Learning to Play Network Games

Amy Greenwald

Ph.D. Candidate
Courant Institute of Mathematical Sciences
New York University

MIT

Invited Talk
April 8, 1999

Contributions

- Nash equilibrium is not always the outcome of (highly) rational learning
- Nash equilibrium arises via low-rationality learning

Applications

- Santa Fe Bar Problem [Greenwald, Mishra, Parikh]
- Shopbot Economics [Greenwald and Kephart]

Overview

Part I Game Theory in Computer Science

- Learning in Network Games
  - Negative Result via Theory
  - Positive Results via Simulation

Part II Computer Science in Game Theory

- Shopbots and Pricebots

Learning in Network Games

Question

What is the outcome of learning in network games?

Assumption

"we assume that the operating point of the network is a Nash equilibrium" [Shenker, 1991]

Conclusions

- Nash equilibrium is not always the outcome of (highly) rational learning
- Nash equilibrium arises via low-rationality learning

What Network Games?

- Players: host machines
- Strategies: transmission rates and routes
- Utilities: function of network delay and loss

Why Network Games?

Approaches

1. Network Standards
   - achieves network efficiency
   - assumes infinite politeness
     - on any timeout, halve the current TCP window size [Van Jacobson, 1988]

2. Game Theory
   - invulnerable to selfish users
   - assumes infinite knowledge

Learning in Network Games

Santa Fe Bar Problem

There is one bar in Santa Fe. Every Thursday night, there is (good) live music at the bar. So every week, the city dwellers must decide whether or not to go to the bar. They would all like to go to the bar, but only if there is an available seat. Otherwise, they prefer to stay at home.

[Arthur 1991]

New York City Bar Problem

An Example

Cournot Best-Reply Dynamics

1. Rationality play best-replies to beliefs

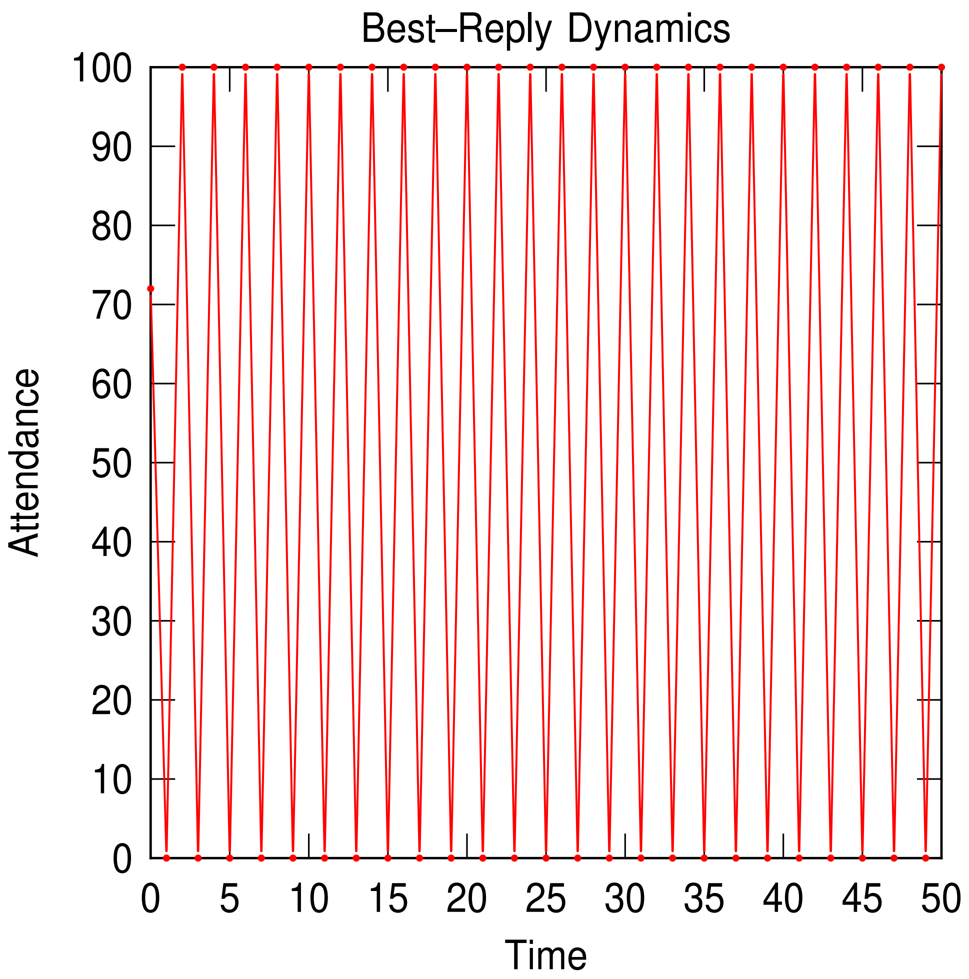
2. Beliefs today's beliefs = yesterday's outcome

Paradoxical Outcome

If bar is undercrowded, then everyone goes
$\Rightarrow$ but then bar is overcrowded

If bar is overcrowded, then everyone stays home
$\Rightarrow$ but then bar is undercrowded

Simulations



Diagonalization

Propositional Logic

$\forall \Gamma, \exists T$ *s.t.* T is consistent iff Γ has a pure strategy Nash equilibrium. [Parikh, 1998]

Axioms

$G \vee H \quad U \vee O$

$U \Rightarrow \neg H \quad O \Rightarrow \neg G \quad G \Rightarrow \neg U \quad H \Rightarrow \neg O$

