General-sum games

- You could still play a minimax strategy in general-sum games
  - I.e., pretend that the opponent is only trying to hurt you
- But this is not rational:

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

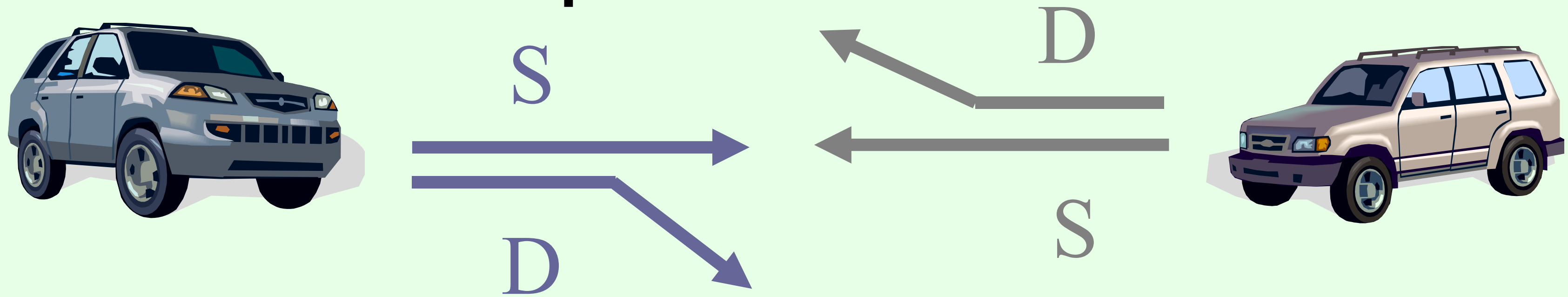
- If Column was trying to hurt Row, Column would play Left, so Row should play Down
- In reality, Column will play Right (strictly dominant), so Row should play Up
- Is there a better generalization of minimax strategies in zero-sum games to general-sum games?

Nash equilibrium

[Nash 50]

- A vector of strategies (one for each player) is called a strategy profile
- A strategy profile $(\sigma_1, \sigma_2, \ldots, \sigma_n)$ is a Nash equilibrium if each σ_i is a best response to σ_{-i}
  - That is, for any i, for any σ_i', $u_i(\sigma_i, \sigma_{-i}) \geq u_i(\sigma_i', \sigma_{-i})$
- Note that this does not say anything about multiple agents changing their strategies at the same time
- In any (finite) game, at least one Nash equilibrium (possibly using mixed strategies) exists [Nash 50]
- (Note - singular: equilibrium, plural: equilibria)

Nash equilibria of "chicken"



	D	S
D	0, 0	-1, 1
S	1, -1	-5, -5

- (D, S) and (S, D) are Nash equilibria
  - They are pure-strategy Nash equilibria: nobody randomizes
  - They are also strict Nash equilibria: changing your strategy will make you strictly worse off
- No other pure-strategy Nash equilibria

Nash equilibria of "chicken"…

	D	S
D	0, 0	-1, 1
S	1, -1	-5, -5

- Is there a Nash equilibrium that uses mixed strategies? Say, where player 1 uses a mixed strategy?
- Recall: if a mixed strategy is a best response, then all of the pure strategies that it randomizes over must also be best responses
- So we need to make player 1 indifferent between D and S
- Player 1's utility for playing D = $-p^c_S$
- Player 1's utility for playing S = $p^c_D - 5p^c_S = 1 - 6p^c_S$
- So we need $-p^c_S = 1 - 6p^c_S$ which means $p^c_S = 1/5$
- Then, player 2 needs to be indifferent as well
- Mixed-strategy Nash equilibrium: ((4/5 D, 1/5 S), (4/5 D, 1/5 S))
  - People may die! Expected utility -1/5 for each player

The presentation game

Audience		Presenter: *Put effort into presentation (E)*	Presenter: *Do not put effort into presentation (NE)*
	Pay attention (A)	4, 4	-16, -14
	Do not pay attention (NA)	0, -2	0, 0

- Pure-strategy Nash equilibria: (A, E), (NA, NE)
- Mixed-strategy Nash equilibrium:

  ((1/10 A, 9/10 NA), (4/5 E, 1/5 NE))
  - Utility 0 for audience, -14/10 for presenter
  - Can see that some equilibria are strictly better for both players than other equilibria, i.e. some equilibria Pareto-dominate other equilibria

