CMSC 474, Game Theory

6b. Evolutionary Games

Dana Nau

University of Maryland

Evolutionary Stability (from Lecture 3)

- Concept from evolutionary biology
- Population of various species (kinds of organisms)
  - Relative "fitness" causes their proportions to grow or shrink
- Pure strategy ⇔ a particular species
  - Everything that might affect evolutionary fitness of this kind of organism
    - size, aggressiveness, sensory abilities, intelligence, …
- Mixed strategy = relative proportions of each species
- *Fitness* of a species = expected payoff from interacting with a random member of the population
  - High payoff ⇒ more likely to reproduce ⇒ proportion is likely to grow
  - Low payoff ⇒ less likely to reproduce ⇒ proportion is likely to shrink

Replicator Dynamic

- Assume an infinite population, continuous time, n possible strategies
- Population profile $\mathbf{x} = (x_1, \ldots, x_n)$, where x_i = proportion of strategy i
  - $0 \le x_i \le 1$ and $x_1 + \ldots + x_n = 1$
- Strategy i's *fitness* is $f_i \ge 0$
  - Expected number of offspring of each individual with strategy i
  - f_i high $\Rightarrow$ more likely to reproduce $\Rightarrow$ x_i likely to increase
  - f_i low $\Rightarrow$ less likely to reproduce $\Rightarrow$ x_i likely to decrease
- *Replicator dynamic*: rate of change is proportional to the difference between f_i and the average fitness $f_{avg} = \sum_j x_j f_j$
  - $dx_i / dt = x_i(f_i - f_{avg})$
  - Has some appealing mathematical properties
  - Does well at reflecting changes in animal populations

Evolutionary Stability (from Lecture 3)

- Infinite population, all of species 1: $x_1 = 1$
  - Invasion of species 2 at level δ: $x_1 = 1-\delta$ and $x_2 = \delta$
- For each species' fitness, use its expected payoff
  - $f_1 = u_1 = (1-\delta)a + \delta b$
  - $f_1 = u_2 = (1-\delta)c + \delta d$
- If $f_1 < f_2$ then species 1's proportion will shrink and species 2's will grow
  - $f_1 < f_{avg} < f_2$
  - $dx_1 / dt = x_1 (f_1 - f_{avg}) < 0$ and $dx_2 / dt = x_2 (f_2 - f_{avg}) > 0$
- If $f_1 > f_2$ then species 1's proportion will grow and species 2's will shrink
  - $f_1 > f_{avg} > f_2$
  - $dx_1 / dt = x_1 (f_1 - f_{avg}) > 0$ and $dx_2 / dt = x_2 (f_2 - f_{avg}) < 0$
- Species 1 is evolutionarily stable against species 2 if $\exists\, \varepsilon > 0$ such that for all $\delta < \varepsilon$, $f_1 > f_2$
  - $a > c$, or $a = c$ and $b > d$

	1	2
1	a	b
2	c	d

Hawk-Dove Game (from Lecture 3)

- 2 animals contend for a piece of food
- Each animal may be either a hawk (H) or a dove (D)
  - Prize is worth 6 to each; fighting has a cost of 5
- Hawk meets dove ⇒ hawk gets prize without a fight
  - Hawk gets 6, dove gets 0
- 2 doves meet ⇒ split the prize without a fight
  - Each dove gets 3
- 2 hawks meet ⇒ they fight, each has probability 0.5 of getting the prize
  - Hawk's expected payoff is $-5 + 0.5(6) = -2$
- Unique Nash equilibrium (s,s), where $s = \{(0.6, \text{H}), (0.4, \text{D})\}$
  - i.e., 60% hawks, 40% doves
- s is an ESS
  - For *every* strategy $t \neq s$
    - $u(s,s) = u(t,s)$ and $u(s,t) > u(t,t)$

	H	D
H	–2	6
D	0	3

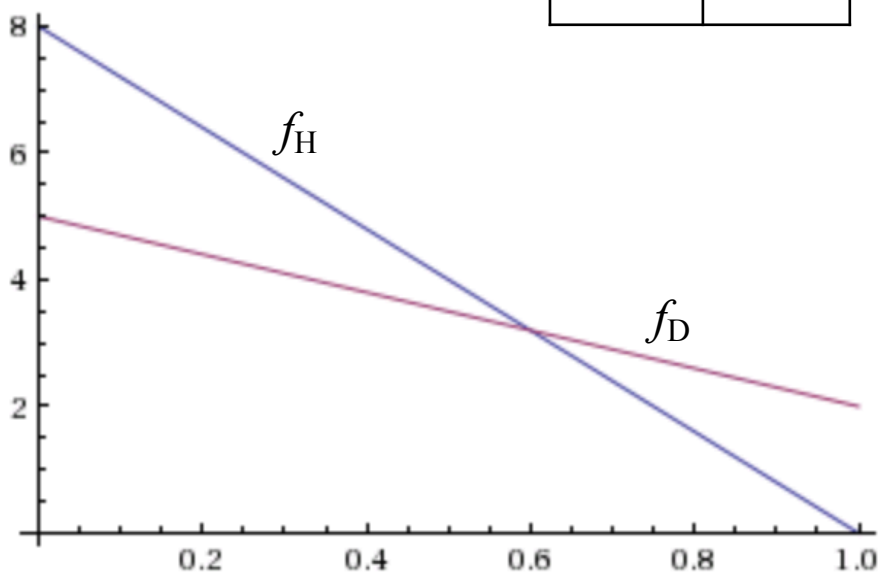
	s	t
s	$u(s,s)$	$u(s,t)$
t	$u(t,s)$	$u(t,t)$

Replicator Dynamic for Hawk-Dove Game

- Positive affine transformation of payoff matrix
- Population profile $\mathbf{x} = (x_H, x_D)$
  - so $x_D = 1 - x_H$
- Replicator dynamic: $dx_i/dt = x_i(f_i - f_{avg})$
- $f_H = u_H = 0\,x_H + 8(1-x_H) = 8 - 8\,x_H$
- $f_H = u_D = 2\,x_H + 5(1-x_H) = 5 - 3\,x_H$
- If $x_H < 0.6$ then $f_H > f_{avg} > f_D$
  - $dx_H/dt > 0$
  - $dx_D/dt < 0$
- If $x_H > 0.6$ then $f_H < f_{avg} < f_D$
  - $dx_H/dt < 0$
  - $dx_D/dt > 0$
- If $x_H = 0.6$ then $f_H = f_{avg} = f_D$
  - $dx_H/dt = dx_D/dt = 0$

	H	D
H	–2	6
D	0	3

	H	D
H	0	8
D	2	5



Rock-Paper-Scissors

- Unique Nash equilibrium
  - $s = \{(^1/_3, \text{R}), (^1/_3, \text{P}), (^1/_3, \text{S})\}$
- Evolutionarily stable?

	R	P	S
R	0	–1	1
P	1	0	-1
S	–1	1	0

- Replace each u_1 with $u_1/2 + ½$
  - Why is this OK?
  - Why do we want to do it?

	R	P	S
R	0.5	0	1
P	1	0.5	0
S	0	1	0.5

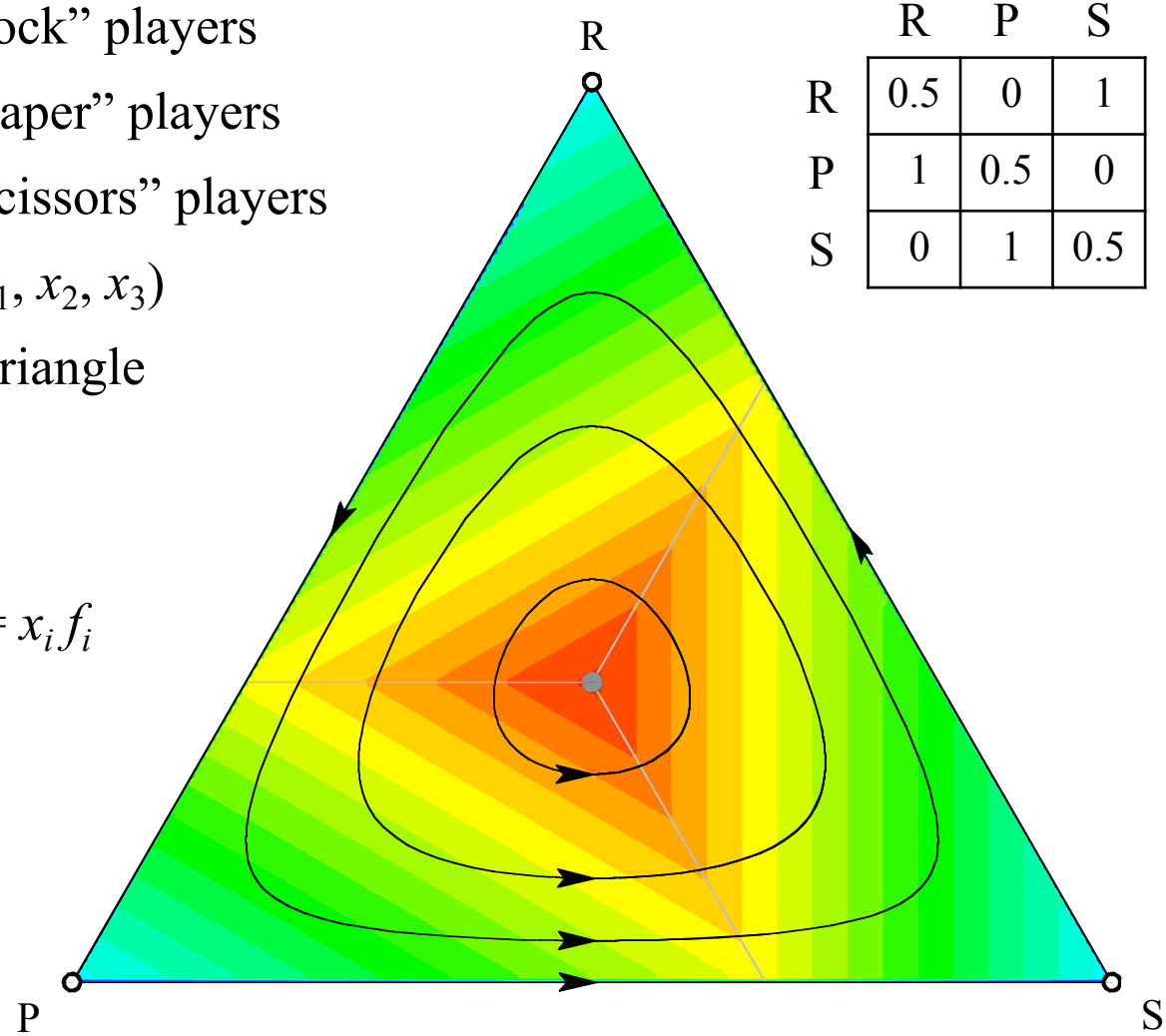
- Let t be any strategy other than s
  - $a = c$ and $b = d$
  - Not evolutionarily stable

	s	t
s	0.5	0.5
t	0.5	0.5

Rock-Paper-Scissors

- Consider an infinite population of Rock-Paper-Scissors players
  - x_1 = proportion of "rock" players
  - x_2 = proportion of "paper" players
  - x_3 = proportion of "scissors" players
- Population profile $\mathbf{x} = (x_1, x_2, x_3)$
- Each $\mathbf{x}$ is a point on the triangle
- Fitness $f_i = u_i$
- Replicator dynamic:
  - $dx_i/dt = x_i\,(f_i - f_{avg}) = x_i f_i$
- Oscillation
  - $t = 0.5$: closed orbits