Theorems

$U \Rightarrow \neg H \Rightarrow G \Rightarrow \neg U$

$O \Rightarrow \neg G \Rightarrow H \Rightarrow \neg O$

Liar Paradox

A: B is true
B: A is not true

$A \Rightarrow B \Rightarrow \neg A$

Rational Learning Paradox

Nash Equilibrium

- play best-replies to strategies

Convergence to Nash Equilibrium

1. Rationality play best-replies to beliefs
2. Predictivity beliefs converge to strategies

There is NO Rational Learning in SFBP

- play does not converge to Nash equilibrium in SFBP
- if players are rational, they cannot be predictive

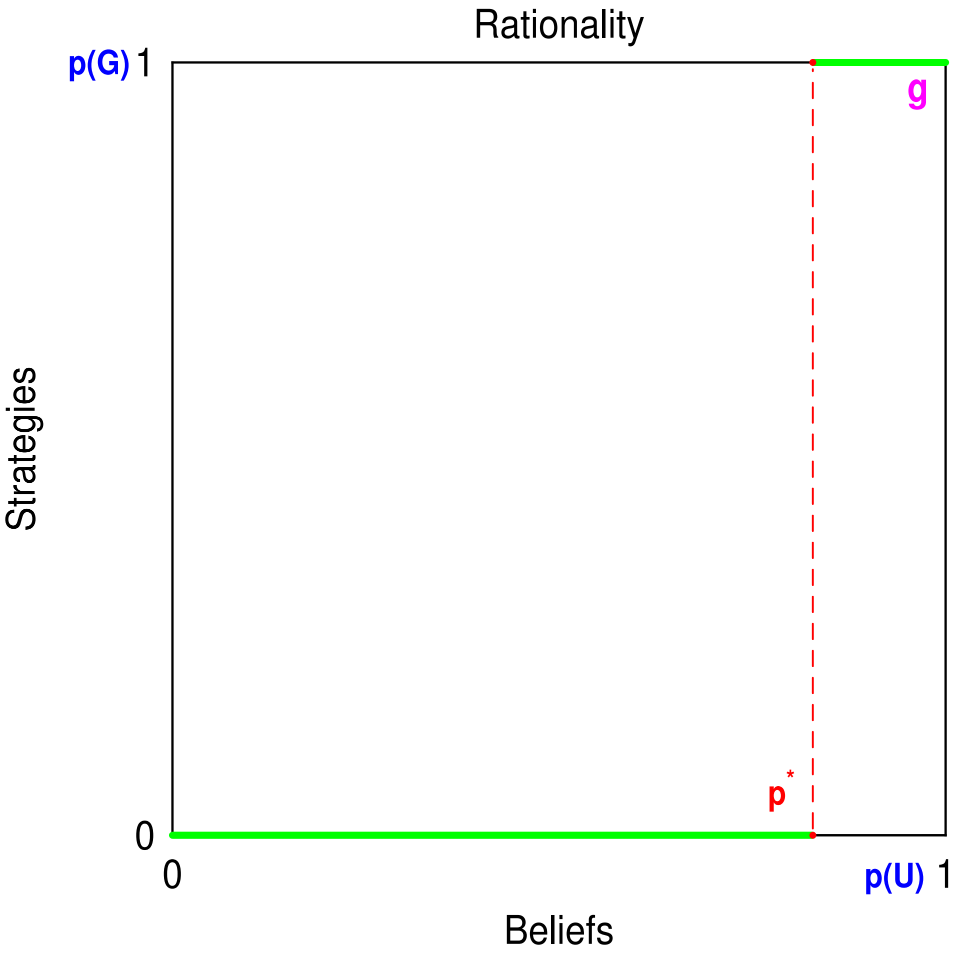
Belief-Based Learning

- Events $\mathcal{E} = \{\text{undercrowded, overcrowded}\}$
- Beliefs $\{p_n^t\}$: sequence of subjective probabilities, or beliefs, of player n that bar is undercrowded
- Strategies $\{p_0^t\}$: sequence of objective probabilities, or realization of (mixed) strategies

Rationality

Definition 1

Rational players play best-replies to beliefs.



Predictivity

Definition 2

An algorithm is predictive iff $\lim_{t\to\infty} |\, p_n^t - p_0^t \,| = 0$.

Definition 3

A set of players $\mathcal{N}$ reaches consensus iff $\forall n, m \in \mathcal{N}$, $\lim_{t\to\infty} |\, p_n^t - p_m^t \,| = 0$.

Lemma 1

If players are predictive, they reach consensus.

Key Lemma

Lemma 2

If the players are rational and predictive, then beliefs converge to p^*.

Proof $\forall n,\ p_n^t \to p^*$ as $t \to \infty$. Suppose not.

Case 1 $\exists m$ *s.t.* $p_m^t > p^*$ i.o.

$\Rightarrow$	$\forall n, p_n^t > p^*$ i.o.	by Consensus Lemma
$\Rightarrow$	$\forall n, s_n^t = 1$ i.o.	by Rationality
$\Rightarrow$	$\forall n, p_0^t = 0 \leq p^* < p_n^t$ i.o.	Contradicts Predictivity

Case 2 $\exists m$ *s.t.* $p_m^t < p^*$ i.o.

Corollary 2

If the players are rational and predictive, then strategies converge to p^*.

Proof

$$
\begin{array}{rll}
 & p_0^t \approx p_n^t,\ \forall n & \text{by Predictivity} \\
\Rightarrow & p_n^t \to p^*,\ \forall n & \text{by Key Lemma} \\
\Rightarrow & p_0^t \to p^* &
\end{array}
$$

Negative Result

Theorem 1

If players are rational and predictive, then $p_0^t \not\to p$ as $t \to \infty$, for all but finitely many p.

