geq \max_{q_n \in Q_n} \pi_n(q_n, p_n^t)$.

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 $t+1$, he assumes that the outcome that is realized during round t will again be the outcome of round $t+1$, and consequently plays a best-reply to the outcome of round t.

Definition 3.2 A given player n is said to employ *best-reply dynamics* in SFBP iff for all times t, player n assumes that

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

and moreover, player n plays only best-replies to these beliefs. In other words, if player n utilizes best-reply dynamics, then $q_n^t \in \arg\max_{q_n \in Q_n} \pi_n(q_n, s^t)$.

Assume that all players employ best-reply dynamics. If at time t, $s^t \leq c$, then $p_n^{t+1} = 1$ 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 $p_n^{t+2} = 0$ for all n, to which the best response at time $t+2$ is pure strategy $s_n^{t+2} = 0$. Now, it follows that $s^{t+2} \leq c$ and $p_n^{t+3} = 1$. This pattern repeats itself indefinitely, generating oscillatory behavior that is far from equilibrium. A similar argument arises in the case in which $s^t > c$. The following remark captures this intuition.

Remark 3.3 In the uniform repeated SFBP, best-reply dynamics do not converge: *i.e.*, $\forall n$, $\lim_{t,t' \to \infty} | p_n^t - p_n^{t'} | \neq 0$ and $\forall n$, $\lim_{t,t' \to \infty} | s^t - s^{t'} | \neq 0$.

Recall that 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 [6, 11]. Consistent with this result, the serially undominated set obtained in SFBP is not a singleton – on the contrary, it includes all the strategies of the game – and moreover, best-reply dynamics do not converge. Stability results regarding best-reply dynamics have been achieved, however, for special cases within the class of ordered externality games, some of which resemble non-uniform versions of SFBP [6].

4 A Preliminary Result

In what follows, we generalize our observation pertaining to best-reply dynamics. Specifically, we argue that if players are rational and if they learn according to Bayes' rule, their strategies do not converge to equilibrium behavior. In fact, this result is not contingent on the assumption of Bayesian learning and is readily applicable to any predictive belief-based learning mechanism. We arrive at this negative result by first deriving a seemingly positive result, namely that beliefs can *sometimes* converge to Nash equilibrium, assuming rationality and predictivity. Later we show that in practice sometimes amounts to *almost never*.

Definition 4.1 A learning algorithm is said to be *predictive* iff it generates a sequence of beliefs $\{p_n^t\}$ for player n *s.t.* $\lim_{t\to\infty} |\, p_n^t - p_0^t \,| = 0$.

In words, if player n utilizes a predictive learning algorithm, then the difference between player n's subjective beliefs p_n^t and the objective probabilities p_0^t converges to zero. This definition does not require that the objective probabilities themselves converge, only that player n's subjective beliefs approach the objective probabilities.

Definition 4.2 A set of players $\mathcal{N}$ is said to reach *consensus* iff for all $n, m \in \mathcal{N}$, $\lim_{t\to\infty} |\, p_n^t - p_m^t \,| = 0$.

Lemma 4.3 *If all players within set $\mathcal{N}$ are predictive, then they reach consensus.*

Proof 4.4 By the triangle inequality, $|\,p_n^t - p_m^t\,| \leq |\,p_n^t - p_0^t\,| + |\,p_0^t - p_m^t\,|$, for all $n, m \in \mathcal{N}$, and for all t. Taking limits and applying the definition of predictivity, it follows immediately that $\lim_{t\to\infty} |\,p_n^t - p_m^t\,| = 0$. □

Definition 4.5 A player n is *rational* iff he plays only best-replies to his beliefs p_n^t.

The following theorem states that in the uniform version of the repeated SFBP, whenever players exhibit rationality and predictivity, beliefs converge to p^*. It follows by predictivity that the objective probabilities, which correspond to the players' joint strategies, must converge to p^* as well. Thus, rational players who play best-replies to their beliefs, ultimately play best-replies to actual strategies: *i.e.*, play converges to Nash equilibrium. This seemingly positive result is contested in later sections.

Theorem 4.6 *In the uniform repeated* SFBP*, if players are rational and predictive, then* $\lim_{t\to\infty} |\,p_n^t - p^*| = 0$.

