Learning and Games

Price of Anarchy and Game Dynamics

Éva Tardos, Cornell

Learning and Games

Price of Anarchy and Game Dynamics

Lecture 1:

- What are games, and Nash equilibrium of simple games
- And what is learning

A few simple games:

Nash equilibrium of the game

Coordination:

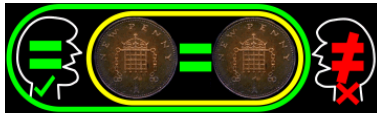
		M	S
1/2	M	2 / 2	0 / 0
1/2	S	0 / 0	1 / 1

Prisoner's dilemma:

	C	D
H	0 / 0	0.1 / -1.1
T	-1.1 / 0.1	-1 / -1

Matching pennies:

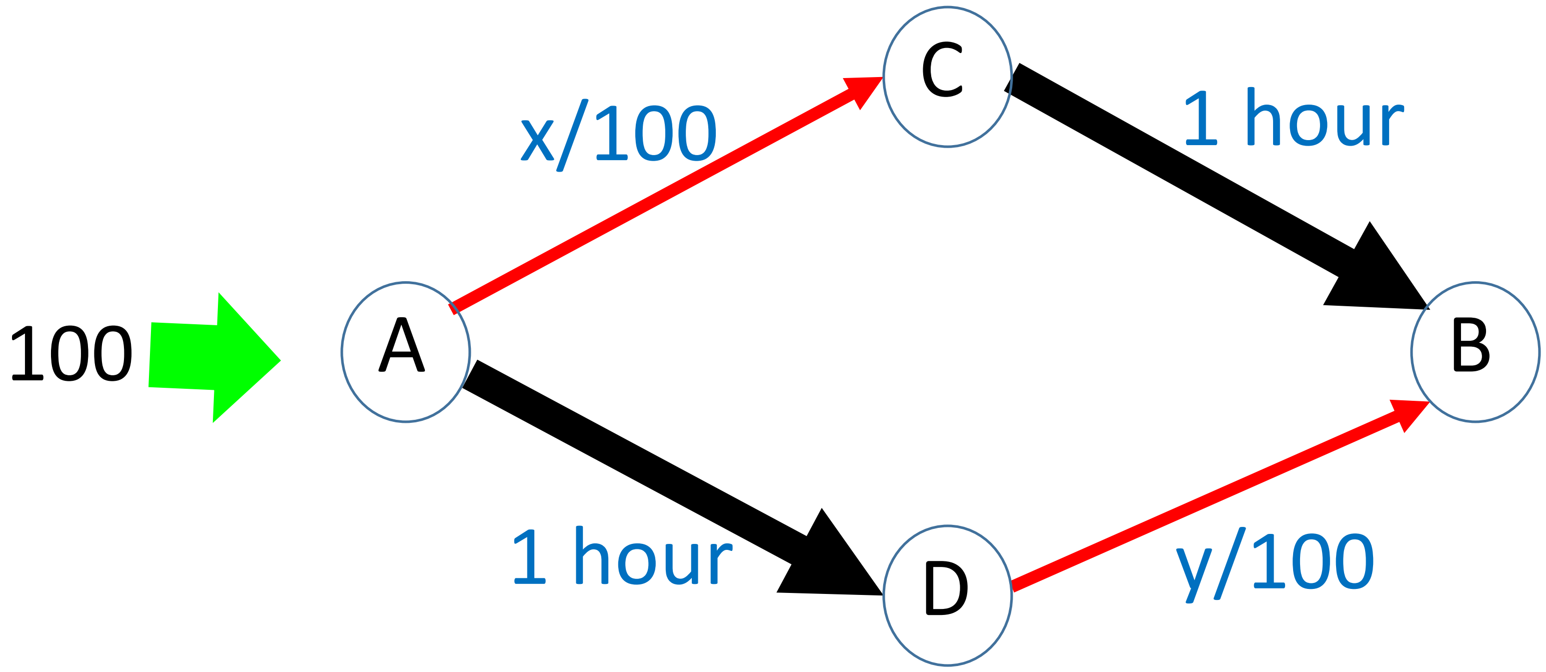
		H	T
1/2	H	-1 / 1	1 / -1
1/2	T	1 / -1	-1 / 1



Rock-Paper-Scissor:

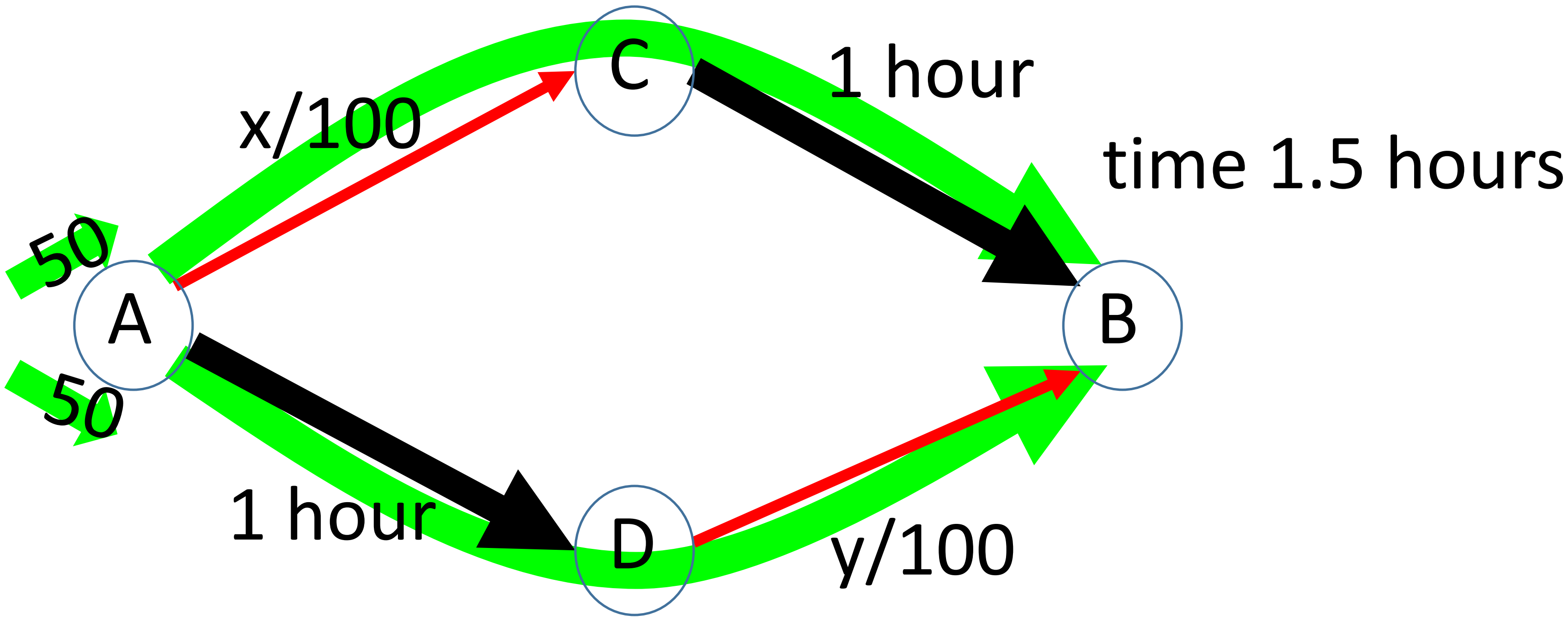
		1/6		5/6
		R	P	S
1/6	R	-9 / -9	1 / -1	-1 / 1
5/6	P	-1 / 1	-9 / -9	1 / -1
	S	1 / -1	-1 / 1	-9 / -9

