CS540 Introduction to Artificial Intelligence
Lecture 21

Young Wu
Based on lecture slides by Jerry Zhu and Yingyu Liang

August 5, 2019

Monte Carlo Tree Search

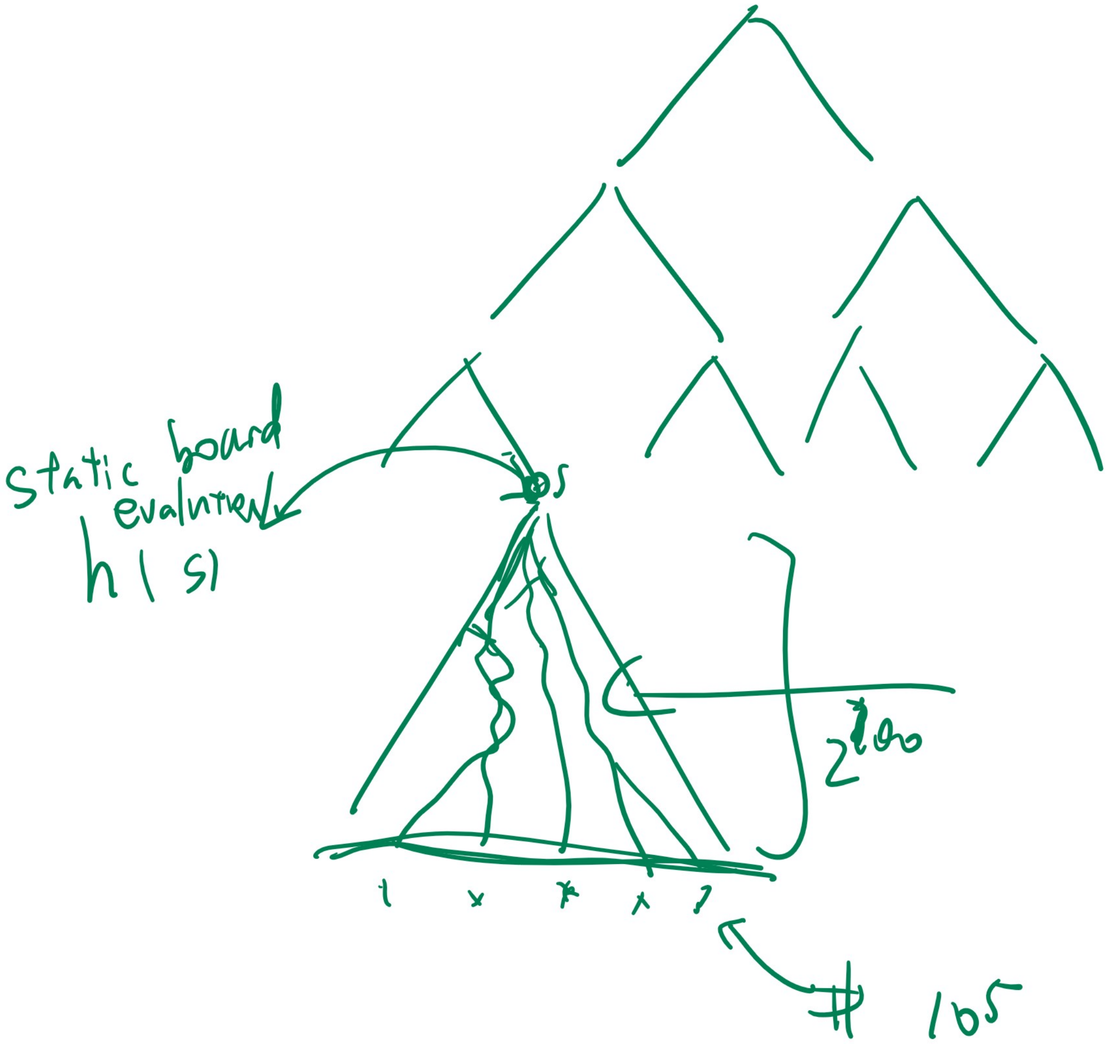
Discussion

Lecture 19

- Simulate random games by selecting random moves for both players.
- Exploitation by keeping track of average win rate for each successor from previous searches and picking the successors that lead to more wins.
- Exploration by allowing random choices of unvisited successors.