	R	P	S
R	0.5	0	1
P	1	0.5	0
S	0	1	0.5

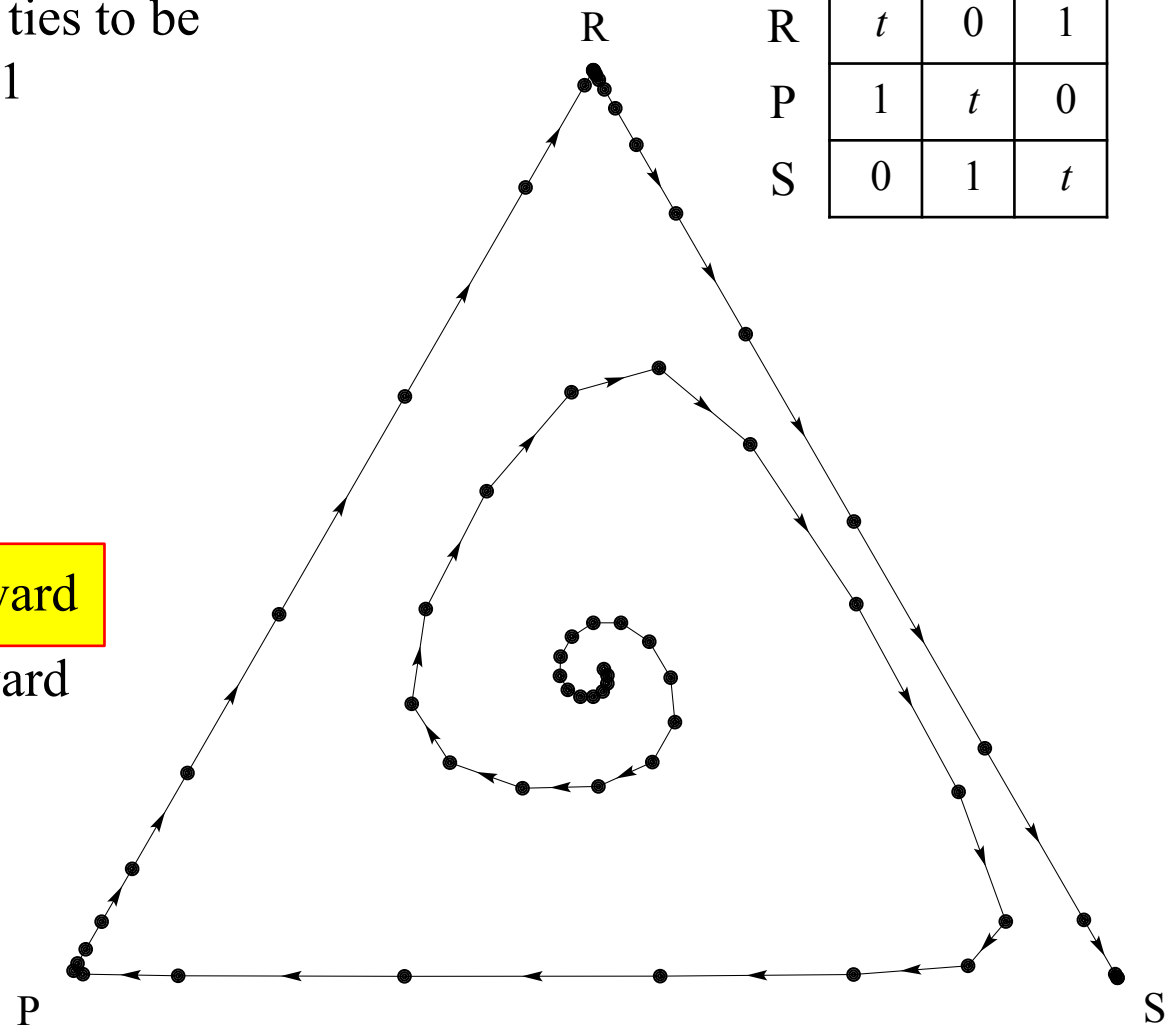


Rock-Paper-Scissors

- Generalize
  - Allow the payoff for ties to be a fixed value $0 < t < 1$

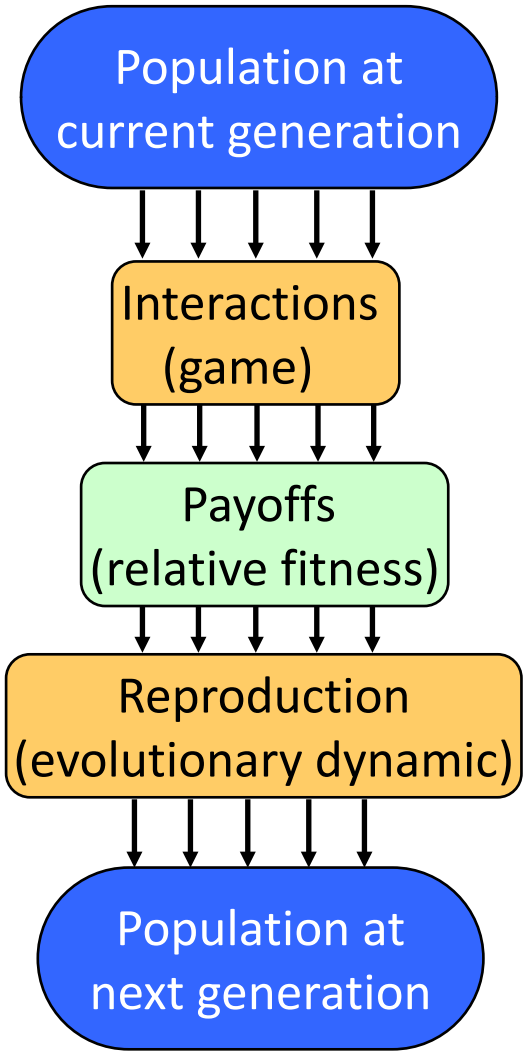
	R	P	S
R	t	0	1
P	1	t	0
S	0	1	t

- Oscillation
  - $t = 0.5$: closed orbits
  - $t < 0.5$: converge inward
  - $t > 0.5$: diverge outward



Discrete Replicator Dynamic

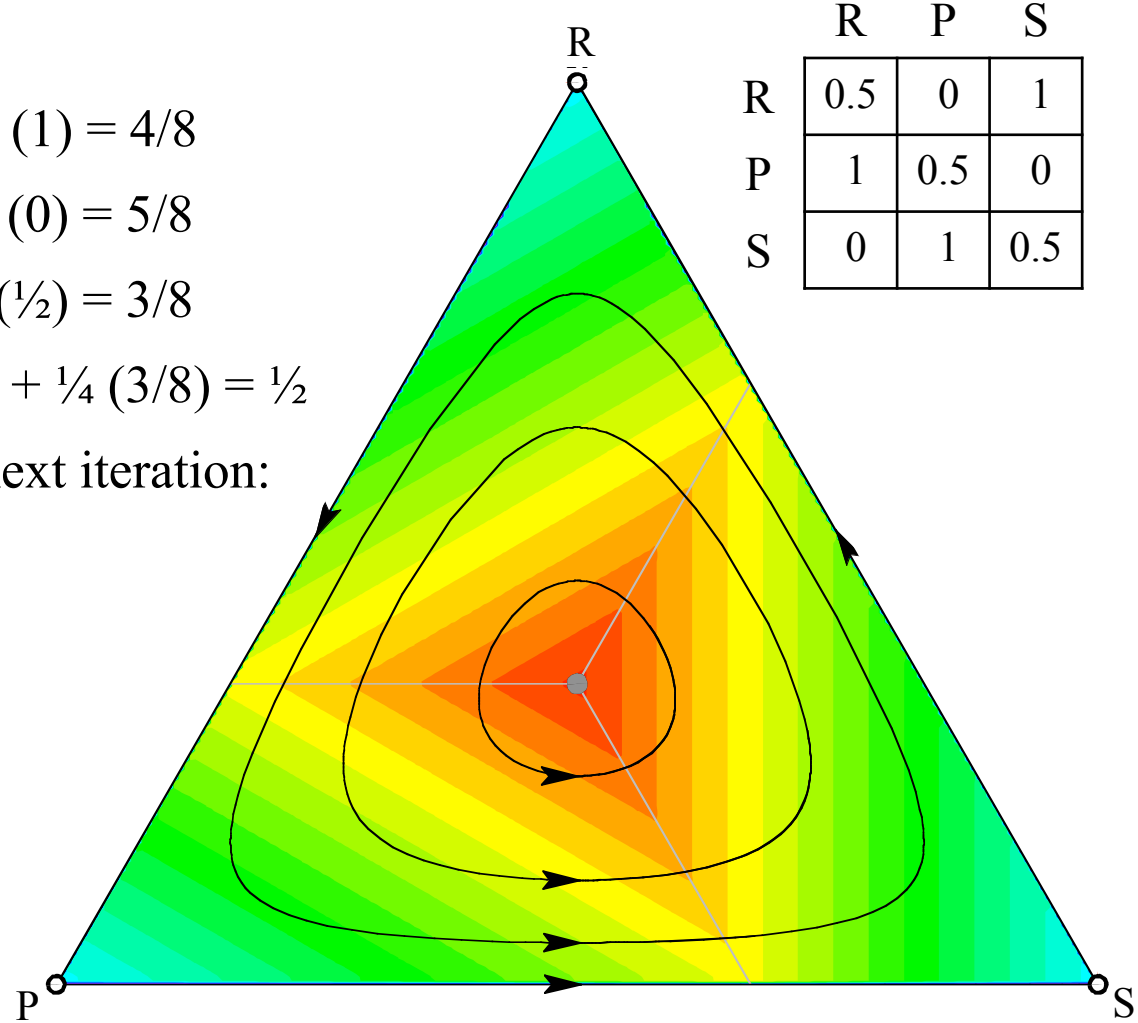
- Discrete simulation
  - sequence of discrete **stages** or **generations**
    - as in a repeated game
  - At each stage
    - Each agent interacts with one or more neighbors
    - Calculate each agent's "fitness" based on results of interactions
    - Update strategies based on fitness
- Discrete version of replicator dynamic:
    - $x_i^{(j+1)} = x_i^{(j)} f_i^{(j)} / f_{avg}^{(j)}$
  - $x_i^{(j)}$ is the proportion of strategy i at generation j



Example

- Rock-paper-scissors, discrete replicator dynamic
- Population profile $\mathbf{x} = (x_R, x_P, x_S) = (½, ¼, ¼)$
- For each i, fitness $f_i = u_i(\mathbf{x})$
  - $f_R = ½\,(½) + ¼\,(0) + ¼\,(1) = 4/8$
  - $f_P = ½\,(1) + ¼\,(½) + ¼\,(0) = 5/8$
  - $f_S = ½\,(0) + ¼\,(1) + ¼\,(½) = 3/8$
  - $f_{avg} = ½\,(4/8) + ¼\,(5/8) + ¼\,(3/8) = ½$
- Population proportions at next iteration:
  - $x_R = ½\,(½) / ½ = ½$
  - $x_P = ¼\,(5/8) / ½ = 5/16$
  - $x_S = ¼\,(3/8) / ½ = 3/16$
- Slightly different than discrete case

