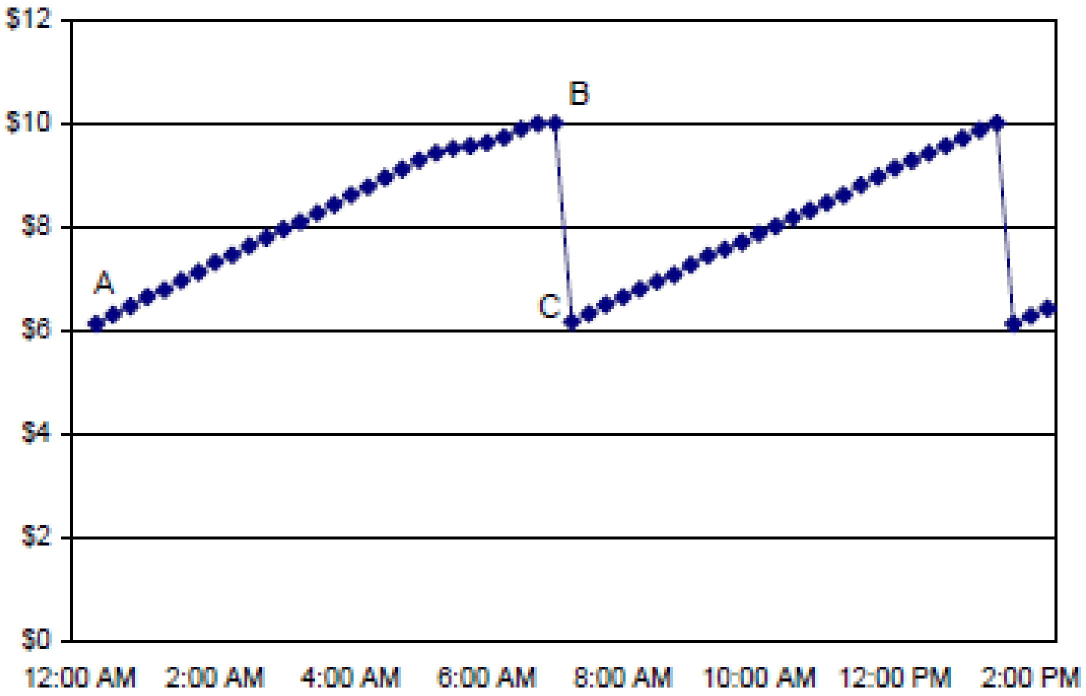
CS 407

Lecture 25: Learning in Games I

The Story So Far…

- Equilibrium in:
  - Normal form games
  - Extensive form games
  - BGP
  - Auctions
  - Mechanisms
  - …

How to get to equilibrium?



(a) 14 hours

Normal-Form Game

- A game in normal form consists of:
  - Set of players $N = \{1, \ldots, n\}$
  - Set of actions for each player A_i
  - Utility functions $u_i: A \to \mathbb{R}$
    - That is, if each j ∈ N plays the action $a_j \in A_j$, the utility of player i is $u_i(a_1, \ldots, a_n)$

Best Response Dynamics

- Start with an arbitrary joint action a
- Repeat Forever:
  - Choose one player i to update strategy
    - Need to ensure each player will get chosen eventually
  - Compute $u_i(a', a_{-i})$ for each action a'
  - Set $a_i \in argmax_{a'} u_i(a', a_{-i})$

The prisoner’s dilemma

	Cooperate	Defect
Cooperate	-1,-1	-9,0
Defect	0,-9	-6,-6

Convergence

- At some point best response dynamics may stop changing
- That is we reach an a such that $\forall i, a_i = BR_i(a)$
- We say that best response dynamics converge to a.
- Observations:
  - If we converge, a is a Pure Nash Equilibrium
  - We might not converge even if a is a Pure Nash Equilibrium???