Monte Carlo Tree Search Diagram

Discussion



Upper Confidence Bound

Discussion

- Combine exploitation and exploration by picking sucessors using upper confidence bound for tree.

$$\operatorname{argmax}_{s} \quad \frac{w_s}{n_s} + c\sqrt{\frac{\log t}{n_s}}$$

- w_s is the number of wins after successor s, and n_s the number of simulations after successor s, and t is the total number of simulations.
- Similar to the UCB algorithm for MAB.

Alpha GO Example

Discussion

- MCTS with $> 10^5$ playouts.
- Deep neural network to compute SBE.

Normal Form of Sequential Games

Discussion

Lecture 20

L

		L	R
R	U		
	D		

- Sequential games can have normal form too, but the solution concept is different.
- Nash equilibria of the normal form may not be a solution of the original sequential form game.

Non-credible Threat Example, Part I

Quiz (~~Graded~~)

Participation

- Country A can choose to Attack or Not attack country B. If country A chooses to Attack, country B can choose to Fight back or Escape. The costs are the largest for both countries if they fight, but otherwise, A prefers attacking (and B escaping) and B prefers A not attacking. What are the Nash equilibria?
- A: (A, F)
- B: (A, E)
- C: (N, F)
- D: (N, E)
- ~~E: (N)~~

not credible.

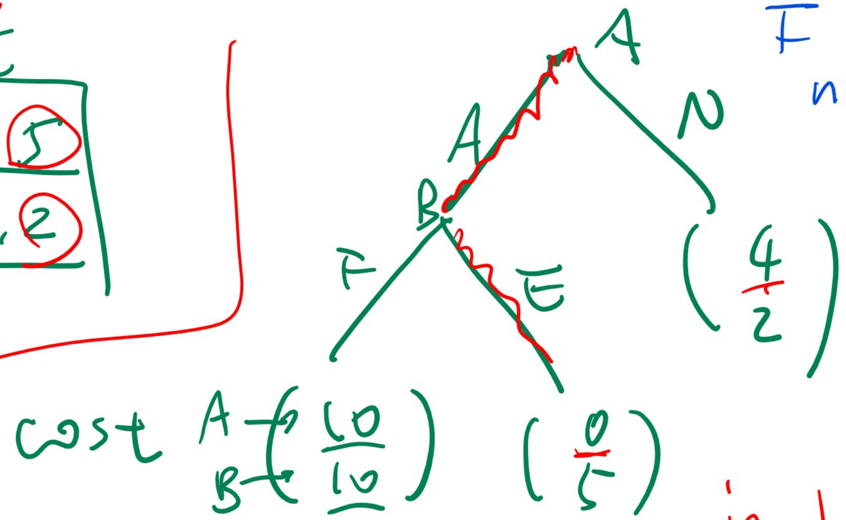
in L20

B

	F	E
A	10,10	0,5
N	4,2	4,2

A

Nash equilibria



$\left(\frac{10}{10}\right)$ $\left(\frac{0}{5}\right)$ $\left(\frac{4}{2}\right)$

F is a non credible threat.