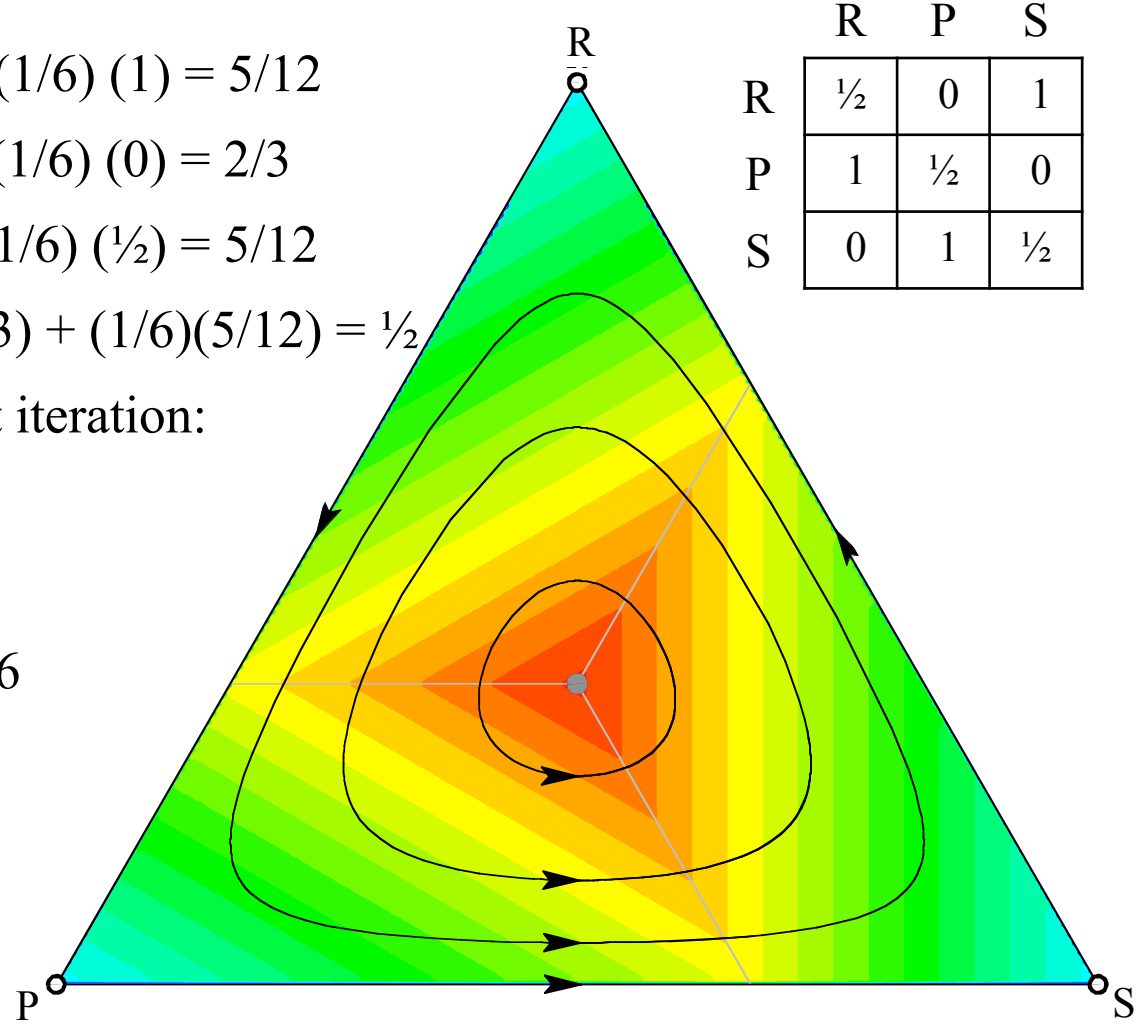
	R	P	S
R	0.5	0	1
P	1	0.5	0
S	0	1	0.5



Example

- Rock-paper-scissors, discrete replicator dynamic
- Population profile $\mathbf{x} = (x_R, x_P, x_S) = (^1/_2, ^1/_3, ^1/_6)$
- For each i, fitness $f_i = u_i(\mathbf{x})$
  - $f_R = (½)\ (½) + (1/3)\ (0) + (1/6)\ (1) = 5/12$
  - $f_P = (½)\ (1) + (1/3)\ (½) + (1/6)\ (0) = 2/3$
  - $f_S = (½)\ (0) + (1/3)\ (1) + (1/6)\ (½) = 5/12$
  - $f_{avg} = (½)(5/12) + (1/3)(2/3) + (1/6)(5/12) = ½$
- Population proportions at next iteration:
  - $x_R = ½\ (5/12)\ /\ ½ = 5/12$
  - $x_P = (1/3)\ (2/3)\ /\ ½ = 4/9$
  - $x_S = (1/6)\ (5/12)\ /\ ½ = 5/36$
- Sanity check
  - $5/12 + 4/9 + 5/36 = 1$

	R	P	S
R	½	0	1
P	1	½	0
S	0	1	½



Cultural Evolution

- Use EGT to model evolution of cultural characteristics
  - Strategies ⇔ possible behaviors
  - Reproduction ⇔ cultural transmission
    - Humans imitate others, learn from others
    - Successful strategies have a higher probability of being adopted by others
- Several possible models of evolutionary dynamics
  - Replicator, others

Game-Theoretic Modeling

- Human interactions are very complicated
- Game-theoretic models deliberately omit large amounts of detail
  - All of the factors that might lead humans to perform various actions
  - What the actions do physically when they are performed
- Map everything into payoff values that represent an abstract notion of the outcomes' desirability

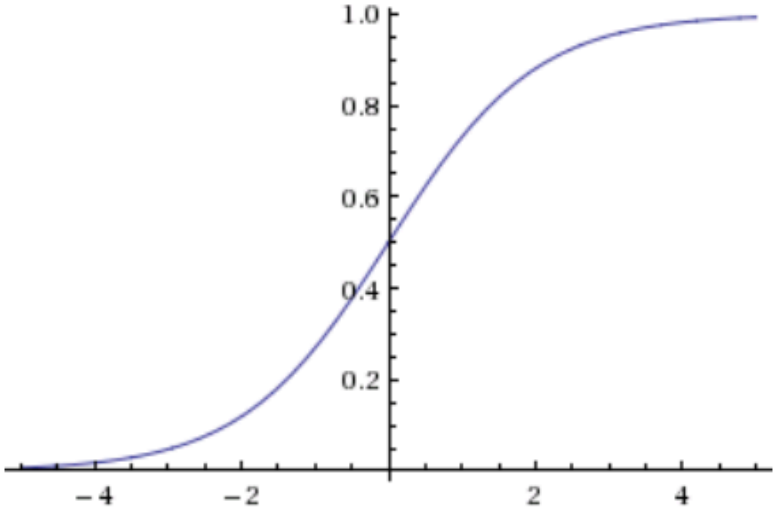
- Not always easy to get the model right
  - In most research papers on EGT, the authors spend a lot of time explaining the justification for their model
- The model is highly simplified
  - Can't give exact numeric predictions of what would happen in real life
- But can provide explanations of the underlying dynamics
  - Establish support for causal relationships

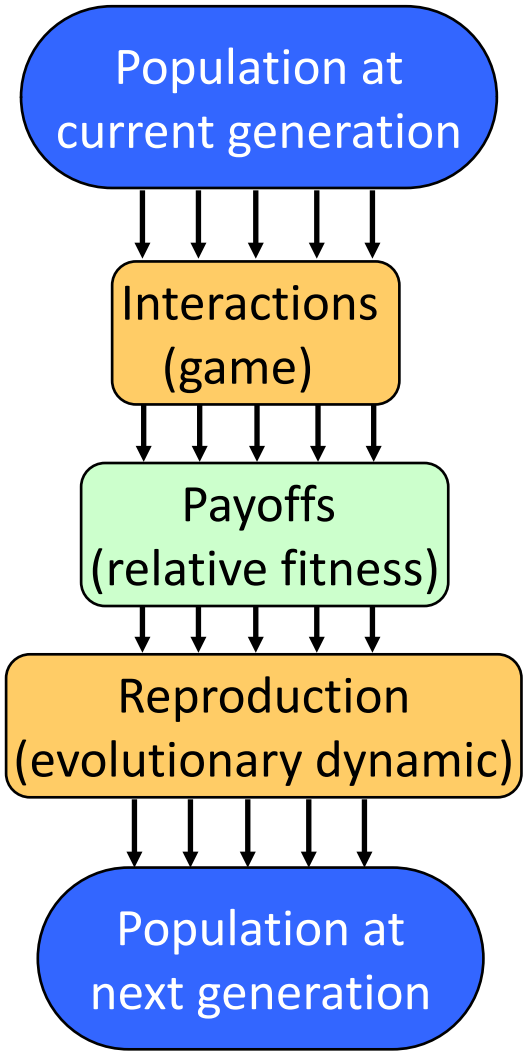
Evolutionary Dynamic

- Model of learning by imitating others
  - Evidence that this does well at modeling how behaviors spread when people copy others' behavior
- loop
  - Compute payoff for each agent
  - For each agent *i*
    - Compare payoffs of *i* and others
    - Decide what strategy *i* will use next time
- Several ways to decide what strategy to use next time
  - Imitate the better, Fermi rule, …
- Strategy's proportion grows if it does better than average
  - But different growth rate than with the replicator dynamic

Fermi Rule

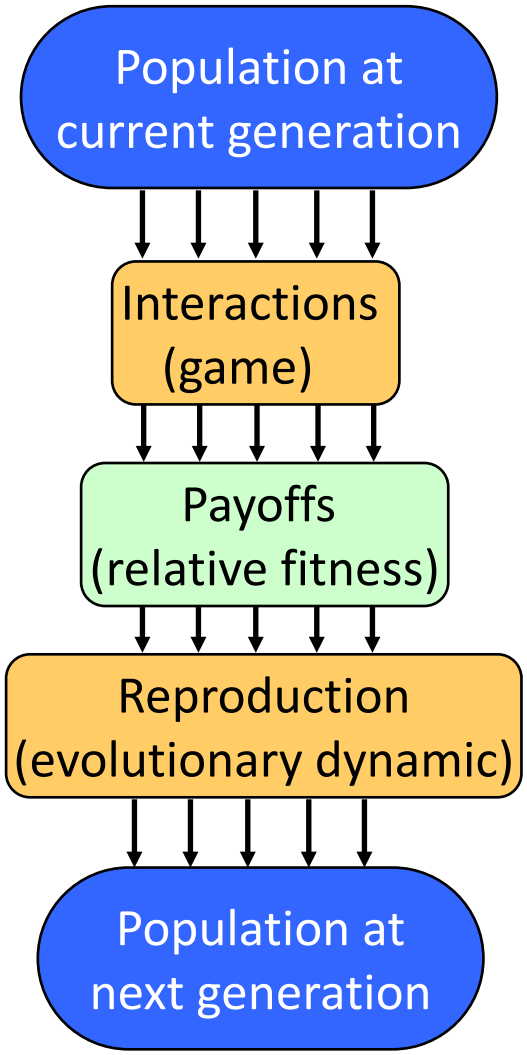
- Adapted from statistical mechanics
  - Individuals interact using their current strategies
    - Get their payoffs
  - Each individual compares its payoff to that of a randomly chosen neighbor
  - Pr[switch to neighbor's strategy] = $1/(1 + e^{s(r - r')})$
    - r, r' = individual's and neighbor's payoffs
    - $s \geq 0$ is the **selection strength**





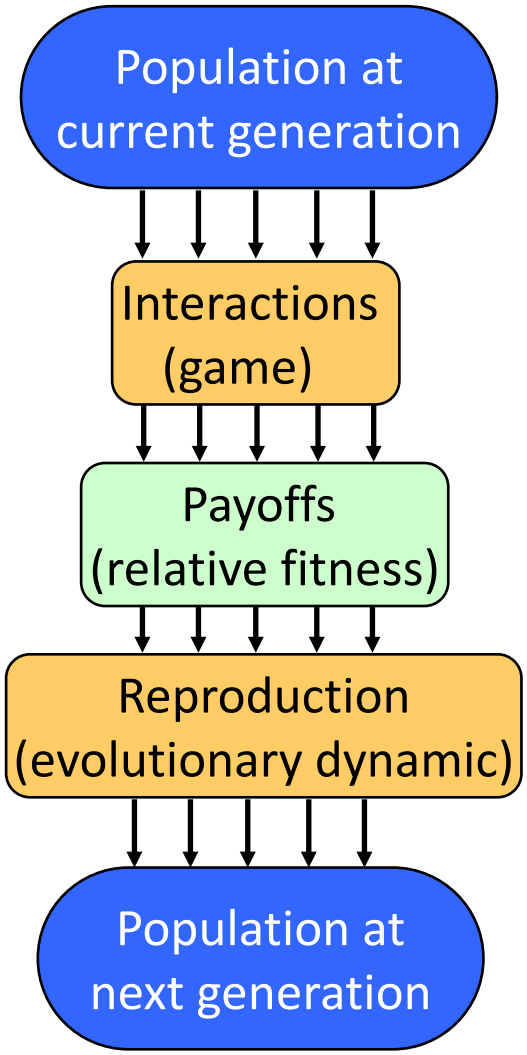
Imitate the Better

- At stage j, let A_j = {all agents at stage j}
- To build A_{j+1} , do the following steps k times:
  - Randomly choose 2 agents in A_j
  - Let a be the one that got the higher payoff (or choose a at random if both got the same payoff)
  - Add to A_{j+1} an agent that uses a's strategy



Exploration

- In biological reproduction, mutation occurs quite rarely
  - Game-theoretic biological models often omit it
- In cultural evolution, something analogous happens more frequently:
  - Individuals try out new behaviors at random
    - Traulsen *et al*. Exploration dynamics in evolutionary games. *Proc. Nat. Acad. Sci.*, 2009
- Modify replicator or imitate-the-better or Fermi:
  - For each agent in the next generation, a small probability μ that it will choose a strategy i at random from $\{1, \ldots, n\}$
  - regardless of whether i was successful in the current generation
  - regardless of whether any agent is currently using i