The “equilibrium selection problem”

- You are about to play a game that you have never played before with a person that you have never met
- According to which equilibrium should you play?
- Possible answers:
  - Equilibrium that maximizes the sum of utilities (social welfare)
  - Or, at least not a Pareto-dominated equilibrium
  - So-called focal equilibria
    - “Meet in Paris” game - you and a friend were supposed to meet in Paris at noon on Sunday, but you forgot to discuss where and you cannot communicate. All you care about is meeting your friend. Where will you go?
  - Equilibrium that is the convergence point of some learning process
  - An equilibrium that is easy to compute
  - …
- Equilibrium selection is a difficult problem

Stackelberg (commitment) games

	L	R
L	1, -1	3, 1
R	2, 1	4, -1

- Unique Nash equilibrium is (R,L)
  - This has a payoff of (2,1)

Commitment

	L	R
L	(1,-1)	(3,1)
R	(2,1)	(4,-1)

- What if the officer has the option to (credibly) announce where he will be patrolling?
- This would give him the power to “commit” to being at one of the buildings
  - This would be a pure-strategy Stackelberg game

Commitment…

	L	R
L	(1,-1)	(3,1)

- If the officer can commit to always being at the left building, then the vandal's best response is to go to the right building
  - This leads to an outcome of (3,1)

Committing to mixed strategies

	L	R
L	(1,-1)	(3,1)
R	(2,1)	(4,-1)

- What if we give the officer even more power: the ability to commit to a mixed strategy
  - This results in a mixed-strategy Stackelberg game
  - E.g., the officer commits to flip a weighted coin which decides where he patrols

Committing to mixed strategies is more powerful

	L	R
L	(1,-1)	(3,1)
R	(2,1)	(4,-1)

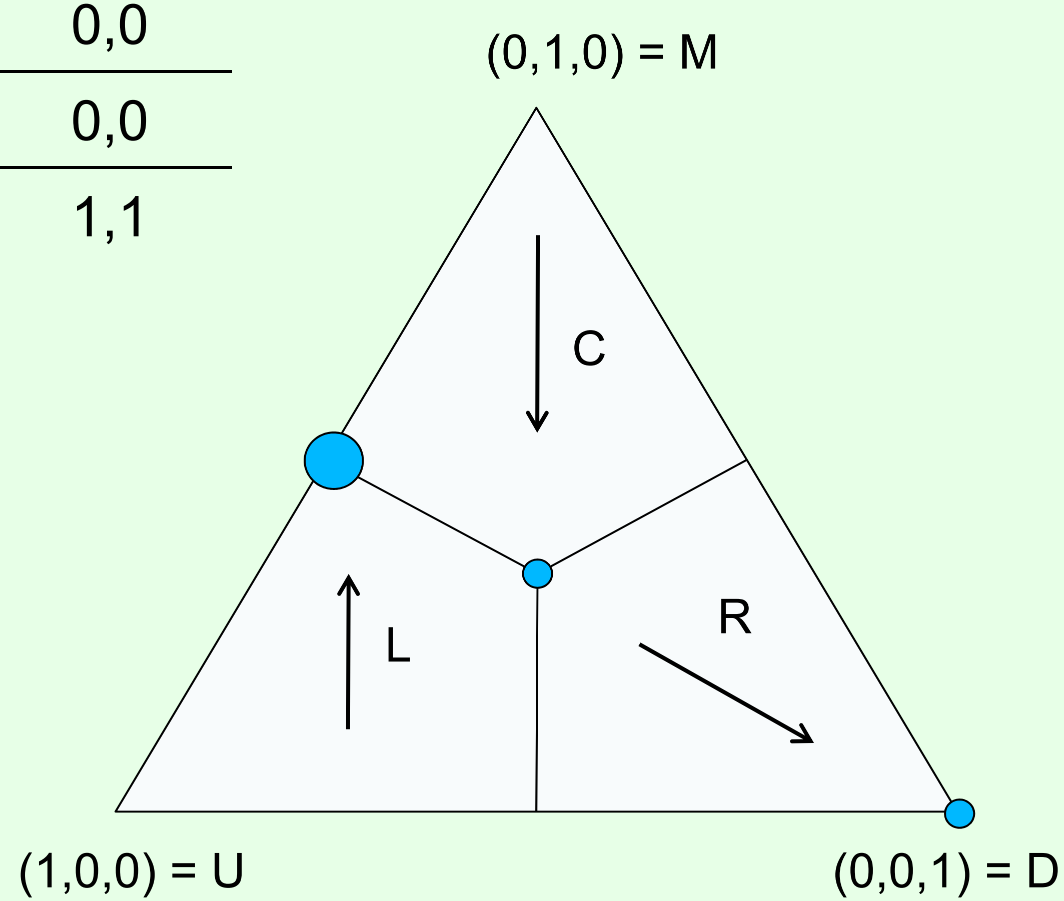
- Suppose the officer commits to the following strategy: {(.5+ε)L,(.5- ε)R}
  - The vandal’s best response is R
  - As ε goes to 0, this converges to a payoff of (3.5,0)