Example: 100 travelers from A to B

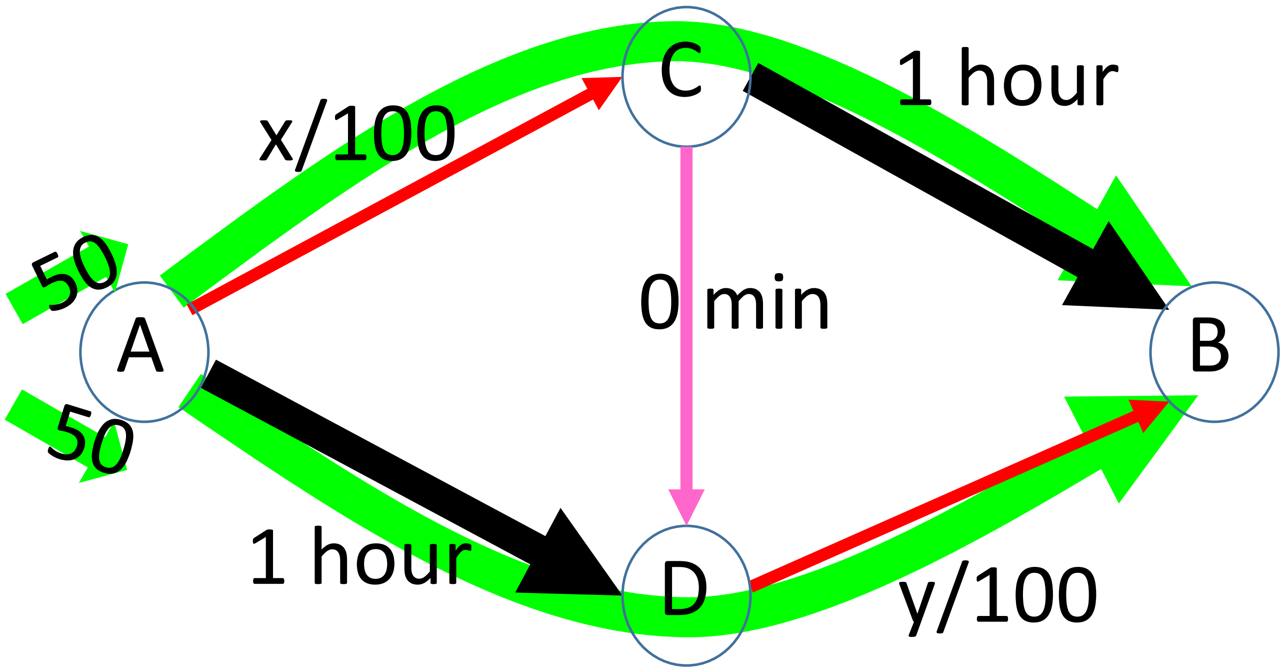


time as a function of congestion x or y

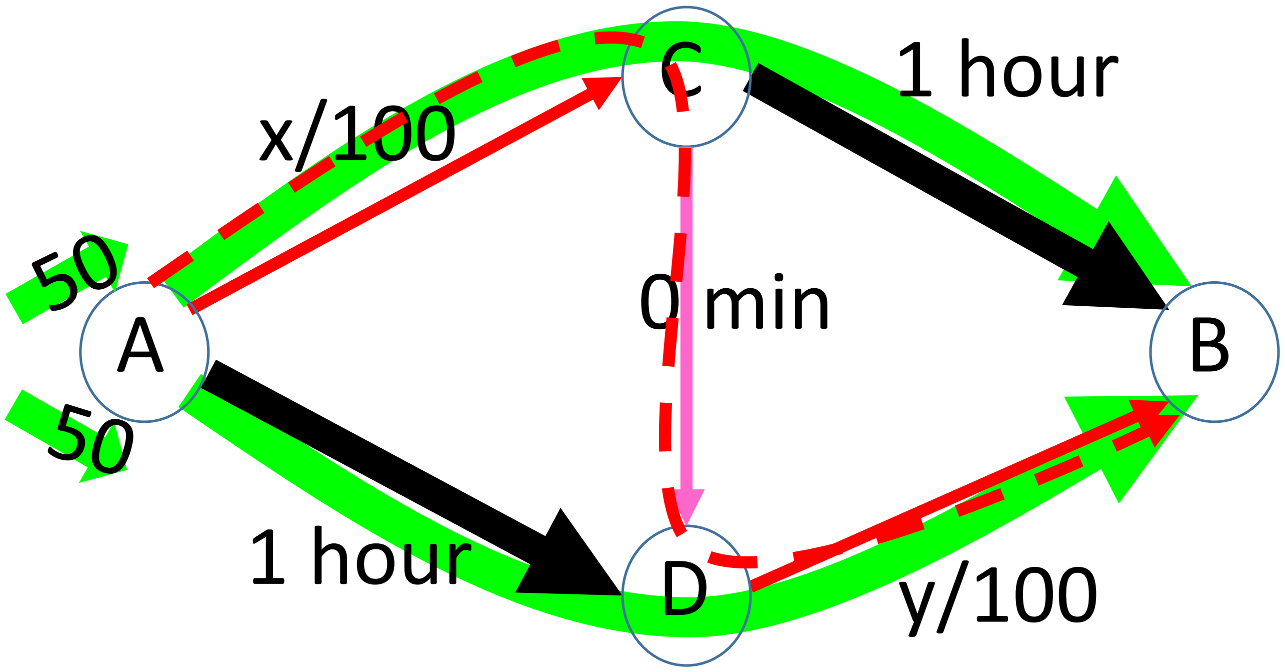
Example: flow equilibrium with 100 travelers



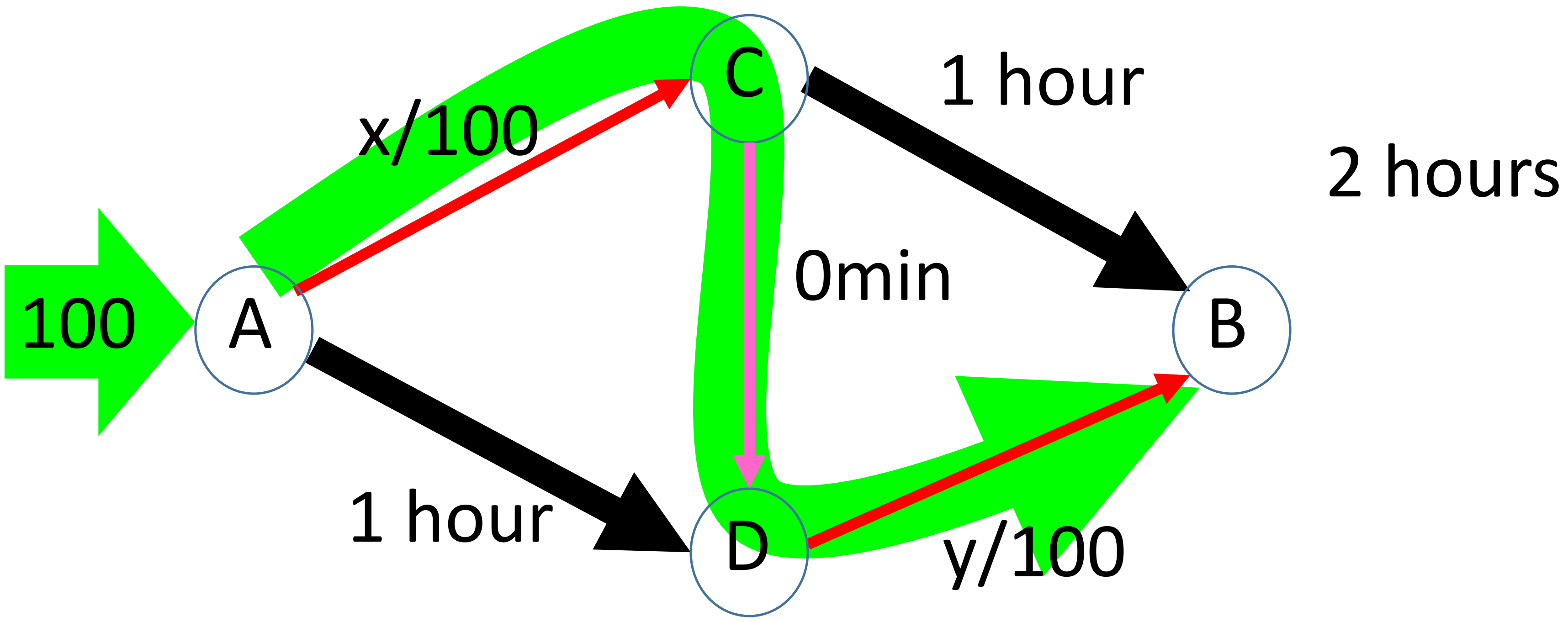
Add a new edge



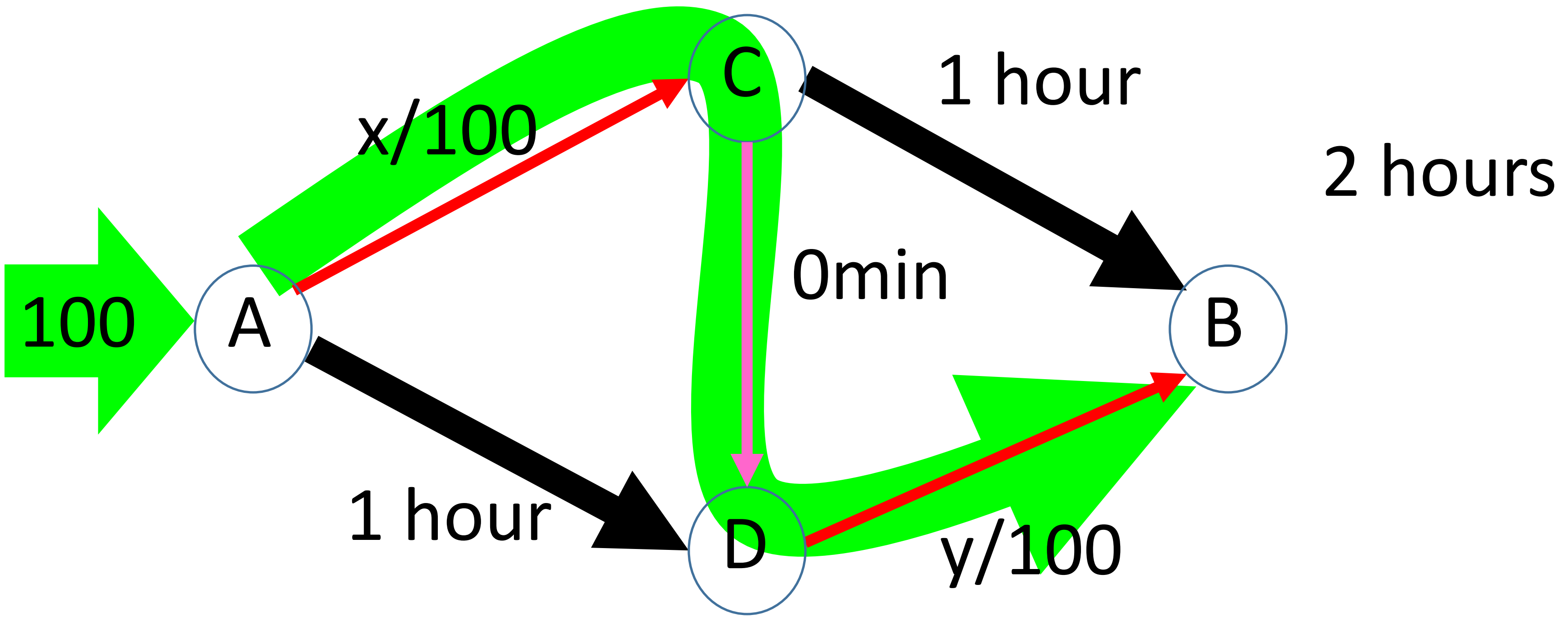
Not equilibrium!



Equilibrium

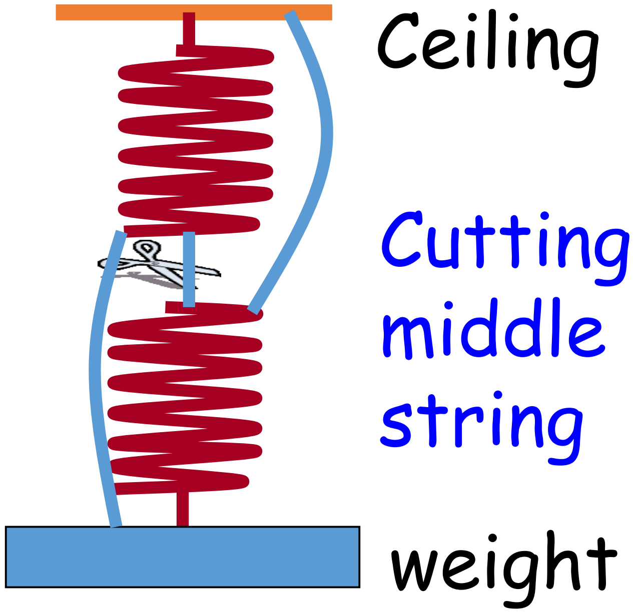


Braess' Paradox



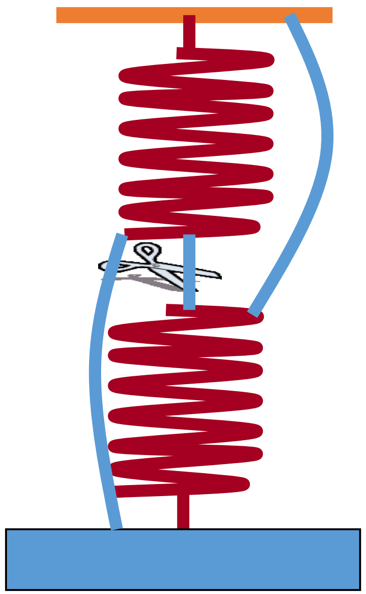
Paradox: players optimize their own flow, yet total not optimal?