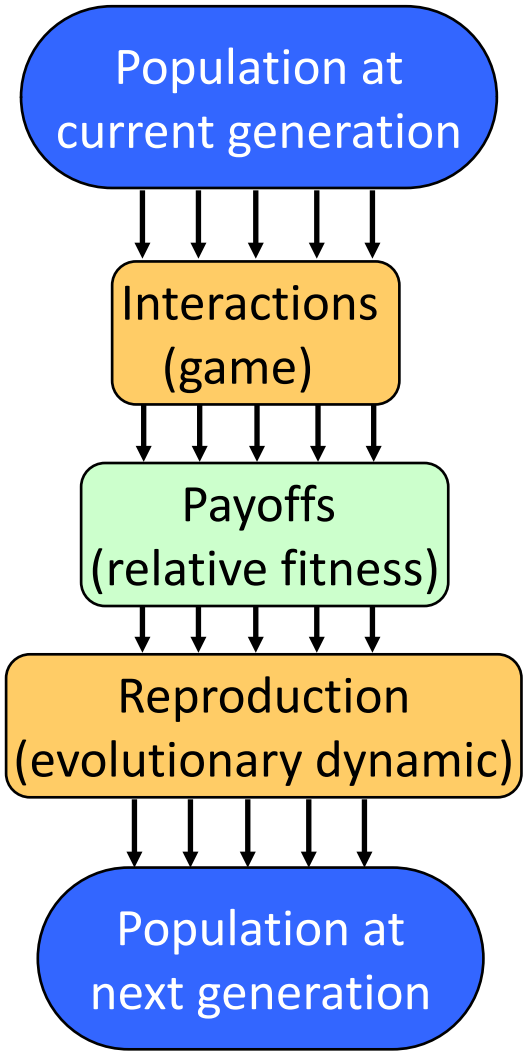
Imitate-the-Better Dynamic

- For the replicator dynamic, we have a formula for calculating $x_i^{(j+1)}$ from proportions and fitness at stage j

- Formula for imitate-the-better is more complicated
- Special case: just two strategies
  - Each agent chooses two agents a, b at random
  - For each strategy i, $\Pr[a \text{ uses } i] = \Pr[b \text{ uses } i] = x_i$

$$
\begin{aligned}
\Pr[\text{choose } i] = {} & \Pr[a \text{ and } b \text{ both use } i] \\
& + \Pr[a \text{ uses } i,\ b \text{ doesn't},\ r_a > r_b] \\
& + \tfrac{1}{2} \Pr[a \text{ uses } i,\ b \text{ doesn't},\ r_a = r_b] \\
& + \Pr[b \text{ uses } i,\ a \text{ doesn't},\ r_b > r_a] \\
& + \tfrac{1}{2} \Pr[b \text{ uses } i,\ a \text{ doesn't},\ r_a = r_b] \\
= {} & x_i^2 + 2 \textstyle\sum_{j \neq i} x_i x_j \left(\Pr[r_i > r_j] + \tfrac{1}{2} \Pr[r_i = r_j]\right)
\end{aligned}
$$

- You don't need to remember the formula



Situation-Dependent Risk Preferences

Recall this slide from my 1st lecture:

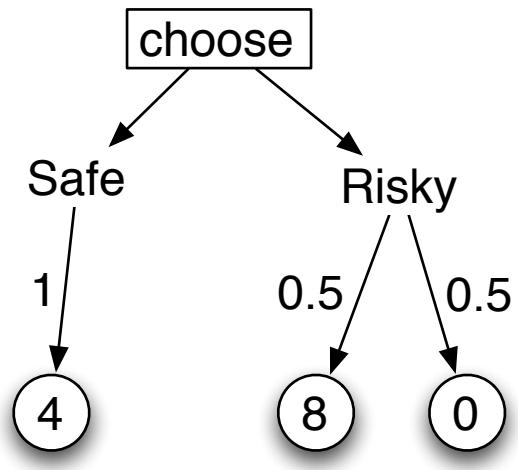
- Kahneman & Tversky, 1979
  - http://www.econport.org/econport/request?page=man_ru_advanced_prospect
- A: 100% chance of receiving $3000
- B: 80% chance of getting $4000; 20% chance of getting nothing
  - EV(A) = $3000 < EV(B) = $3200, but most people would choose A
  - For prospects involving gains, we're **risk-averse**
- C: 100% chance of losing $3000
- D: 80% chance of losing $4000; 20% chance of losing nothing
  - EV(C) = –$3000 > EV(D) = –$3200, but most people would choose D
  - For prospects involving losses, we're **risk-prone**

Human risk preferences are **situation-dependent**

Decision Making Under Risk

- Why might we have such preferences?

- Let's look at a simple repeated lottery game
- At each stage, agents make choices between two lotteries
  - The ***safe*** lottery: guaranteed reward of 4
  - The ***risky*** lottery: Pr[0] = ½ ; Pr[8] = ½
- Two pure (deterministic) strategies:
  - *S*: always choose the safe lottery
  - *R*: always choose the risky lottery



Roos, Carr, & Nau. Evolution of state-dependent risk preferences. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2010.

Roos & Nau. Risk preference and sequential choice in evolutionary games. *Advances in Complex Systems*, 2010.

Lottery Game, Replicator Dynamic

- At each stage, each strategy's average payoff is 4

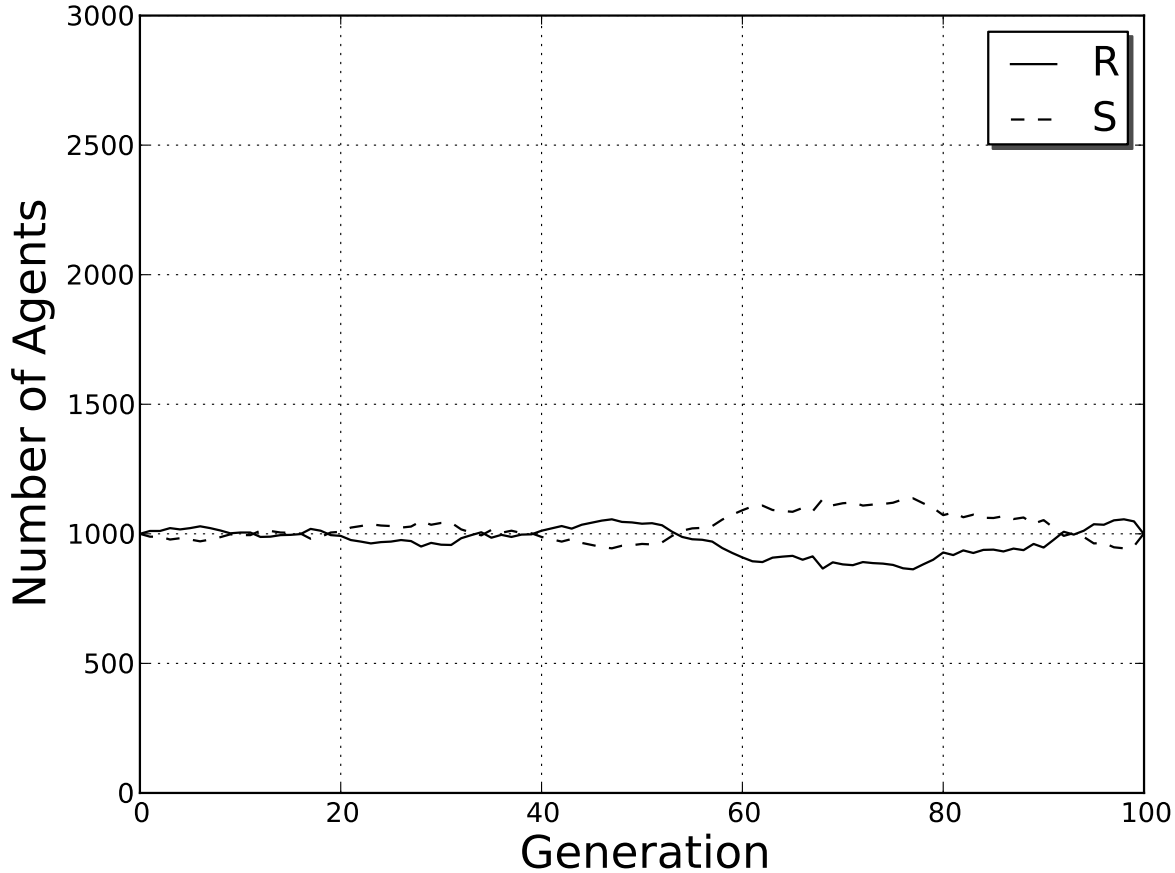
$$x_i^{(j+1)} = x_i^{(j)} f_i^{(j)} / f_{avg}^{(j)} = x_i^{(j)} (4/4) = x_i^{(j)}$$

- For infinite population, each strategy's proportion stays constant

- For finite population, should stay roughly constant
  - Why only roughly?

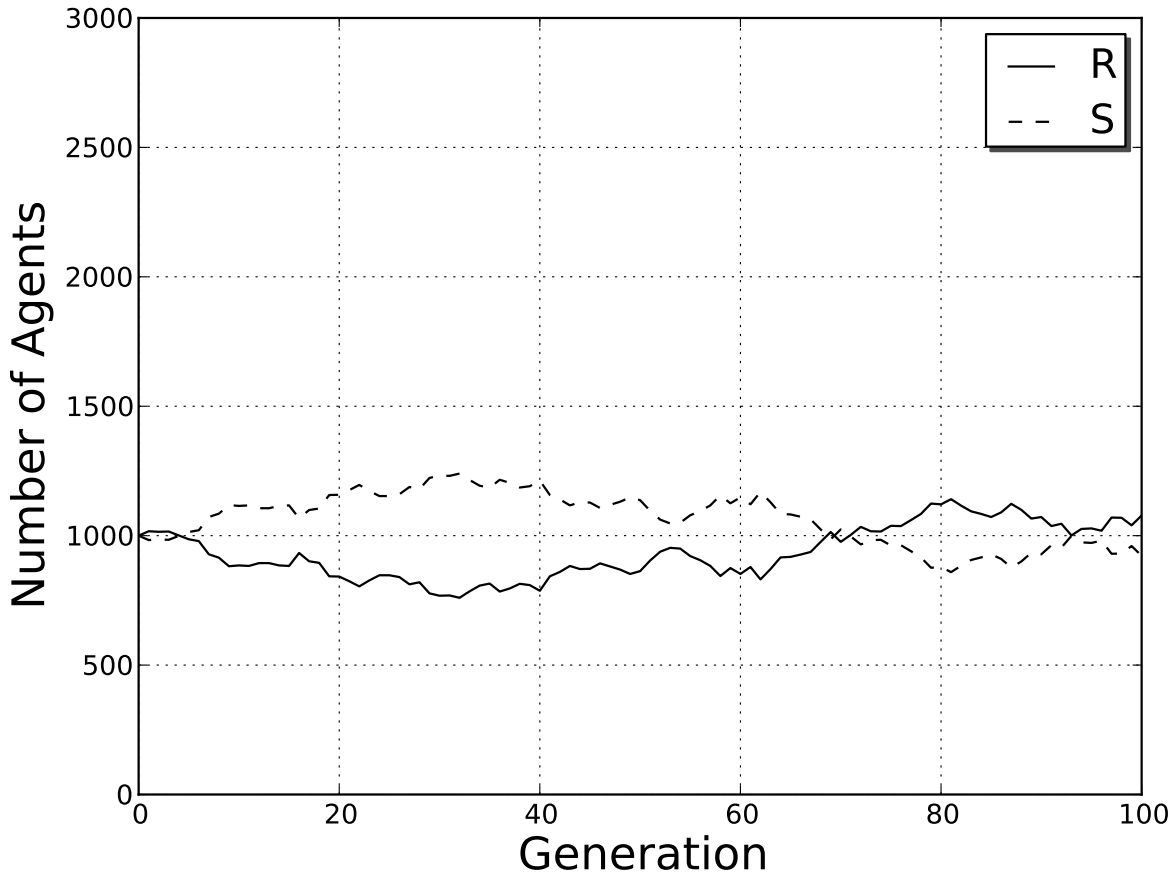
- Verified by simulation for *S* and *R* →