Stackelberg games in general

- One of the agents (the leader) has some advantage that allows her to commit to a strategy (pure or mixed)
- The other agent (the follower) then chooses his best response to this

Visualization

	L	C	R
U	0,1	1,0	0,0
M	4,0	0,1	0,0
D	0,0	1,0	1,1



Easy polynomial-time algorithm for two players

- For every column t, we solve separately for the best mixed row strategy (defined by p_s) that induces player 2 to play t
- maximize $\Sigma_s\, p_s\, u_1(s, t)$
- subject to

  for any t', $\Sigma_s\, p_s\, u_2(s, t) \geq \Sigma_s\, p_s\, u_2(s, t')$

  $\Sigma_s\, p_s = 1$
- (May be infeasible)
- Pick the t that is best for player 1

(a particular kind of) Bayesian games

leader utilities

2	4
1	3

follower utilities (type 1)

1	0
0	1

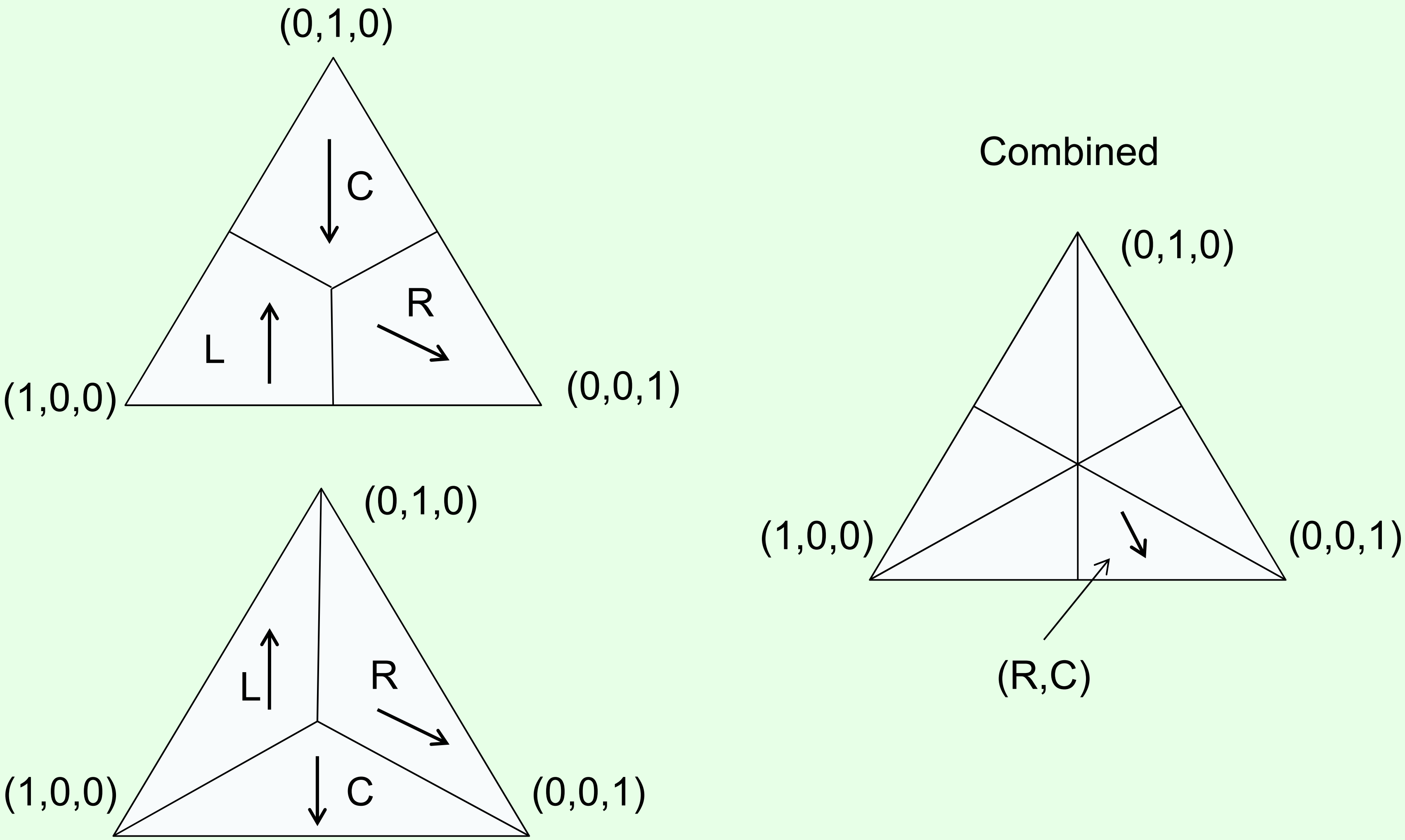
probability .6

follower utilities (type 2)

1	0
1	3

probability .4

Multiple types - visualization



Solving Bayesian games

- There's a known MIP for this[1]
- Details omitted due to the fact that its rather nasty.
  - The main trick of the MIP is encoding a exponential number of LP's into a single MIP
  - Used in the ARMOR system deployed at LAX

[1] Paruchuri et al. Playing Games for Security: An Efficient Exact Algorithm for Solving Bayesian Stackelberg Games

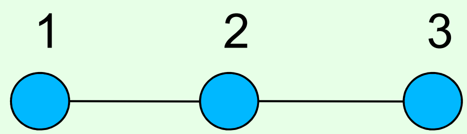
(In)approximability

- (# types)-approximation: optimize for each type separately using the LP method. Pick the solution that gives the best expected utility against the entire type distribution.

- Can't do any better in polynomial time, unless P=NP
  - Reduction from INDEPENDENT-SET

