RN, Chapter 17.6– 17.7

Decisions with Multiple Agents: Game Theory & Mechanism Design

Thanks to R Holte

Decision Theoretic Agents

- Introduction to Probability [Ch13]
- Belief networks [Ch14]
- Dynamic Belief Networks [Ch15]
- Single Decision [Ch16]
- Sequential Decisions [Ch17]
- Game Theory + Mechanism Design [Ch17.6 – 17.7]

Outline

- Game Theory
  - Motivation: Multiple agents
  - Dominant Action
  - Strategy
  - Prisoner's Dilemma
    - Domain Strategy Equilibrium; Paretto Optimum; Nash Equilibrium
  - Mixed Strategy (Mixed Nash Equilibrium)
  - Iterated Games
- Mechanism Design
  - Tragedy of the Commons
  - Auctions
  - Price of Anarchy
  - Combinatorial Auctions

Framework

- Make decisions in Uncertain Environments
  So far: due to “random” (benign) events
- What if **due to OTHER AGENTS ?**
- Alternating move, complete information, . . .
  $\Rightarrow$ 2-player games
  (use minimax, alpha-beta, … to find optimal moves)
- But
  - simultaneous moves
  - partial information
  - stochastic outcomes
- Relates to
  - auctions (frequency spectrum, . . .)
  - product development / pricing decisions
  - national defense

  Billions of $$s, 100,000's of lives, . . .

Simple Situation

1. Candy is worth \$5 to Buyer
2. Candy costs Seller \$1.50 to make
3. "Discount" only if Buyer puts name on mailing list... automatically giving Seller \$0.10, even if no sale

- Two players: *Buyer*, *Seller*
  - *Seller:* discount (ML + ask \$2) or fullPrice (ask \$4)
  - *Buyer:* yes or no

	Buyer: yes	*Buyer:* no
Seller: discount	B=3; S=0.6	B=0; S=0.1
Seller: fullPrice	B=1; S=2.5	B=0; S=0.0

- What should *Buyer* do?
  Seller is either discount or fullPrice
  - If *Seller:*discount, then
    *Buyer:*yes is better (3 vs 0)
  - If *Seller:*fullPrice, then
    *Buyer:*yes is better (1 vs 0)

  So clearly *Buyer* should play yes !
  ... For *Buyer*, yes dominates no

Simple Situation, con't

	Buyer: yes	*Buyer:* no
Seller: discount	B=3; S=0.6	B=0; S=0.1
Seller: fullPrice	B=1; S=2.5	B=0; S=0.0

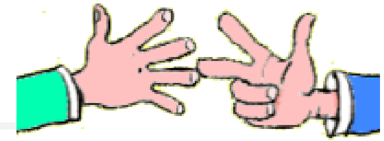
- What should *Seller* do?
  As *Buyer* will play yes, either
  - *Seller*:discount ⇒ 0.6
  - *Seller*:fullPrice ⇒ 2.5

  So *Seller* should play fullPrice

- Note: If *Buyer*:no, then
  Seller should play discount : 0.1 vs 0.0
  ... so what... NOT going to happen!

- Not “zero-sum" game
- Usually not so easy ...

Two-Finger Morra



- Two players: *O, E*
  - *O* plays 1 or 2
  - *E* plays 1 or 2

  simultaneously

- Let f = O+E be TOTAL #
- If f is $\begin{Bmatrix}\text{odd}\\ \text{even}\end{Bmatrix}$, then $\begin{Bmatrix}\text{O}\\ \text{E}\end{Bmatrix}$ collects *$\$f$* from other

  aka Inspection Game; Matching Pennies; . . .
- Payoff matrix:

	O: one	*O:* two
E: one	E=2; O=-2	E=-3; O=3
E: two	E=-3; O=3	E=4; O=-4

- What should *E* do? ... *O* do?

  No fixed single-action works ...

Player Strategy

	O: one	*O:* two
E: one	E=2; O=-2	E=-3; O=3
E: two	E=-3; O=3	E=4; O=-4

- Pure Strategy $\Rightarrow$ deterministic action
  - Eg, *O* plays two
- Mixed Strategy
  - Eg, [0.3 : one; 0.7 : two]
- Strategy Profile $\equiv$ strategy of EACH player
  - Eg,

$$\left\langle \begin{matrix} O & [0.3: one;\ 0.7: two] \\ E & [0.9: one;\ 0.1: two] \end{matrix} \right\rangle$$

- 0-sum game:
  - Player#1's gain = Player#2's loss
  - Not always true... *Buyer/Seller*!

    Sometimes. . .
    - single action-pair can BENEFIT BOTH, or
    - single action-pair can HURT BOTH !

	Buyer: yes	*Buyer:* no
Seller: discount	B=3; S=0.6	B=0; S=0.1
Seller: fullprice	B=1; S=2.5	B=0; S=0.0

Notes on Framework

- In *Seller/Buyer*:

	Buyer: yes	*Buyer:* no
Seller: discount	B=3; S=0.6	B=0; S=0.1
Seller: fullprice	B=1; S=2.5	B=0; S=0.0

FIXED STRATEGY is optimal:

$$\left\langle \begin{array}{ll} Buyer & [1.0 : yes;\ 0.0 : no] \\ Seller & [0.0 : discount;\ 1.0 : full\,\Pr ice] \end{array} \right\rangle$$

- Can eliminate any row that is DOMINATED by another, for each player
- No FIXED STRATEGY is optimal for Morra:

	O: one	*O:* two
E: one	E=2; O=-2	E=-3; O=3
E: two	E=-3; O=3	E=4; O=-4

- Can have >2 options for each player
- Different action sets, for different players

Prisoner's Dilemma

- Alice, Bob arrested for burglary
  ... interrogated separately
  - If BOTH testify: *A, B* each get -5 (5 years)
  - If BOTH refuse: *A, B* each get -1
  - If *A* testifies but *B* refuses: *A* gets 0, *B* gets -10
  - If *B* testifies but *A* refuses: *B* gets 0, *A* gets -10

	A: testify	*A:* refuse
B: testify	*A*= -5; *B*= -5	*A*= -10; *B*= 0
B: refuse	*A*= 0; B= -10	*A*= -1; *B*= -1

- Price of oil in Oil Cartel
  Disarming around the world
  ...

