Game-Theoretic Learning

Amy Greenwald
with David Gondek, Amir Jafari, and Casey Marks
Brown University

ICML Tutorial I

July 4, 2004

Overview

1. Introduction to Game Theory
2. Regret Matching Learning Algorithms
   - Regret Matching Learns Equilibria
3. Machine Learning Applications

Introduction to Game Theory

1. General-Sum Games
    - Nash Equilibrium
    - Correlated Equilibrium
2. Zero-Sum Games
    - Minimax Equilibrium

Regret Matching Learning Algorithms

1. Regret Variations
    - No Φ-Regret Learning
    - External, Internal, and Swap Regret
2. Sufficient Conditions for No Φ-Regret Learning
    - Blackwell's Approachability Theorem
    - Gordon's Gradient Descent Theorem
      - Potential Function Argument
3. Expected and Observed Regret Matching Algorithms
    - Polynomial and Exponential Potential Functions
    - External, Internal, and Swap Regret
4. No Φ-Regret Learning Converges to Φ-Equilibria

   So Φ-Regret Matching Learns Φ-Equilibria

Machine Learning Applications

1. Online Classification
2. Offline Boosting

Game Theory and Economics

- Perfect Competition agents are price-takers
- Monopoly single entity commands all market power
- Game Theory payoffs in a game are jointly determined by the strategies of all players

Knowledge, Rationality, and Equilibrium

Assumption

Players are rational: i.e., optimizing wrt their beliefs.

Theorem

Mutual knowledge of rationality and common knowledge of beliefs is sufficient for the deductive justification of Nash equilibrium. (Aumann and Brandenburger 95)

Question

Can learning provide an inductive justification for equilibrium?

Dimensions of Game Theory

- zero-sum vs. general-sum
- simultaneous vs. sequential-move
  - deterministic vs. stochastic transitions
- cooperative vs. non-cooperative
- one-shot vs. repeated

Learning in Repeated Games

Rational Learning in Repeated Games

- An Iterative Method of Solving a Game
  Robinson 51

- Rational Learning Leads to Nash Equilibrium
  Kalai and Lehrer 93

- Prediction, Optimization, and Learning in Repeated Games
  Nachbar 97

Low-Rationality Learning in Repeated Games

Evolutionary Learning

No-Regret Learning

- No-external-regret learning converges to minimax equilibrium
- No-internal-regret learning converges to correlated equilibrium
- No-Φ-regret learning does not converge to Nash equilibrium, $\forall \Phi$

One-Shot Games

1. General-Sum Games
   - Nash Equilibrium
   - Correlated Equilibrium
2. Zero-Sum Games
   - Minimax Equilibrium

An Example

Prisoners' Dilemma

	C	D
C	$4, 4$	$0, 5$
D	$5, 0$	$1, 1$

C: Cooperate

D: Defect

Unique Nash, Correlated, and "Minimax" Equilibrium (D, D)

Normal Form Games

A normal form game is a 3-tuple $\Gamma = (I, (A_i, r_i)_{i \in I})$, where

- I is a set of players
- for all players $i \in I$,
  - a set of actions A_i with $a_i \in A_i$
  - a reward function $r_i : A \to \mathbb{R}$, where $A = \prod_{i \in I} A_i$

Normal form games are also called strategic form, or matrix, games.

Notation

Write $a = (a_i, a_{-i}) \in A$ for $a_i \in A_i$ and $a_{-i} \in A_{-i} = \prod_{j \neq i} A_j$.
Write $q = (q_i, q_{-i}) \in Q$ for $q_i \in Q_i$ and $q_{-i} \in Q_{-i} = \prod_{j \neq i} Q_i$,
where $Q_i = \{q_i \in \mathbb{R}^{A_i} | \sum_j q_{ij} = 1 \text{ \& } q_{ij} \geq 0, \forall j\}$.

Nash Equilibrium

Preliminaries

$$\mathbb{E}[r_i(q)] = \sum_{a \in A} q(a) r_i(a), \quad \text{where } q(a) = \prod_i q_i(a_i) \tag{1}$$

$$\mathsf{BR}_i(q) \equiv \mathsf{BR}_i(q_{-i}) = \{q_i^* \in Q_i \mid \forall q_i \in Q_i,\ \mathbb{E}[r_i(q_i^*, q_{-i})] \geq \mathbb{E}[r_i(q_i, q_{-i})]\} \tag{2}$$

Definition

A Nash equilibrium is an action profile q^* s.t. $q^* \in \mathsf{BR}(q^*)$.

Theorem [Nash 51]

Every finite strategic form game has a mixed strategy Nash equilibrium.