- For adversarially chosen types, cannot decide in polynomial time whether it is possible to guarantee positive utility, unless P=NP

Reduction from independent set



leader utilities

	A	B
a_l^1	1	0
a_l^2	1	0
a_l^3	1	0

follower utilities (type 1)

	A	B
a_l^1	3	1
a_l^2	0	10
a_l^3	0	1

follower utilities (type 2)

	A	B
a_l^1	0	10
a_l^2	3	1
a_l^3	0	10

follower utilities (type 3)

	A	B
a_l^1	0	1
a_l^2	0	10
a_l^3	3	1

Value

	L	R
L	(1,-1)	(3,1)
R	(2,1)	(4,-1)

- We define three ratios:
  - Value of Pure Commitment (VoPC) 3/2
  - Value of Mixed Commitment (VoMC) 3.5/2
  - Mixed vs Pure (MvP) 3.5/3

Related concepts

- Value of mediation[1]
  - Ratio of combined utility of all players between:
    - Best correlated equilibrium
    - Best Nash equilibrium
- Price of anarchy[2]
  - Measures loss of welfare lost due to selfishness of the players
  - Ratio between:
    - Worst Nash equilibrium
    - Optimal social welfare
- One difference is that we are only concerned with player 1’s utility

1 Ashlagi et al. 08 2 Koutsoupias and Papadimitriou 99

Normal form (2×2)

	L	R
U	ϵ,1	1,0
D	0,0	1-ϵ ,1

VoPC = VoMC = ∞

Normal form (2×2) continued

	L	R
U	ϵ,1	1,0
D	0,0	2ϵ ,1

MvP = ∞

Summary of value results

Game type	VoPC	VoMC	MvP
Normal-form (2×2)	∞ (ISD)	∞ (ISD)	∞ (ISD)
Symmetric normal-form (3×3)	∞ (ISD)	∞ (ISD)	∞ (ISD)
Extensive-form (3 leaves)	∞ (SPNE)	∞ (SPNE)	∞ (SPNE)
Security games (2×2)	∞ (UNash)	∞ (UNash)	∞ (UNash)
Atomic Selfish routing			
Linear costs (n players)	n (ISD)	n (ISD)	n (ISD)
Quadratic costs (n players)	n^2 (ISD)	n^2 (ISD)	n^2 (ISD)
k-nomnial costs (n players)	n^k (ISD)	n^k (ISD)	n^k (ISD)
Arbitrary costs (2 players)	∞ (ISD)	∞ (ISD)	∞ (ISD)

