

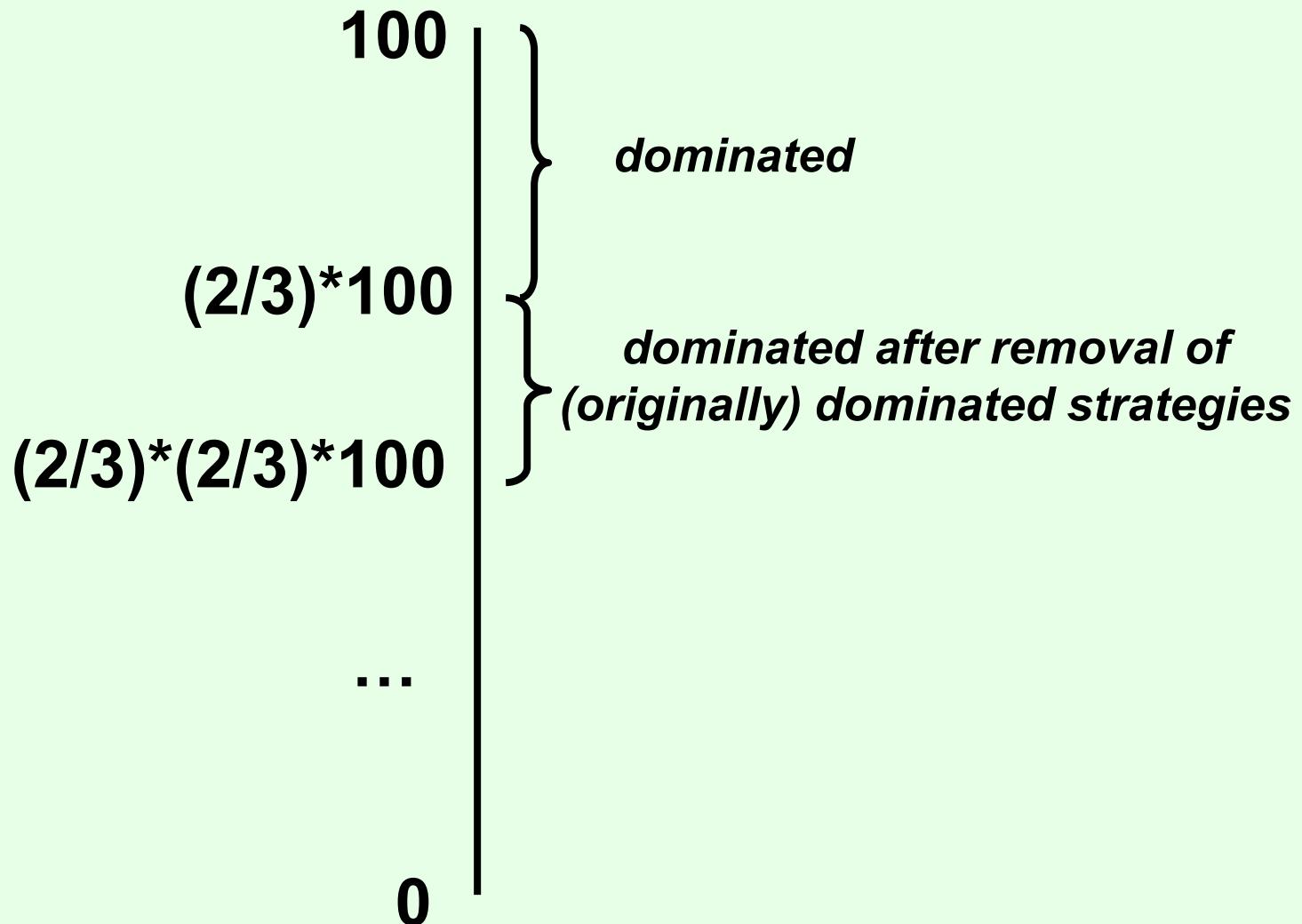
# Learning in games

Vincent Conitzer  
[conitzer@cs.duke.edu](mailto:conitzer@cs.duke.edu)

# “2/3 of the average” game

- Everyone writes down a number between 0 and 100
- Person closest to 2/3 of the average wins
- Example:
  - A says 50
  - B says 10
  - C says 90
  - Average(50, 10, 90) = 50
  - 2/3 of average = 33.33
  - A is closest ( $|50-33.33| = 16.67$ ), so A wins

# “2/3 of the average” game revisited



# Learning in (normal-form) games

- Approach we have taken so far when playing a game: just compute an optimal/equilibrium strategy
- Another approach: **learn** how to play a game by
  - playing it many times, and
  - updating your strategy based on experience
- Why?
  - Some of the game's utilities (especially the other players') may be **unknown** to you
  - The other players may **not** be playing an equilibrium strategy
  - Computing an optimal strategy can be **hard**
  - Learning is what **humans** typically do
  - ...
- Learning strategies ~ strategies for the repeated game
- Does learning converge to equilibrium?