Prisoner's Dilemma, con't

	A: testify	*A:* refuse
B: testify	*A*= -5; *B*= -5	*A*= -10; *B*=0
B: refuse	*A*=0; B= -10	*A*= -1; *B*= -1

- What should *A* do?
  B is either testify or refuse
  - If *B*:testify, then
    A:testify is better (-5 vs -10)
  - If *B*:refuse, then
    A:testify is better (0 vs -1)

  So clearly *A* should play testify !
  $\Rightarrow$ testify is DOMINANT strategy (for *A*)

- What about *B* ?

Prisoner's Dilemma, III

	A: testify	*A:* refuse
B: testify	*A*= -5; *B*= -5	*A*= -10; *B*=0
B: refuse	*A*=0; B= -10	*A*= -1; *B*= -1

- What should *B* do?
  Clearly *B* show testify also (same argument)

- So ⟨ *A* : testify; *B* : testify ⟩
  is **Dominant Strategy Equilibrium**
  w/payoff: *A* = -5, *B* = -5

- ... but consider ⟨ *A* : refuse; *B* : refuse ⟩
  Payoff *A* = -1, *B*= -1 is better for BOTH!
  - jointly preferred outcome occurs when each chooses individually worse strategy

Why not ⟨ *A*:refuse, *B*:refuse ⟩?

- ⟨ *A*:refuse, *B*:refuse ⟩ is not "equilibrium":
  if *A* knows that *B*:refuse, then *A*:testify !
  (payoff ⟨0 , -10 ⟩, not ⟨-5, -5 ⟩)
  Ie, player *A* has incentive to change!
- Strategy profile S is **Nash equilibrium** iff
  ∀ player P,
  P would do worse if deviated from S[P],
  when all other players follow S
- **Thrm: Every game has ≥ 1 Nash Equilibrium !**
- Every dominant strategy equilibrium is Nash
  but … ∃ Nash Equil. even if no dominant!
  ... i.e., ∃ rational strategies even if no dominant strategy!

Pareto Optimal

	A: testify	*A:* refuse
B: testify	*A*= -5; *B*= -5	*A*= -10; *B*=0
B: refuse	*A*=0; B= -10	*A*= -1; *B*= -1

- $\langle$ *A* : refuse; *B* : refuse $\rangle$ is **Pareto Optimal** as
  $\neg\exists$ strategy where
  - $\geq$ 1 players do better,
  - 0 players do worse
- $\langle$ *A* : testify; *B* : testify $\rangle$ is NOT Pareto Optimal

DVD vs CD

Example with
no dominant strategies…

- Acme: video game Hardware
  Best: video game Software
- Both WIN if both use DVD
  Both WIN if both use CD

	A: dvd	*A:* cd
B: dvd	*A*= 9; *B*= 9	*A*= -4; *B*=-1
B: cd	*A*=-3; B= -1	*A*= 5; *B*= 5

- NO dominant strategies
- 2 Nash Equilibria: 〈 dvd, dvd 〉, 〈 cd, cd 〉
  (If 〈 dvd, dvd 〉 and *A* switches to cd, then *A* will suffer…)
- Which Nash Equilibrium?
  - Prefer 〈 dvd, dvd 〉 as Pareto Optimal
    (payoff 〈 A = 9; B = 9 〉 better than
    〈 cd, cd〉, w/ 〈 A = 5; B = 5 〉)
  - … but sometimes ≥ 1 Pareto Optimal Nash Equilibrium…

?Pure? Nash Equilibrium

- Morra

	O: one	*O:* two
E: one	E=2; O= -2	E=-3; O=3
E: two	E= -3; O=3	E=4; O=-4

- No PURE strategy

  (else *O* could predict *E*, and beat it)

- Thrm [von Neumann, 1928]:
  For every 2-player, 0-sum game,
  ∃ OPTIMAL mixed strategy

- Let U(e, o) be payoff to *E* if *E*:e, *O*:o

  (So *E* is maximizing, *O* is minimizing)

Mixed Nash Equilibrium

	O: one	*O:* two
E: one	E=2; O= -2	E=-3; O=3
E: two	E= -3; O=3	E=4; O=-4

- Spse E plays
  [p : one; (1 – p) : two]
  For each FIXED p, *O* plays pure strategy

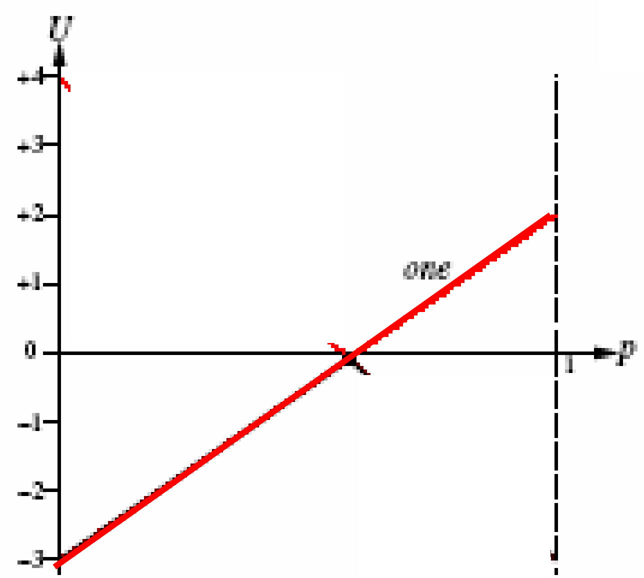
- If *O* plays one, payoff is

  $p \times U(\text{one}, \text{one}) + (1-p) \times U(\text{one}, \text{two})$
  $= p \times 2 + (1-p) \times -3 = \boldsymbol{5p - 3}$

  If *O* plays two, payoff is $\boldsymbol{4 - 7p}$

⇒For each *p*,
O plays

$$\begin{cases} \text{one} & \text{if } 5p - 3 \geq 4 - 7p \\ \text{two} & \text{if } 5p - 3 < 4 - 7p \end{cases}$$



- E can get maximum of { 5p – 3, 4 – 7p } ... largest at p = 7/12

⇒ E should play [7/12 : one; 5/12 : two]
Utility is **–1/12**

What about O?