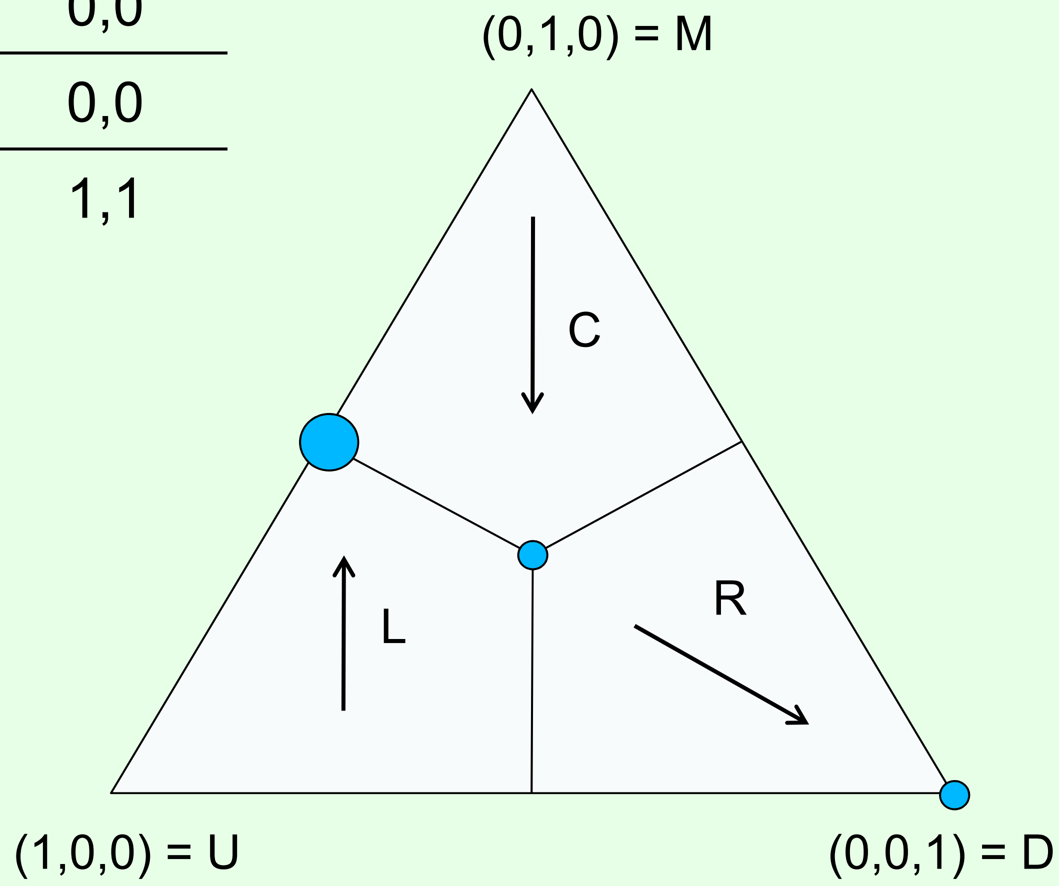
Learning

- Single follower type
- Unknown follower payoffs
- Repeated play: commit to mixed strategy, see follower's (myopic) response