Interpretation

- strategies converge to p^* only very rarely
- by predictivity, beliefs rarely converge to p^*

$\Rightarrow$ play does not converge to Nash equilibrium

Rational Learning Paradox

Corollary 1

No rational learning algorithm is predictive.

Proof Suppose not

$$\begin{array}{lll} & p_0^t \approx p_n^t,\ \forall n & \text{by Predictivity} \\ \Rightarrow & p_n^t \to p^*,\ \forall n & \text{by Key Lemma} \\ \Rightarrow & p_0^t \to p^* & \text{Contradicts Theorem} \end{array}$$

Proof by Counterexample

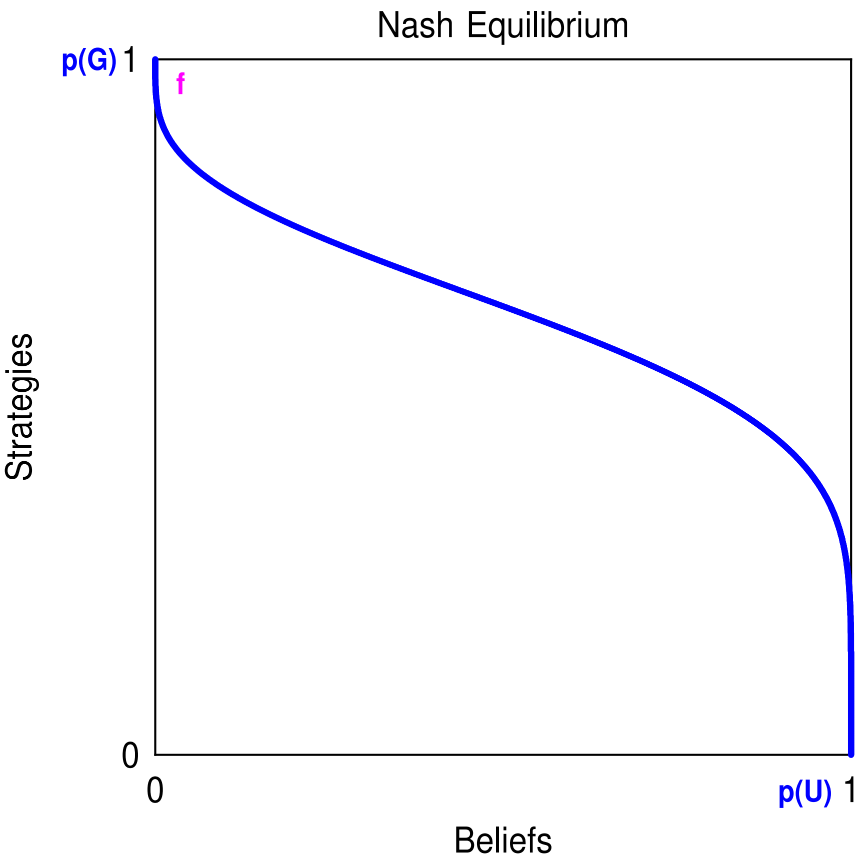
- G optimists hold beliefs $p^* + 1/t$
  - optimists always go to the bar
  - excess capacity $d = c - G$
- H pessimists hold beliefs $p^* - 1/t$
  - pessimists never go to the bar
- I realists hold beliefs p^* at all times
  - realists flip coin of bias q^*

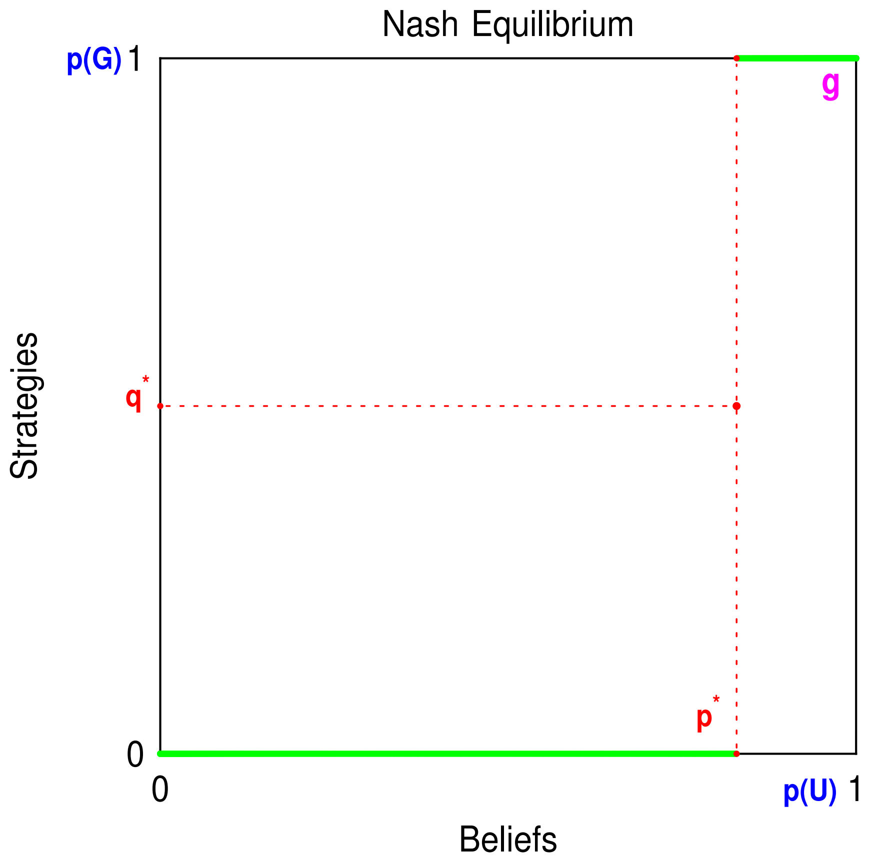
$$f(q^*) = \begin{cases} 0 & \text{if } d < 0 \\ 1 & \text{if } d \geq I \\ \sum_{j=0}^{d} \binom{I}{j} q^{*j} (1 - q^*)^{I-j} & \text{otherwise} \end{cases}$$