Proof 4.7 Suppose not. Two cases arise. First suppose $\exists \epsilon > 0$ *s.t.* $p_n^t > p^* + \epsilon$ infinitely often (i.o.). It follows by Lemma 4.3 that for all m, $p_m^t > p^* + \delta$ i.o., for all $0 < \delta < \epsilon$. Now by rationality, all players play best-replies, which for such t is to go to the bar: *i.e.*, for all n, $s_n^t = 1$ i.o.. This ensures that the bar will be crowded with probability 1, yielding $p_0^t = 0 < p^* + \epsilon < p_n^t$ i.o., which implies that $p_n^t - p_0^t > \epsilon$ i.o., contradicting predictivity. The argument in the second case is analogous. □

It follows immediately from Theorem 4.6 and the definition of predictivity that whenever players are rational and predictive, strategies (objective probabilities), as well as beliefs, must converge to p^*. The question now arises as to whether p^* can indeed be the actual objective probability, for if not, players cannot be both rational and predictive.

5 A First Negative Result

We now present our first negative result, namely that no mechanism of rational, predictive, belief-based learning (including Bayesian updating) gives rise to objective probabilities that converge to p^*, except in rare circumstances. Before making any general claims, we construct an example of one such special p^* for which convergence is possible. We assume that indifferent players (*i.e.*, those players n for which $p_n^t = p^*$) flip a fair coin; for if, on the contrary, players were to flip a biased coin favoring one strategy or another, they would not be truly indifferent between the two alternatives.

Example 5.1 Let $f(t) \to 0$ be a monotonically decreasing function of t: *e.g.*, $f(t) = 1/t$.

- Suppose G players (the optimists) hold beliefs $p^* + f(t)$. These players' beliefs converge to p^*, but by rationality, these players *always* go to the bar.
- Let H players (the pessimists) hold beliefs $p^* - f(t)$. These players' beliefs also converge to p^*, but by rationality, these players *never* go to the bar.
- Let I players (the realists) hold beliefs exactly p^* at all times t. These players are indifferent between going to the bar and not going, so they flip a fair coin.

Given that players' beliefs converge to p^*, we now consider the conditions under which the players' strategies (*i.e.*, actual probabilities) also converge to p^*. Let the excess capacity of the bar $d = c - G$ for the I indifferent players, after accomodating the G players who go to the bar in every period. Suppose indifferent players go to the bar iff their coin flips show heads. In this scenario, the probability p that the bar is uncrowded is the probability that with I flips of a fair coin,[16] at most d heads appear:

$$p = \begin{cases} 0 & \text{if } d < 0 \\ 1 & \text{if } d \geq I \\ \frac{1}{2^I} \sum_{j=0}^{d} \binom{I}{j} & \text{otherwise} \end{cases} \tag{2}$$

[16]Mathematically, this result holds in the more general case when players flip a coin of bias q. In

Now as $t \to \infty$, $p_n^t \to p^*$ and $p_0^t \to p$ (in fact, $p_0^t = p$, for all t). Thus, if $p^* = p$, then both beliefs and strategies converge to p^*. $\square$

Using the layout of Example 5.1, it is possible to describe all possible values of p in Equation 2. At fixed time t, let G denote the number of players who definitely go to the bar; let H denote the number of players who definitely stay at home; and let I denote the number of players who are indifferent and therefore flip a fair coin in deciding whether or not to attend the bar. The following set F describes all the realizable probabilities under these circumstances:

$$F = \{p \mid \exists\, G, H, I \in \{0, \ldots, N\} \ s.t.\ p \text{ is defined by Equation 2}\}$$

F is a finite set since there are only finitely many possible values of G, H, and I. The following theorem states that objective probabilities cannot possibly converge to p^*, if $p^* \notin F$.

Theorem 5.2 *In the uniform repeated* SFBP, *if players are rational and predictive, then* $\lim_{t\to\infty} |\, p_0^t - p^* | \neq 0$, *unless* $p^* \in F$, *provided indifferent players flip a fair coin.*

Proof 5.3 Suppose to the contrary that $p^* \notin F$, but $\lim_{t\to\infty} |\, p_0^t - p^* | = 0$. Let $\delta = \min\{d(p^*, x) \mid x \in F\}$. Note that $\delta > 0$, since $p^* \notin F$. By the assumption of convergence, $\exists T \ s.t.\ \forall t > T, |\, p_0^t - p^* | < \delta$. But now, since $p_0^t \in F$, it follows that $d(p^*, F) < \delta$. Contradiction. $\square$

In SFBP, assuming a bar of capacity c, if players are rational and predictive, then strategies can only converge to p^* if p^* happens to be an element of the finite

particular, Equation 2 becomes

$$p = \begin{cases} 0 & \text{if } d < 0 \\ 1 & \text{if } d \geq I \\ \sum_{j=0}^{d} \binom{I}{j} q^j (1-q)^{I-j} & \text{otherwise} \end{cases} \qquad (1)$$

We do not present this case, however, since this assumption is more difficult to justify.

set F. Thus, it is only on rare occasions that players exhibit both rationality and predictivity, such that both beliefs and strategies converge to p^*: *i.e.*, play converges to Nash equilibrium. In general, play does not converge to Nash equilibrium in SFBP.