General-Sum Games

Battle of the Sexes

	B	F
B	$2, 1$	$0, 0$
F	$0, 0$	$1, 2$

Stag Hunt

	C	D
C	$2, 2$	$0, 1$
D	$1, 0$	$1 + \epsilon, 1 + \epsilon$

Coordination Game

	L	C	R
T	$3, 3$	$0, 0$	$0, 0$
M	$0, 0$	$2, 2$	$0, 0$
B	$0, 0$	$0, 0$	$1, 1$

Shapley Game

	L	C	R
T	$0, 0$	$1, 0$	$0, 1$
M	$0, 1$	$0, 0$	$1, 0$
B	$1, 0$	$0, 1$	$0, 0$

Correlated Equilibrium

Chicken

	L	R
T	6,6	2,7
B	7,2	0,0

CE

	L	R
T	1/2	1/4
B	1/4	0

$$\max 12\pi_{TL} + 9\pi_{TR} + 9\pi_{BL} + 0\pi_{BR}$$

subject to

$$\pi_{TL} + \pi_{TR} + \pi_{BL} + \pi_{BR} = 1$$
$$\pi_{TL}, \pi_{TR}, \pi_{BL}, \pi_{BR} \geq 0$$

$$\begin{aligned}
6\pi_{L|T} + 2\pi_{R|T} &\geq 7\pi_{L|T} + 0\pi_{R|T} \\
7\pi_{L|B} + 0\pi_{R|B} &\geq 6\pi_{L|B} + 2\pi_{R|B} \\
6\pi_{T|L} + 2\pi_{B|L} &\geq 7\pi_{T|L} + 0\pi_{B|L} \\
7\pi_{T|R} + 0\pi_{B|R} &\geq 6\pi_{T|R} + 2\pi_{B|R}
\end{aligned}$$

Correlated Equilibrium

Definition

An action profile $q^* \in Q$ is a correlated equilibrium iff for all strategies $a_i' \in A_i$, if $q(a_i) > 0$,

$$\sum_{a_{-i} \in A_{-i}} q(a_{-i}|a_i)\, r_i(a_i, a_{-i}) \quad \geq \quad \sum_{a_{-i} \in A_{-i}} q(a_{-i}|a_i)\, r_i(a_i', a_{-i}) \tag{3}$$

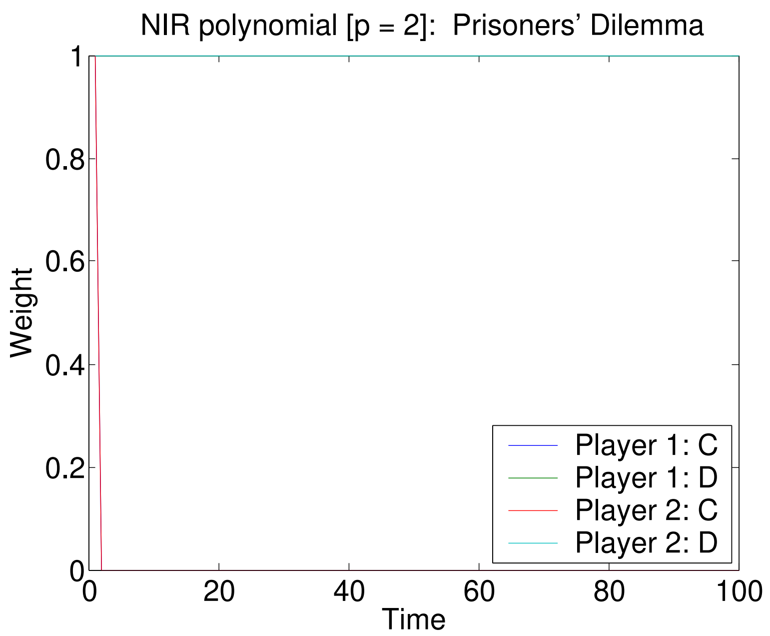
Observe

Every Nash equilibrium is a correlated equilibrium $\Rightarrow$

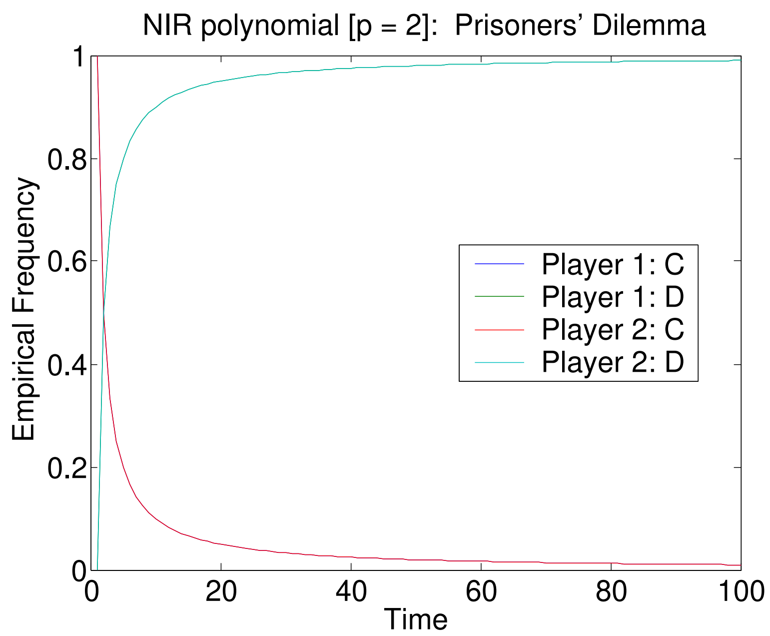
Every finite normal form game has a correlated equilibrium.