if $f(q^*) = p^*$, beliefs and strategies converge to p^*

therefore, play converges to Nash equilibrium

Graphical Proof

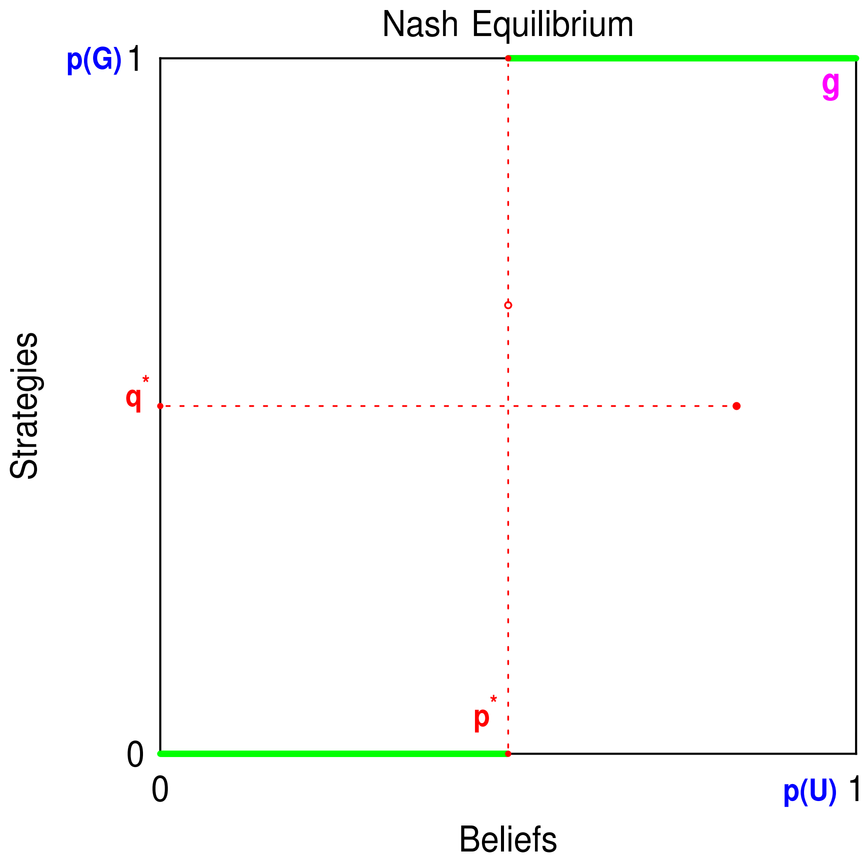




Fixed Point

$f(q^*) = p^*$ p^* is realization of q^*

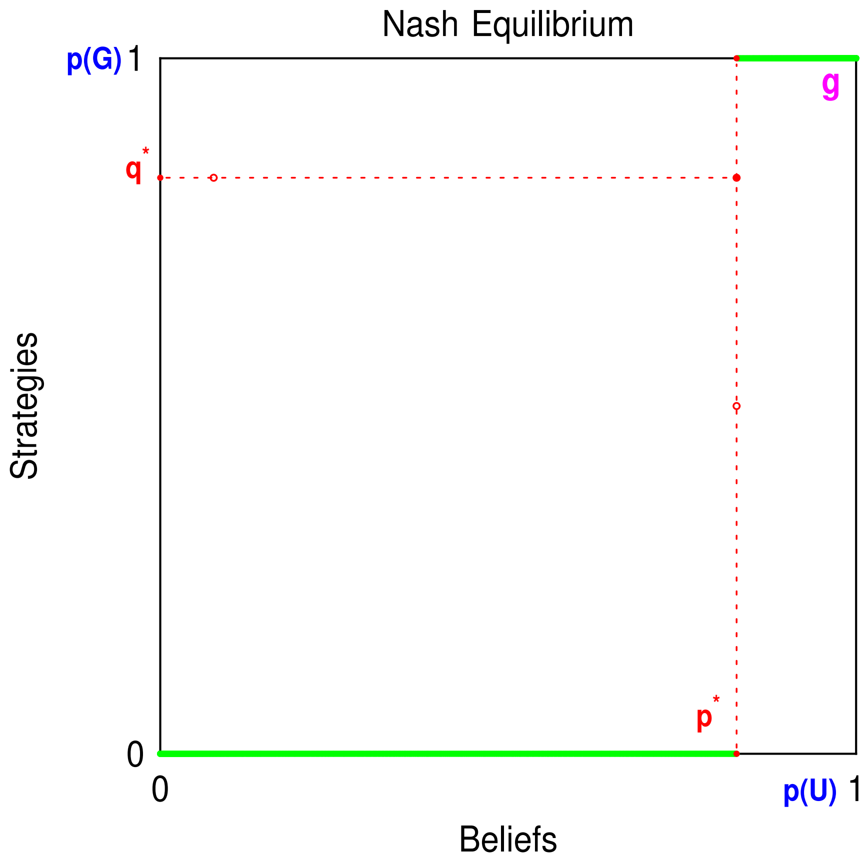
$g(p^*) = q^*$ q^* is best-reply to p^*



Non-Fixed Point

$f(q^*) \neq p^*$

$g(p^*) \neq q^*$



Non-Fixed Point

$f(q^*) \neq p^*$

$g(p^*) \neq q^*$

Related Work

1. Kalai and Lehrer [1993]

   ◦ Harsanyi Hypothesis is sufficient for prediction

2. Foster and Young [1998]

   ◦ rationality and predictivity are incompatible in 2×2 games of incomplete information, with unique mixed strategy equilibria