Rock Paper Scissors

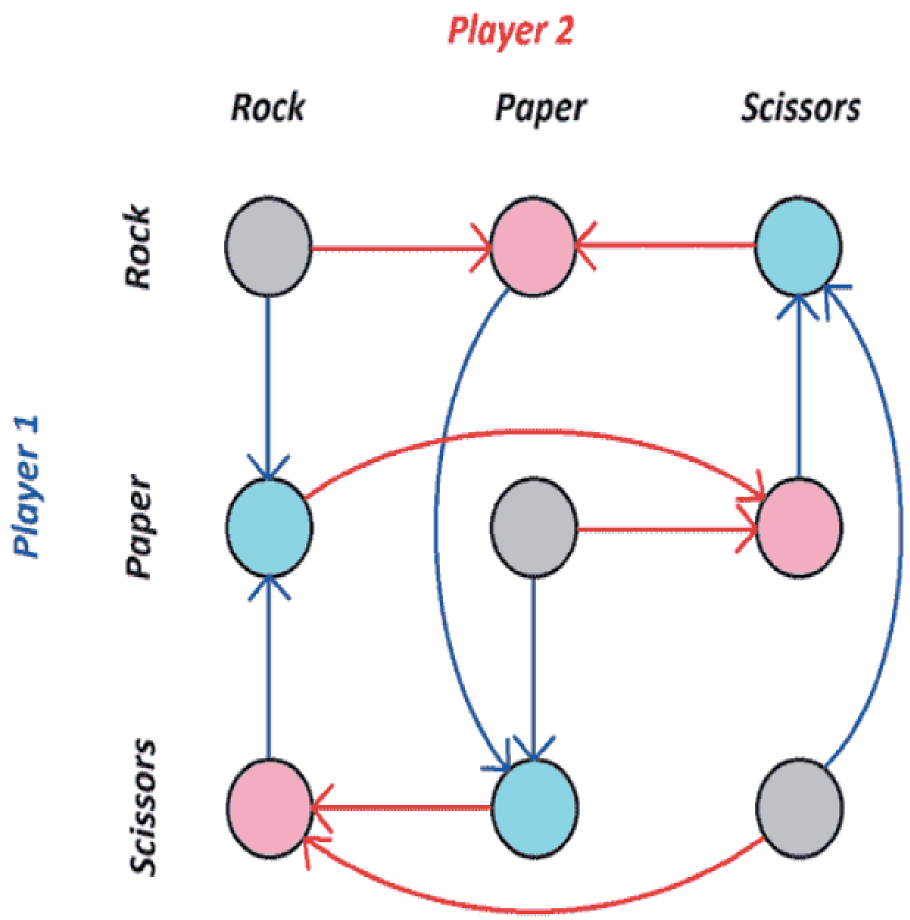
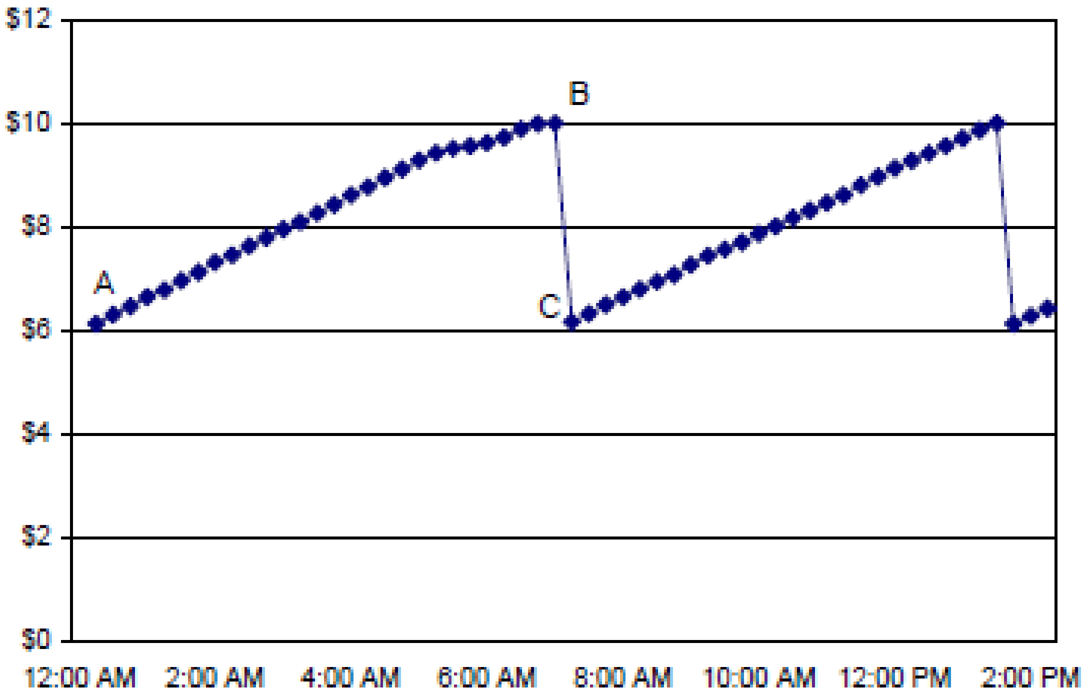