- Could get similar results for any strategy that mixes *S* and *R*



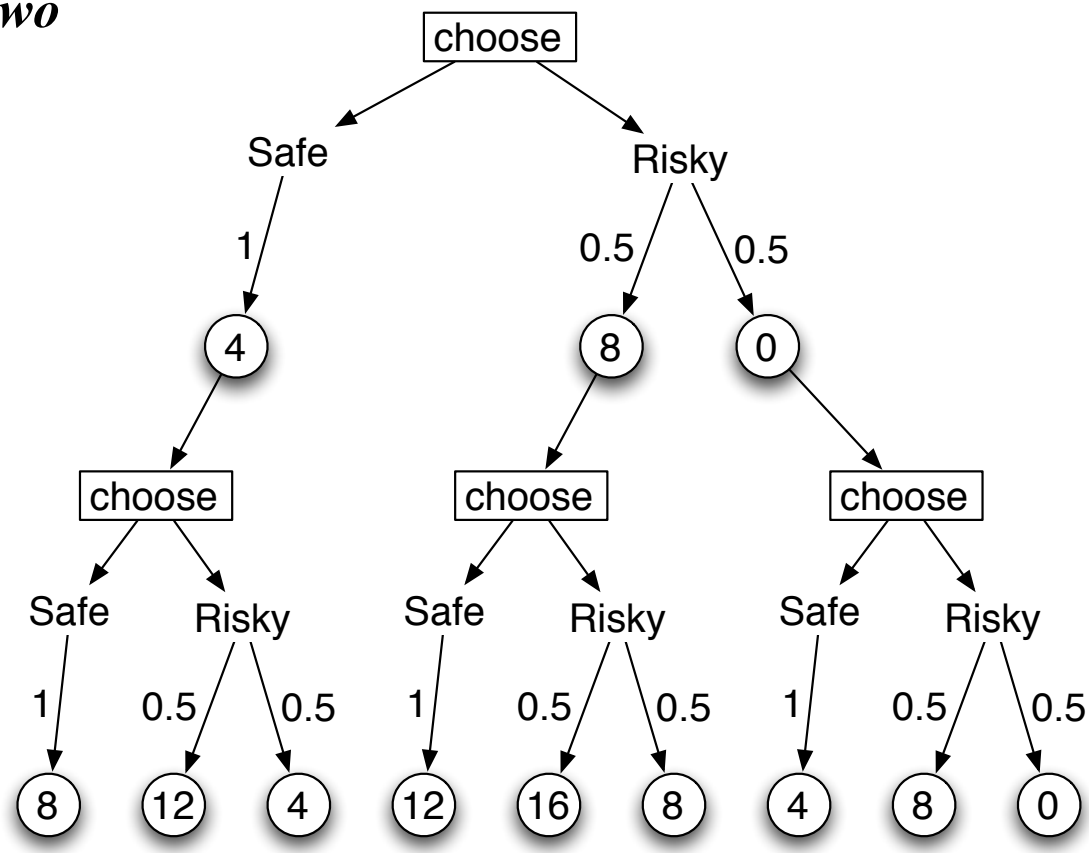
Lottery Game, Imitate-the-Better Dynamic

- Pick any two agents, and let *s* and *t* be their strategies
- Regardless of what *s* and *t* are, each agent has equal probability of getting a higher payoff than the other
- Like before, each strategy's proportion should stay roughly constant
- Verified by simulation for *S* and *R* →
- Could get similar results for any strategy that's a mixture of *S* and *R*



Double Lottery Game

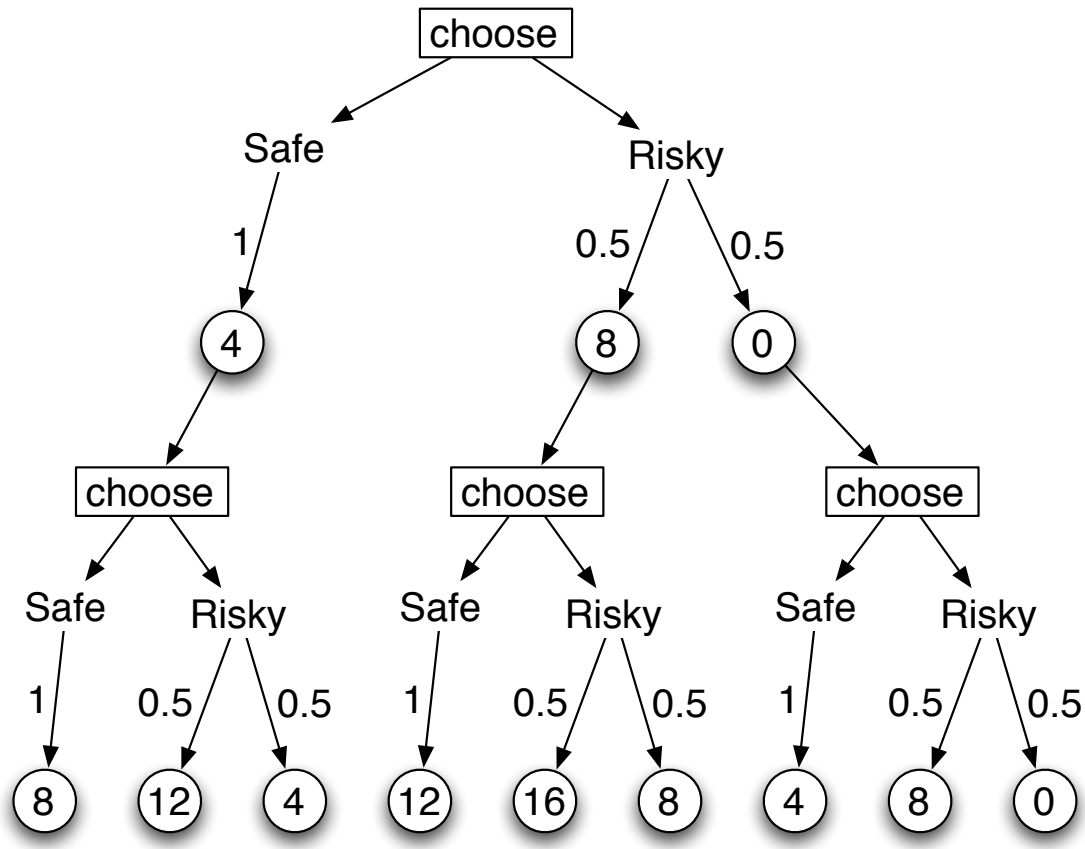
- At each stage, agents make ***two*** rounds of lottery choices
  1. Choose between the safe lottery and the risky lottery, get a payoff
  2. Choose between the safe lottery and the risky lottery again, and get an additional payoff



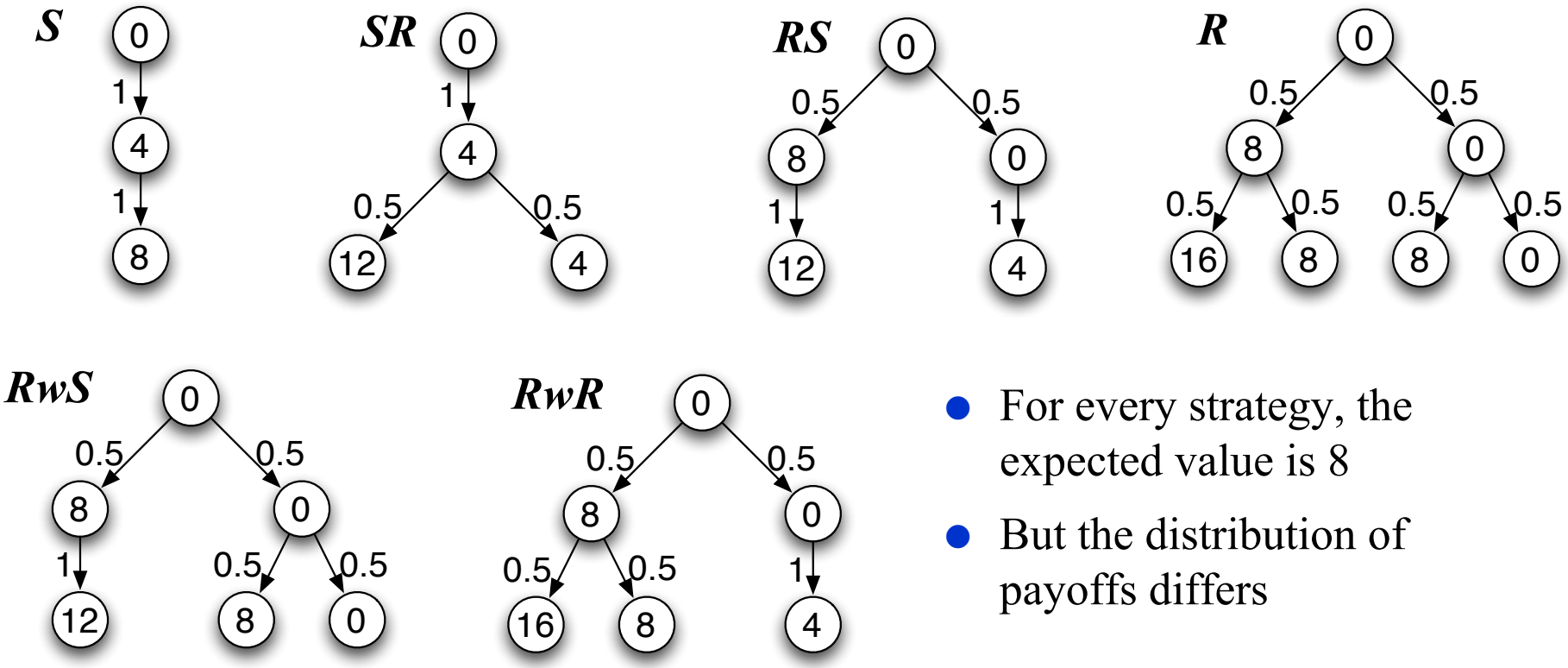
Double Lottery Game

There are 6 pure strategies:

- *S*: choose Safe both times
- *SR*: 1st time choose Safe
  2nd time choose Risky
- *RS*: 1st time choose Risky
  2nd time choose Safe
- *R*: Risky both times
- *RwS*: 1st time choose Risky
  - 2nd time: if 1st time was a win (payoff 8), then Safe, otherwise Risky
- *RwR*: 1st time choose Risky
  - 2nd time: if 1st time was a win (payoff 8), then Risky, otherwise Safe



Distribution of Payoffs for Each Strategy



- For every strategy, the expected value is 8
- But the distribution of payoffs differs

	S	*SR*		*RS*		*R*			*RwS*			*RwR*		
Payoff	**8**	**4**	**12**	**4**	**12**	**0**	**8**	**16**	**0**	**8**	**12**	**4**	**8**	**16**
Probability	1	½	½	½	½	¼	½	¼	¼	¼	½	½	¼	¼

No skew

Skewed left

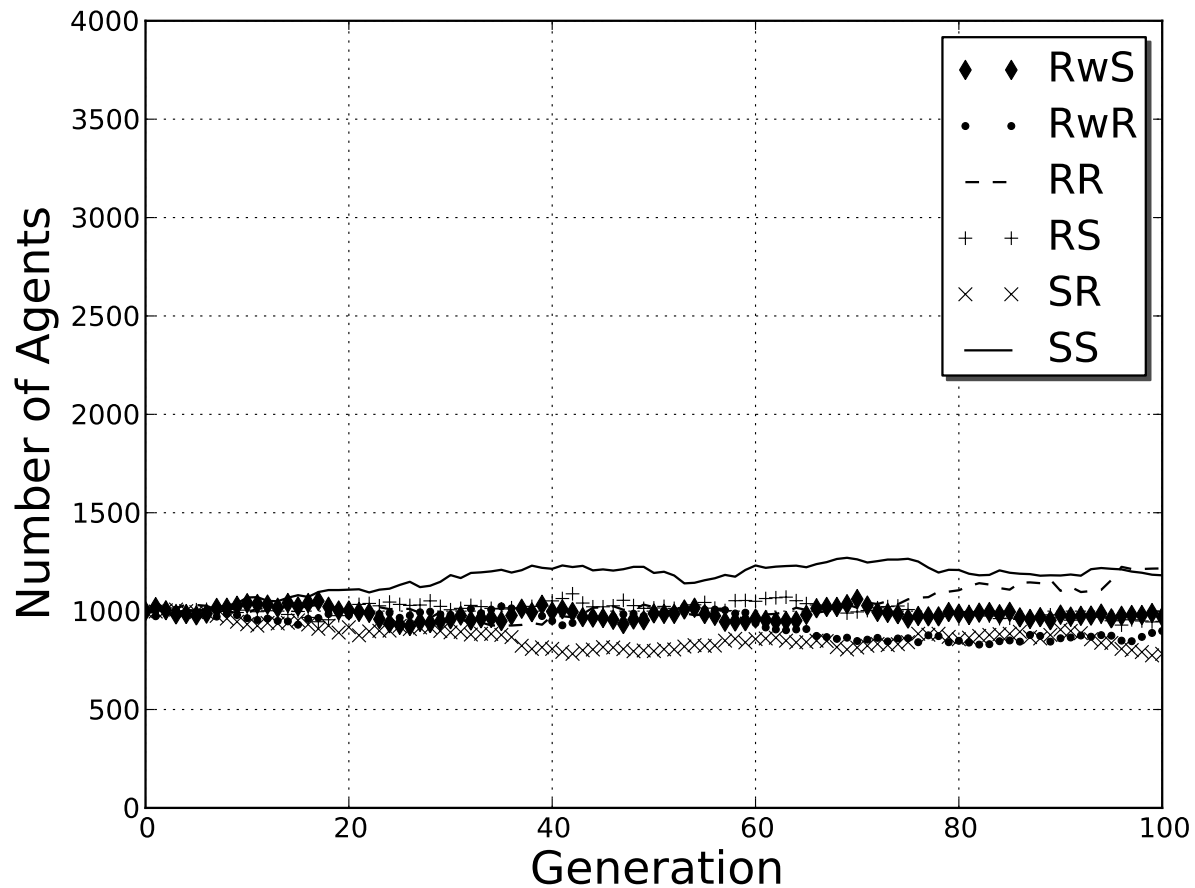
Skewed right

Double Lottery Game, Replicator Dynamic

- Suppose we start out with each strategy having proportion 1/6
- For every strategy s_i expected payoff is $r_i = 8$, so $r_{avg} = 8$
- Discrete replicator dynamic:

$$x_i^{(j+1)} = x_i^{(j)} f_i^{(j)} / f_{avg}^{(j)}$$
$$= x_i^{(j)} (8/8) = x_i^{(j)}$$