Homework (optional)

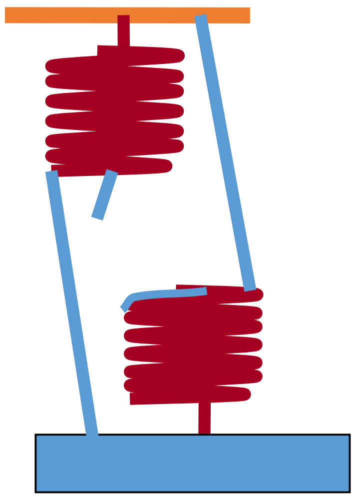


- What will happen to the weight? Goes up or down?
- And what does this have to do with what we talked about so far?

Braess paradox in springs (aside)



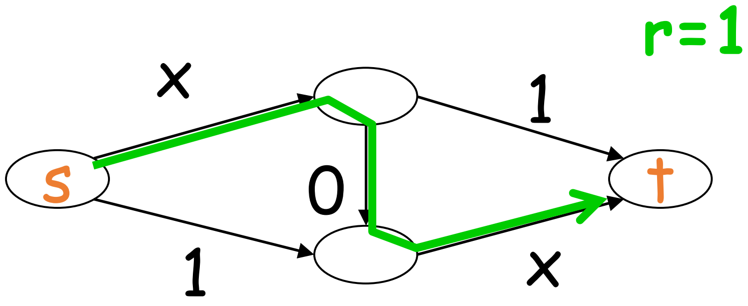
Cutting middle string



makes the weight rise

power flow along springs

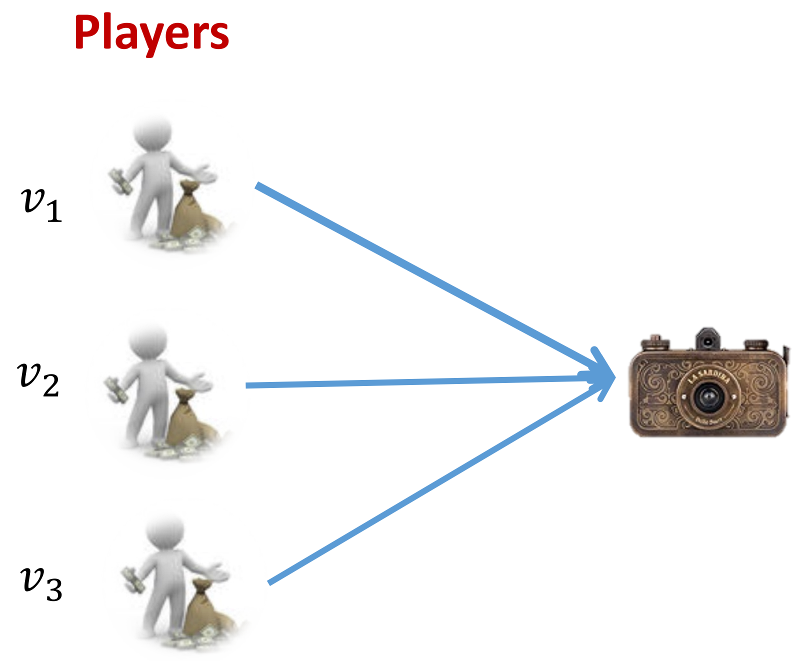
Flow=power; delay=distance



Single Item Auctions

- Second price = Vickrey auction
- First price
- All pay

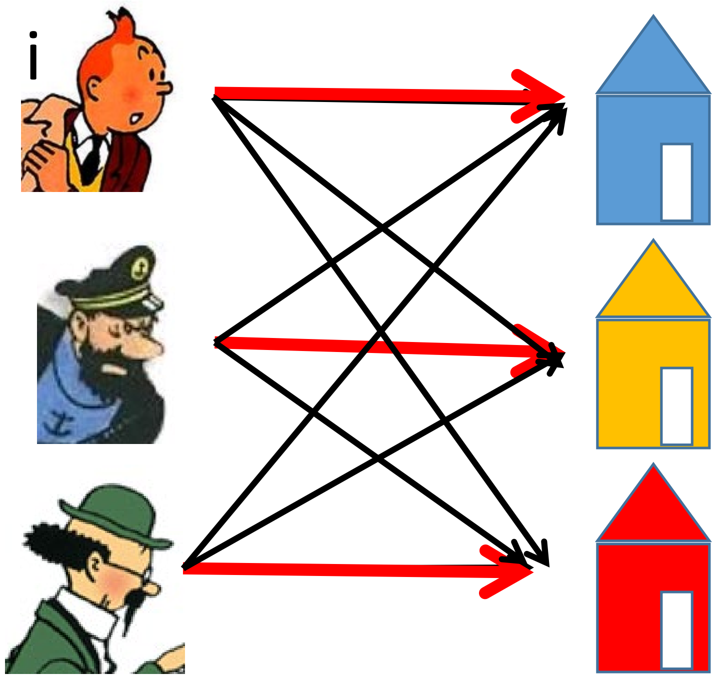
Or some mix of these



Winner is the bidder with highest bid.

Versions determine the payment.

Multiple items (e.g. unit demand bidders)



Value if i gets subset S is $v_i(S)$

for example: $v_i(S) = \max_{j \in S} v_{ij}$

Optimum is max value matching!

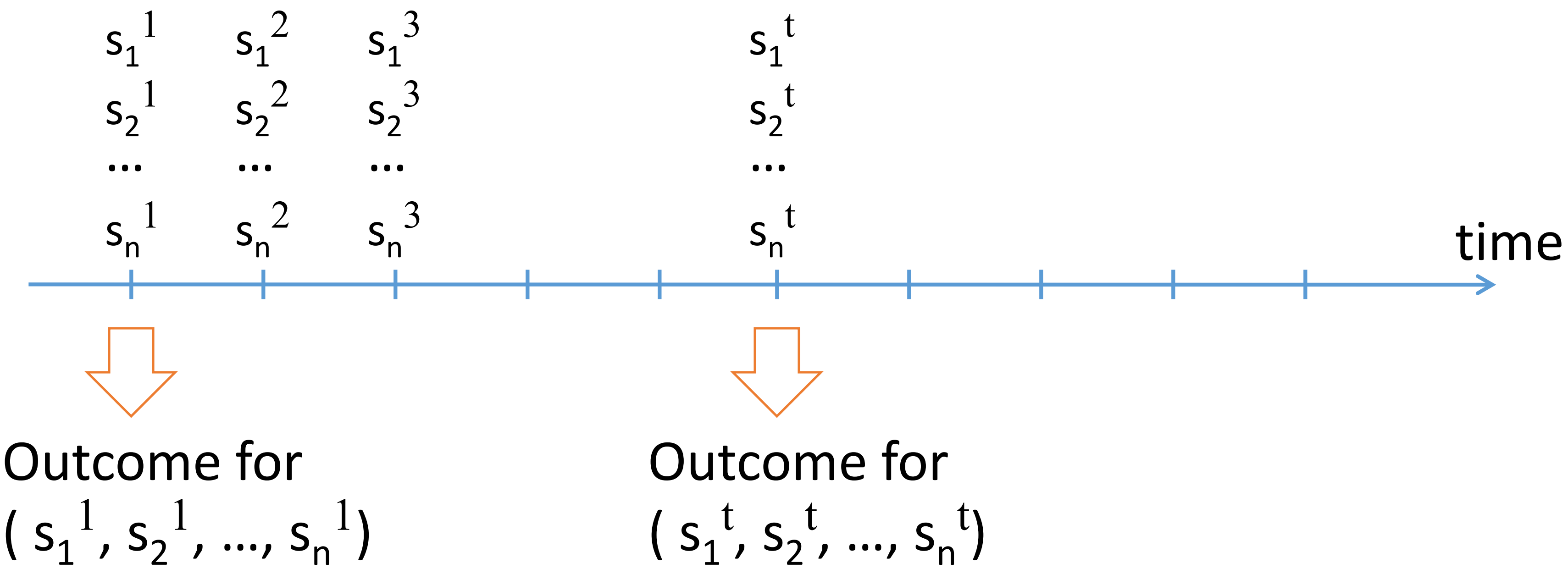
$$\max_{M^*} \sum_{ij \in M^*} v_{ij}$$

Extension also if $v_i(A)$ submodular function of set A

Also for diminishing value of added items:

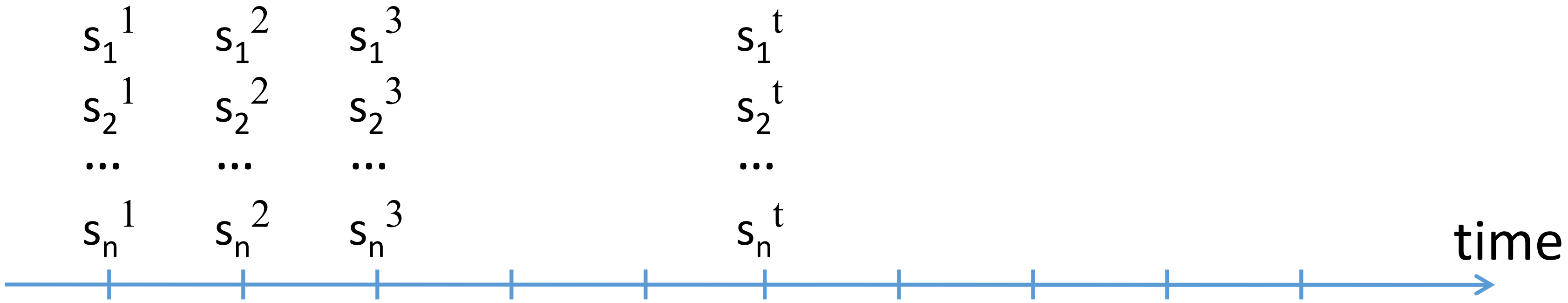
$$A \subset B \Rightarrow \; v_i(A + x) - v_i(A) \geq v_i(B + x) - v_i(B)$$

Repeated games



- Assume same game each period
- Player's value/cost additive over periods

Learning in games



Maybe here they don’t know how to play, who are the other players, …

By here they have a better idea…

Outcome of Learning in Repeated Game

- What is learning?
- Does learning lead to finding Nash equilibrium?