Image: Rafael Pass

Best Response Dynamics can Cycle



(a) 14 hours

Potential Functions

- A function $\Phi(a)$ is a potential function if

$$u_i(a_i', a_{-i}) > u_i(a_i, a_{-i})$$

$$\Rightarrow$$

$$\Phi(a_i', a_{-i}) > \Phi(a_i, a_{-i})$$

The prisoner's dilemma

	Cooperate	Defect
Cooperate	-1,-1	-9,0
Defect	0,-9	-6,-6

Potential Games

- Theorem: If a game has a potential function then best response dynamics converges to a pure Nash equilibrium (assuming players only change strategies when their utility improves).

- Proof: At each step $\Phi(a)$ increases and there are a finite number of a

Examples of Potential Games

- Routing games
- Cost-sharing games
- Power control in cellular networks
- Cournot Competition
  - Each firm chooses a quantity of goods q_i to produce
  - Sale price is determined by $P(\sum_i q_i)$
  - Utility $\mathrm{q_i}(P(\sum_i q_i) - c)$
  - Potential Function $\Phi(q) = (\Pi_i \mathrm{q_i})(P(\sum_i q_i) - c)$

Fictitious Play (for 2 players)

- Initialize historical aggregate strategy w_{-i}
  - Typically $w_{-i} = (0, \ldots, 0)$
  - But can use other values to get desired behavior
    - Useful on Homework!
- Repeat Forever:
  - Update $w_{-i} \mathrel{+}= s_{-i}$ with the most recent opponent strategy
  - Compute average opponent strategy $\overline{w}_{-i} = w_{-i}/(\sum w_{-i})$
  - Compute $u_i(a', \overline{w}_{-i})$ for each action a'
  - Set $a_i \in argmax_{a'} \, u_i(a', \overline{w}_{-i})$
    - Or if desired $s_i \in \Delta(argmax_{a'} \, u_i(a', \overline{w}_{-i}))$

The prisoner's dilemma

	Cooperate	Defect
Cooperate	-1,-1	-9,0
Defect	0,-9	-6,-6

Coordination Game

	Left	Right
Up	1,1	0,0
Down	0,0	1,1

Matching Pennies

	Heads	Tails
Heads	1,-1	-1,1
Tails	-1,1	1,-1

Possible Kinds of Convergence

- "Steady State"
  - $s^{t_0} = s \Rightarrow s^t = s \; \forall \; t > t_0$
- "Empirical Distribution"
  - $\lim_{t \to \infty} \overline{w}^t = s$
- "Payoffs"
  - $\lim_{t \to \infty} u(s^t) = v$

Fictitious Play Convergence

Theorem: If a is a steady state of fictitious play then a is a Nash equilibrium

Proof:

- $\lim_{t \to \infty} \overline{w}^t = a$
- $a_i \in argmax_{a'} \, u_i(a', a_{-i})$

Fictitious Play Convergence

Theorem: If the empirical distribution of fictitious play converges then it converges to a Nash equilibrium

Proof:

- $\lim_{t\to\infty} \overline{w}^t = s$
- $support(s_i) \subseteq argmax_{a'}\, u_i(a', s_{-i})$

Coordination Game

	Left	Right
Up	1,1	0,0
Down	0,0	1,1

Shapley Game

	R	P	S
R	0,0	0,1	1,0
P	1,0	0,0	0,1
S	0,1	1,0	0,0

Fictitious Play Convergence

Theorem: The empirical distribution of fictitious play converges in the following four classes of games

- Zero-sum games
- “Dominance Solvable” Games
  - i.e. iterated elimination of strictly dominated strategies leads to a unique equilibrium
- Potential Games
- The game has two players and “generic” payoffs
  - Take any game and add a small random noise to each payoff
  - Sufficient condition: all payoffs are unique

Key Takeaways

- Learning Algorithms:
  - Best Response Dynamics
  - Fictitious Play
- Other important ideas:
  - Convergence
  - Potential Function