- Like before, each strategy's proportion stays roughly constant
- Verified by simulation
  - 6 strategies, start with 1000 of each



Double Lottery Game, Imitate-the-Better Dynamic

- For imitate-the-better, do the following *k* times:
  - Choose two agents *a* and *b*, and compare their payofs
    - Reproduce the one that got a higher payoff
    - If they got the same payoff, choose either of them at random
- Suppose *a* uses *S*, and *b* uses *SR*
  - Pr[*a* gets 8, *b* gets 4] = ½ → reproduce *a*
  - Pr[*a* gets 8, *b* gets 12] = ½ → reproduce *b*
- Thus *a* and *b* are equally likely to be reproduce
  - Also true if *a* and *b* are any of *S, R, SR,* and *RS*

	a	*b*												
	S	*SR*		*RS*		*R*			*RwS*			*RwR*		
Payoff	**8**	**4**	**12**	**4**	**12**	**0**	**8**	**16**	**0**	**8**	**12**	**4**	**8**	**16**
Probability	1	½	½	½	½	¼	½	¼	¼	¼	½	½	¼	¼

No skew

Skewed left

Skewed right

Double Lottery Game, Imitate-the-Better Dynamic

- Suppose *a* uses *S* and *b* uses *RwS*
  - Pr[*b* gets 0] = ¼ → reproduce *a*
  - Pr[*b* gets 8] = ¼ → *a* and *b* equally likely to reproduce
  - Pr[*b* gets 12] = ½ → reproduce *b*
- Thus
  - Pr[*a* reproduces] = ¼ + ½ (¼) = 0.375
  - Pr[*b* reproduces] = ½ + ½ (¼) = 0.625
- When *RwS* and *S* are paired, *RwS* is more likely to reproduce

	a								*b*					
	S	*SR*		*RS*		*R*			*RwS*			*RwR*		
Payoff	**8**	**4**	**12**	**4**	**12**	**0**	**8**	**16**	**0**	**8**	**12**	**4**	**8**	**16**
Probability	1	½	½	½	½	¼	½	¼	¼	¼	½	½	¼	¼
	No skew								Skewed left			Skewed right		

Double Lottery Game, Imitate-the-Better Dynamic

- In general:
  - If *RwS* is paired with *S*, *R*, or *RwR*, then *RwS* is more likely to reproduce
- For all other pairs of strategies, both of them are equally likely to reproduce

	S	*SR*		*RS*		*R*			*RwS*			*RwR*		
Payoff	**8**	**4**	**12**	**4**	**12**	**0**	**8**	**16**	**0**	**8**	**12**	**4**	**8**	**16**
Probability	1	½	½	½	½	¼	½	¼	¼	¼	½	½	¼	¼

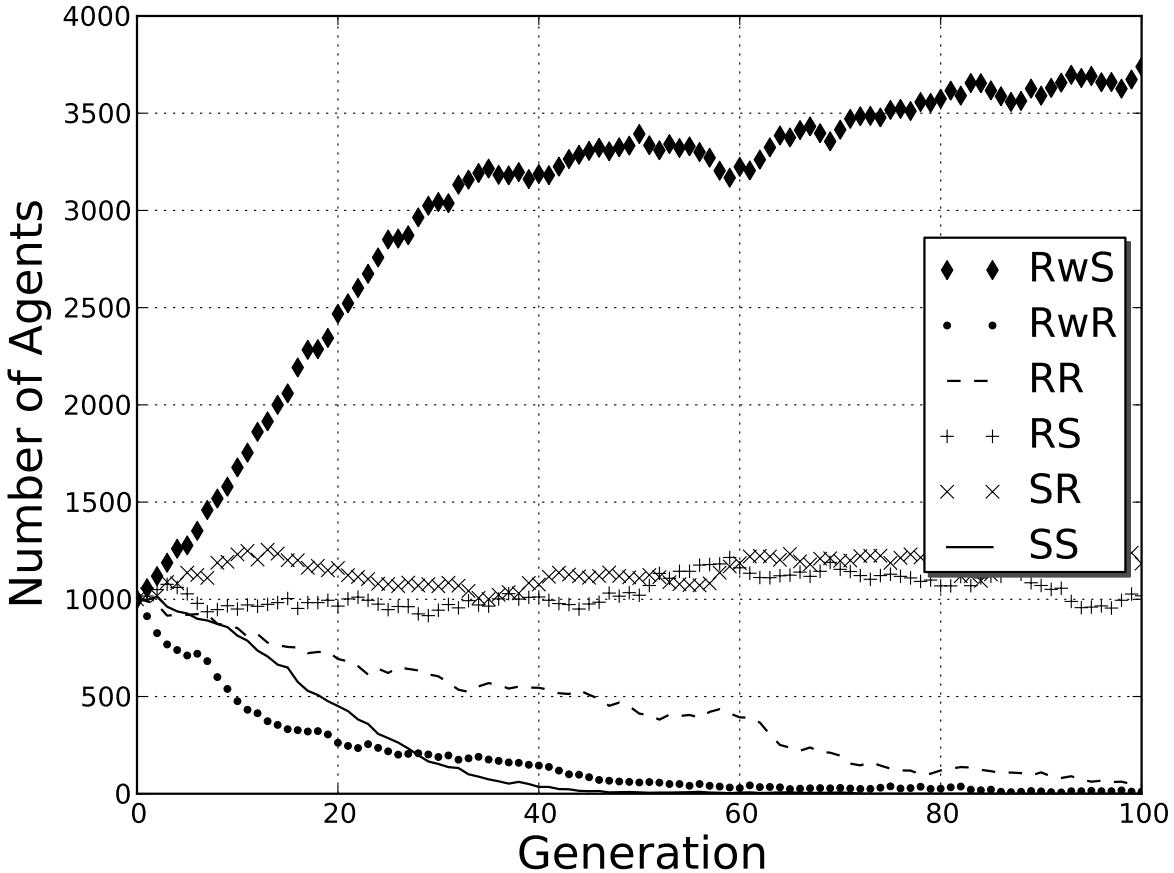
No skew

Skewed left

Skewed right

Double Lottery Game, Imitate-the-Better Dynamic

- *RwS* has an advantage whenever it's paired with *S, R*, or *RwR*
  - For all other pairings, neither strategy has an advantage
- Proportion of *RS* and *SR* should stay roughly constant
- *S, R*, and *RwR* should gradually go extinct
- *RwS*'s proportion should gradually approach 2/3
- Verified by simulation →
  - Why do you think *R* decreases more slowly than *S* and *RwR*?



Real-World Analogy to RwS

- Recall what RwS does:
  - Plays it safe when it's ahead
  - Takes risks when it's behind

- Example from American football
  - ***Hail Mary pass***: a very long forward pass with only a small chance of success, made in desperation when the clock is running out

Discussion

- Double lottery game
  - *RwS* has an evolutionary advantage if we use the imitation dynamic
- Possible reason why state-dependent risk preferences might spread
  - But not conclusive

- Didn't really model much about culture
  - No interactions among agents
- How to model that?

Modeling Cooperation among Groups

- How to care for a shared resource?
  - Natural resources: air, water, grazing land, …
  - Infrastructure: highways, bridges, streets, sidewalks, communication networks, …
- *Public Good Game (PGG)*
  - n-player generalization of Prisoner's Dilemma
- Each individual is asked to contribute an amount c
    - Represents the cost of maintaining or not over-exploiting the resource
  - *Cooperate*: contribute c
  - *Defect*: contribute nothing
- The sum of all contributions is multiplied by a factor $b > 1$
    - Represents the benefit that the resource provides
  - Resulting amount is distributed equally among everyone

Selsley Common, Gloucestershire

Modeling Cooperation among Groups

- *Public Good Game (PGG)*
- Each individual is asked to contribute an amount c
  - *Cooperate*: contribute c
  - *Defect*: contribute nothing
- The sum of all contributions is multiplied by a factor $b > 1$
  - Resulting amount is distributed equally among everyone

- Payoff matrix?
  - Payoffs aren't determined by 2-player interactions
  - Determined by interaction among entire population
    - $n \times n$ rather than 2×2
- But can look at each individual i's payoff as a function of what everyone else does

Modeling Cooperation among Groups

- Expected payoffs if n is very large
  - If everyone else plays C
    - u(C, everyone else plays C) = $bcn/n - c = bc - c$
    - u(D, everyone else plays C) = $bc(n-1)/n \approx bc$
  - If everyone else plays D
    - u(C, everyone else plays D) = $c/n - c \approx -c$
    - u(D, everyone else plays D) = 0
  - If proportion p cooperate, $1-p$ defect
    - $u(C, \{(p,C), (1-p, D)\}) \approx bc(pn)/n - c = bcp - c$
    - $u(D, \{(p,C), (1-p, D)\}) \approx bc(pn)/n = bcp$
- In all cases, u(C,…) ≈ u(D,…) – c
- Evolutionary game (without exploration)
  - Evolve to 100% defectors, everyone's utility is 0

Which evolutionary dynamic?

Modeling Cooperation among Groups

- Expected payoffs if n is very large
  - If everyone else plays C
    - u(C, everyone else plays C) = $bcn/n - c = bc - c$
    - u(D, everyone else plays C) = $bc(n-1)/n \approx bc$
  - If everyone else plays D
    - u(C, everyone else plays D) = $c/n - c \approx -c$
    - u(D, everyone else plays D) = 0
  - If proportion p cooperate, $1-p$ defect
    - $u(C, \{(p,C), (1-p, D)\}) \approx bc(pn)/n - c = bcp - c$
    - $u(D, \{(p,C), (1-p, D)\}) \approx bc(pn)/n = bcp$
- In all cases, u(C,…) ≈ u(D,…) – c
- Evolutionary game (without exploration)
  - Evolve to 100% defectors, everyone's utility is 0

Which evolutionary dynamic?

What happens with exploration?

How to Discourage Defection?

- Evidence that punishment is important in the emergence and maintenance of cooperation in societies
  - Humans are willing to pay a cost to punish deviations from cooperative norms
  - Punishment can establish and maintain cooperative norms in collective action and cooperation games
  - Different cultures have different propensities to punish deviations from their societal norms

- Two kinds of punishment
  - Institutional
  - Individual
- We'll look at individual

Modeling Punishment in the PGG

- At each generation:
  - The contribution phase, then a *punishment phase*:
    - Each individual may pay an amount λ to reduce a defector's payoff by an amount $\rho > \lambda$
  - Each strategy involves two choices:
    - whether to defect
    - whether to punish defectors

- Problem:
  - Punishing lowers punisher's payoff by λ
    - Evolve to nearly 100% non-punishers
  - Without punishment, defectors have higher payoff then cooperators
    - Evolve to nearly 100% defectors

Better Model

- Hilbe & Traulsen, 2012. Emergence of responsible sanctions without second order free riders, antisocial punishment or spite. *Nature Scientific Reports*.