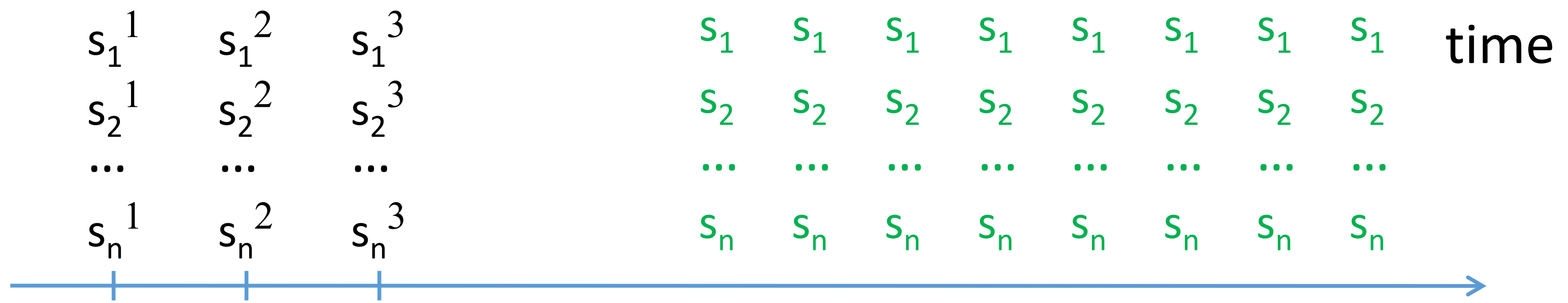
Brown’51 and Robinson’51:

- fictitious play = best respond to past history of other players: best response to assumption that the other player will choose a random strategy from the past uniformly.

Goal: “pre-play” as a way to learn to play Nash.

Robinson’51: Two-player 0-sum game, fictitious play does converge to Nash

Stable fictitious play: Nash equilibrium



Nash equilibrium: Stable actions s with no incentive to switch to any alternate strategy s_i':

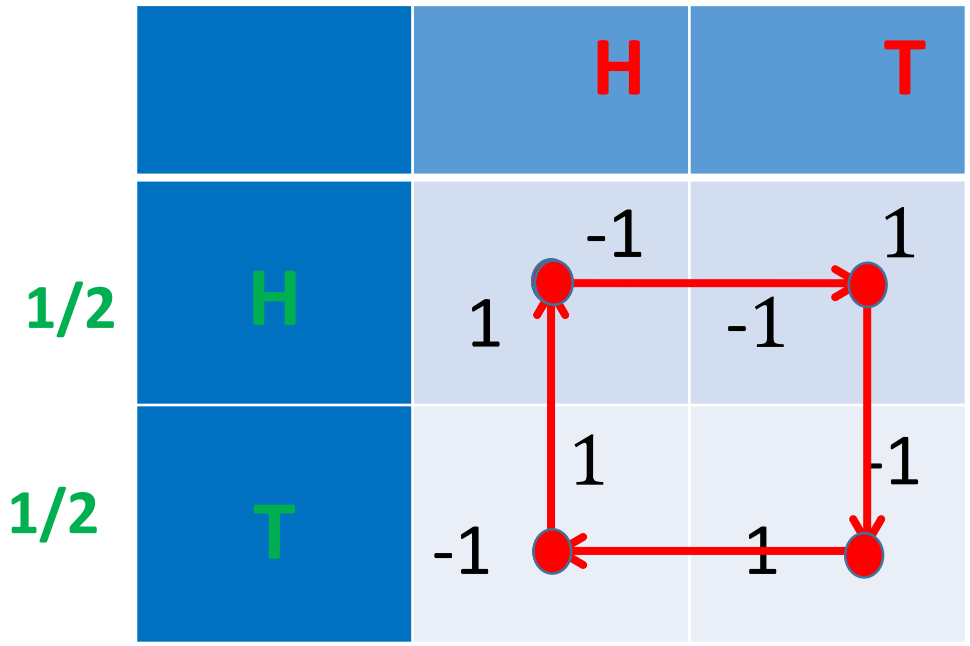
$$c_i(s_i', s_{-i}) \geq c_i(s)$$

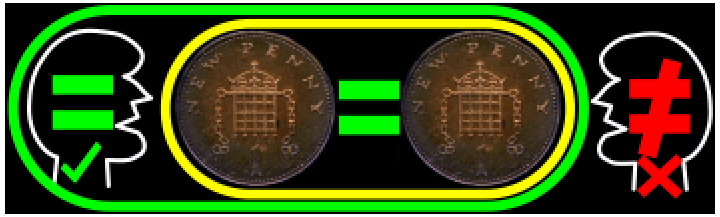
Cost for player i with action vector s

Cost for player i with action s_i' for i and s for all others

No regret

Fictitious play for Matching Pennies





G sees (H,T) history	R sees (H,T) history		resulting play
(0,0)	(0,2)	→	(H,H)
(1,0)	(1,2)	→	(H,H)
(2,0)	(2,2)	→	(H,T)
(2,1)	(3,2)	→	(H,T)
(2,2)	(4,2)	→	(T,T)

…

Result: Distribution is Nash

But cycles

Excercises:

If fictitious play converges (in the time average), does this imply that the outcome Nash?

a. Suppose fictitious play converges to strategy vector s. After a while each play i chooses a fixed pure strategy s_i. Prove that s is Nash. yes

b. Suppose in a 2 person game, the history of fictitious play of play i converges to a mix of σ_i (probability distribution of his strategies) for both players. Prove that the product of mixed strategies $\sigma_1 \times \sigma_2$ is a mixed Nash equilibrium. yes

c. Can you extend this to more players? Depends what we mean.

Fictitious play in coordination game

	A (1/2)	B (1/2)
A	1^+, 1	0, 0
B	0, 0	1, 1^+

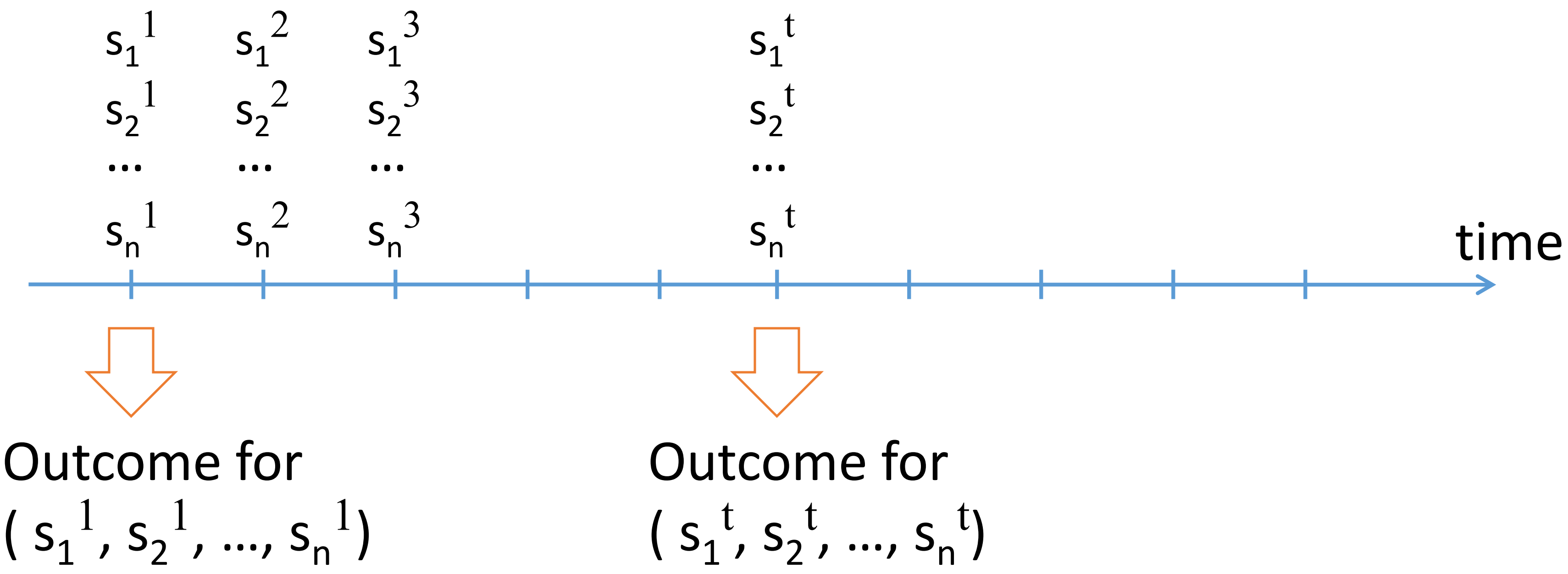
Start (A,B)

A sees	B sees		Play
(1,0)	(1,0)	→	(B,A)
(1,1)	(1,1)	→	(A,B)
(2,1)	(1,2)	→	(B,A)
…		…	

Theorem [Miyasawa'61]: Fictitious play distributions converges to Nash in 2-player 2 strategy games.

d. Suppose the mixed strategy vector σ both players (or all players). Does this imply that the distribution vector σ a Nash equilibrium? no

No-regret without stability: learning



All players i have not much incentive to switch to any fixed alternate strategy s_i':

In costs: $\sum_t c_i(s^t) \le \sum_t c_i(s_i', s_{-i}^t)$ +small regret

In values: $\sum_t v_i(s_i', s_{-i}^t) \le \sum_t v_i(s^t)$ + small regret

Fictitious play can have large regret!

	A (1/2)	B (1/2)
A	1^+, 1	0, 0
B	0, 0	1, 1^+

Start (A,B)

A sees	B sees		Play
(1,0)	(1,0)	→	(B,A)
(1,1)	(1,1)	→	(A,B)
(2,1)	(1,2)	→	(B,A)
…	…		

Resulting payoff for each play is 0!

Regret for player 1: $0 = \sum_{t=1}^{T} v_1(s^t) \ll \sum_{t=1}^{T} v_1(A, s_{-i}) = \frac{T}{2}$

Learning in Repeated Game 2

Smoothed fictitious play: randomize between similar payoffs.

- Fictitious play = best respond to past history of other player

$$argmin_x \sum_{\tau=1}^{t-1} c_i(x, s_{-i}^{\tau})$$

- Multiplicative weights: play prob. distribution $\sigma(x)$

$$argmin_\sigma \sum_{\tau=1}^{t} E_{x\sim\sigma}(c_i(x, s_{-i}^{\tau})) - \nu\, \mathrm{H}(\sigma)$$

where $\nu > 0$ and $H(\sigma) = -\sum_x \sigma(x) \log \sigma(x)$

- Follow the perturbed leader: chose a random r_x,

select $argmin_x[-r_x + \sum_{\tau=1}^{t-1} c_i(x, s_{-i}^{\tau})]$

Fictitious play and no regret

Fictitious play = best respond to past history of other players

$$s_i^t = argmin_x \sum_{\tau=1}^{t-1} c_i(x, s_{-i}^\tau)$$

Magic enhancement of Fictitious play with response included

$$s_i^t = argmin_x \sum_{\tau=1}^{t} c_i(x, s_{-i}^\tau)$$

Theorem 1: Magic fictitious play has no regret.