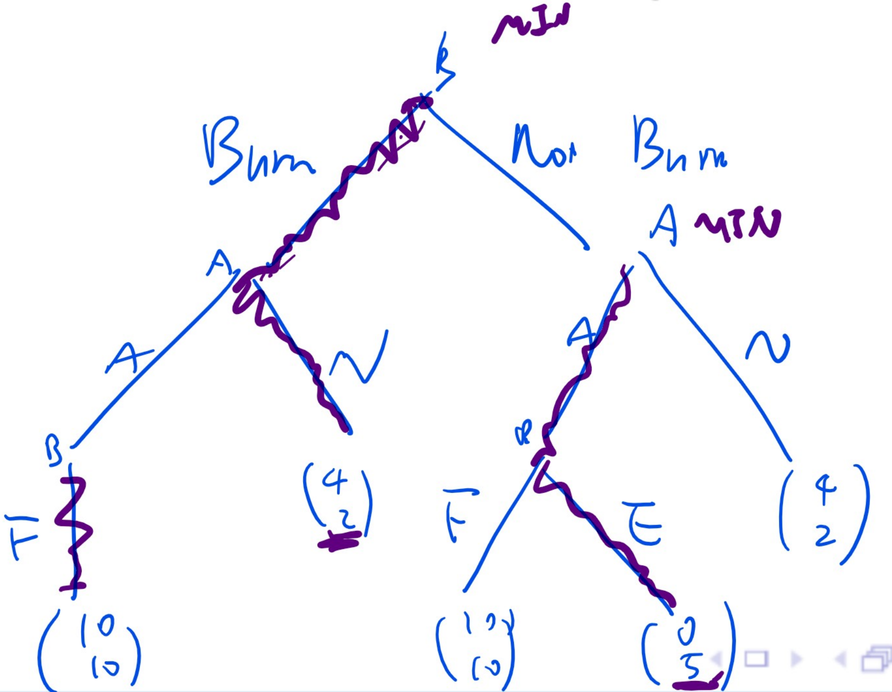
in L19

"solution" to sequential

Non-credible Threat Example, Part II

Quiz (Graded)

- What if country B can burn the bridge at the beginning of the game so that it cannot choose to escape?



Prisoner's Dilemma

Discussion

- A simultaneous move, non-zero-sum, and symmetric game is a prisoner's dilemma game if the Nash equilibrium state is strictly worse for both players than another state.

–	C	D
C	(x, x)	$(0, y)$
D	$(y, 0)$	$(1, 1)$

- C stands for Cooperate and D stands for Defect (not Confess and Deny). Both players are MAX players. The game is PD if $y > x > 1$. Here, (D, D) is the only Nash equilibrium and (C, C) is strictly better than (D, D) for both players.

Wage Competition, Version I

Quiz (Participation)

- Assume the productivity of the applicant is 20 dollars per hour, and in case of a tie in the offers, the applicant randomly picks each company with probability a half. What should the companies offer?

NE

- A: (18, 18)
- B: (19, 19)
- C: (20, 20)
- D: (19, 18)
- E: (18, 19)

SPE

solution

seq level

A

MAX

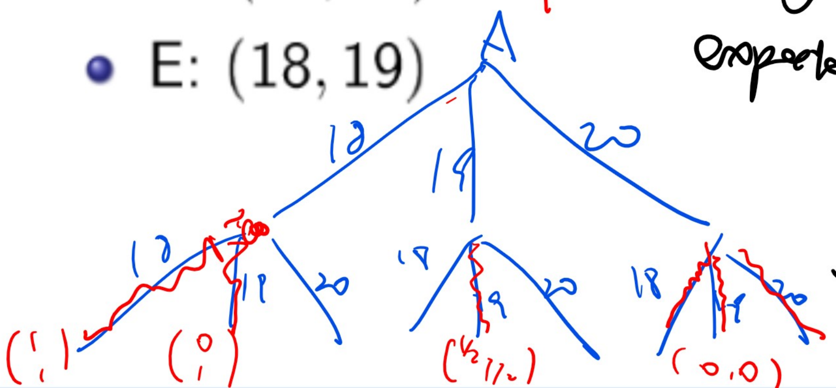
⇓

expected

	B		
A	18	19	20 MAX
18	(1, 1)	(0, 1)	(0, 0)
19	(1, 0)	$(\frac{1}{2}, \frac{1}{2})$	(0, 0)
20	(0, 0)	(0, 0)	(0, 0)

$\frac{1}{2} \cdot 2 \cdot \quad \frac{1}{2} \cdot 0$

Max 1



Wage Competition, Version II

Quiz (Participation)

- Assume the productivity of the applicant is 20 dollars per hour, and in case of a tie in the offers, the applicant pick company 1. What should the companies offer?
- A: $(18, 18)$
- B: $(19, 19)$
- C: $(20, 20)$
- D: $(19, 18)$
- E: $(18, 19)$