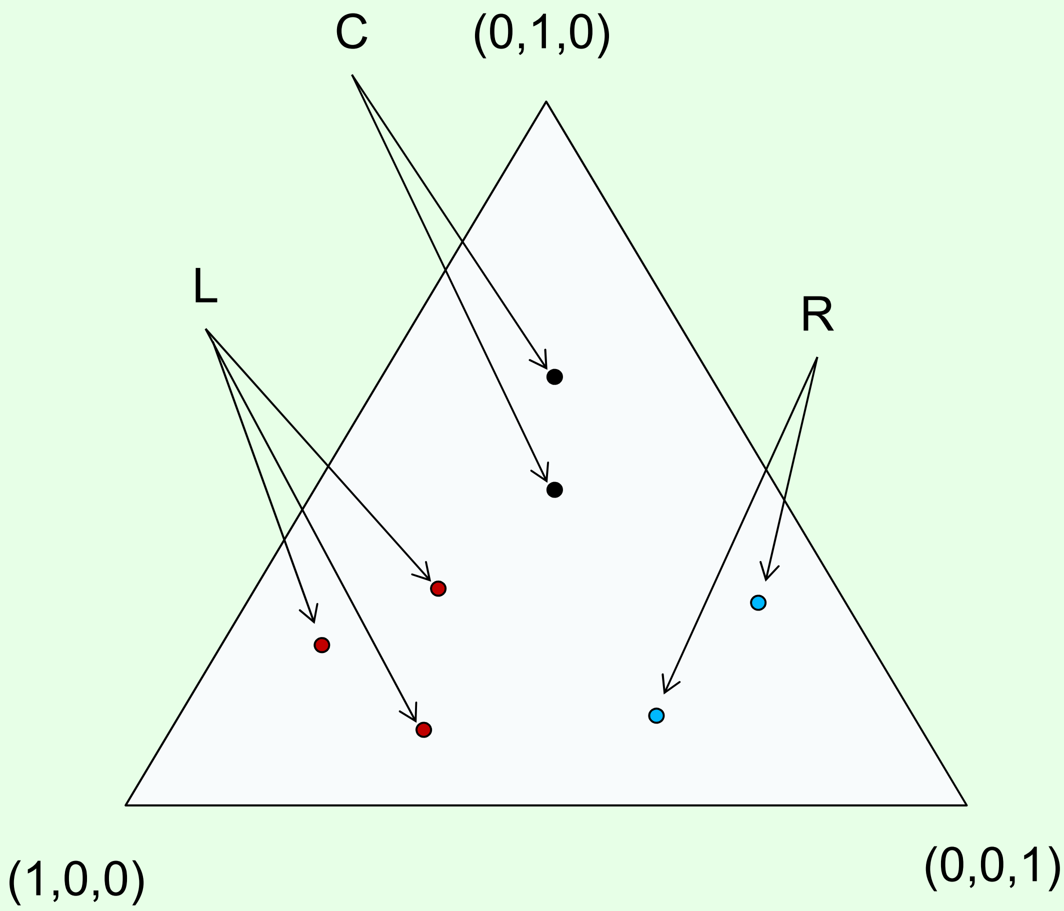
	L	R
U	1,?	3,?
D	2,?	4,?

Visualization

	L	C	R
U	0,1	1,0	0,0
M	4,0	0,1	0,0
D	0,0	1,0	1,1



Sampling



Three main techniques in the learning algorithm

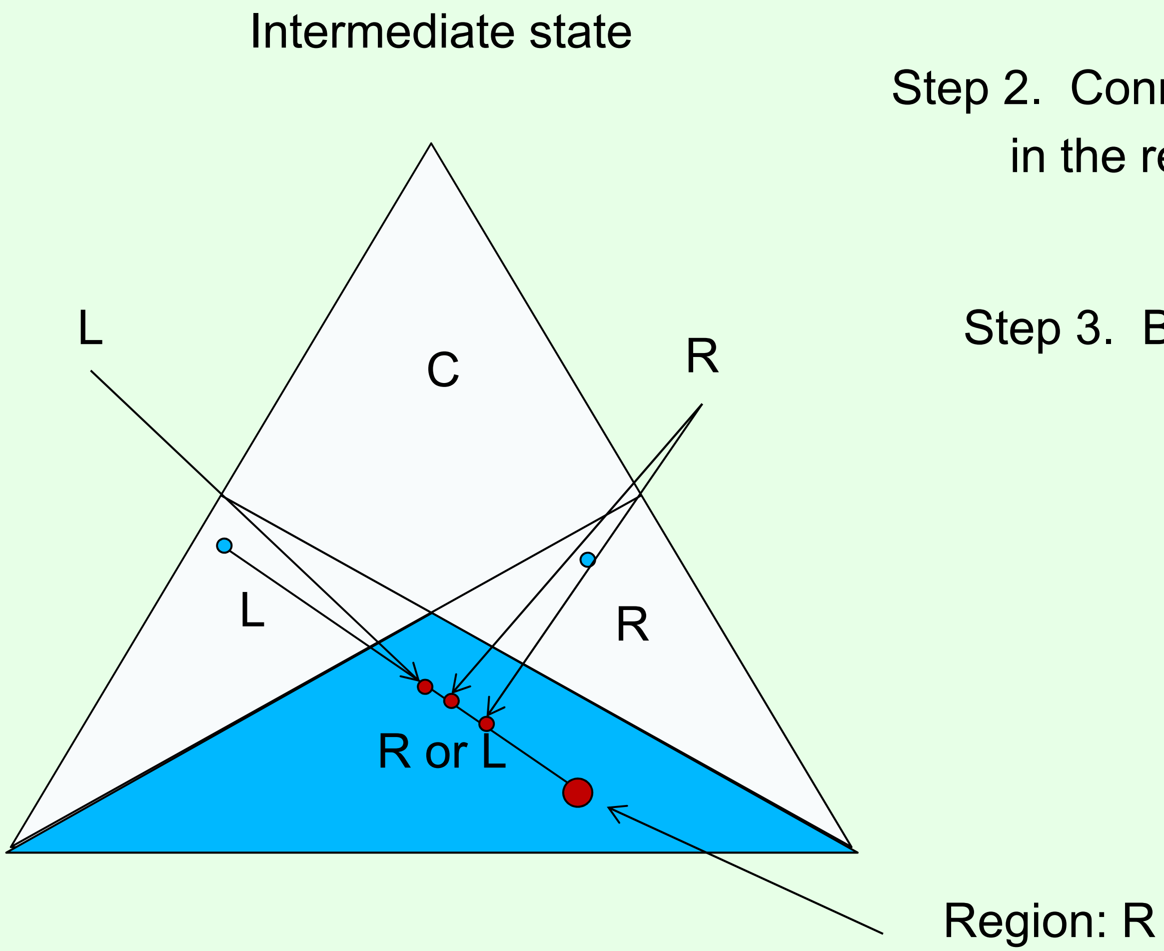
- Find one point in each region (using random sampling)
- Find a point on an unknown hyperplane
- Starting from a point on an unknown hyperplane, determine the hyperplane completely