Proof: by induction we claim that

By choice of $\boldsymbol{s_i^t}$

$$\sum_{\tau=1}^{t} c_i(s^\tau) \le \sum_{\tau=1}^{t} c_i\left(\boldsymbol{s_i^t}, s_{-i}^\tau\right) \le \min_x \sum_{\tau=1}^{t} c_i(x, s_{-i}^\tau)$$

IH ⇓ with $x = s_i^t$

$$\sum_{\tau=1}^{t} c_i(s^\tau) = \sum_{\tau=1}^{t-1} c_i(s^\tau) + c_i(s^t) \le \sum_{\tau=1}^{t-1} c_i\left(s_i^t, s_{-i}^\tau\right) + c_i(s^t)$$

QED

Follow the perturbed leader has small regret (Theorem)

Follow the perturbed leader: chose a random r_x,

$$\text{select } argmin_x[-r_x + \sum_{\tau=1}^{t-1} c_i(x, s_{-i}^{\tau})]$$

Step 1: Magic Follow the perturbed leader has regret at most $\max_x r_x$

$$\text{select } argmin_x[-r_x + \sum_{\tau=1}^{t} c_i(x, s_{-i}^{\tau})]$$

Proof: as before

$$\sum_{\tau=1}^{t} c_i(s^{\tau}) - r_{s_i^1} \le \sum_{\tau=1}^{t} c_i(\boldsymbol{s_i^t}, s_{-i}^{\tau}) - r_{s_i^t} \le \min_x \sum_{\tau=1}^{t} c_i(x, s_{-i}^{\tau}) - r_x$$

IH

$$\sum_{\tau=1}^{t} c_i(s^{\tau}) - r_{s_i^1} = \sum_{\tau=1}^{t-1} c_i(s^{\tau}) - r_{s_i^1} + c_i(s^t) \le \sum_{\tau=1}^{t-1} c_i(s_i^t, s_{-i}^{\tau}) - r_{s_i^t} + c_i(s^t)$$

QED

Real follow the perturbed leader

Let r_x random: number of coins till you get H, if probability of H is ϵ

So $E(r_x) = \frac{1}{\epsilon}$ Also, for n strategies $E(\max_x r_x) = O(\frac{\log n}{\epsilon})$

Step 2: if $\max c_i(s) \leq 1$, then in any one step, the probability that magic perturbed follow the leader makes a different choice than real $\leq \epsilon$

Alternate way to flip the coins.

Start with r_x=1 all x

While more than one x possible

Take largest x, and flips its coin.
If H: x is eliminated.

When one x left: flip coins for x till H

If ≠H, then adding $c_i(x, s^t_{-i})$ or not makes no difference, prob=$1 - \epsilon$

Follow perturbed leader: small regret

Assuming we always follow magic version: regret at most $\max_x r_x$

- Expected value $E(\max_x r_x) \leq O(\frac{\log n}{\epsilon})$
- Cost from a step we don't follow the magic version at most 1

  So expected total cost of such steps at most ϵT
- Total regret at most

$$\sum_\tau^t c_i(s^t) \leq \min_x \sum_\tau^t c_i\left(x, s_i^t\right) + \epsilon T + O(\frac{\log n}{\epsilon})$$

Theorem: Select $\epsilon = \sqrt{\frac{\log n}{T}}$ then resulting regret at most $O(\sqrt{\mathrm{T} \log n})$

Exercise

Improved analysis of follow the perturbed leader

a. Dependence on T is very unfortunate: would much prefer bound of

$$\sum_{\tau}^{t} c_i(s^t) \leq (1+\epsilon)\min_x \sum_{\tau}^{t} c_i\left(x, s_i^t\right) + O(\frac{\log n}{\epsilon})$$

Is this also true?

b. when strategies are path s to t: there are exponentially many path! Can we add randomness r_e on the edges? And have $r_P = \sum_{e \in P} r_e$?

Smoothed fictitious play 2: Multiplicative weight?

- Multiplicative weights: play prob. distribution $\sigma(x)$

$$argmin_\sigma \sum_{\tau=1}^{t} E_{x\sim\sigma}(c_i(x, s_{-i}^{\tau})) - \nu H(\sigma)$$

where $\nu > 0$ and $H(\sigma) = -\sum_x \sigma(x) \log \sigma(x)$

Theorem: Multiplicative weight with rewards and $\alpha = 1 - \epsilon$ achieves (for a player with n strategies):

$$argmax_\sigma \sum_{\tau=1}^{t} E_{x\sim\sigma}(r_i(x, s_{-i}^{\tau})) + \nu H(\sigma)$$

Multiplicative weights (rewards)'

Reinforcement learning = reinforce actions that worked well in the past

sequence of play $s^1, s^2, \ldots, s^t$

Focus on player i:

Randomized strategy: weight/value of strategy x: w_x

probability of playing action x is p_x=$w_x / \sum_{s_i} w_{s_i}$

Update $w_x \leftarrow w_x \alpha^{c_i(x, s_{-i}^t)}$ for some $\alpha < 1$

Multiplicative weight update (MWU) or Hedge [Freund and Schapire'97]

Multiplicative weights and smoothed fictitious play

Theorem

- Smoothed fictitious play with entropy = Multiplicative weight update (with $\alpha = e^{-1/\nu}$)

Smoothed Fictitious Play:

$$argmix_{\sigma} \sum_t E_{x\sim\sigma}(c_i(x, s^t_{-i})) - \nu\, \mathrm{H}(\sigma)$$

Multiplicative weight:

probability of playing action x is $p_x = w_x / \sum_{s_i} w_{s_i}$

Update $w_x \leftarrow w_x \alpha^{c_i(x, s^t_{-i})}$

Proof:

Proof of equivalence (sketch)

Smoothed Fictitious Play:

$$argmin_\sigma \sum_t E_{x\sim\sigma}(c_i(x, s_{-i}^t)) - \nu\, \mathrm{H}(\sigma)$$

Let q_x probability of playing x, and use $C(x) = \sum_t c_i(x, s_{-i}^t)$

$$\min F(q) = \sum_x q_x C(x) - \nu q_x \ln q_x$$

Minimized when all partial derivatives are the same

$\Delta_{q_x}(F) = \mathrm{C(x)} - \nu \ln q_x - \nu$ so $\mathrm{C(x)}/\nu - \ln q_x = const$

So $q_x = \exp\left(\frac{C(x)}{\nu}\right) / \exp(\text{const}) = \alpha^{C(x)} * \exp(const)$

$\alpha = e^{-1/\nu}$

Detour: Multiplicative weight is no regret

- Use regards not costs with n strategies

$$\sum_\tau r_i(s^\tau) \geq (1-\epsilon)\max_x \sum_\tau r_i(x, s_{-i}^\tau) - \frac{\log n}{\epsilon}$$

- Assume $0 \leq r_i(s^\tau) \leq 1$
- Multiplicative weight
  - $p_x = w_x / \sum_{s_i} w_{s_i}$
  - Update $w_x \leftarrow w_x \alpha^{c_i(x, s_{-i}^t)}$ now $\alpha > 1$, e.g., $\alpha = \exp(1+\epsilon)$

Detour: Buy and Hold investment

W wealth, n stocks to invest in, with return rates $(1+\epsilon)^{r_i^t}$ period t with $0 \le r_i^t \le 1$

- All invested in stock i we get: $\mathrm{W_i(t)} = W \prod_t (1+\epsilon)^{r_i^t} = W(1+\epsilon)^{\sum_t r_i^t}$
- Invest equally and hold $(\frac{W}{n}, \dots, \frac{W}{n})$ and hold
- Resulting wealth: $\mathrm{W(t)} = \sum_i \frac{W}{n} \prod_t (1+\epsilon)^{r_i^t} = \frac{W}{n} \sum_i (1+\epsilon)^{\sum_t r_i^t}$

$$\ge \max_i \frac{W}{n} (1+\epsilon)^{\sum_t r_i^t}$$

We get $\log_{1+\epsilon} W(t) \ge \max_i \log_{1+\epsilon} W(1+\epsilon)^{\sum_t r_i^t} - \log_{1+\epsilon} n =$

$$log_{1+\epsilon}(max_i W_i(t)) - \log_{1+\epsilon} n$$

Buy and Hold investment $\Rightarrow$ learning

Connection: if W=1 and you use $x_1^t, \dots, x_n^t$ to invest at time t you get

$$\boxed{\log_{1+\epsilon} W'(t)} = \log_{1+\epsilon} \left(W'(t-1) \sum_i x_i^t (1+\epsilon)^{r_i^t}\right)$$

$(1+\epsilon)^r$ convex in r

$$= \log_{1+\epsilon} W'(t-1) + \log_{1+\epsilon} \sum_i x_i^t (1+\epsilon)^{r_i^t} \le \log_{1+\epsilon} W'(t-1) + \log_{1+\epsilon} \sum_i x_i^t (1+\epsilon r_i^t)$$

$$= \log_{1+\epsilon} W'(t-1) + \log_{1+\epsilon}(1+\epsilon \textstyle\sum_i x_i^t r_i^t) = \log_{1+\epsilon} W'(t-1) + \frac{\ln(1+\epsilon \sum_i x_i^t r_i^t)}{\ln(1+\epsilon)}$$

$$\le \log_{1+\epsilon} W'(t-1) + \frac{\epsilon \sum_i x_i^t r_i^t}{\ln(1+\epsilon)} \le \boxed{\frac{\epsilon}{\ln(1+\epsilon)} \sum_t \sum_i x_i^t r_i^t}$$

$1 + x \le e^x$

$\log_{1+\epsilon} W'$ is a lower bound on reward of learner!!

Buy and Hold investment ⇒ learning

Buy all and hold as a learning strategy, so we get

$$x_i^t = \frac{(1+\epsilon)^{\sum_t r_i^t}}{\sum_j (1+\epsilon)^{\sum_j r_j^t}}$$

From previous

Good investment

The result:

$$\sum_i \sum_t x_i^t r_i^t \geq \frac{ln(1+\epsilon)}{\epsilon} \log_{1+\epsilon} W(T) \geq \frac{\ln(1+\epsilon)}{\epsilon} (\max_i \log_{1+\epsilon} W_i(T) - log_{1+\epsilon} n)$$

$$= \frac{\ln(1+\epsilon)}{\epsilon} (\max_i \sum_t r_i^t - log_{1+\epsilon} n) \geq (1-\epsilon) \max_i \sum_i r_i^t - \frac{\ln n}{\epsilon}$$

Outcome with no-regret learning in games

Limit distribution σ of play (strategy vectors s=$(s_1, s_2, \dots, s_n)$)

- all players i have no regret for all strategies x

$$E_{s\sim\sigma}\big(c_i(s)\big) \le E_{s\sim\sigma}(c_i(x, s_{-i}))$$

Hart & Mas-Colell: Long term average play is (coarse) correlated equilibrium

Players update independently, but correlate on shared history

Correlated equilibrium vs Nash equilibrium

- No-regret learning $\rightarrow$ coarse correlated equilibrium exists. No need for the fixed point proof of Nash…
- Coarse correlated equilibria form a convex set!