Wage Competition, Version III

Quiz (Participation)

- Assume the productivity of the applicant is 20 dollars per hour, and in case of a tie in the offers, the applicant pick company 2. What should the companies offer?
- A: $(18, 18)$
- B: $(19, 19)$
- C: $(20, 20)$
- D: $(19, 18)$
- E: $(18, 19)$

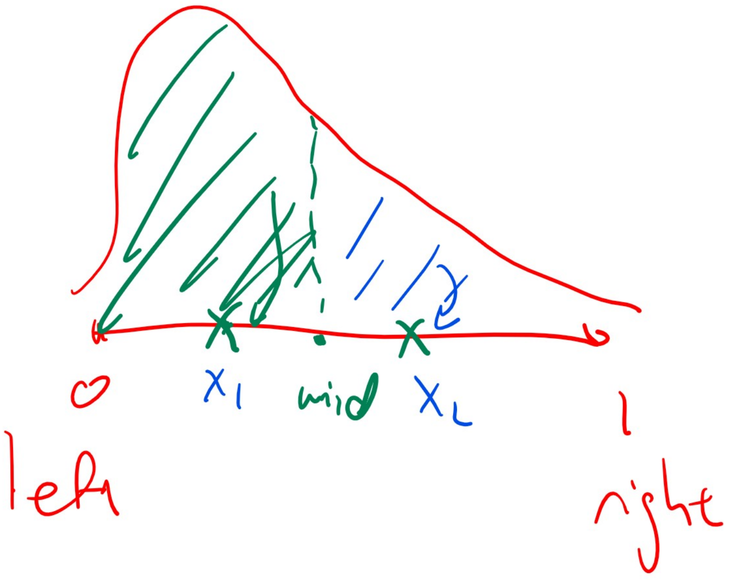
Median Voter Theorem, Part I

Quiz (Participation)

- Voters are distributed according to density function $f(x)$ on the one dimensional political spectrum $x \in [0,1]$. Each voter votes for the politician closer to his or her own position (randomly pick one in case of a tie). Two politicians choose positions x_1 and x_2 trying the maximize the number of votes. $+|x_1 - \hat{x}_1|$

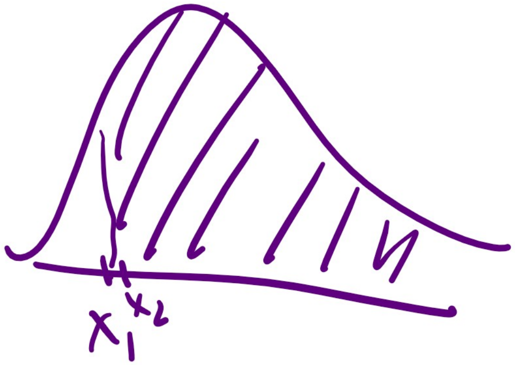
where is NE actions?

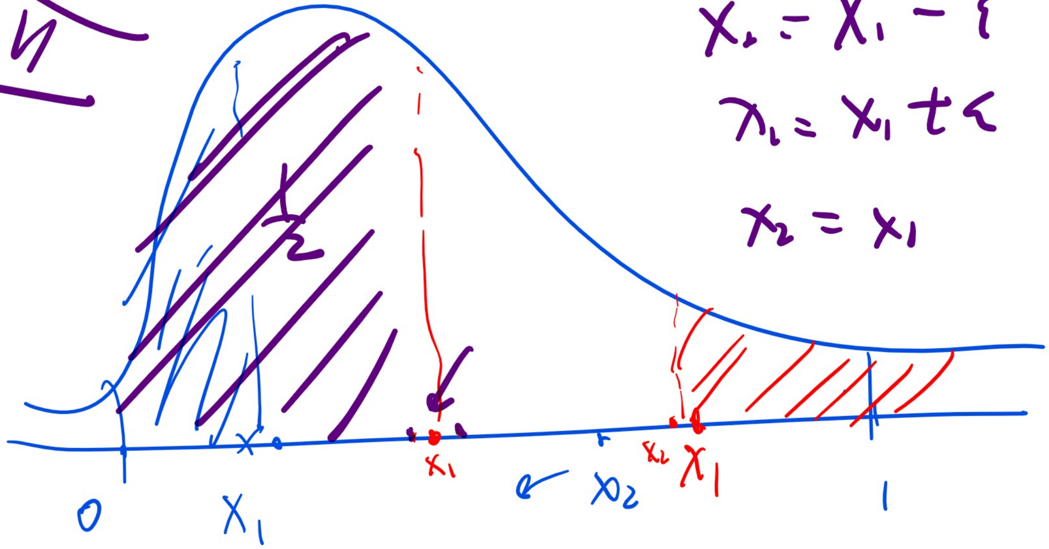
- A: Midpoint: $\frac{1}{2}$
- B: Mean: $\int_0^1 xf(x)\,dx$
- [✓ C: Median: $m : \int_0^m f(x)\,dx = \frac{1}{2}$]
- D: Mode: $\max_{x \in [0,1]} f(x)$