# Iterated best response

- In the first round, play something arbitrary
- In each following round, play a best response against what the other players played in the **previous** round
- If all players play this, it can converge (i.e., we reach an equilibrium) or cycle

0, 0	-1, 1	1, -1
1, -1	0, 0	-1, 1
-1, 1	1, -1	0, 0

*rock-paper-scissors*

-1, -1	0, 0
0, 0	-1, -1

*a simple congestion game*

- **Alternating best response**: players alternatingly change strategies: one player best-responds each odd round, the other best-responds each even round

# Fictitious play [Brown 1951]

- In the first round, play something arbitrary
- In each following round, play a best response against the **empirical distribution** of the other players' play
  - I.e., as if other player randomly selects from his past actions
- Again, if this converges, we have a Nash equilibrium
- Can still fail to converge...

0, 0	-1, 1	1, -1
1, -1	0, 0	-1, 1
-1, 1	1, -1	0, 0

*rock-paper-scissors*

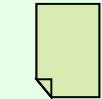
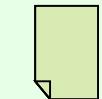
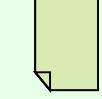
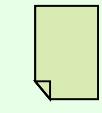
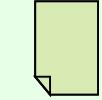
-1, -1	0, 0
0, 0	-1, -1

*a simple congestion game*

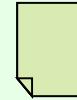
# Fictitious play on rock-paper-scissors

			
	0, 0	-1, 1	1, -1
	1, -1	0, 0	-1, 1
	-1, 1	1, -1	0, 0

Row



Column



30% R, 50% P, 20% S

30% R, 20% P, 50% S

# Does the empirical distribution of play converge to equilibrium?

- ... for iterated best response?
- ... for fictitious play?

3, 0	1, 2
1, 2	2, 1

# Fictitious play is guaranteed to converge in...

- Two-player zero-sum games [Robinson 1951]
- Generic 2x2 games [Miyasawa 1961]
- Games solvable by iterated strict dominance [Nachbar 1990]
- Weighted potential games [Monderer & Shapley 1996]
- **Not in general** [Shapley 1964]
- But, fictitious play always converges to the set of  $\frac{1}{2}$ -approximate equilibria [Conitzer 2009; more detailed analysis by Goldberg, Savani, Sørensen, Ventre 2011]

# Shapley's game on which fictitious play does not converge

- starting with  $(U, M)$ :

0, 0	0, 1	1, 0
1, 0	0, 0	0, 1
0, 1	1, 0	0, 0

# “Teaching”

- Suppose you are playing against a player that uses a strategy that eventually learns to best-respond
- Also suppose you are very patient, i.e., you only care about what happens in the long run
- How will you (the row player) play in the following repeated games?

4, 4	3, 5
5, 3	0, 0

1, 0	3, 1
2, 1	4, 0

- Note relationship to optimal strategies to commit to
- There is some work on learning strategies that are in **equilibrium** with each other [Brafman & Tennenholtz AIJ04]

# Evolutionary game theory

- Given: a symmetric game

	dove	hawk
dove	1, 1	0, 2
hawk	2, 0	-1, -1

Nash equilibria: (d, h), (h, d), ((.5, .5), (.5, .5))

- A large population of players plays this game, players are randomly matched to play with each other
- Each player plays a pure strategy
  - Fraction of players playing strategy  $s = p_s$
  - $p$  is vector of all fractions  $p_s$  (the **state**)
- Utility for playing  $s$  is  $u(s, p) = \sum_{s'} p_{s'} u(s, s')$
- Players **reproduce** at a rate that is proportional to their utility, their offspring play the same strategy
  - Replicator dynamic**
- $dp_s(t)/dt = p_s(t)(u(s, p(t)) - \sum_{s'} p_{s'} u(s', p(t)))$
- What are the **steady states** of this?