3. Nachbar [1997]

   ◦ no Bayesian learning algorithm converges to Nash equilibrium unless players' initial beliefs coincide with Nash equilibrium

Positive Results

Assumptions

1. Low-rationality

2. Non-Bayesian Learning
   - Additive Updating
     - Roth and Erev [1996]
     - Greenwald and Foster [1997]
   - Multiplicative Updating
     - Freund and Schapire [1995]
   - No External Regret Learning
     - Foster and Vohra [1993]
   - No Internal Regret Learning
     - Foster and Vohra [1997]
     - Hart and Mas-Colell [1997]

Multiplicative Updating

$$\rho_j^t = \sum_{x=0}^{t} \pi_j^x$$

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

where

- $i, j \in S$ are strategies
- π_i^t is the payoff of strategy i at time t
- w_i^t is the weight of strategy i at time t
- $\beta > 0$ is the learning rate

Simulations

1. Santa Fe Bar Problem

   (*a.k.a.* One-bar Problem)

   ◦ 100 agents and 1 bar of capacity 60

   ◦ 100 agents and 1 bar of varying capacity

   ◦ 100 agents and 1 bar in naive setting

   ◦ 100 agents with non-uniform values of α

2. New York City Bar Problem

   (Two-bar Problem)

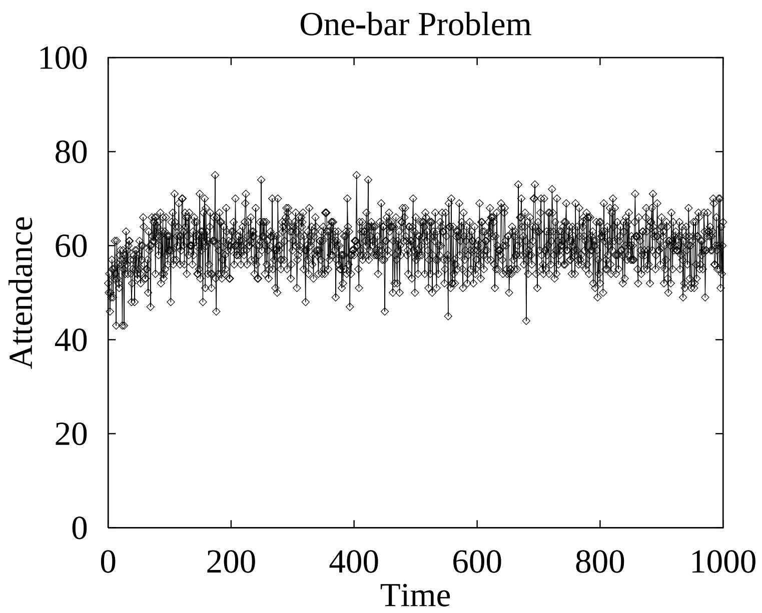
   ◦ 100 agents and 2 bars of capacity 40

   ◦ 100 agents and 2 bars in naive setting

   ◦ 100 agents with non-uniform values of α

One-bar Problem

Multiplicative Updating

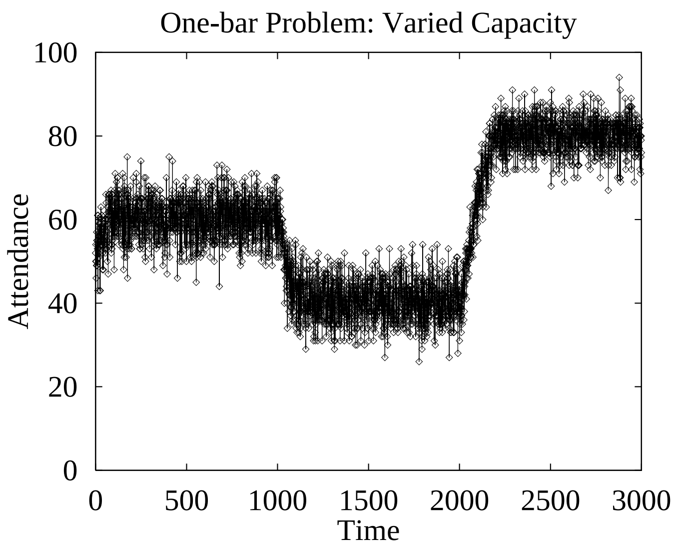


Properties of the Solution

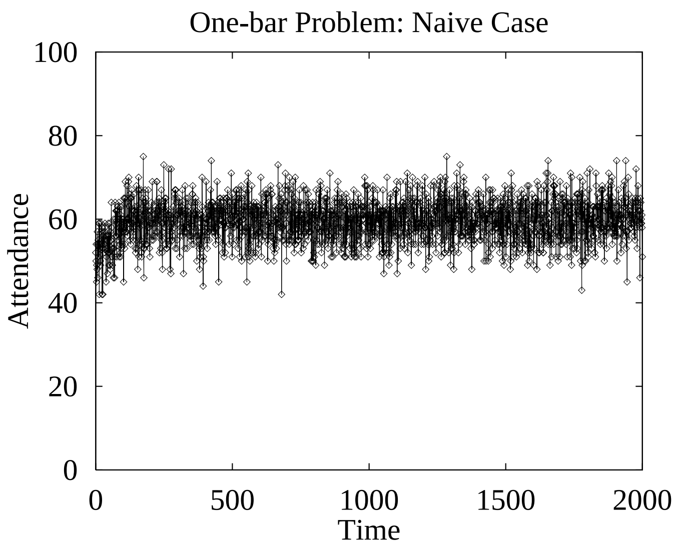
- Mixed strategy Nash equilibrium $\approx$ (.4, .6)
- Play approaches neighborhood of (.4, .6)
  - $\beta = .01 \Rightarrow (.392,.608) - (.404,.596)$
- Fair solution agents attend 60% of the time