Finding a point on an unknown hyperplane

Step 1. Sample in the overlapping region

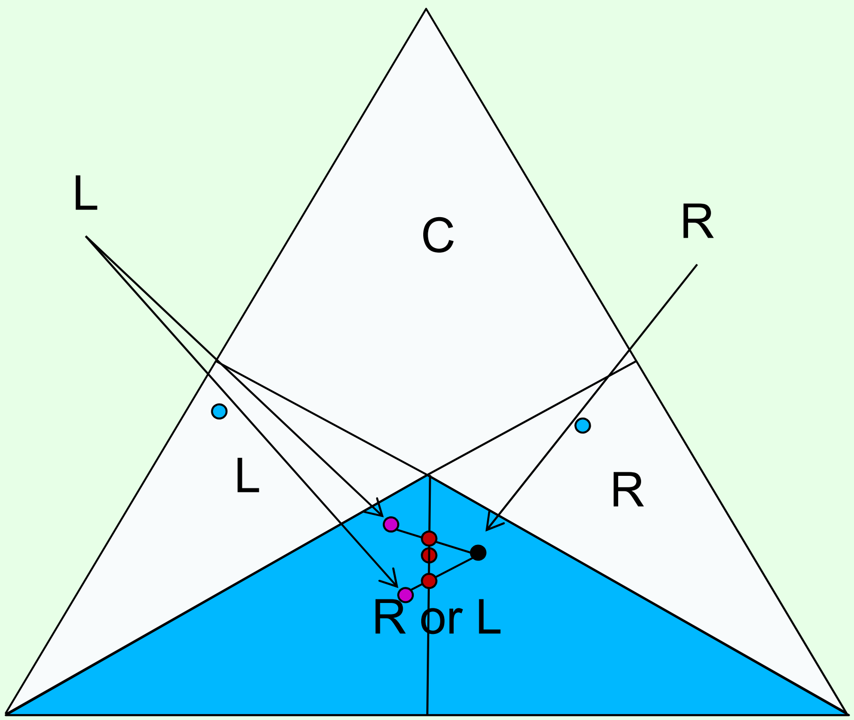
Step 2. Connect the new point to the point in the region that doesn't match

Step 3. Binary search along this line



Determining the hyperplane

Intermediate state



Step 1. Sample a regular d-simplex centered at the point

Step 2. Connect d lines between points on opposing sides

Step 3. Binary search along these lines

Step 4. Determine hyperplane (and update the region estimates with this information)

Bound on number of samples

Theorem. *Finding all of the hyperplanes necessary to compute the optimal mixed strategy to commit to requires* $O(Fk \log(k) + dLk^2)$ *samples*

- F depends on the size of the smallest region
- L depends on desired precision
- k is the number of follower actions
- d is the number of leader actions