Median Voter Theorem, Part II

Quiz (Participation)





$x_2 \to x_1 + \varepsilon$

if x_1 = median

$x_2 = x_1 - \varepsilon \Rightarrow$ less than $\frac{1}{2}$ votes

$x_2 = x_1 + \varepsilon \Rightarrow$ ____

$x_2 = x_1 \Rightarrow \frac{1}{2}$ votes.

Penalty Kick, Part I

Quiz (Participation)

- The kicker (ROW) and the goalie (COL) choose L, C, R simultaneously. The following table is the estimated probability of scoring the goal given the actions. Kicker maximizes the probability and goalie minimizes the probability. Find all mixed strategy Nash.

–	L	C	R
L	0.6	0.9	0.9
C	1	0.4	1
R	0.9	0.9	0.6

Penalty Kick, Part II

Quiz (Participation)

$L \rightarrow \frac{2}{5}0.4 + \frac{1}{5}0 + \frac{2}{5}0.1 = \frac{1}{5}$

$C \rightarrow \frac{2}{5}\cdot 0.1 + \frac{1}{5}0.6 + \frac{2}{5}\cdot 0.1$

$= \frac{1}{5}$

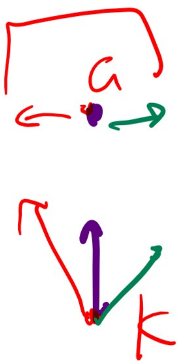
Row →

–	L	C	R
L	0.6	0.9	0.9
C	1	0.4	1
R	0.9	0.9	0.6

COL ×	L	C	R
$\frac{2}{5}$ L	0.4	0.1	0.1
$\frac{1}{5}$ C	0	0.6	0
$\frac{2}{5}$ R	0.1	0.1	0.4

ROW COL

- A: $\left(\cancel{\left(\frac{1}{3}L, \frac{1}{3}C, \frac{1}{3}R\right)}, \left(\frac{1}{3}L, \frac{1}{3}C, \frac{1}{3}R\right)\right)$
- B: $\boxed{\left(\left(\frac{2}{5}L, \frac{1}{5}C, \frac{2}{5}R\right), \left(\frac{1}{3}L, \frac{1}{3}C, \frac{1}{3}R\right)\right)}$
- C: $\left(\cancel{\left(\frac{1}{3}L, \frac{1}{3}C, \frac{1}{3}R\right)}, \left(\frac{2}{5}L, \frac{1}{5}C, \frac{2}{5}R\right)\right)$
- D: $\left(\left(\frac{2}{5}L, \frac{1}{5}C, \frac{2}{5}R\right), \left(\frac{2}{5}L, \frac{1}{5}C, \frac{2}{5}R\right)\right)$



Penalty Kick, Part III

Quiz (Participation)

Col : $L \to \frac{1}{3}0.4 + \frac{1}{3}\cdot 0. + \frac{1}{3}\cdot 0.1 = \frac{0.5}{3}$

$C \to \frac{1}{3}\cdot 0.1 + \frac{1}{3}\cdot 0.6 + \frac{1}{3}\cdot 0.1 = \frac{0.8}{3}$

$\frac{1}{3}$ $\frac{1}{3}$ $\frac{1}{3}$

p q $1-p-q$

L = C = R

—	L	C	R
L	0.6	0.9	0.9
C	1	0.4	1
R	0.9	0.9	0.6

$\frac{1}{3}$ $\frac{1}{3}$ $\frac{1}{3}$

$br_{col}(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) = C$

Row is indifferent ⇒ any mix is best response.