One-bar Problem

Varied Capacity

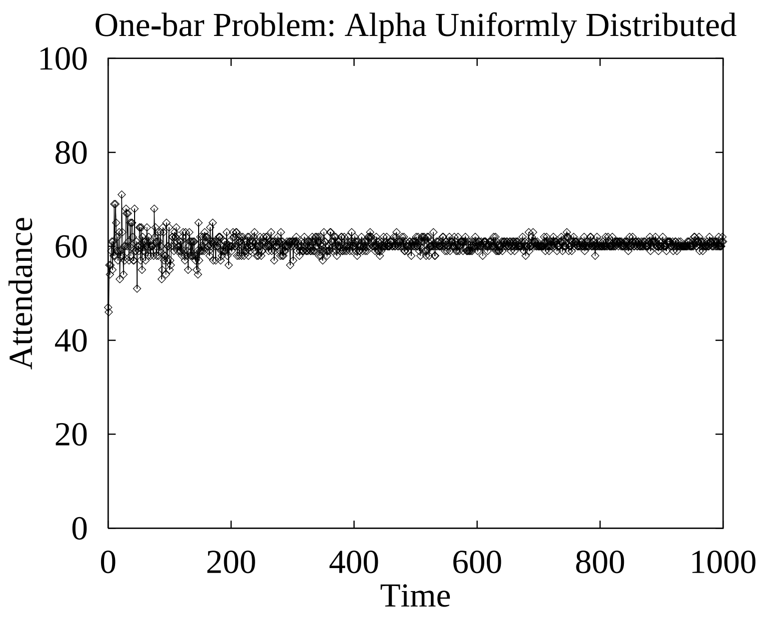


Naive Setting


One-bar Problem: Naive Case
100
80
60
40
20
0
Attendance
0
500
1000
1500
2000
Time

Non-uniform Values of α

Multiplicative Updating

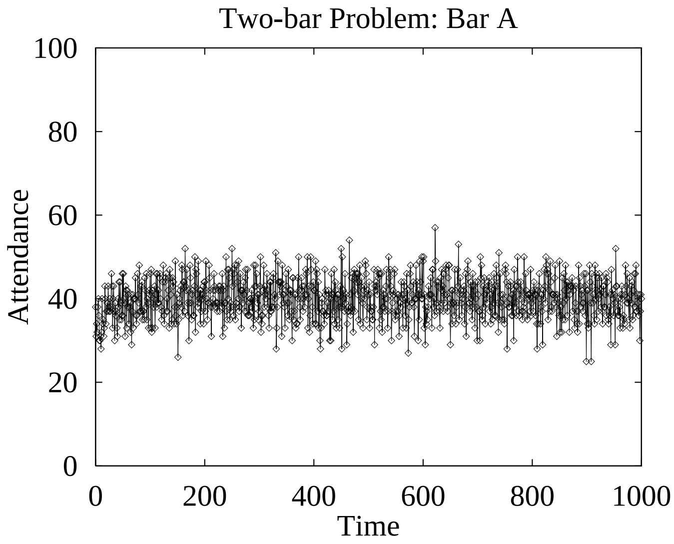


Related Work

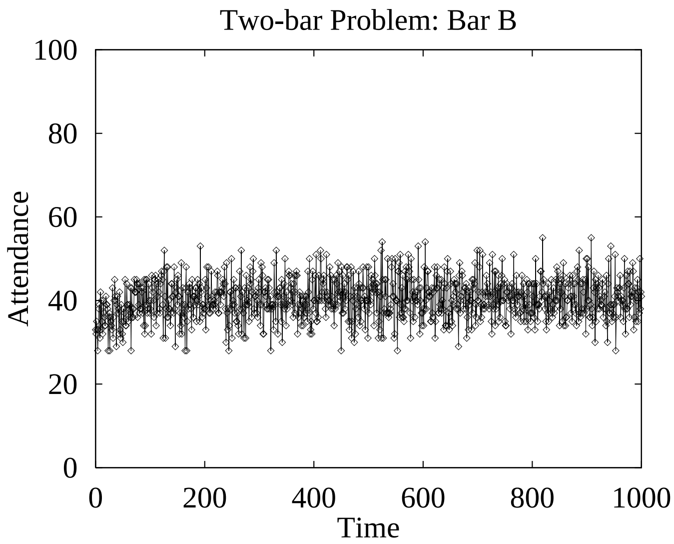
- Friedman [1996]
  - stability results for best-reply dynamics in externality games like non-uniform SFBP

Two-bar Problem

Bar A



Bar B



Conclusions

Santa Fe Bar Problem

- complete information
- homogeneous utilities
- deterministic beliefs
- synchronous play

Nash equilibrium does not always arise via rational learning in network games

[Greenwald, Mishra, and Parikh, 1998]

Network Contexts

- private information
- responsive learning
- dynamic structure
- asynchronous play

Nash equilibrium does not arise via rational learning in network contexts

[Greenwald, Friedman, and Shenker, 1998]

Shopbots and Pricebots

What are Shopbots?