	O: one	*O:* two
E: one	E=2; O= -2	E=-3; O=3
E: two	E= -3; O=3	E=4; O=-4

- Spse *O* plays
  [q : one; (1 – q) : two]

$$\Rightarrow \text{For each q, E plays} \begin{cases} \text{one} & \text{if } 5q - 3 \le 4 - 7q \\ \text{two} & \text{if } 5q - 3 > 4 - 7q \end{cases}$$

$\Rightarrow$ *O* should minimize {5q – 3, 4 – 7q}
... smallest when **q = 7/12**

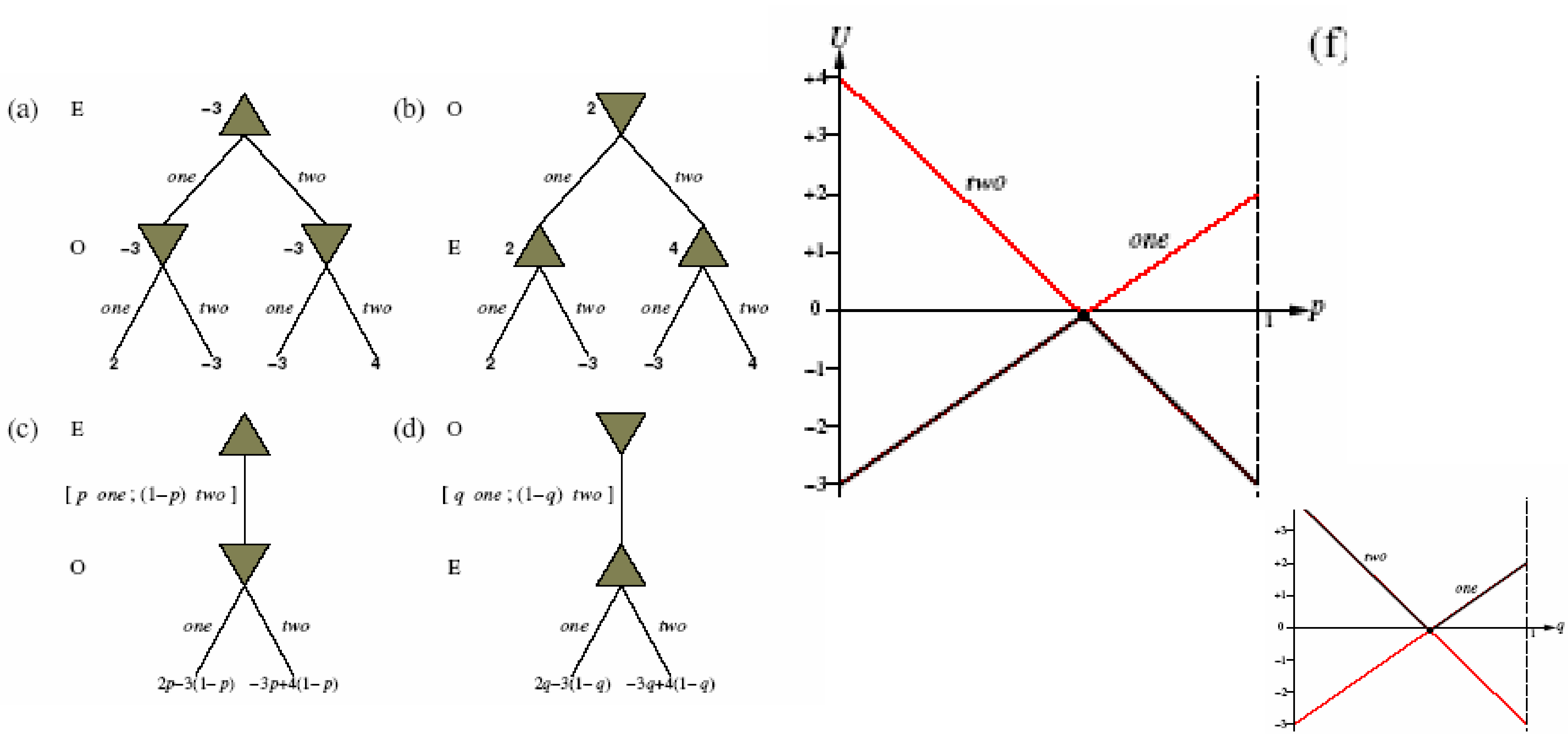
$\Rightarrow$ O should play [7/12 : one; 5/12 : two]
Utility is **-1/12**

- **Maximin equilibrium**... and Nash Equilibrium!

- Coincidence that *O* and *E* have same strategy.

NOT coincidence that utility is same!