Prisoners' Dilemma

Weights

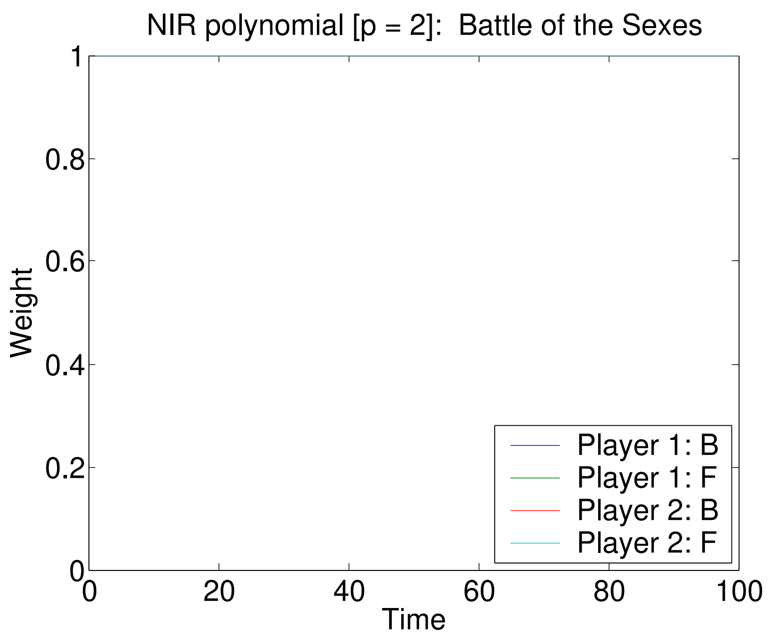


Frequencies

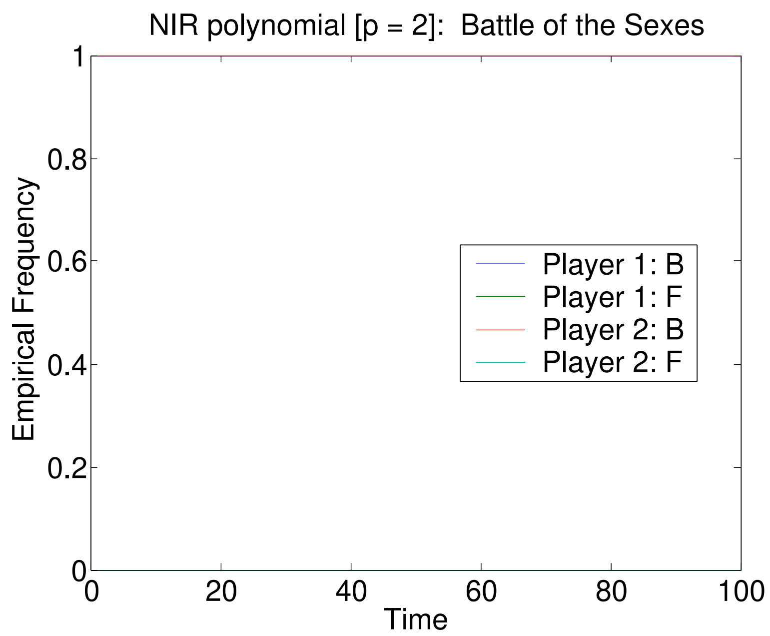


Battle of the Sexes

Weights

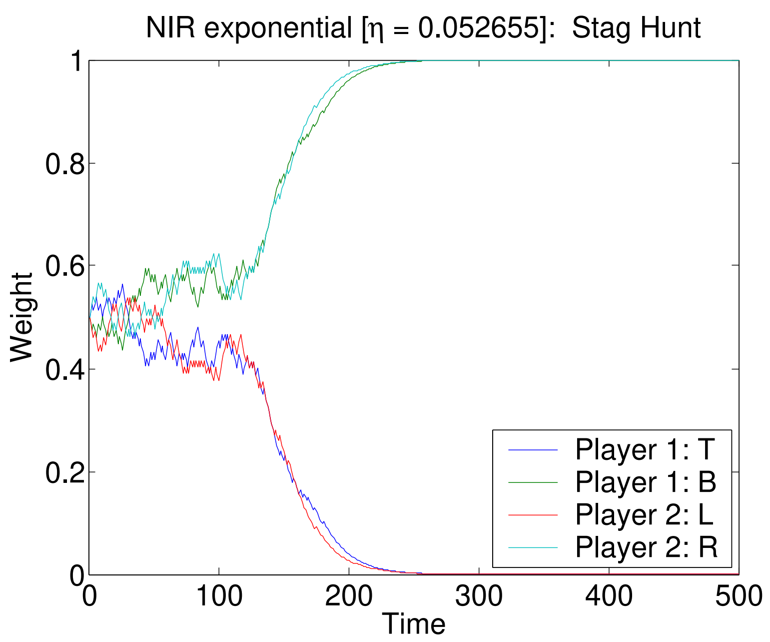


Frequencies

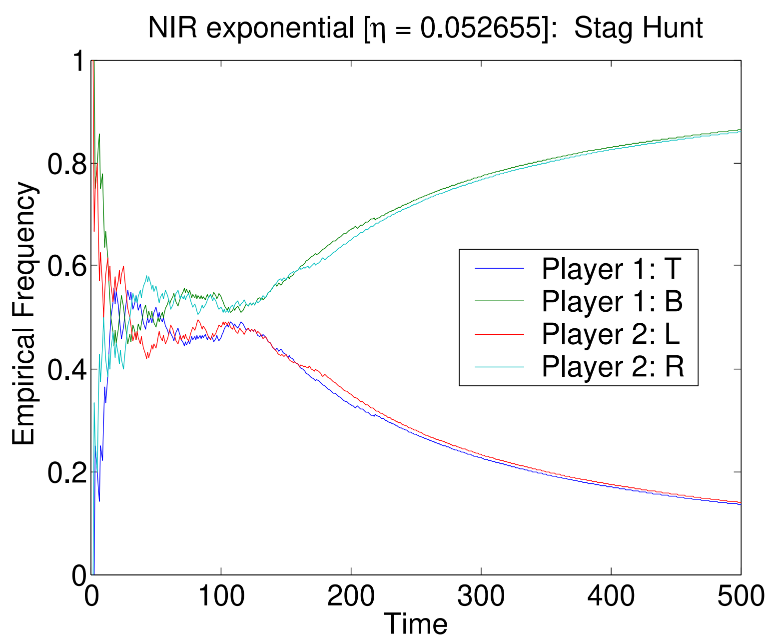


Stag Hunt

Weights

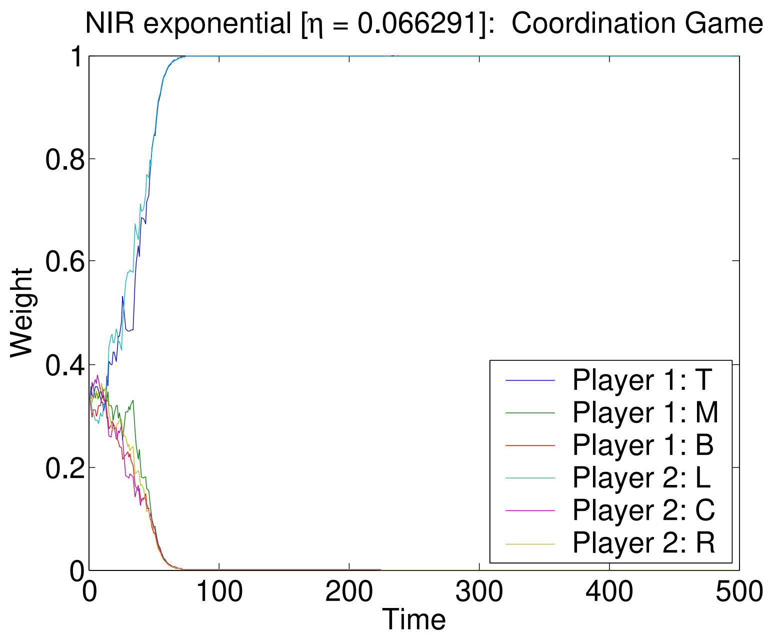


Frequencies

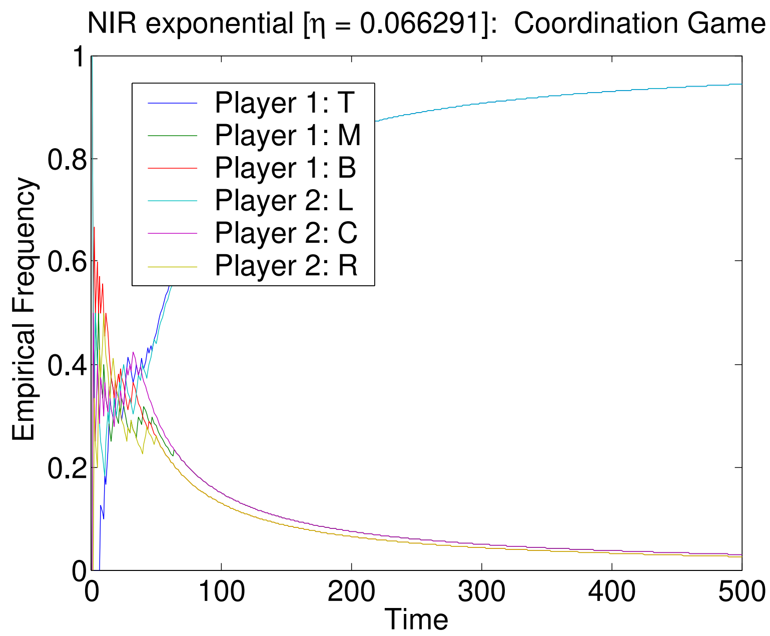


Coordination Game

Weights



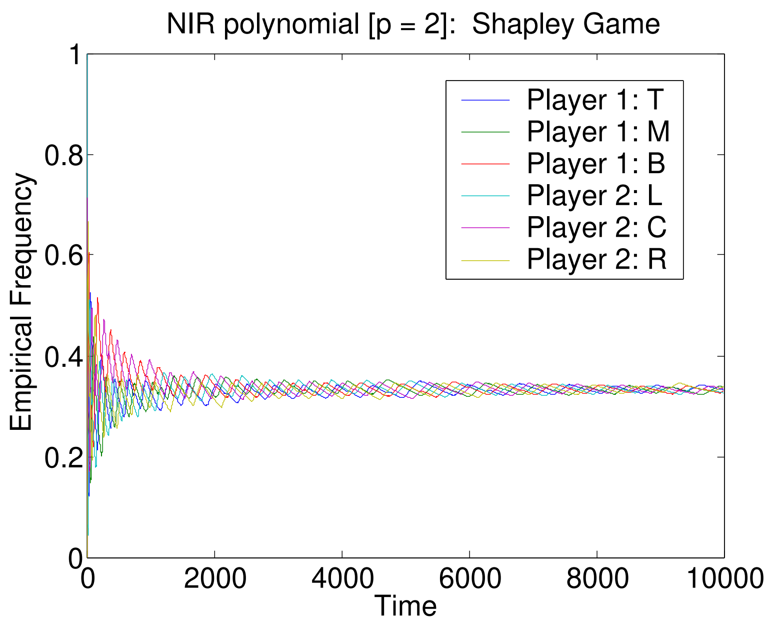
Frequencies

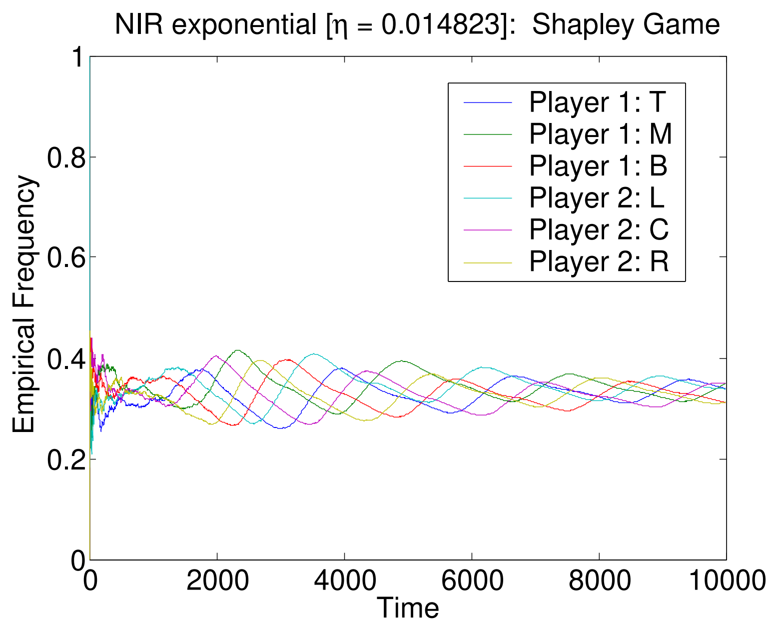


