

# Learning in Repeated Games

Game Theory Course:  
Jackson, Leyton-Brown & Shoham

[illegible]

- ## Learning in Repeated Games

[illegible]

- Game Theory Course: Jackson, Leyton-Brown & Shoham

# Fictitious Play



## Formally

- Maintain counts of opponents actions
  - For every  $a \in A$  let  $w(a)$  be the number of times the opponent has player action  $a$ .
  - Can be initialized to non-zero starting values.
- Assess opponent's strategy using these counts:

$$\sigma(a) = \frac{w(a)}{\sum_{a' \in A} w(a')}$$

- (pure strategy) best respond to this assessed strategy.
  - Break ties somehow.

# Fictitious Play

Example using matching pennies



Round	1's action	2's action	1's beliefs	2's beliefs
0			(1.5,2)	(2,1.5)
1	T	T	(1.5,3)	(2,2.5)
2	T	H	(2.5,3)	(2,3.5)
3	T	H	(3.5,3)	(2,4.5)
4	H	H	(4.5,3)	(3,4.5)
5	H	H	(5.5,3)	(4,4.5)
6	H	H	(6.5,3)	(5,4.5)
7	H	T	(6.5,4)	(6,4.5)
⋮	⋮	⋮	⋮	⋮

## Convergence

If the empirical distribution of each player's strategies converges in fictitious play, then it converges to a Nash equilibrium.

[illegible]

# Fictitious Play

## Convergence

### Theorem

*If the empirical distribution of each player's strategies converges in fictitious play, then it converges to a Nash equilibrium.*

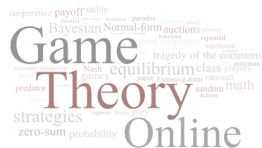
### Theorem

*Each of the following are a sufficient conditions for the empirical frequencies of play to converge in fictitious play:*

- *The game is zero sum;*
- *The game is solvable by iterated elimination of strictly dominated strategies;*
- *The game is a potential game;*
- *The game is  $2 \times n$  and has generic payoffs.*



## Definitions



The **regret** an agent experiences at time  $t$  for not having played  $s$  is  $R^t(s) = \alpha^t - \alpha^t(s)$ .



# No-regret Learning

## Definitions



### Definition (Regret)

The **regret** an agent experiences at time  $t$  for not having played  $s$  is  $R^t(s) = \alpha^t - \alpha^t(s)$ .

### Definition (No-regret learning rule)

A learning rule exhibits **no regret** if for any pure strategy of the agent  $s$  it holds that  $Pr([\liminf R^t(s)] \leq 0) = 1$ .

# No-regret Learning

## Regret Matching

- Example learning rule that exhibits no regret: **Regret Matching**.



# No-regret Learning

## Regret Matching



- Example learning rule that exhibits no regret: **Regret Matching**.
- At each time step each action is chosen with probability proportional to its regret. That is,

$$\sigma_i^{t+1}(s) = \frac{R^t(s)}{\sum_{s' \in S_i} R^t(s')},$$

where  $\sigma_i^{t+1}(s)$  is the probability that agent  $i$  plays pure strategy  $s$  at time  $t + 1$ .

# No-regret Learning

## Regret Matching



- Example learning rule that exhibits no regret: **Regret Matching**.
- At each time step each action is chosen with probability proportional to its regret. That is,

$$\sigma_i^{t+1}(s) = \frac{R^t(s)}{\sum_{s' \in S_i} R^t(s')},$$

where  $\sigma_i^{t+1}(s)$  is the probability that agent  $i$  plays pure strategy  $s$  at time  $t + 1$ .

- Converges to a correlated equilibrium for finite games.