Example 5.4 Consider an instance of SFBP in which players are both rational and predictive. Let $N = I = 10$, and assume $c = 6$. In other words, there are 10 players, all of whom are indifferent and flip a fair coin. According to Equation 2, there exists $p^* \approx .828$ such that the players' strategies (*i.e.*, actual probabilities) converge to p^*. By Theorem 4.6, beliefs also converge to p^*. Thus, in this particular instance of SFBP, if by chance $\alpha = 1 - p^* \approx .172$, then play converges to Nash equilibrium. □

As Theorems 4.6 and 5.2 yield contradictory conclusions in all but finitely many cases, we deduce that together the assumptions of rationality and predictivity are inconsistent: *in general, there is no rational learning in* SFBP.

Corollary 5.5 *In the uniform repeated* SFBP*, players cannot be both rational and predictive, unless* $p^* \in F$.

This concludes the discussion of our first negative result. It was argued that two conditions which together are sufficient for convergence to Nash equilibrium, namely rationality and predictivity, are incompatible in SFBP. A similar analysis appeared in Greenwald [7]. In the next section, a second negative result is derived, which is based on the work of Mishra [12].

6 A Second Negative Result

The negative result presented in the previous section is contingent on the fact that in SFBP rational best-replies are in general deterministic. At fixed time t, if $p_n^t < p^*$, then player n believes that the bar is crowded, so his best-reply is to stay home, or go to the bar with probability 0; on the other hand, if $p_n^t > p^*$, then player n believes that the bar is uncrowded, so his best-reply is to go to the bar with probability 1. Although Theorem 4.6 ensures that players' beliefs converge to p^*, only when beliefs exactly coincide with p^* is player n indifferent between going to the bar and staying at home; at this point, player n may play any mixed strategy q_n^t. The step function that represents these deterministic best-replies is depicted in Figure 1(a). Beliefs p_n^t are plotted along the x-axis, and rational strategies q_n^t are plotted along the y-axis.

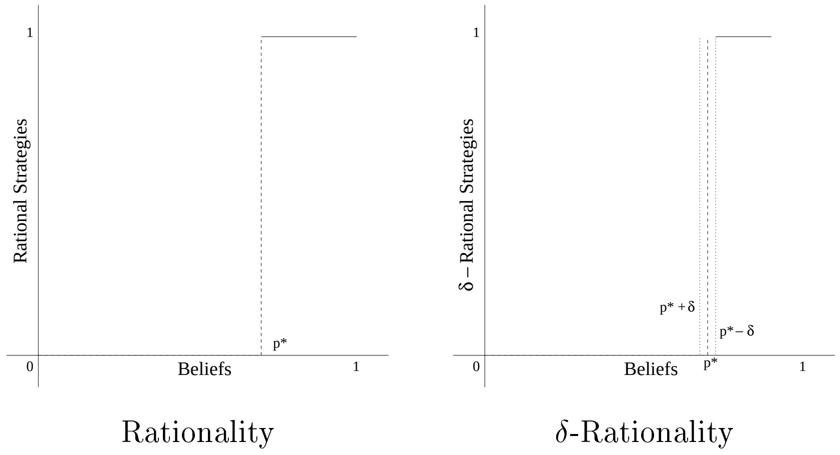

Figure 1: (a)If $p_n^t < p^*$, then player n believes that the bar is crowded, so his best-reply is to stay home, or go to the bar with probability 0; on the other hand, if $p_n^t > p^*$, then player n believes that the bar is uncrowded, so his best-reply is to go to the bar with probability 1. (b) If $p_n^t \leq p^* - \delta$, the δ-best-reply is to go to the bar with probability 0; on the other hand, if $p_n^t \geq p^* + \delta$, the δ-best-reply is to go to the bar with probability 1. In the range $(p^* - \delta, p^* + \delta)$, the precise behavior of the players is unspecified.

This section describes a second negative result based on the following slightly weaker notion of rationality. Here, players' behavior near p^* is not so clear-cut; rather it is unspecified, encouraging players to perhaps experiment with randomized strategies anywhere in the range $(p^* - \delta, p^* + \delta)$, as depicted in Figure 1(b).