Shapley Game: No Internal Regret Learning

Frequencies

Weird Bug

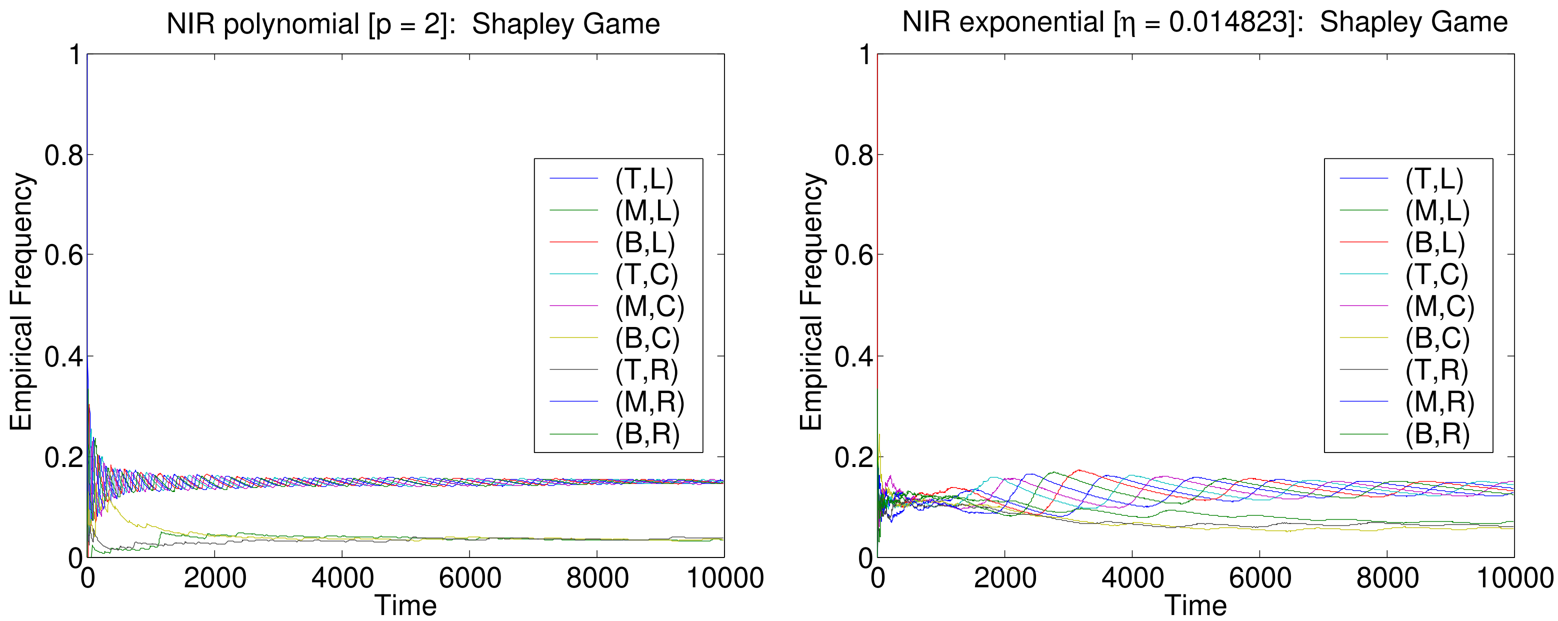




Shapley Game: No Internal Regret Learning

Joint Frequencies

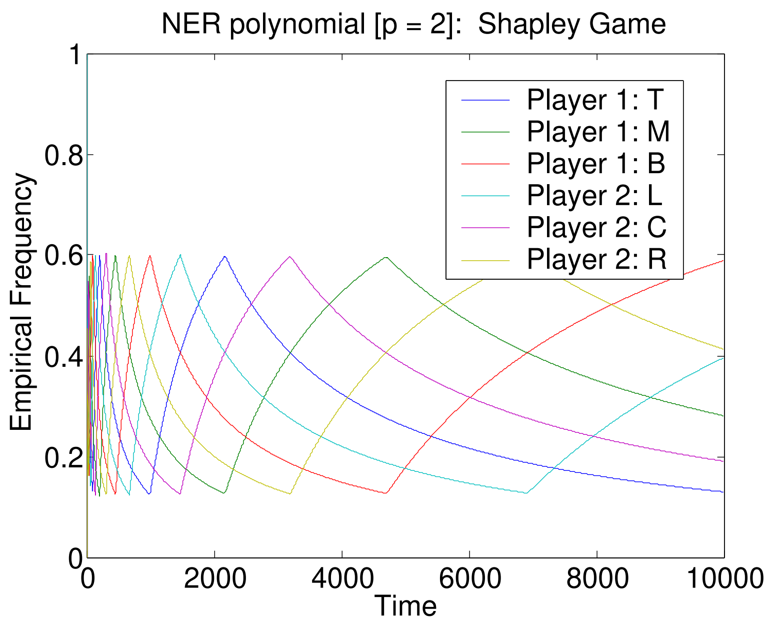
Weird Bug

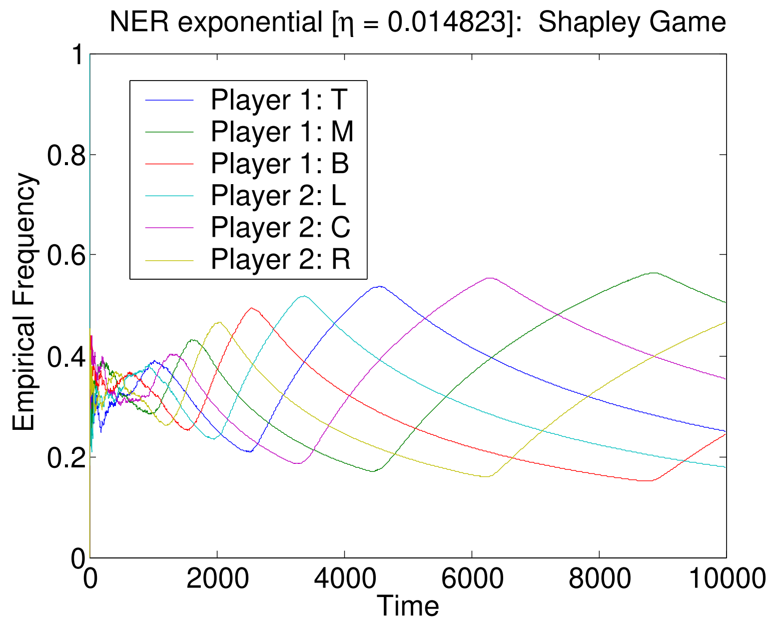


Shapley Game: No External Regret Learning

Frequencies

Weird Bug





Zero-Sum Games

Matching Pennies

	H	T
H	$-1, 1$	$1, -1$
T	$1, -1$	$-1, 1$

Rock-Paper-Scissors

	R	P	S
R	$0, 0$	$-1, 1$	$1, -1$
P	$1, -1$	$0, 0$	$-1, 1$
S	$-1, 1$	$1, -1$	$0, 0$