$L \to \frac{1}{3}\cdot 0.6 + \frac{1}{3}\cdot 0.9 + \frac{1}{3}\cdot 0.9 = \frac{2.4}{3}$

$C \to \frac{1}{3}\cdot 1 + \frac{1}{3}\cdot 0.4 + \frac{1}{3}\cdot 1 = \frac{2.4}{3}$

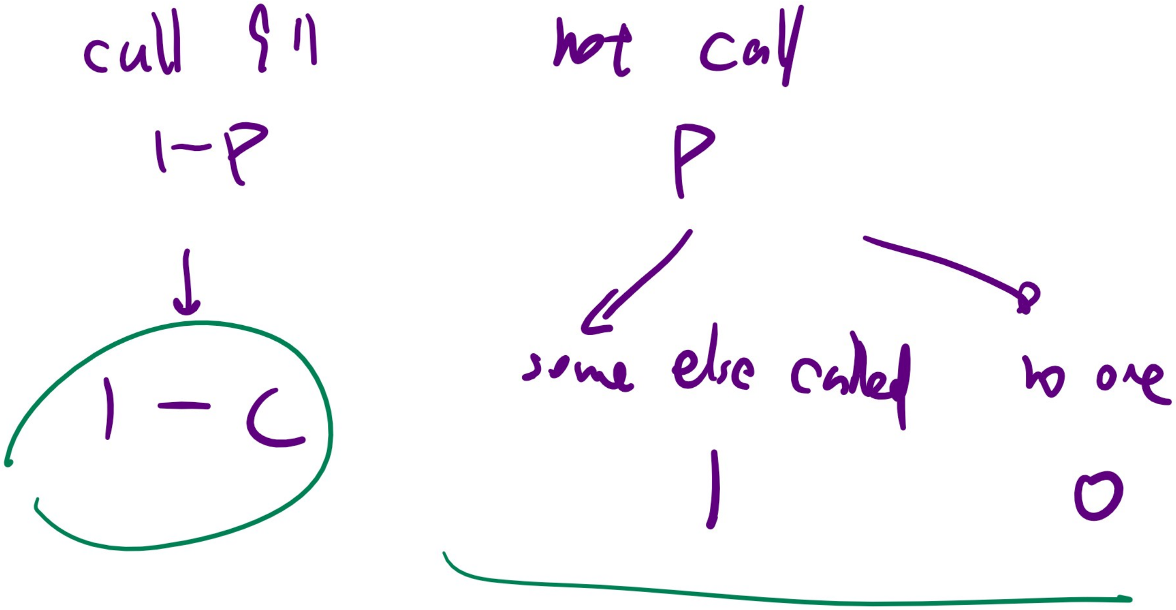
$R \to$ ______ $= \frac{2.4}{3}$

Volunteer's Dilemma, Part I

Quiz (Participation)

- On March 13, 1964, Kitty Genovese was stabbed outside the apartment building. There are 38 witnesses, and no one reported. Suppose the benefit of reported crime is 1 and the cost of reporting is $c < 1$. What is the probability that no one reported?

- A: c
- B: $c^{1/37}$
- [✓ C: $c^{38/37}$]
- D: $c^{1/38}$
- E: $c^{37/38}$



$$p^{38} = c^{38/37}$$

Volunteer's Dilemma, Part II

Quiz (Participation)

<u>call</u> not p

$$\boxed{1 - c = \overset{\text{no one called}}{\underline{p}^{37}} \cdot 0 + (1 - p^{37}) \cdot 1}$$

$$1 - c = 1 - p^{37}$$

$$c = p^{37} \qquad p = c^{1/37}$$

Public Good Game, Part I

Quiz (Participation)

- You received one free point for this question and you have two choices.
- A: Donate the point.
- B: Keep the point.
- Your final grade is the points you keep plus twice the average donation.