Minimax Game Trees for Morra



General Results

- Every 2-player 0-sum game has
  a maximin equilibrium
  ...often a mixed strategy.
- *Thrm: Every Nash equilibrium in 0-sum game is maximin for both players.*

- Typically more complex:
  - when *n* actions, need hyper-planes (not lines)
  - need to remove dominated pure strategies (recursively)
  - use linear programming

Iterated Prisoner Dilemma

	A: testify	*A:* refuse
B: testify	*A*= -5; *B*= -5	*A*= -10; *B*=0
B: refuse	*A*=0; B= -10	*A*= -1; *B*= -1

- If A, B play just once…
  expect each to *testify*,
  … even though suboptimal for BOTH !
- If play MANY times. . .
  Will both refuse, so *BOTH* do better?
- Probably not:
  Suppose play 100 times
  - On R#100, no further repeats, so ⟨ testify, testify ⟩ !
  - On R#99, as R#100 known, again use dominant ⟨ testify, testify ⟩!
  - . . .
  - So sub-optimal all the way down… each gets 500 years!!

Iterated P.D., con't

- Suppose 99% chance of meeting again
  ... not clear which round is last
  ??Co-operation??

- **Perpetual Punishment:**
  - refuse unless other player ever testify
  - As long as both players refuse: $\sum_{t=0}^{\infty} 0.99^t \times (-1) = -100$

  If one player testify:
  - 0 for this round, then -10 forever
  - $\sum_{t=i}^{\infty} 0.99^t \times (-10) = -990$
    (Mutually assured destruction ... both players lose)

$\Rightarrow$ neither player should testify!

$\Rightarrow \langle$ refuse, refuse $\rangle$ at each step!

Iterated P.D., III

- **tit-for-tat**
  - $MyAction_1$ = refuse, then
  - $MyAction_{t+1} = OpponentAction_t$

Works pretty well...

Outline

- Game Theory
  - Motivation: Multiple agents
  - Dominant Action
  - Strategy
  - Prisoner's Dilemma
    - Domain Strategy Equilibrium; Paretto Optimum; Nash Equilibrium
  - Mixed Strategy (Mixed Nash Equilibrium)
  - Iterated Games
- Mechanism Design
  - Tragedy of the Commons
  - Auctions
  - Price of Anarchy
  - Combinatorial Auctions

Mechanism Design: Inverse Game Theory

- Design rules for Agent environment such that

  Agent maximizing OWN utility
  will maximize COLLECTIVE GOOD
- Eg:
  - Design protocols for
  - Internet Trac routers to maximize global throughput
  - auction off cheap airline tickets
  - assign medical intern to hospitals
  - get soccer players to cooperate
- 1990, gov't auctioned off frequencies due to bad design, lost $$ millions!
- Defn: Mechanism
  - set of strategies each agent may adopt
  - outcome rule G determining payoff for any strategy profile of allowable strategies
- Why complicated?

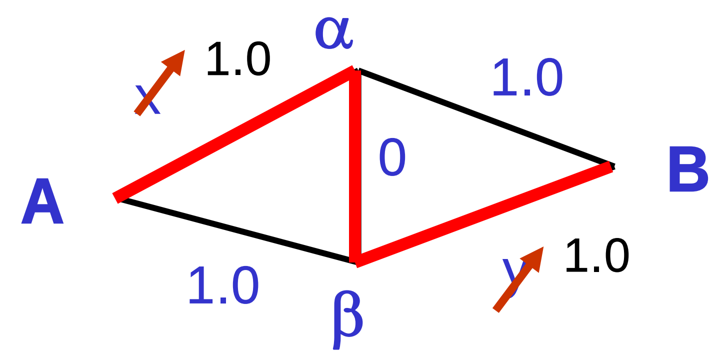
Tragedy of the Commons

- Every farmer can bring livestock to town commons
  $\Rightarrow$ destruction from overgrazing
  . . . negative utility to ALL farmers
- Every individual farmer acted rationally
  - use of commons is free
  - refraining from use won't help, as others will use it anyway

  (use of atmosphere, oceans, . . .)
- Solution: Setting prices
  - ... must explicate external effects on global utility
  - What is correct price?
- Goal: Each agent maximizes global utility

  Impossible for agent, as does not know
  - current state
  - effect of actions on other agents
- First: simplify to deal with simpler decision

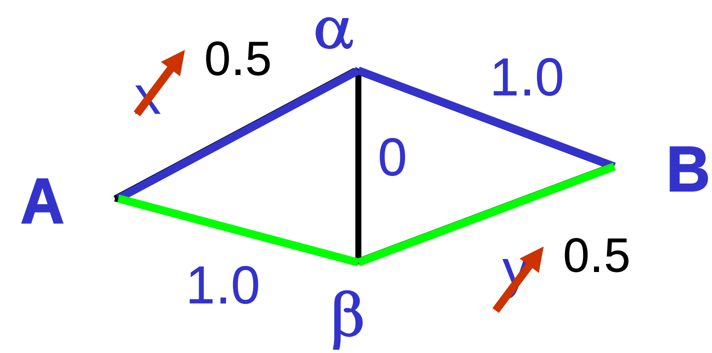
Price of Anarchy



- Many people want to go from A to B
  - Cost of A $\rightarrow$ β is 1; from α $\rightarrow$ B is 1; α $\leftrightarrow$ β is 0
  - Cost from A to α is "% of people on route" $x \in [0,1]$
  - Cost from β to B is "% of people on route" $y \in [0,1]$
- Which path would YOU take?
  - As $x \leq 1$ and $y \leq 1$, clearly A$\rightarrow$ α $\rightarrow$ β $\rightarrow$ B is best (always $\leq$ 2)
- But if EVERYONE takes it, cost $\equiv$ 2
- non Anarchy:
  - [A-M] take A$\rightarrow$ α$\rightarrow$ B
  - [N-Z] take A$\rightarrow$ β $\rightarrow$ B

Everyone pays only 1.5 !