  π_s: probability of strategy vector s

  $\pi_s \geq 0, \sum_s \pi_s = 1$

  $\sum_s \pi_s u_i(s) \geq \pi_s u_i(s'_i, s_{-i})$ for all i, $s'_i \in S_i$ (i has no regret)

  Poly time computable [Roughgarden-Papadimitriou'05, Jiang &Leyton-Brown'11]
- Correlated equilibrium where σ is a product distribution (players choose independently) is a Nash

Plan for today and going forward

- Today: outcome of learning in 0-sum games
- Next: outcome in of learning in congestion games and auctions
- Then: what was is learning better than Nash?

Exercises

1. If all players use one of our no-regret learning algorithms (with regret <<T (such as $O(\sqrt{T})$ or just $o(T)$)

and suppose distribution of the history of play converges to a fixed strategy vector σ.

Does this imply that the distribution vector σ a Nash equilibrium?

Yes: if players update independently, reacting to the same history: it most be product distribution

2. Can probability of play on Cooperate in Prisoner's dilemma remain >0 in a no-regret play?

No: C is a dominated by D: player would have regret if playing C

Correlated equilibrium vs Nash equilibrium

- No-regret learning → coarse correlated equilibrium exists. No need for the fixed point proof of Nash…
- Coarse correlated equilibria form a convex set!

  π_s: probability of strategy vector s

  $\pi_s \geq 0, \sum_s \pi_s = 1$

  $\sum_s \pi_s u_i(s) \geq \sum_s \pi_s u_i(s_i', s_{-i})$ for all i, $s_i' \in S_i$ (i has no regret)

  Poly time computable [Roughgarden-Papadimitriou'05, Jiang &Leyton-Brown'11]

- Correlated equilibrium where σ is a product distribution (players choose independently) is a Nash

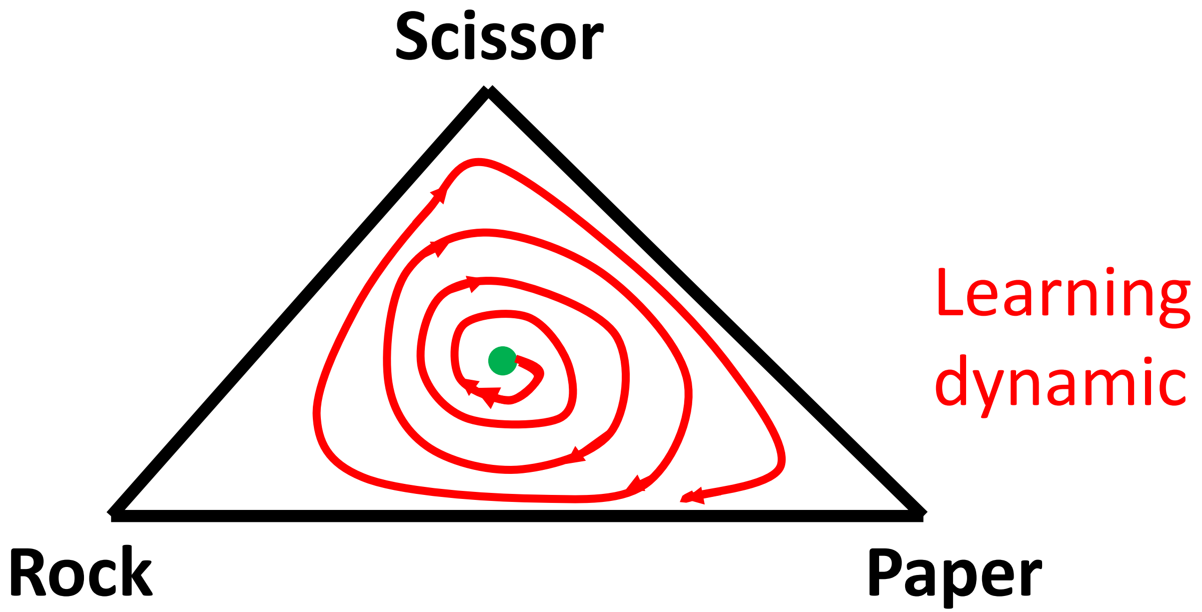
Simple example: rock-paper-scissor

	R	P	S
R	0 0	1 -1	-1 1
P	-1 1	0 0	1 -1
S	1 -1	-1 1	0 0

Nash equilibrium unique

mixed: $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ each

Dynamics of rock-paper-scissor (Shapley)



Nash: $\frac{1}{3}$ $\frac{1}{3}$ $\frac{1}{3}$

	R	P	S
R	-9, -9	-1, 1	1, -1
P	1, -1	-9, -9	-1, 1
S	-1, 1	1, -1	-9, -9

Payoffs/utility

- Doesn't converge
- correlates on shared history
- Payoff better than any Nash!

- Same also with regular RPS

Two person 0-sum games and no-regret learning

p_{xy} probability distribution that is a coarse correlated equilibrium.

- Payoff matrix A, then payoff is $\sum_{xy} p_{xy} A_{xy}$
- Value $\mathrm{v} = \sum_{xy} p_{xy} A_{xy}$

  same as Nash

Theorem: Marginal distributions $\mathbf{q_x} = \sum_{y} p_{xy}$ and $\mathbf{r_y} = \sum_{x} p_{xy}$ for a Nash

Note that we didn't claim: $\mathrm{p_{xy}} \neq \mathrm{q_x} r_y$

Two person 0-sum games (proof)

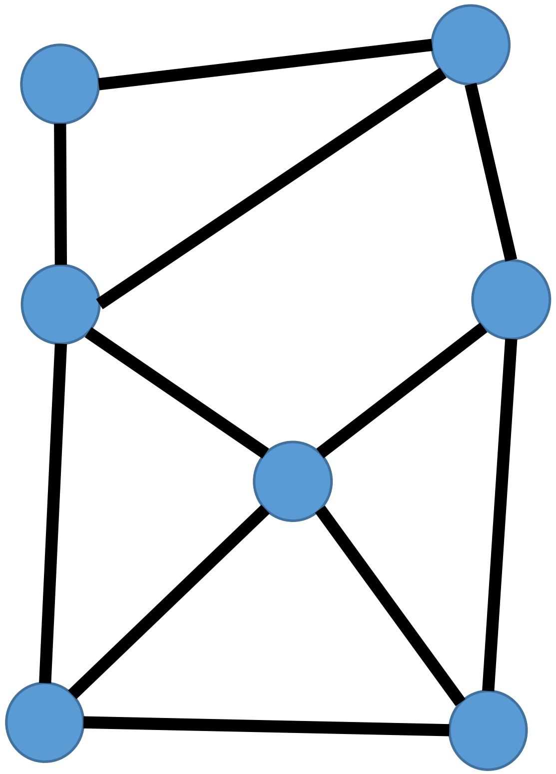
- Matrix A is first player's payoff, so with distribution p_{xy}
  - player 1 gets $\sum_{xy} p_{xy} A_{xy}$ =v
  - Player 2 gets $-\sum_{xy} p_{xy} A_{xy}$=-v
- Marginal distributions $\mathbf{q_x} = \sum_y p_{xy}$ and $\mathbf{r_y} = \sum_x p_{xy}$
- Player 1 has no regret: her value= $v \geq \max_x \sum_y A_{xy} \boldsymbol{r_y}$:

  player 1 getting her best response value to 2's marginal distribution!
- Player 2 has no regret: his loss= v $\leq \min_y \sum_x \boldsymbol{q_x} A_{xy}$
  - player 2 getting his best response value to 1's marginal distribution!

$$v \leq \min_y \sum_x q_x A_{xy} \quad \leq \sum_{xy} q_x A_{xy} r_y \quad \leq \max_x \sum_y A_{xy} p_x \quad \leq v$$

So $\boldsymbol{q}$ and $\boldsymbol{r}$ is Nash, and v is Nash value! … but $p_{xy} \neq \mathbf{r_y q_x}$

Extension to networked 0-sum games



- Two-player 0 sum game on each edge
- Nodes are players, need to play **same strategy** in each game

Theorem [Daskalakis-Papadimitriou ICALP'09] Nash for a convex set, no-regret play converges to Nash (projection to each player)

Proof idea: 2-person game: add RPS with payoff $\pm M$

Next time? Exercise?

No-regret learning as a behavioral model?

- Er'ev and Roth'96

  lab experiments with 2 person coordination game

- Fudenberg-Peysakhovich EC'14

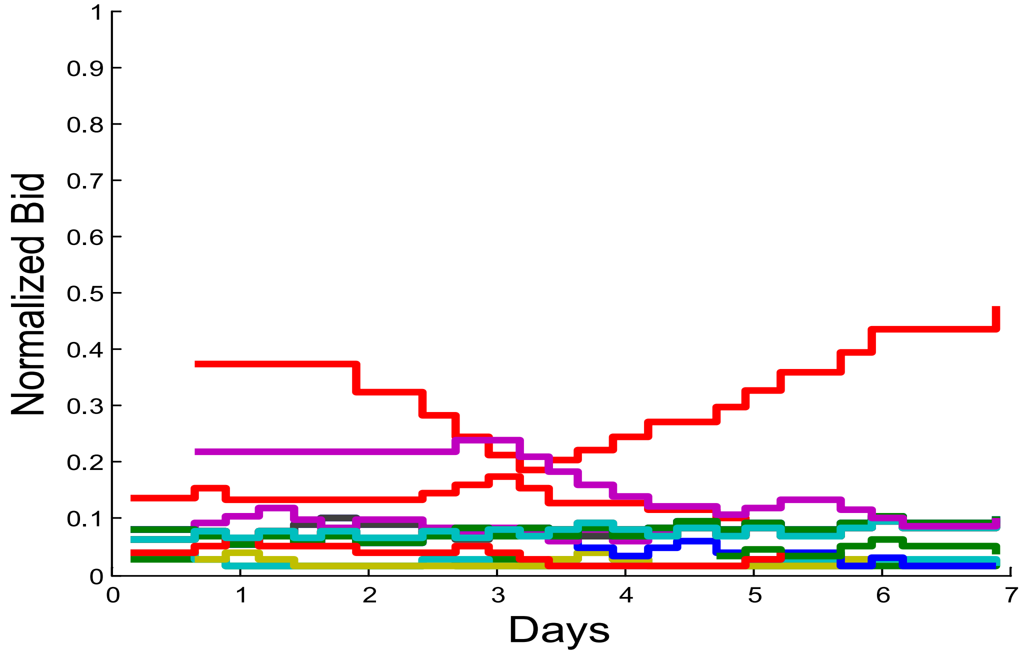
  lab experiments with seller-buyer game