Public Good Game, Part II

Quiz (Participation)

Split or Steal Game

Quiz (Participation)

- Two players choose whether to split or steal a large sum of money, say x dollars. If both choose to split, each player gets $\frac{x}{2}$. If both choose to steal, each player gets 0. If one player chooses to steal, that player gets x. What is a pure strategy Nash equilibrium?
- A: (Split, Split)
- B: (Steal, Split)
- C: (Split, Steal)
- D: (Steal, Steal)

Rubinstein Bargaining Game, Part I

Quiz (Participation)

- There is a cake of size 1. Two kids bargain how to divide the cake for N rounds. The size of the cake is reduced to δ^t after t rounds of bargaining. In round t, if t is odd, kid 1 proposes the division, and kid 2 decides whether to accept or reject, and if t is even, kid 2 proposes the division, and kid 1 decides whether to accept or reject. The game ends when a proposal is accepted, and both kids get 0 if all proposals are rejected. How should the kid 1 propose in round 1? Assume kids accept when indifferent.

Rubinstein Bargaining Game, Part II

Quiz (Participation)

- How should the kid 1 propose in round 1 if $N = 2$? Assume kids accept when indifferent.
- A: $(1, 0)$
- B: $(1 - \delta, \delta)$
- C: $(1 - \delta + \delta^2, \delta - \delta^2)$
- D: $(1 - \delta + \delta^2 - \delta^3, \delta - \delta^2 + \delta^3)$
- E: $\left(\frac{1}{1-\delta}, \frac{\delta}{1-\delta}\right)$

Rubinstein Bargaining Game, Part III

Quiz (Participation)

- How should the kid 1 propose in round 1 if $N = 4$? Assume kids accept when indifferent.
- A: $(1, 0)$
- B: $(1 - \delta, \delta)$
- C: $(1 - \delta + \delta^2, \delta - \delta^2)$
- D: $(1 - \delta + \delta^2 - \delta^3, \delta - \delta^2 + \delta^3)$
- E: $\left(\frac{1}{1-\delta}, \frac{\delta}{1-\delta}\right)$

Rubinstein Bargaining Game, Part IV

Quiz (Participation)

- How should the kid 1 propose in round 1 if $N = \infty$? Assume kids accept when indifferent.
- A: $(1, 0)$
- B: $(1 - \delta, \delta)$
- C: $(1 - \delta + \delta^2, \delta - \delta^2)$
- D: $(1 - \delta + \delta^2 - \delta^3, \delta - \delta^2 + \delta^3)$
- E: $\left(\dfrac{1}{1-\delta}, \dfrac{\delta}{1-\delta}\right)$

Rubinstein Bargaining Game, Part V

Quiz (Participation)

First Price Auction, Version I

Quiz (Participation)

- If the value of an object to you is $v \in [0,1]$, how much should you bid for it in a first-price sealed-bid auction: simultaneous move, highest bidder gets the object and pays the highest bid? Suppose there are n bidders with values uniformly distributed in $[0,1]$.
- A: v
- B: $\frac{1}{2}v$
- C: $\frac{1}{n}v$
- D: $\frac{n-1}{n}v$

First Price Auction, Version II

Quiz (Participation)

Second Price Auction, Version I

Quiz (Participation)

- If the value of an object to you is $v \in [0, 1]$, how much should you bid for it in a second-price sealed-bid auction: simultaneous move, highest bidder gets the object and pays the second-highest bid? Suppose there are n bidders with values uniformly distributed in $[0, 1]$.

- A: v
- B: $\frac{1}{2}v$
- C: $\frac{1}{n}v$
- D: $\frac{n-1}{n}v$

Second Price Auction, Version II

Quiz (Participation)