Price of Anarchy



- Many people want to go from A to B
  - Cost of A $\rightarrow$ β is 1; from α $\rightarrow$ B is 1; α $\leftrightarrow$ β is 0
  - Cost from A to α is "% of people on route" $x \in [0,1]$
  - Cost from β to B is "% of people on route" $y \in [0,1]$
- Which path would YOU take?
  - As $x \leq 1$ and $y \leq 1$, clearly A$\rightarrow$ α $\rightarrow$ β $\rightarrow$ B is best (always $\leq$ 2)
- But if EVERYONE takes it, cost $\equiv$ 2
- non Anarchy:
  - [A-M] take A$\rightarrow$ α$\rightarrow$ B
  - [N-Z] take A$\rightarrow$ β $\rightarrow$ B

  Everyone pays only 1.5 !

Auctions

- Mechanism for selling goods to individuals
  - ("good" $\equiv$ item for sale)
- Single "good"

  Each bidder Q_i has utility v_i for good
  - ... only Q_i knows v_i
- **English Auction**
  - auctioneer increments prices of good,
  - until only 1 bidder remains
  - Bidder w/ highest v_i gets good, at price $b_m + d$
  - (b_m is highest OTHER bid, d is increment)
- Strategy for Q_i:
  - bid current price p if $p \leq \mathrm{v_i}$

English Auction (con't)

- “Dominant” as
  independent of other's strategy
   No need to contemplate other player's strategy

- **Strategy-proof** mechanism:
      players have dominant strategy
      (reveal true incentives)

- but… *High communication costs!*

Sealed Bid Auction

- Each player posts single bid to auctioneer
  - Q_i w/highest bid b_i wins
  - . . . Q_i pays b_i, to get good

Q: Should Q_i bid v_i?

A: Not dominant!

Better is $\min\{ v_i, b_m + \varepsilon \}$

(b_m is max of others)

- Drawbacks:
  - player w/highest v_i might not get good

    ... so seller gets too little!

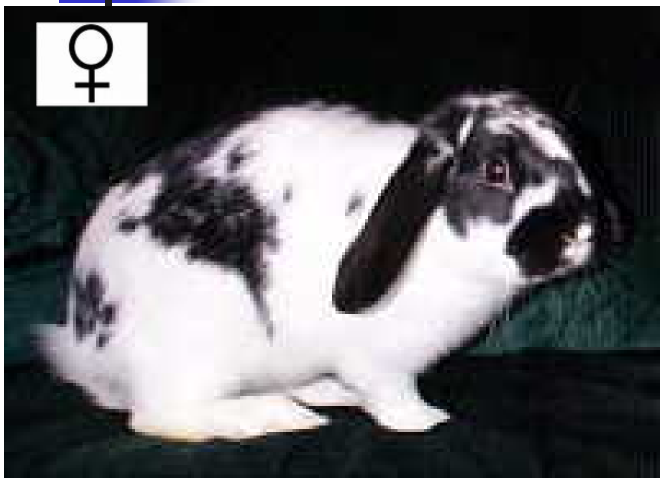
    ... as “wrong” bidder gets good!
  - bidders spend time contemplating others

Sealed-Bid 2nd-Price Auction

- Each player posts single bid to auctioneer
  - Q_i w/highest bid b_i wins
  - … Q_i pays b_m , gets good

  b_m is 2nd highest bid
- Q: Should Q_i bid v_i?
- A: Yes, is dominant!
  - Q_i bids b_i
  - Utility to Q_i is $\begin{cases} v_i - b_m & \text{if } b_i > b_m \\ 0 & \text{otherwise} \end{cases}$
  - $u_i(b_i, b_m) =$ If $v_i - b_m > 0$, any bid winning auction is good
    eg, bid v_i
  - If $v_i - b_m < 0$, any bid losing auction is good
    e.g., bid v_i
  - So v_i is appropriate in all cases
    … is ONLY value appropriate in all cases!
- "Vickrey Auction" (Nobel prize)

Rabbit Auction



Flopsy

Mopsy →

Jack

C1: will pay $5 for any one

C2: will pay $9 for a breeding pair

(Flopsy and one of the others)