Definition

$\sum_{k \in I} r_k(a) = 0$, for all $a \in A$

$\sum_{k \in I} r_k(a) = c$, for all $a \in A$, for some $c \in \mathbb{R}$

Zero-Sum Games: Pure Actions

Two Players $m_{kl} \equiv M(k,l) = r_1(k,l) = -r_2(k,l)$

- **Maximizer** $k^* \in \arg\max_{k \in A_1} \min_{l \in A_2} m_{kl}$
  $v(k^*) = \max_{k \in A_1} \min_{l \in A_2} m_{kl}$

- **Minimizer** $l^* \in \arg\min_{l \in A_2} \max_{k \in A_1} m_{kl}$
  $v(l^*) = \min_{l \in A_2} \max_{k \in A_1} m_{kl}$

Example

	L	R
T	1	2
B	4	3

Zero-Sum Games: Mixed Actions

Two Players $M(p,l) = \sum_{k \in A_1} p(k) M(k,l)$
$M(k,q) = \sum_{l \in A_2} q(l) M(k,l)$

- **Maximizer** $p^* \in \arg\max_{p \in Q_1} \min_{l \in A_2} M(p,l)$
  $v(p^*) = \max_{p \in Q_1} \min_{l \in A_2} M(p,l)$

- **Minimizer** $q^* \in \arg\min_{q \in Q_2} \max_{k \in A_1} M(k,q)$
  $v(q^*) = \min_{q \in Q_2} \max_{k \in A_1} M(k,q)$

Example

	L	R
T	$+1$	-1
B	-1	$+1$

Minimax Theorem

von Neumann 28

Theorem

In two player, zero-sum games, the minimax value equals the maximin value.