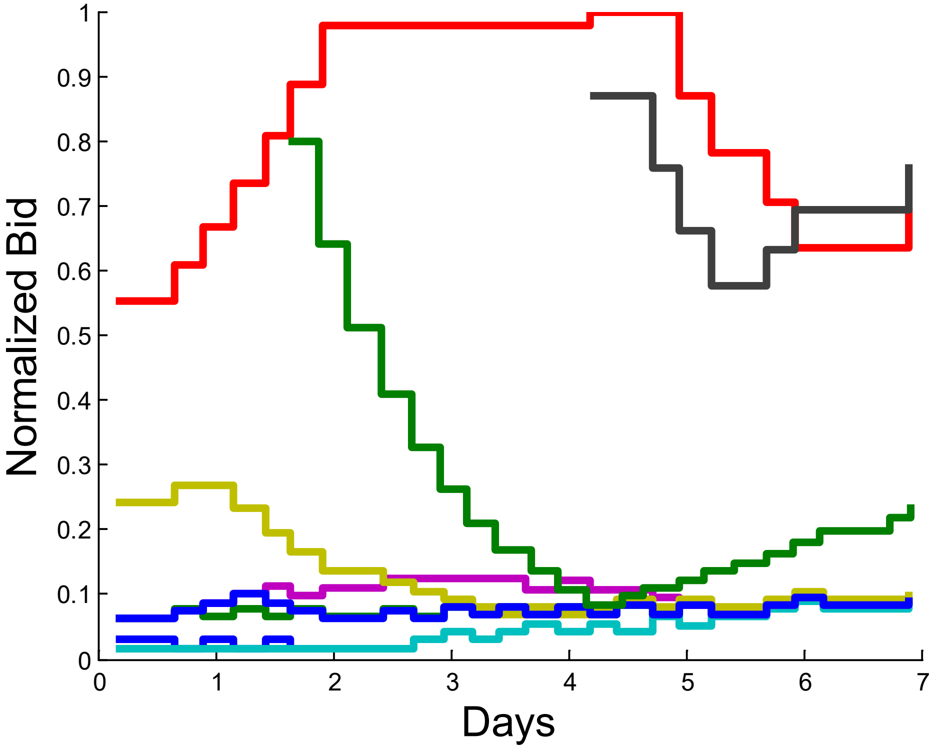
  recency biased learning

- Nekipelov-Syrgkanis-Tardos EC'15

  Bidding data on Bing-Ad-Auctions

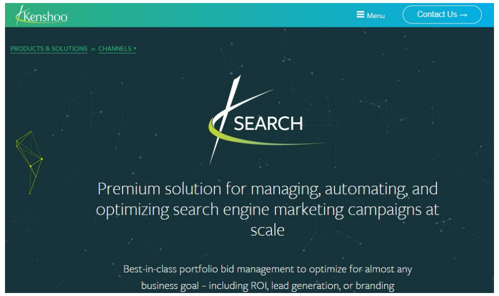
Behavior is far from stable

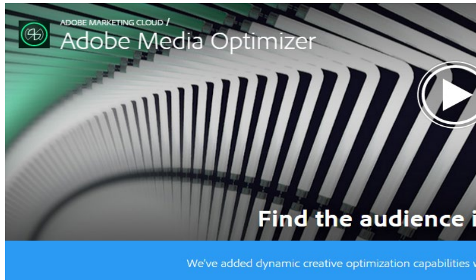




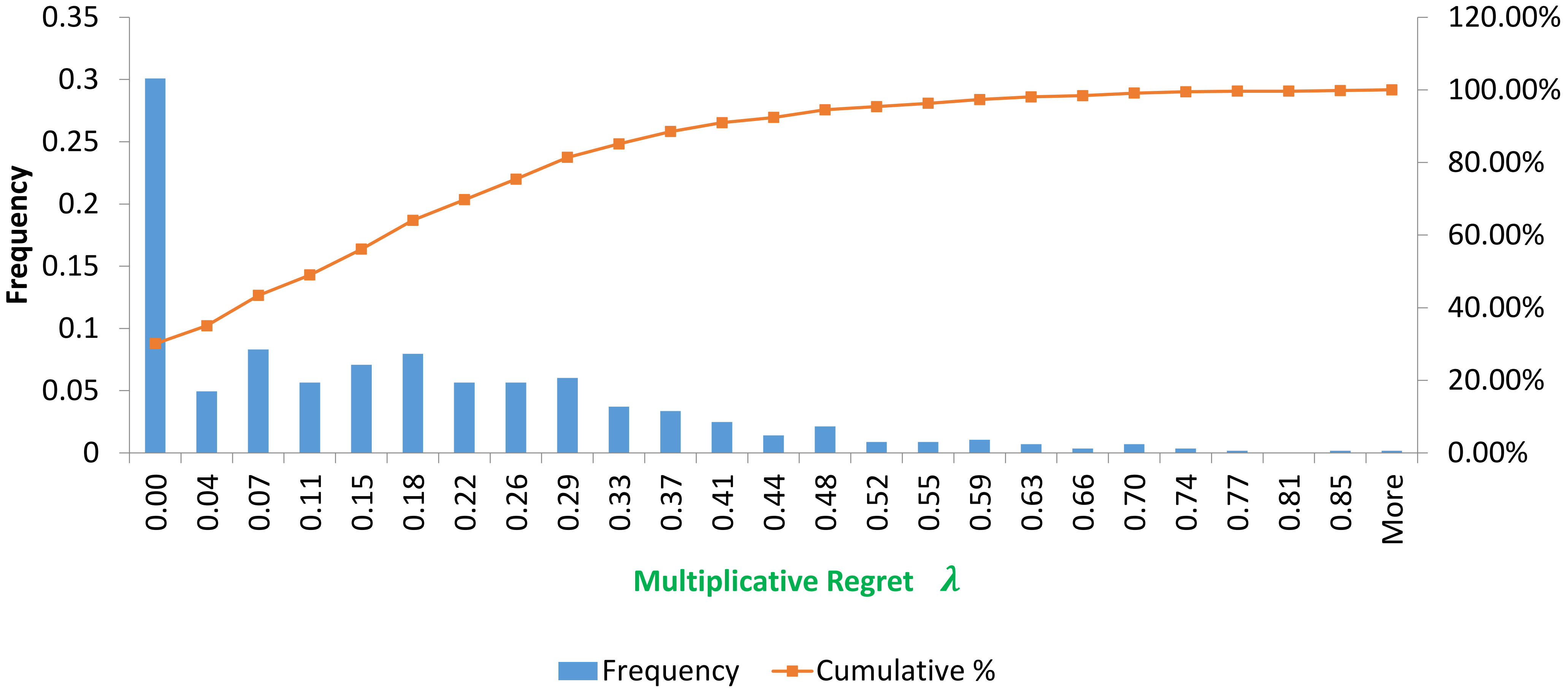
Bing search advertisement bid

Bidders use sophisticated bidding tools

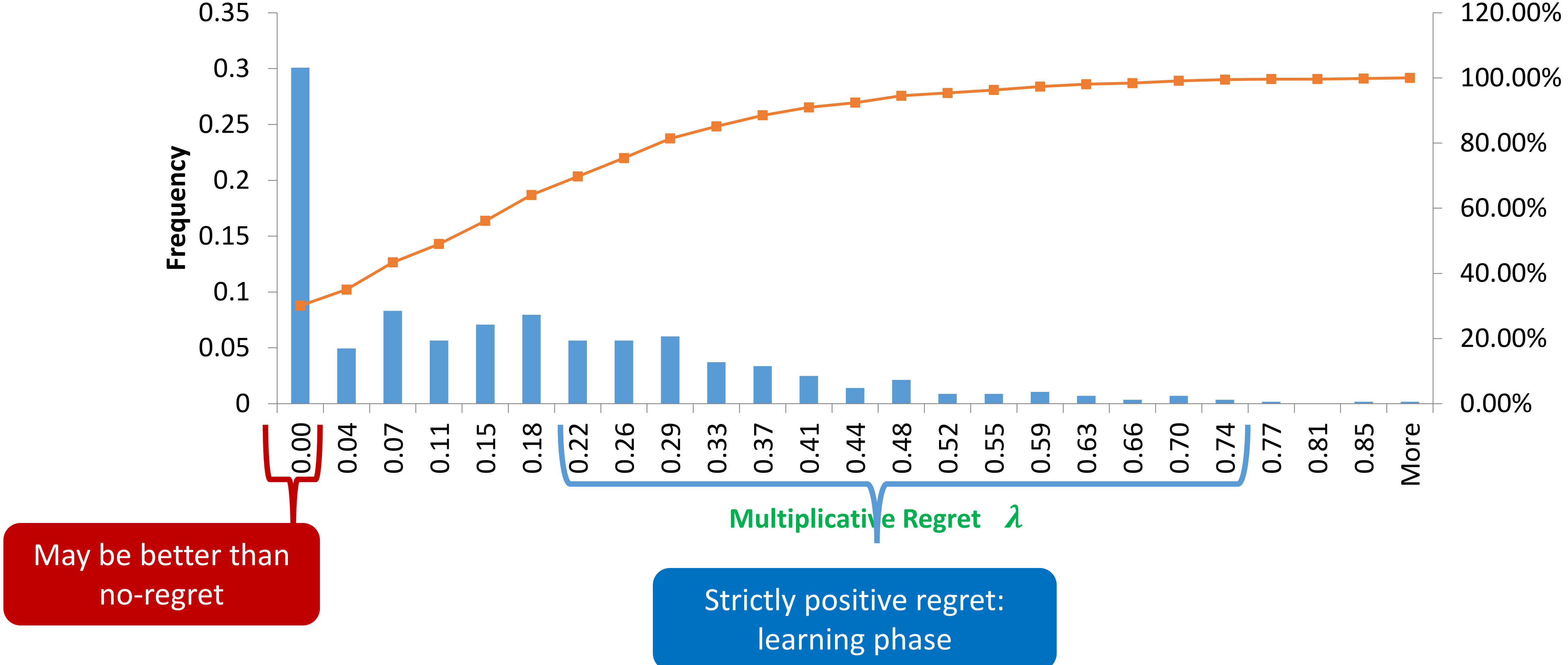




Distribution of smallest rationalizable multiplicative regret



Distribution of smallest rationalizable multiplicative regret



What can we say about learning outcome?

Limit distribution σ of play (strategy vectors s=$(s_1, s_2, \ldots, s_n)$)

- all players i have no regret for all strategies x

$$E_{s\sim\sigma}\left(c_i(s)\right) \leq E_{s\sim\sigma}(c_i(x, s_{-i}))$$

Hart & Mas-Colell: Long term average play is (coarse) correlated equilibrium

How good are coarse correlated equilibria??