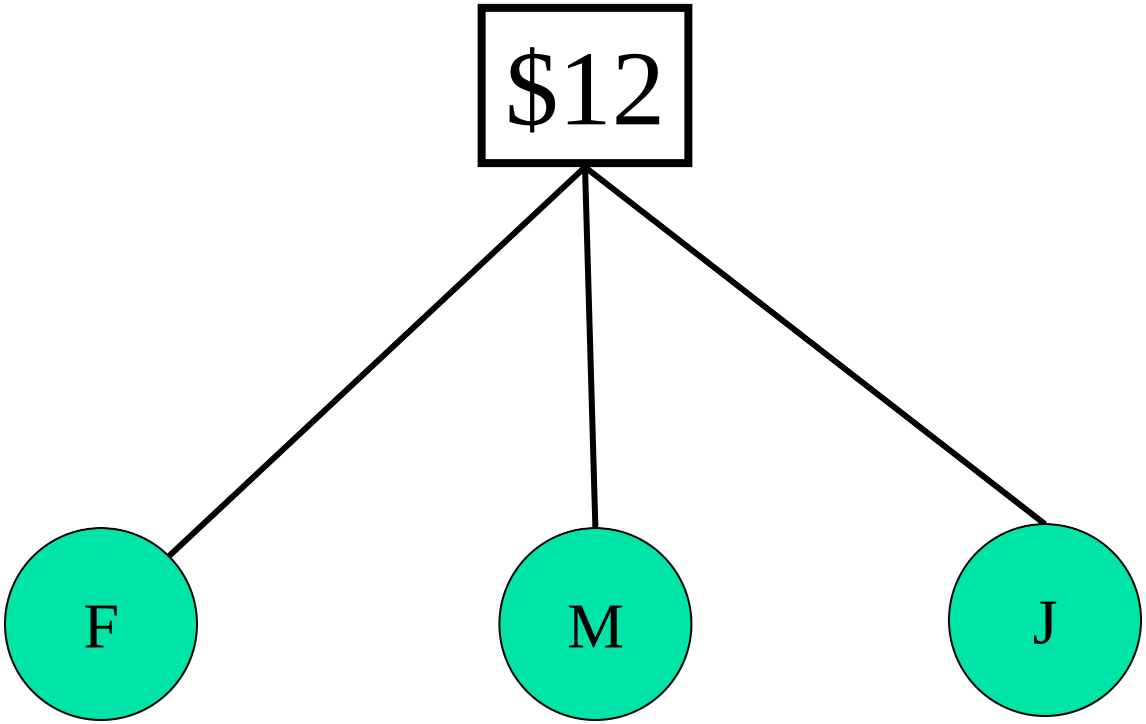
C3: will pay $12 for all three

Combinatorial Auctions

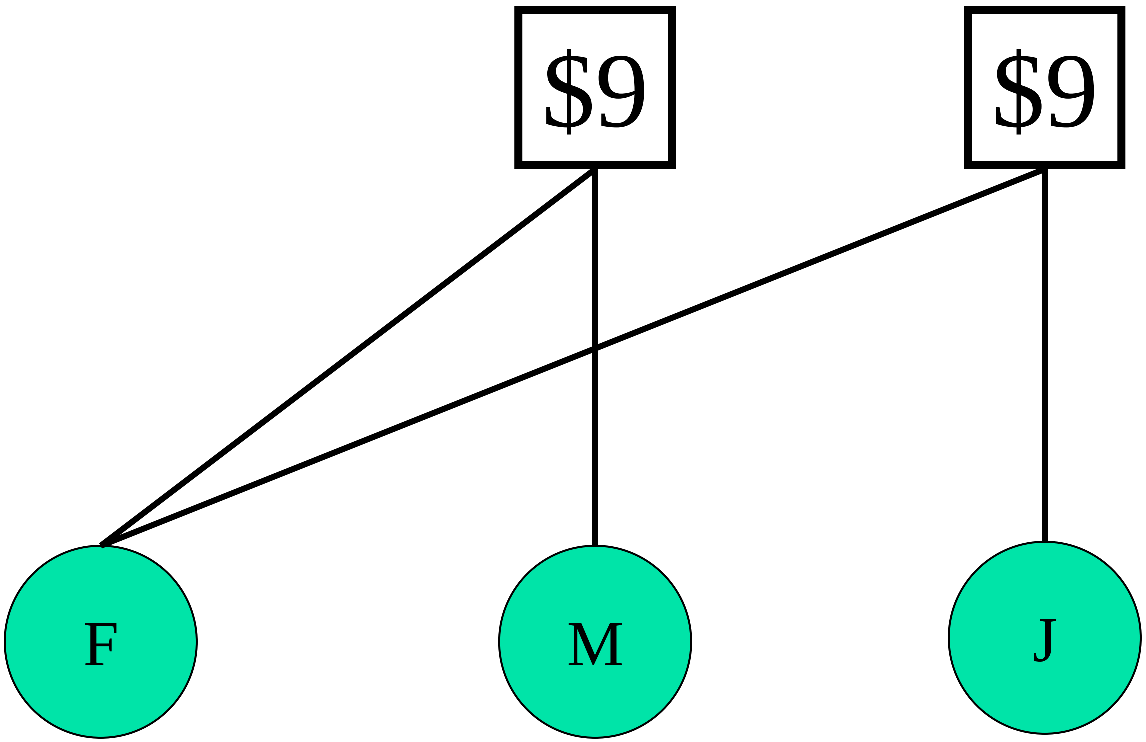
- Auction all items simultaneously
- Bid specifies a price and a set of items
  (“all or nothing”)
- Exclusive-OR: use “dummy item” representing the bidder
- Number of Rounds
  - Multi-round or **Single-round**
- Number of Units (per item)
  - **1 unit** vs Many units
- Number of Items
  - 1 item vs **Many items**