- `www.acses.com`
- `www.shopper.com`
  - compares 1,000,000 prices on 100,000 products

Shopbots Today

- shopbots specialize in collecting and distributing price and quality information

Pricebots Tomorrow

- automated agents that set prices in attempt to maximize profits for sellers, just as shopbots seek to minimize costs for buyers

Diagonalization

Axioms

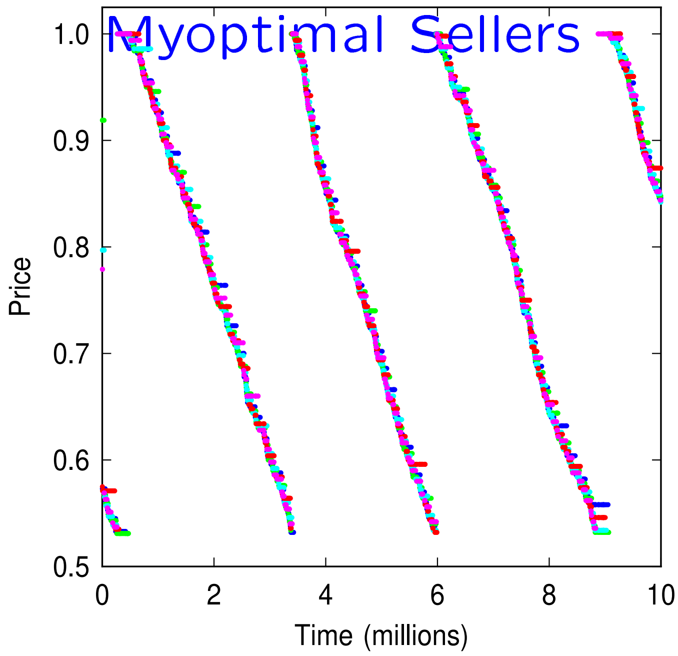
$$H_1 \vee M_1 \vee L_1 \qquad\qquad H_2 \vee M_2 \vee L_2$$

$$\begin{array}{ccccccc}
H_1 & \Rightarrow & \neg H_2 & \qquad & H_2 & \Rightarrow & \neg H_1 \\
H_1 & \Rightarrow & \neg L_2 & & H_2 & \Rightarrow & \neg L_1 \\
M_1 & \Rightarrow & \neg H_2 & & M_2 & \Rightarrow & \neg H_1 \\
M_1 & \Rightarrow & \neg M_2 & & M_2 & \Rightarrow & \neg M_1 \\
L_1 & \Rightarrow & \neg M_2 & & L_2 & \Rightarrow & \neg M_1 \\
L_1 & \Rightarrow & \neg L_2 & & L_2 & \Rightarrow & \neg L_1
\end{array}$$

Theorem

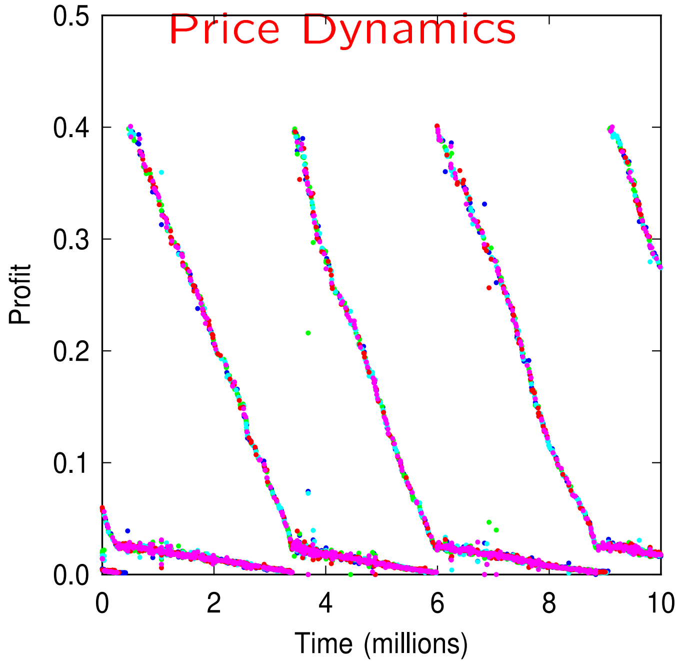
There is no pure strategy Nash equilibrium.