Easy Direction $v(p^*) \leq v(q^*)$

- analogous to weak duality in linear programming

Hard Direction $v(q^*) \leq v(p^*)$

- akin to strong duality in linear programming

Proof of Easy Direction

Observe

M	l		l^*
k			$*$
			$\geq$
k^*	$*$	$\leq$	$*$

Therefore, $v(k^*) = M(k^*, l) \leq M(k, l^*) = v(l^*)$

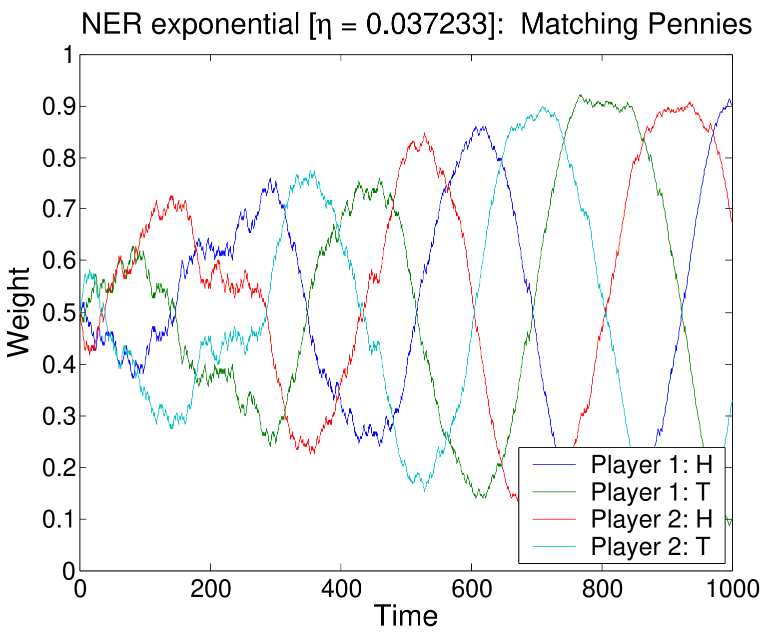
Proof of Hard Direction

Corollary of the existence of no-external-regret learning algorithms