		Number of Items	
		Single	Multiple
Number of Units	Single		
	Multiple		

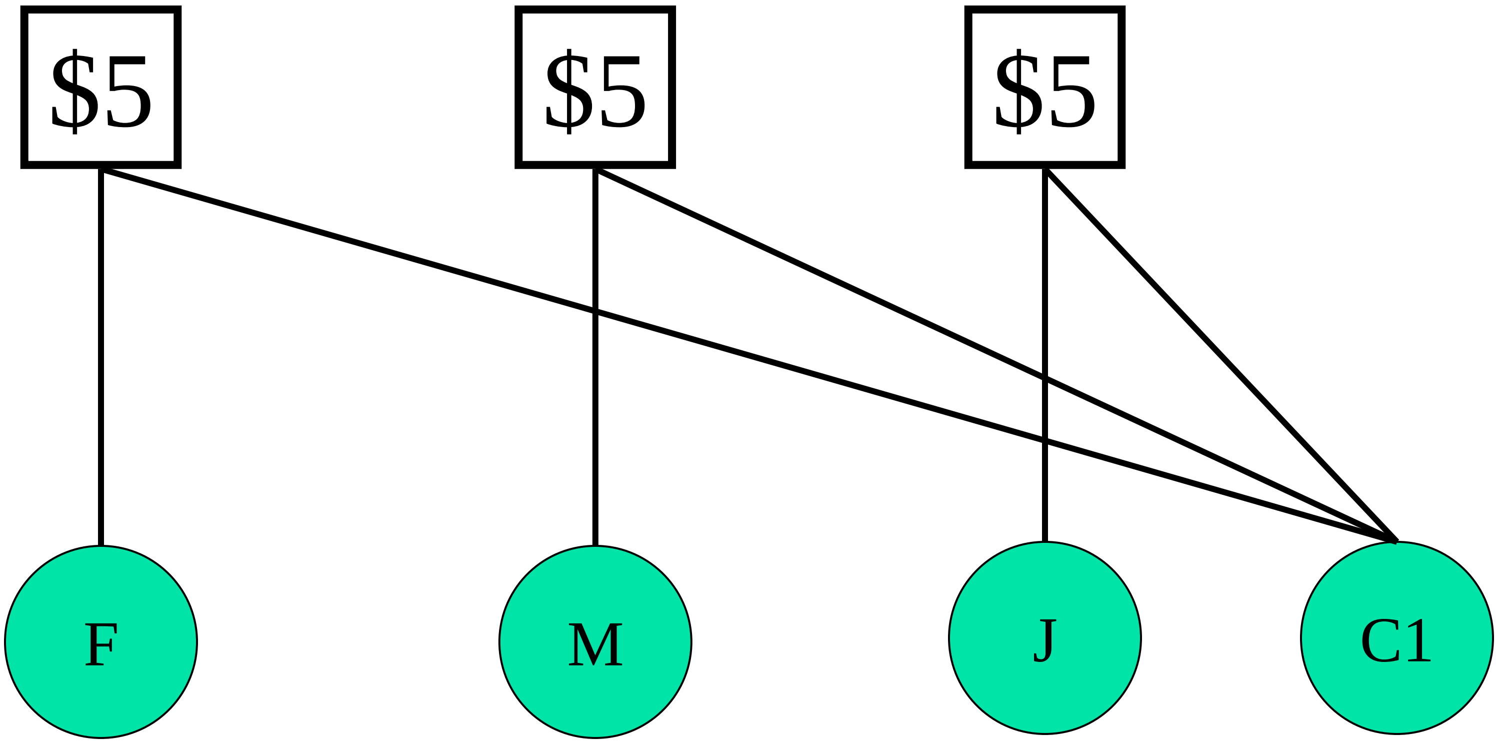
$12 for all three



$9 for a breeding pair



$5 for any one



Applications

- Airport gates
  - Gate in YEG at 2pm &&
  - Gate in YYZ at 6pm
- Parcels of land
  - 4 adjacent beach-front parcels, for 1 hotel
- FCC spectrum auctions
- Goods distribution routes
- eBay
- ...

Winner Determination

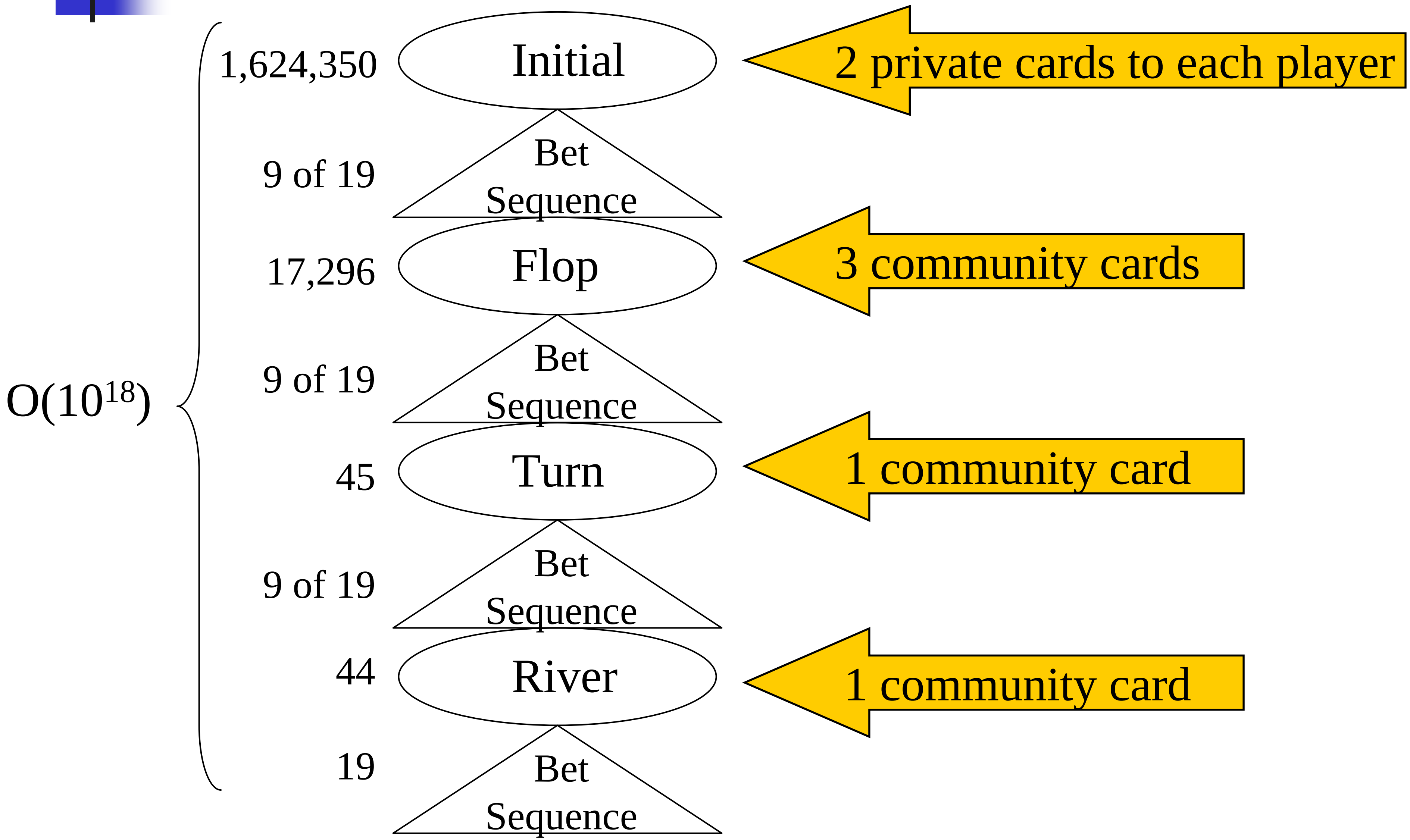
- Problem: how to determine who wins ?
- Choose a set of bids that
  - are feasible (disjoint) and
  - maximize the auctioneer’s profit.
- NP-complete (set packing problem)

How Should Players Interact?