- *Punishment reputation*: how likely that others will punish
  - Enables one to infer Cooperating has higher payoff than Defecting
- *Information level*: i = Pr[know what the punishment reputation is]
- Cooperation strategies:

  C: cooperate; $\mathbf{O_C}$: opportunistically cooperate;
  D: defect; $\mathbf{O_D}$: opportunistically defect

  - Opportunistic choices depend on punishment reputation
- Punishment strategies:

  R (Responsible): punish defectors; **S** (Spiteful): punish everyone;
  A (Antisocial): punish cooperators; **N** (Non-punisher): punish no one

- Result: evolve stable proportions of cooperators/defectors and punishers/nonpunishers

Better Model

- Hilbe & Traulsen, 2012. Emergence of responsible sanctions without second order free riders, antisocial punishment or spite. *Nature Scientific Reports*.

- *Punishment reputation*: how likely that others will punish
  - Enables one to infer Cooperating has higher payoff than Defecting
- *Information level*: i = Pr[know what the punishment reputation is]
- Cooperation strategies:

  C: cooperate; $\mathbf{O_C}$: opportunistically cooperate;
  D: defect; $\mathbf{O_D}$: opportunistically defect

  - Opportunistic choices depend on punishment reputation
- Punishment strategies:

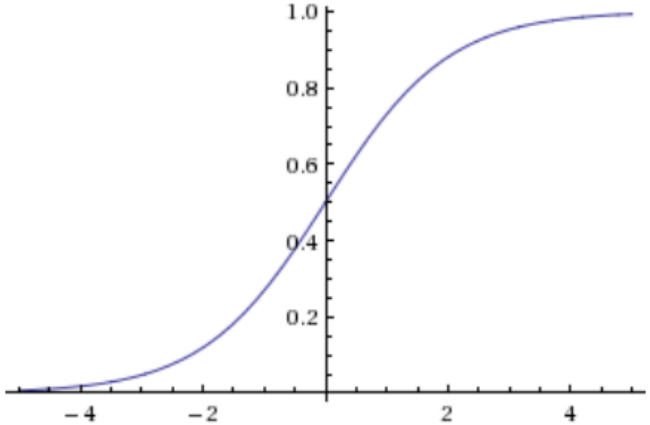
  R (Responsible): punish defectors; **S** (Spiteful): punish everyone;
  A (Antisocial): punish cooperators; **N** (Non-punisher): punish no one

- Result: evolve stable proportions of cooperators/defectors and punishers/nonpunishers

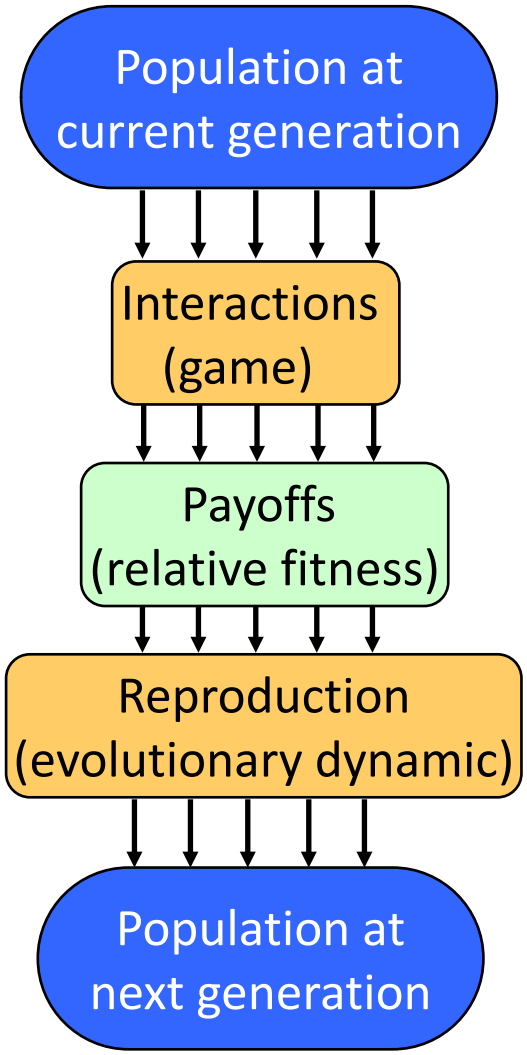
Which evolutionary dynamic?

Evolutionary Dynamic

- They used the Fermi rule
  - Each individual compares its payoff to that of a randomly chosen neighbor
  - Pr[switch to neighbor's strategy] = $1/(1 + e^{s(r - r')})$
    - r, r' = individual's and neighbor's payoffs
    - $s \geq 0$ is the *selection strength*



- *with exploration*:
  - Small probability μ of choosing a strategy at random from {all pure strategies}



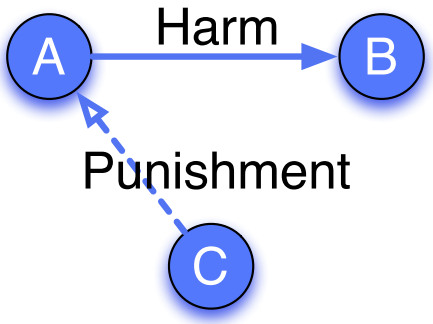
Third-Party Punishment

- **Poll (part 1):**
- You're in a big city
- You go into a large store you haven't visited before
- You see someone shoplifting. What do you do?
  - Try to ignore the incident?
  - Confront the shoplifter?
  - Report it to someone who works in the store?

Third-Party Punishment

- **Poll (part 2):**
- You're in a small "mom & pop" store in your home neighborhood
- You see someone shoplifting. What do you do?
  - Try to ignore the incident?
  - Confront the shoplifter?
  - Report it to someone who works in the store?

Third-Party Punishment (3PP)



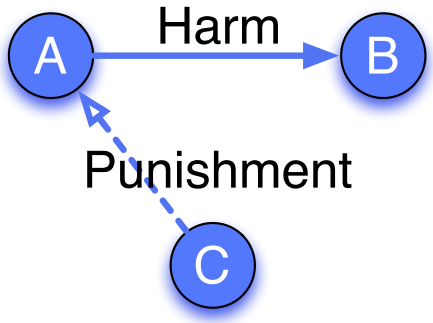
- Individual C (uninvolved third party) punishes individual A for harm that A has caused to B
  - Puzzling both psychologically and game-theoretically
  - Involves a cost to C, and gives C no direct benefit

- Empirical evidence that humans (and some animals) do it
- Other empirical studies in which it did not occur

Our Model

- Basic Questions
  - Under what conditions does 3PP occur?
  - How does this relate to other cultural characteristics?

- Model: similar to Hilbe & Traulsen's, with three main modifications:
  - 2-party interactions instead of PGG
  - Punishment by third party
  - Environmental/structural factors related to cultural characteristics

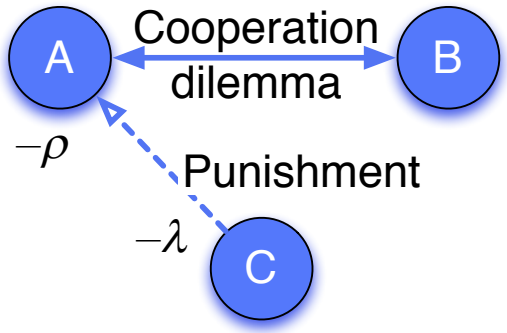


Modeling Third-Party Punishment

Basically the Prisoner's Dilemma

- **Interaction phase**: randomly choose pairs of individuals to play a cooperation dilemma
  - $c > 0$: cost of cooperating with other
  - $b > c$: benefit to self if other cooperates
- **Punishment phase**:
  - For each individual A that interacted, randomly choose an uninvolved neighbor C who may choose to punish A
  - Pay λ to reduce A's payoff by ρ
- Punishment reputation, information level, cooperation strategies, punishment strategies, evolutionary dynamic
  - Like Hilbe & Traulsen's

	Cooperate	Defect
Cooperate	$b - c, b - c$	$-c, b$
Defect	$b, -c$	$0, 0$



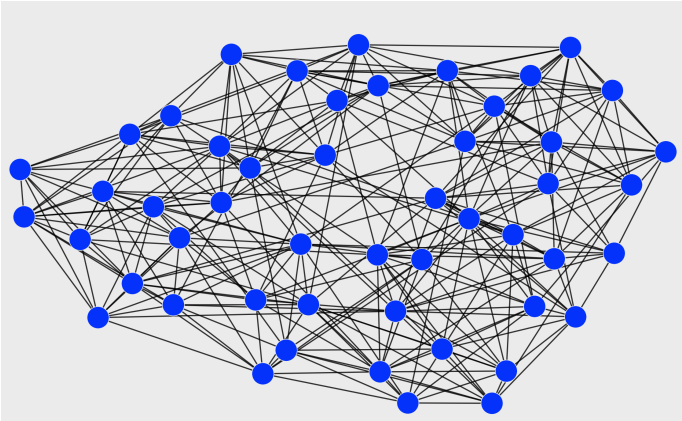
Cultural Characteristics

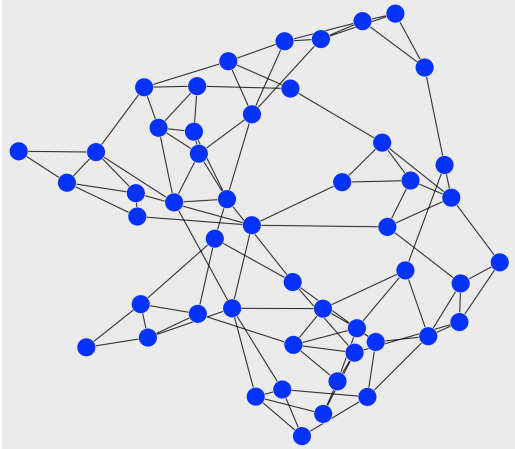
- **Strength of social ties**
  - How strongly are you attached to your neighbors?
- **Mobility**
  - How frequently do people move elsewhere?

- Do this by adding a network structure

Games with Population Structure

- Network of individuals
  - Each individual is at one of the nodes
  - Edges represent social connections
    - Possibilities for interaction and cultural transmission





Watts-Strogatz small-world networks

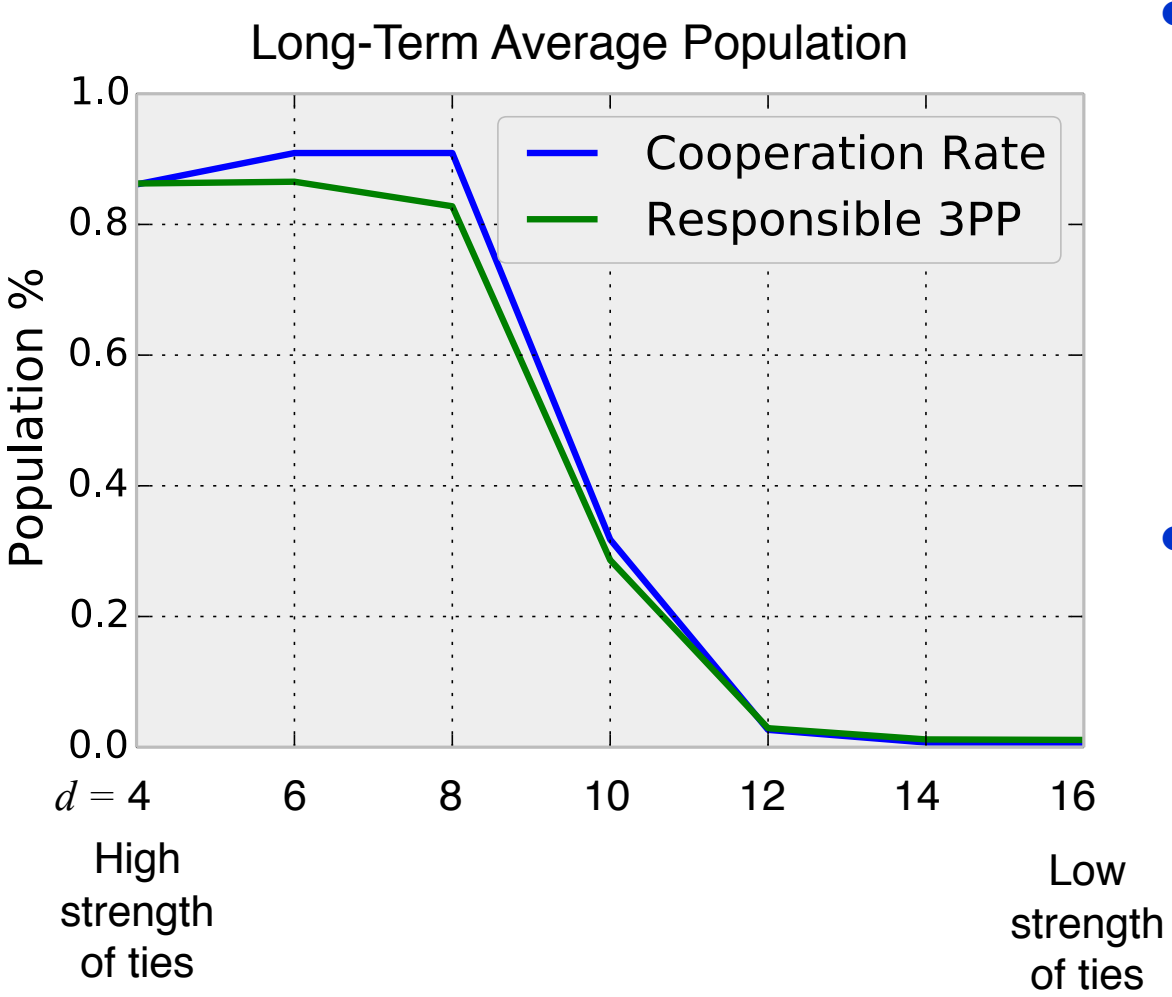
Strength-of-Ties

- How to measure *strength of ties* between humans?
  - How often individuals interact with each other during a period of time
- Suppose
  - individuals interact with their neighbors on the graph
  - individuals have a limited amount of interactions per time period
- Nodes with many neighbors
  - few interactions with each neighbor