Myoptimal Sellers



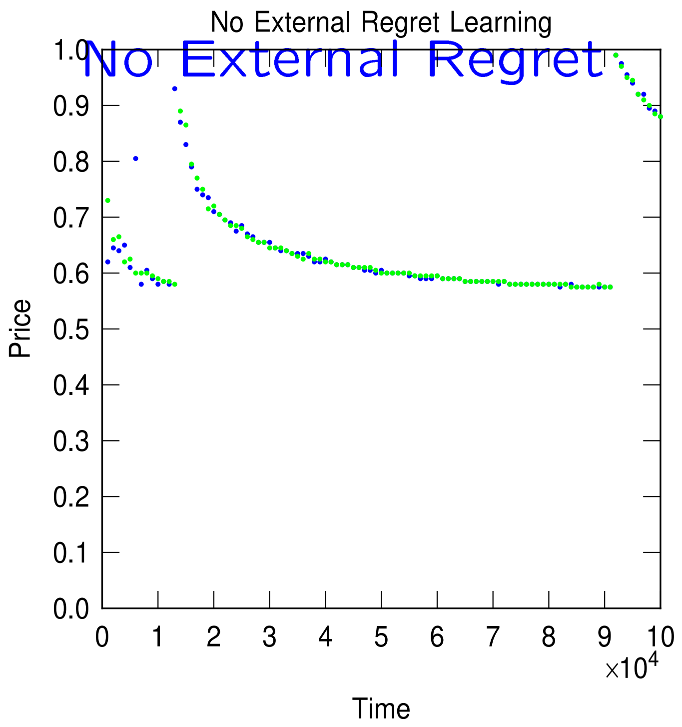
5 MY sellers, w_B=0.75

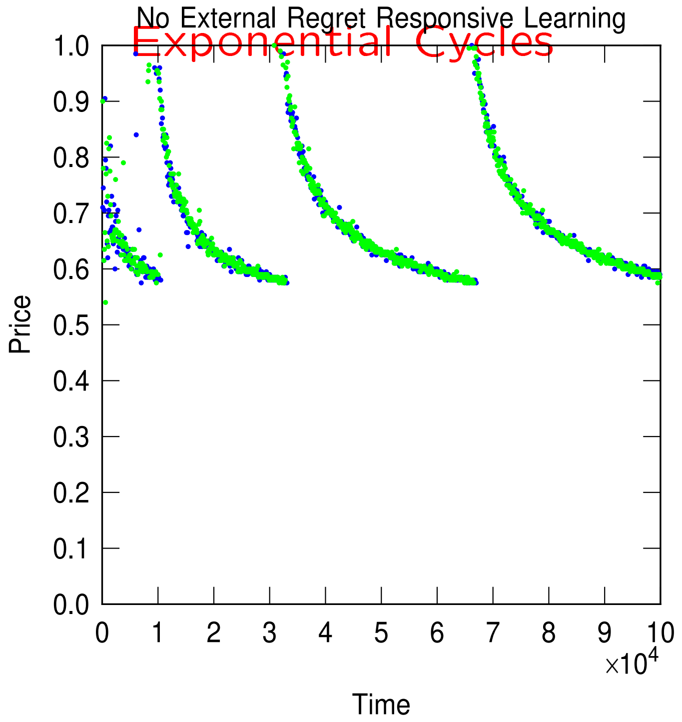
Price Dynamics



5 MY sellers, w_B=0.75

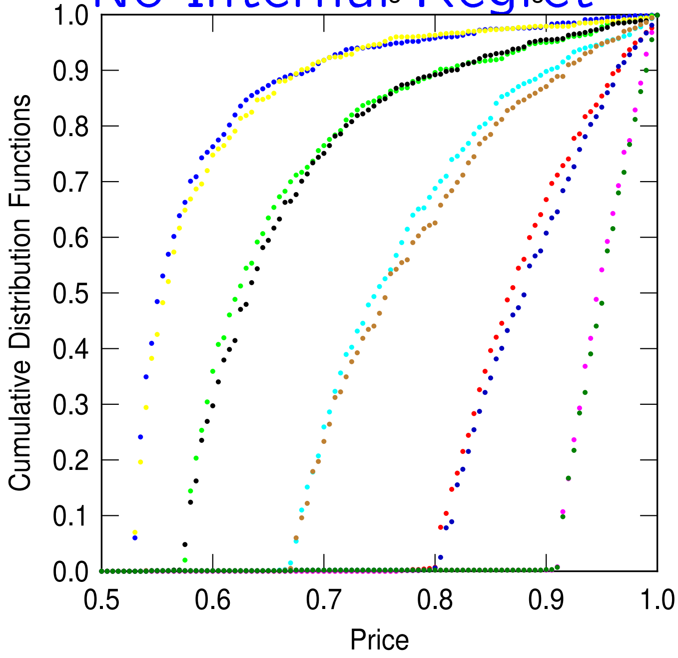
Profit Dynamics



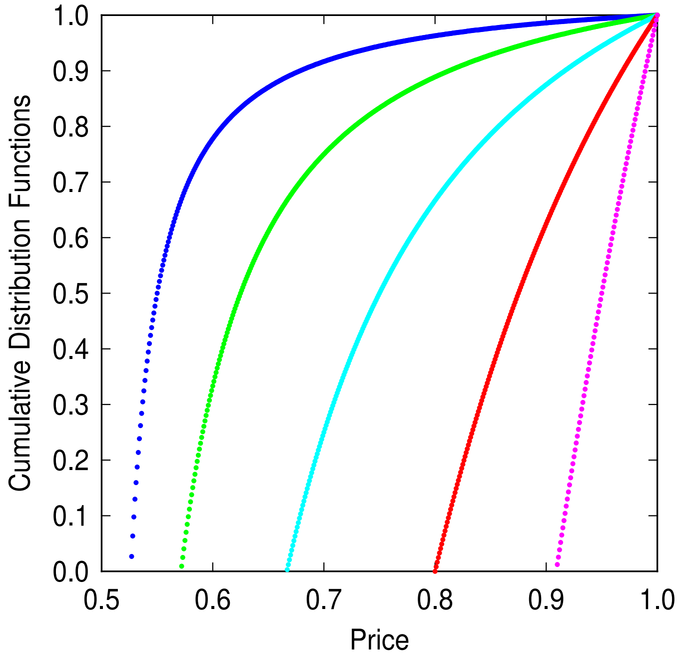


Responsive Learning

No Internal Regret



No Internal Regret



Nash Equilibrium

Contributions

- Nash equilibrium is not always the outcome of (highly) rational learning
- Nash equilibrium arises via low-rationality learning

Applications

- Santa Fe Bar Problem [Greenwald, Mishra, Parikh]
- Shopbot Economics [Greenwald and Kephart]

References

http://www.cs.nyu.edu/amygreen

http://www.research.ibm.com/infoecon