# Stability

	dove	hawk
dove	1, 1	0, 2
hawk	2, 0	-1, -1

- A steady state is **stable** if slightly perturbing the state will not cause us to move far away from the state
- E.g. everyone playing dove is not stable, because if a few hawks are added their percentage will grow
- What about the mixed steady state?
- Proposition: every stable steady state is a Nash equilibrium of the symmetric game
- Slightly stronger criterion: a state is **asymptotically stable** if it is stable, and after slightly perturbing this state, we will (in the limit) return to this state

# Evolutionarily stable strategies

- Now suppose players play **mixed** strategies
- A (single) mixed strategy  $\sigma$  is **evolutionarily stable** if the following is true:
  - Suppose all players play  $\sigma$
  - Then, whenever a very small number of **invaders** enters that play a different strategy  $\sigma'$ ,
  - the players playing  $\sigma$  must get strictly **higher** utility than those playing  $\sigma'$  (i.e.,  $\sigma$  must be able to **repel invaders**)
- $\sigma$  will be evolutionarily stable if and only if for all  $\sigma'$ 
  - $u(\sigma, \sigma) > u(\sigma', \sigma)$ , or:
  - $u(\sigma, \sigma) = u(\sigma', \sigma)$  and  $u(\sigma, \sigma') > u(\sigma', \sigma')$
- Proposition: every evolutionarily stable strategy is asymptotically stable under the replicator dynamic

## Invasion (1/2)

	Dove	Hawk
Dove	1, 1	0, 2
Hawk	2, 0	-1, -1

- Given: population  $P_1$  that plays  $\sigma = 40\%$  Dove,  $60\%$  Hawk
- Tiny population  $P_2$  that plays  $\sigma' = 70\%$  Dove,  $30\%$  Hawk **invades**
- $u(\sigma, \sigma) = .16*1 + .24*2 + .36*(-1) = .28$  but  
 $u(\sigma', \sigma) = .28*1 + .12*2 + .18*(-1) = .34$
- $\sigma'$  (initially) grows in the population; invasion is **successful**

## Invasion (2/2)

	Dove	Hawk
Dove	1, 1	0, 2
Hawk	2, 0	-1, -1

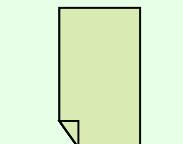
- Now  $P_1$  plays  $\sigma = 50\% \text{ Dove}, 50\% \text{ Hawk}$
- Tiny population  $P_2$  that plays  $\sigma' = 70\% \text{ Dove}, 30\% \text{ Hawk}$  **invades**
- $u(\sigma, \sigma) = u(\sigma', \sigma) = .5$ , so second-order effect:
- $u(\sigma, \sigma') = .35*1 + .35*2 + .15*(-1) = .9$  but  
 $u(\sigma', \sigma') = .49*1 + .21*2 + .09*(-1) = .82$
- $\sigma'$  shrinks in the population; invasion is **repelled**

# Evolutionarily stable strategies

[Price and Smith, 1973]

- A strategy  $\sigma$  is **evolutionarily stable** if the following two conditions both hold:
  - (1) For all  $\sigma'$ , we have  $u(\sigma, \sigma) \geq u(\sigma', \sigma)$  (i.e., **symmetric Nash equilibrium**)
  - (2) For all  $\sigma' (\neq \sigma)$  with  $u(\sigma, \sigma) = u(\sigma', \sigma)$ , we have  $u(\sigma, \sigma') > u(\sigma', \sigma')$

# Rock- Paper- Scissors

			
	0, 0	-1, 1	1, -1
	1, -1	0, 0	-1, 1
	-1, 1	1, -1	0, 0

- Only one Nash equilibrium (Uniform)
- $u(\text{Uniform}, \text{Rock}) = u(\text{Rock}, \text{Rock})$
- No ESS

# The standard $\Sigma_2^P$ -complete problem

*Input: Boolean formula  $f$  over variables  $X_1$  and  $X_2$*

*Q: Does there exist an assignment of values to  $X_1$ , such that for every assignment of values to  $X_2$ ,  $f$  is true?*

# The ESS problem

*Input:* symmetric 2-player normal-form game.

*Q:* Does it have an evolutionarily stable strategy?

(Hawk-Dove: yes. Rock-Paper-Scissors: no. Safe-Left-Right: no.)