Outcome of learning in games: cost minimization

- Finite set of players 1,…,n
- strategy sets S_i for player i:
- Resulting in strategy vector: s=$(s_1, \ldots, s_n)$ for each $s_i \in S_i$
- Cost of player i: $c_i(s)$ or $c_i(s_i, s_{-i})$

  Pure Nash equilibrium if $c_i(s) \leq c_i(s_i', s_{-i})$ for all players and all alternate strategies $s'_i \in S_i$

- Social welfare: $\text{cost}(s) = \sum_i c_i(s)$

  Optimum: $OPT = \min_s \sum_i c_i(s)$

Quality of Learning Outcome

Price of Anarchy [Koutsoupias-Papadimitriou'99]

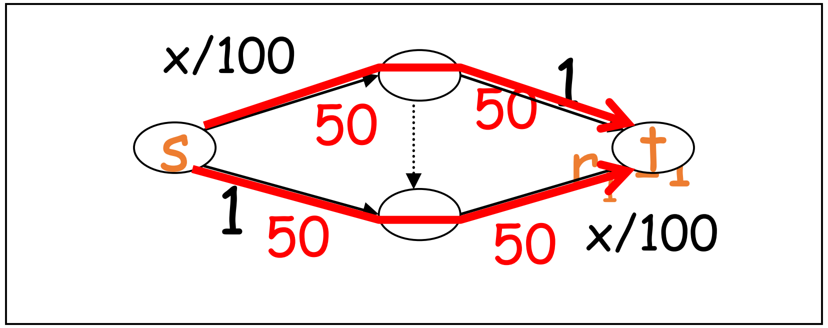
$$PoA = \max_{s\ Nash} \frac{cost(s)}{Opt}$$

Assuming **no-regret learners** in fixed game: [Blum, Hajiaghayi, Ligett, Roth'08, Roughgarden'09]

$$PoA = \lim_{T\to\infty} \frac{\sum_{t=1}^{T} cost(s^t)}{T\ Opt}$$

Example: Model of Routing Game

- A directed graph G = (V,E)
- source–sink pairs s_i,t_i for i=1,..,k

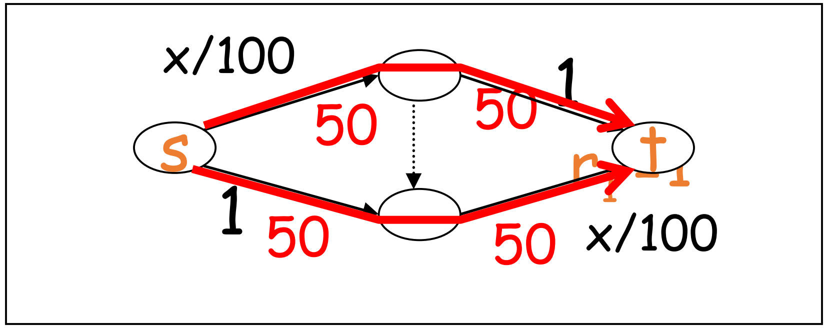


- Goal minimum delay:

  delay adds along path

  edge-cost/delay is a function $c_e(\cdot)$ of the load on the edge *e*

Delay Functions

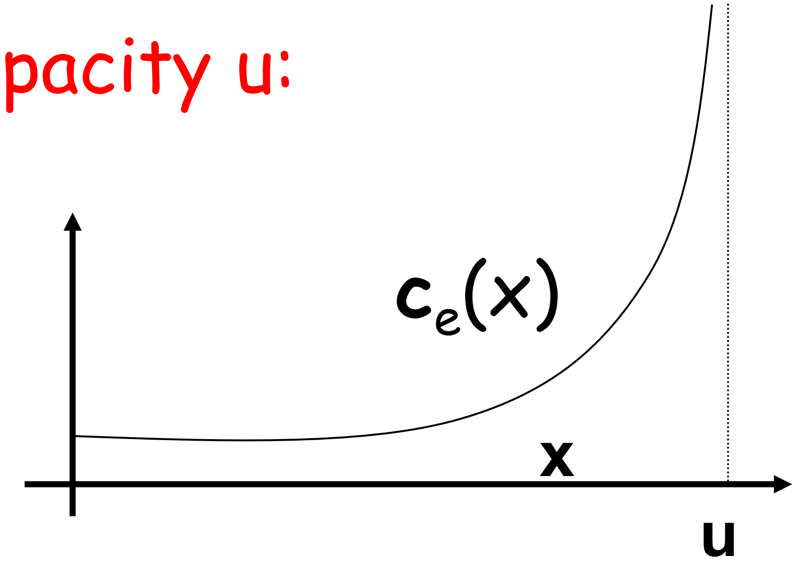


Assume $c_e(x)$ continuous and monotone increasing in load x on edge

No capacity of edges for now

Example to model capacity u:

$c_e(x) = a/(u-x)$



Goal's of the Game: min delay

Personal objective: minimize

$c_P(f)$ = sum of ~~delays~~ Costs of edges along **P** (wrt. flow **f**)

$$c_P(f) = \sum_{e \in P} c_e(f_e)$$

Overall objective:

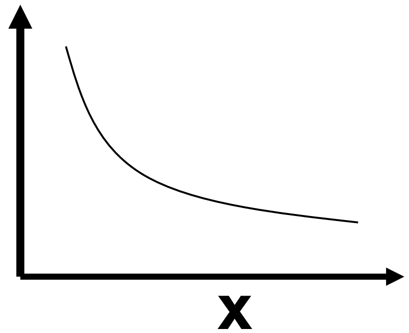
$C(f)$ = total ~~delay~~ Cost - of a flow **f**: $= \sum_P f_P \cdot c_P(f)$

= - social welfare

or total/average delay

Also:

$C(f) = \sum_e f_e \cdot c_e(f_e)$



Price of Anarchy: proof technique [Roughgarden'09]

- What we can work with:

$$\text{Optimum } s^* = (s_1^*, s_2^*, \dots, s_n^*)$$

$$\text{Nash: } s = (s_1, s_2, \dots, s_n)$$

- What we know:

$$c_i(s) \leq c_i(s_i', s_{-i}) \text{ for all i and all } s_i' \in S_i$$

Use it for all players and sum

$$c(s) = \sum_i c_i(s) \leq \sum_i c_i(s_i^*, s_{-i})$$

Proof smooth games

Nash property gave us (s is Nash, s* optimum)

$$c(s) = \sum_i c_i(s) \le \sum_i c_i(s_i^*, s_{-i})$$

Game is smooth if for some μ<1 and λ>0 and all s and s*

$$\sum_i c_i(s_i^*, s_{-i}) \le \lambda c(s^*) + \mu c(s)$$ (λ,μ)-smooth

If Opt <<cost(s), some player will want to deviate to s_i^*

Theorem: (λ,μ)-smooth game ⇒

Price of anarchy at most $\lambda/(1-\mu)$

Learning and price of anarchy (in smooth games)

Use approx no-regret learning:

$\sum_t c_i(s^t) \leq (1+\epsilon) \sum_t c_i(s_i^*, s_{-i}^t) + R$ for all players

A cost minimization game is (λ,μ)-smooth (λ > 0; μ< 1):

$\sum_t \sum_i c_i(s_i^*, s_{-i}^t) \leq \lambda \sum_t Opt + \mu \sum_t c(s^t)$

A approx. no-regret sequence s^t has

$$\frac{1}{T} \sum_t c(s^t) \leq \frac{(1+\epsilon)\lambda}{1-(1+\epsilon)\mu} \mathrm{Opt} + \frac{\mathrm{n}}{\mathrm{T}(1-(1+\epsilon)\mu)} \mathrm{R}$$

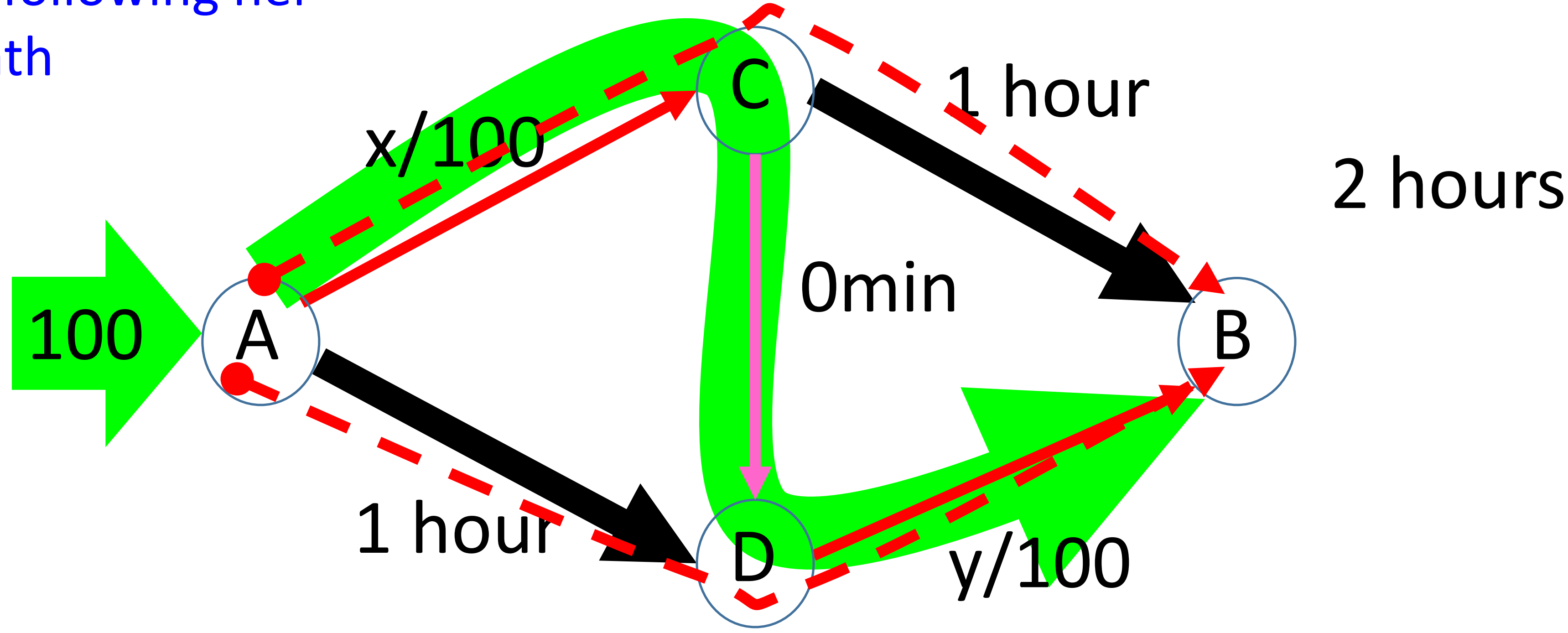
Note the convergence speed! $R = \frac{\log d}{\epsilon}$, so error

$$\frac{\mathrm{n}}{\mathrm{T}} \cdot \frac{log\ d}{\epsilon(1-(1+\epsilon)\mu)}$$

Foster, Li, Lykouris, Sridharan, T, NIPS'16

Equilibrium

Each use must not regret not following her optimal path



+1 flow: $f_e + 1$

No regret inequality for flow

- f_e Nash flow on edge e, P path used by Nash, Q path used by opt

No regret =

$$\sum_{e \in P} c_e(f_e) \le \sum_{e \in P \cap Q} c_e(f_e) + \sum_{e \in Q \setminus P} c_e(f_e + 1)$$

- Without the +1 nonatomic flow: assumes +1 is too small to really make a difference

easier to work with.... See more next time