    → *low* strength-of-ties with each neighbor
- Nodes with few neighbors
  - many interactions with each neighbor

    → *high* strength-of-ties with each neighbor
- So measure strength-of-ties as $1/d$
  - where d = number of neighbors

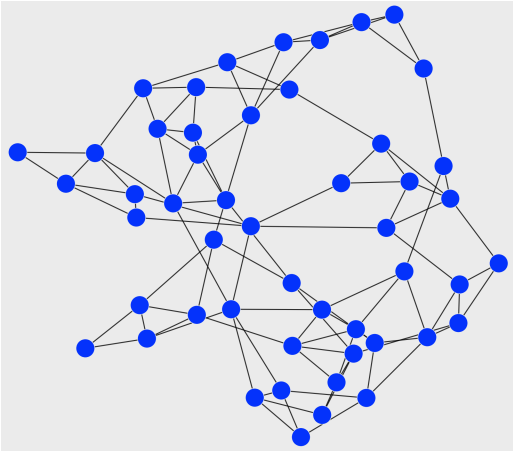
Strength of Ties and Third-Party Punishment



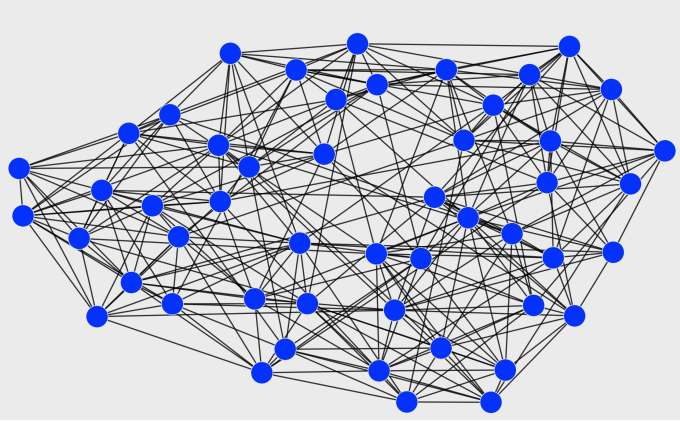
- High strength of ties:
  - High cooperation, low defection
  - High proportion of responsible 3PP strategies, but not much actual punishment
- Low strength of ties:
  - Low cooperation, high defection
  - High proportion of nonpunishers

Explanation

- With high strength-of-ties, there can be small groups of agents that interact primarily with each other
  - If their strategies are to cooperate and to punish responsibly, they can
    - achieve high payoffs
    - induce neighbors to adopt their strategy
  - Create a local environment in which agents are encouraged to cooperate and punish responsibly



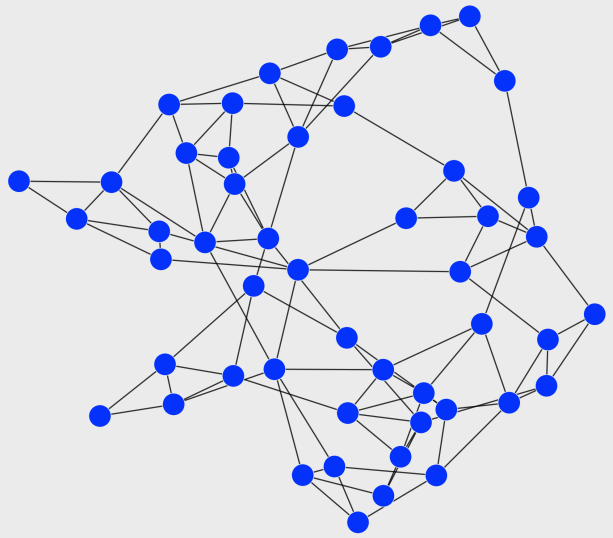
- With low strength-of-ties, such agents will interact with each other only occasionally
  - can't maintain high payoffs, eventually switch to other strategies



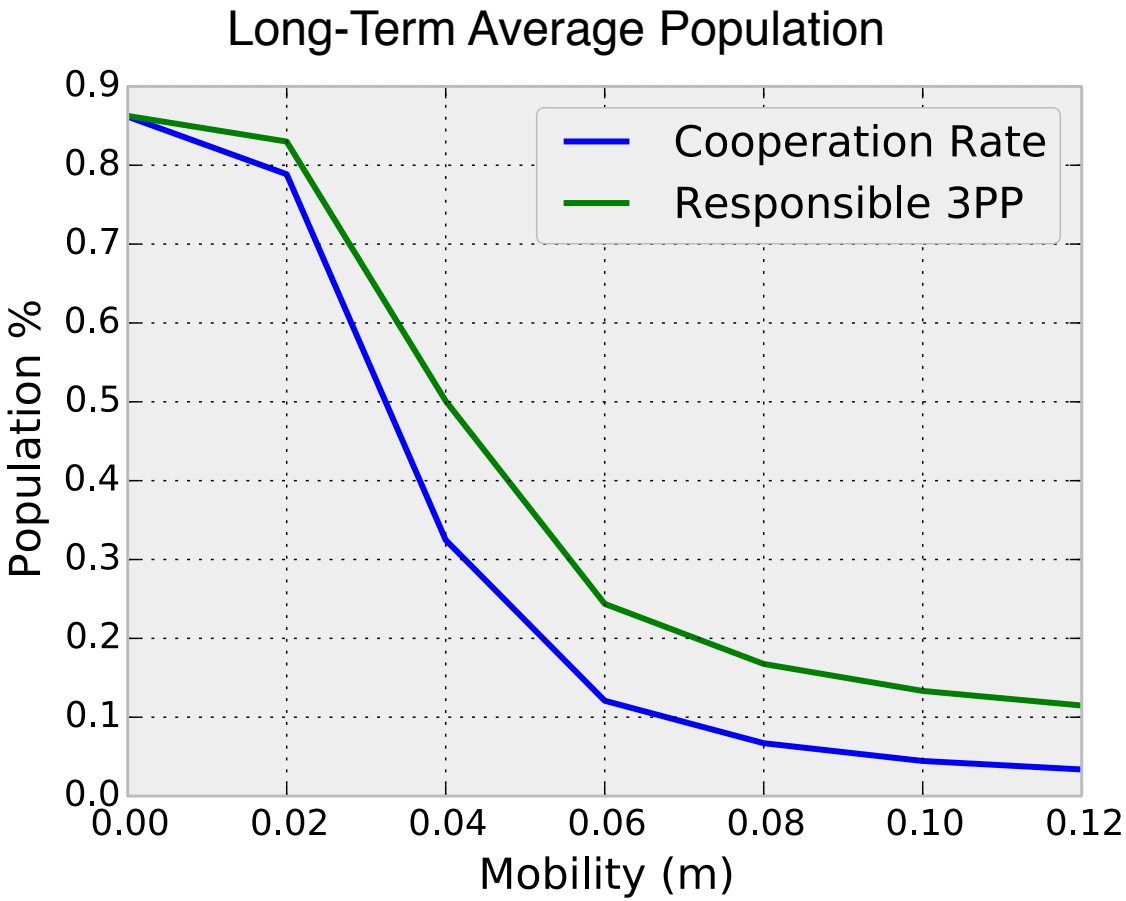
Mobility

- Degree to which humans change their location (social network position within a population)
- Individualistic cultures tend to have very high mobility
  - People can easily exit their social groups
- Collectivistic cultures tend to have low mobility
  - Not as easy to exit social group

- Implementation of Mobility:
  - At each iteration, individuals may switch positions with other randomly chosen individuals, with fixed probability *m*
  - High value for *m* ⇔ high mobility



Evolution of 3PP Requires Low Mobility



- Low mobility:
  - High cooperation, low defection
  - High proportion of responsible 3PP strategies, but not much actual punishment
- High mobility:
  - Low cooperation, high defection
  - High proportion of nonpunishers

- High mobility → successful clusters of 3PP individuals won't last, because the individuals move away

Discussion

- Provide support for causal relationship:
  - Evolution of responsible 3PP requires "punishment reputation" plus some societal constraints
    - High strength-of-ties, low mobility
  - With them, 3PP can emerge as a trait that benefits those who do it
  - Without them, 3PP can't be sustained or uphold cooperation

- More generally
  - Example of how an interdisciplinary team can accomplish something that none of us could have done individually
  - Individual-level interactions + different structural factors
    ➔ differences in evolved culture
  - We're doing a number of studies along this line

Publication

PROCEEDINGS OF THE ROYAL SOCIETY B | BIOLOGICAL SCIENCES

High strength-of-ties and low mobility enable the evolution of third-party punishment

Patrick Roos[1,3]↑, Michele Gelfand[2], Dana Nau[1,3,4] and Ryan Carr[1,3]

- Author Affiliations

[1]Department of Computer Science, University of Maryland, College Park, MD, USA
[2]Department of Psychology, University of Maryland, College Park, MD, USA
[3]Institute for Advanced Computer Studies, University of Maryland, College Park, MD, USA
[4]Institute for Systems Research, University of Maryland, College Park, MD, USA
e-mail: roos@cs.umd.edu

Abstract

As punishment can be essential to cooperation and norm maintenance but costly to the punisher, many evolutionary game-theoretic studies have explored how direct punishment can evolve in populations. Compared to direct punishment, in which an agent acts to punish another for an interaction in which both parties were involved, the evolution of third-party punishment (3PP) is even more puzzling, because the punishing agent itself was not involved in the original interaction. Despite significant empirical studies of 3PP, little is known about the conditions under which it can evolve. We find that punishment reputation is not, by itself, sufficient for the evolution of 3PP. Drawing on research streams in sociology and psychology, we implement a

News & Events

Game theory shows why some communities punish bad neighbors

Friday, December 6, 2013

COLLEGE PARK, MD—You're shopping for holiday gifts when you spot someone pocketing a nice pair of leather gloves. What do you do?

A new study by University of Maryland researchers appearing this week in the *Proceedings of the Royal Society B* predicts

Why Some Communities Police Themselves, While Others Don't

Familiarity and social ties influence our decision to intervene on behalf of others.

MIKE RIGGS | @mikeriggs | Dec 16, 2013 | 5 Comments

Share on Facebook Tweet

Greater Greater Washington

The Washington, DC area is great. But it could be **greater**.

LINKS

Breakfast links: Saving the city

by Chad Maddox • December 17, 2013

Self-policing DC: UMD researchers suggest that people with strong ties to a neighborhood are more likely to intervene on another's behalf as a "responsible third party punisher", which may mean less self-policing in transient cities like DC. (Atlantic Cities)

Gotham City Chronicle

BATKID SAVES CITY

Hooded hero nabs Riddler, rescues damsel in distress

Caped crusader crushes crime

Photo by Shawn on Flicker.

Comments Add a comment »

"Responsible Third-Party Punisher" sounds like a great idea for a caption for t-shirts.

by Frank IBC on Dec 17, 2013 9:54 am • link • report