Definition 6.1 Let $\delta > 0$. A player n is *δ-rational* iff he plays only best-replies to his beliefs in the range $[0, p_n^t - \delta]$ and $[p_n^t + \delta, 1]$, but plays arbitrarily in the range $(p_n^t - \delta, p_n^t + \delta)$.

The proof of Theorem 4.6 is easily extended to show that if players are δ-rational and predictive, then their beliefs converge to near p^*. But δ-rationality allows for the possibility that players experiment with non-deterministic strategies not only when beliefs precisely equal p^*, but even when beliefs are in the neighborhood of p^*. In other words, δ-rationality and predictivity imply δ-indifference.

The next theorem states that no means of δ-rational, predictive learning gives rise to objective probabilities that converge to p^*, unless the capacity of the bar fortuitously lies between $N/2 - k_1\sqrt{N}$ and $N/2 + k_2\sqrt{N}$, for certain $k_1, k_2 > 0$, provided that δ-indifferent players flip a fair coin. This result is explained intuitively as follows. By Theorem 4.6 (extended to the case of δ-rational players), beliefs converge to p^*: *i.e.*, players are eventually δ-indifferent. By assumption, the players flip a fair coin. Thus, attendance at the bar is likely to be near $N/2$. The theorem states that unless the capacity of the bar also happens to be near $N/2$, δ-rational learning is also ineffectual in SFBP.

Theorem 6.2 *Let* $0 < \delta < \min\{(1-\alpha) - (1-\alpha)/e, \alpha - \alpha/e\}$. *In the uniform repeated* SFBP, *assuming players are δ-rational and predictive,* $\lim_{t\to\infty} |\, p_0^t - p^*| > \delta$, *provided that δ-indifferent players flip a fair coin,*[17] *and* $c \leq N/2 - k_1\sqrt{N}$ *where* $k_1 = 1 + \ln(1/(1-\alpha))$, *or* $c \geq N/2 + k_2\sqrt{N}$ *where* $k_2 = [3 + 3\ln(1/\alpha)]/2$.

Proof 6.3 Theorem 4.6 (in the case of δ-rational players) states that beliefs converge to within δ of p^*. Thus, all players are eventually δ-indifferent, from which point on they flip a fair coin. It follows that attendance at the bar is eventually binomially

[17] Mathematically, this result holds for arbitrary probabilities p_n, p_m *s.t.* $|\, p_n - p_m\,| < \epsilon$, for small values of $\epsilon > 0$, 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.

distributed $\sim S(N, 1/2)$. Two distinct cases arise, depending on the capacity of the bar. This proof utilizes the multiplicative variant of the Chernoff bound [2].

Case 6.3.1 Assume $c \leq N/2 - k_1\sqrt{N} = N/2 - \sqrt{[1 + \ln(1/(1-\alpha))]N}$. In this case,

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

Therefore, $p^* - p_0^t > \delta$. Contradiction.

Case 6.3.2 Assume $c \geq N/2 + k_2\sqrt{N} = N/2 + \sqrt{[3 + 3\ln(1/\alpha)][N/2]}$. In this case,

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

Therefore, $p_0^t - p^* > \delta$. Contradiction. □

This theorem shows that even for δ-rational players, unless the capacity of the bar is near $N/2$, rational and predictive (*i.e.*, Nash) behavior is impossible.

Acknowledgements

The authors would like to thank Ron Even and Dean Foster for many 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.

References

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

[2] 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.

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

[4] D. Foster and P. Young. When rational learning fails. Working Paper, 1998.

[5] E. Friedman. Dynamics and rationality in ordered externality games. *Games and Economic Behavior*, 16:65–76, 1996.

[6] E. Friedman. Learnability in a class of non-atomic games arising on the Internet. *Working Paper*, 1998.

[7] A. Greenwald. *Learning to Play Network Games*. PhD thesis, Courant Institute of Mathematical Sciences, New York University, New York, May 1999.

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

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

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

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

[12] B. Mishra. A tale of two bars: A negative result for the Santa Fe bar problem. Working Paper, 1998.

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

[14] R. Parikh. A connection between Nash equilibria and propositional logic. Working Paper, 1998.

[15] B. Russell. Letter from Russell to Frege. In *From Frege to Gödel*, pages 124–125, Cambridge, 1867. Harvard University Press.

[16] T. Schelling. *Micromotives and Macrobehavior.* W.W. Norton & Company, New York, 1978.

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