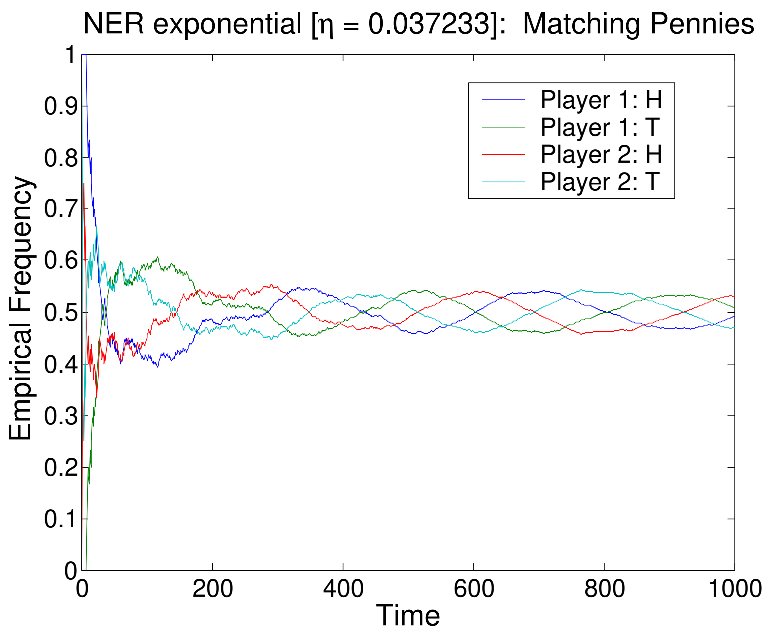
Freund & Schapire 96

Matching Pennies

Weights

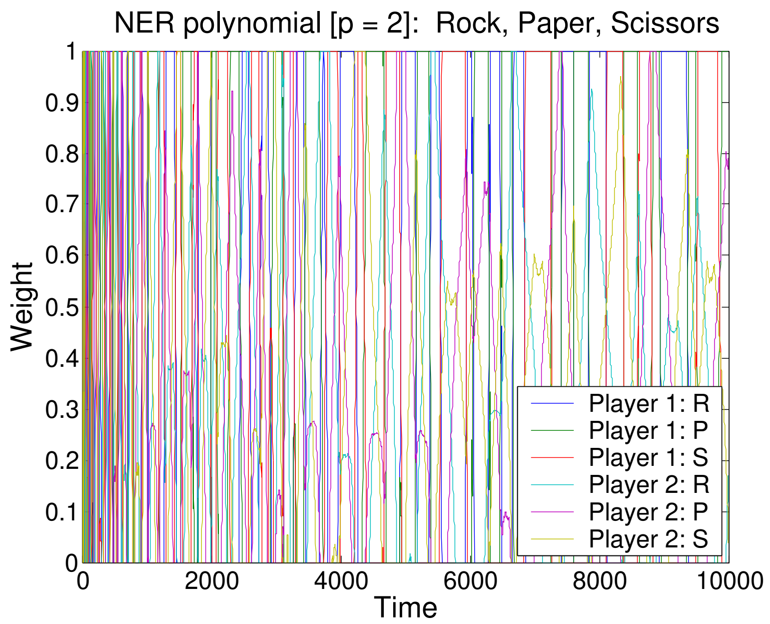


Frequencies

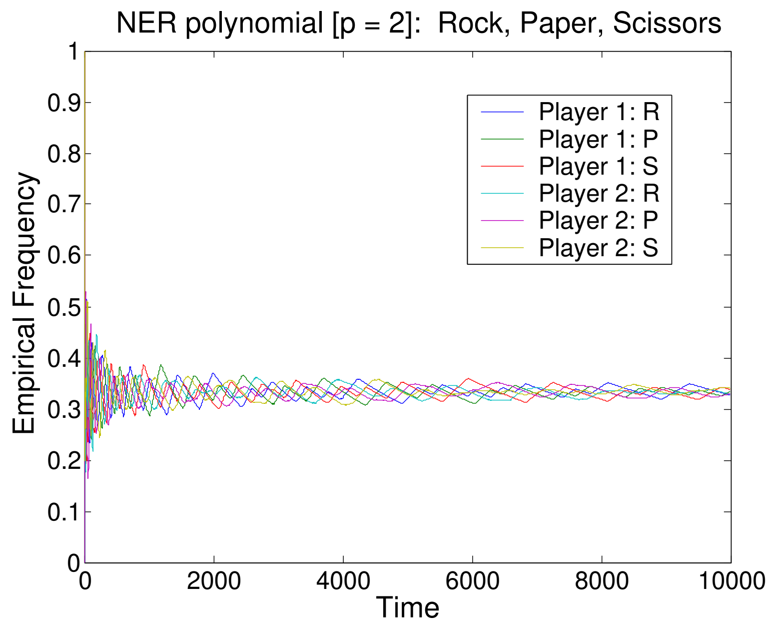


Rock-Paper-Scissors

Weights



Frequencies



Summary of Part I

“A little rationality goes a long way” [Hart 03]

No-Regret Learning

- No-external regret learning converges to minimax equilibrium
- No-internal regret learning converges to correlated equilibrium