- Strategy
  - Dominant Strategy Equilibrium
  - Pareto Optimum
  - Nash Equilibrium
  - Mixed Strategy
- Prisoner's Dilemma, Iterated Games
- Mechanism Design
  - Non trivial (Tragedy of the Commons... of Anarchy)
  - Auctions: English, Sealed Bid, Vickrey
  - Combinatorial Auction

Bonus Material: *Poker*

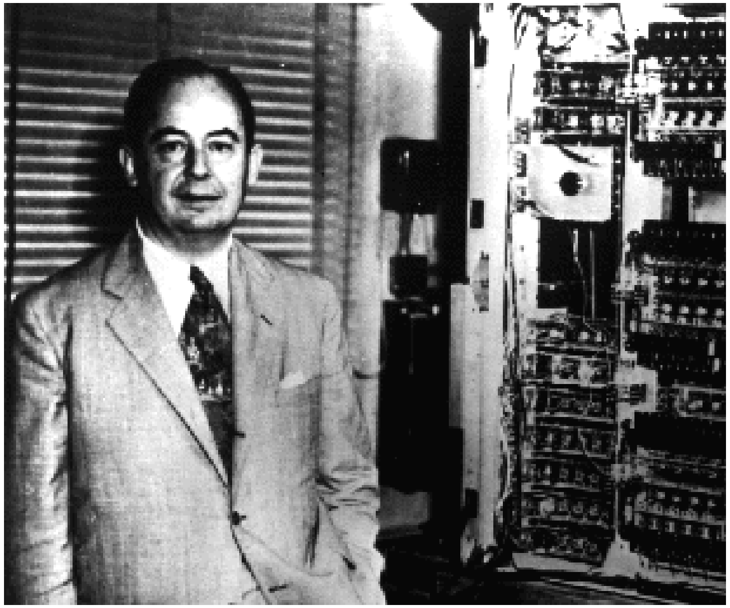
2-player, limit, Texas Hold'em

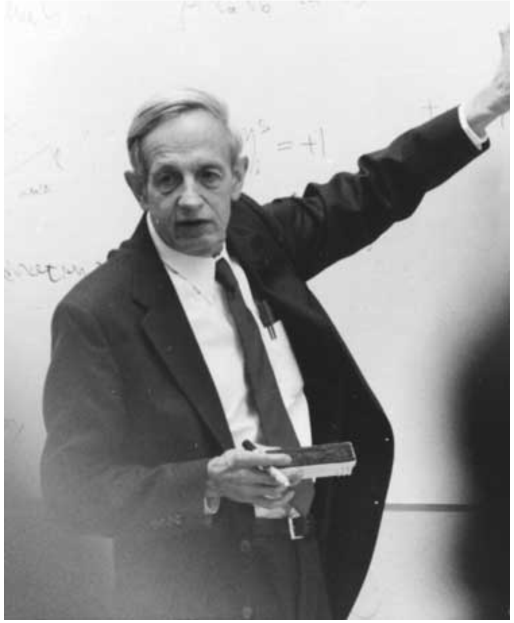


The Challenges

- Large game tree
- Stochastic element
- Imperfect information
  - during hand, and after
- Variable number of players (2–10)
- Aim is not just to win,
  but to maximize winnings
  - Need to exploit opponent weaknesses

Game-Theoretic Approach





Linear Programming

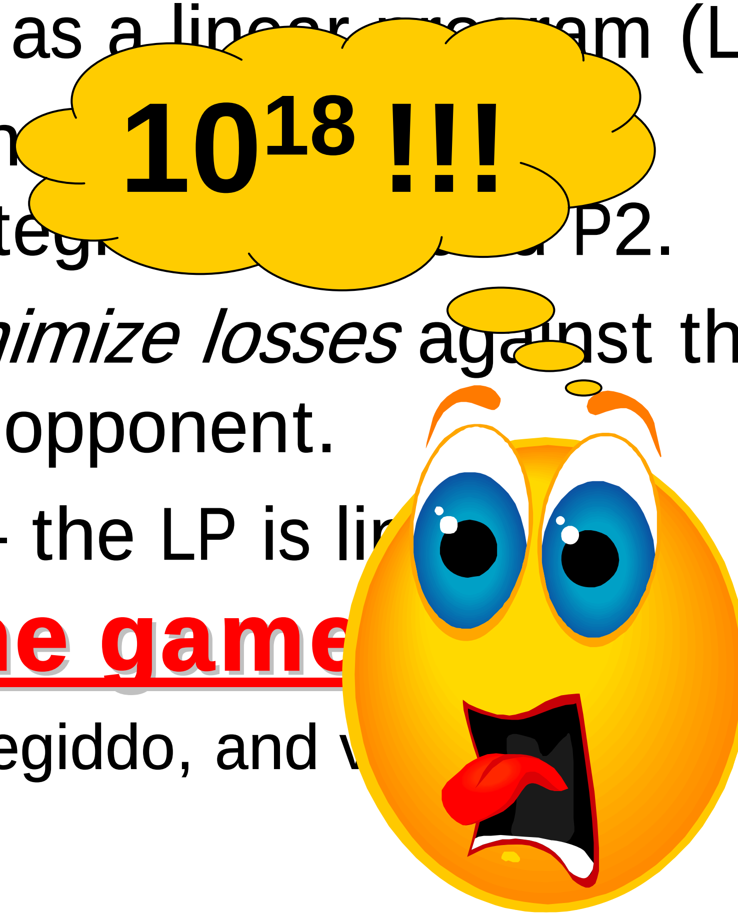
- 2-player, 0-sum game with chance events, mixed strategies, and imperfect information can be formulated as a linear program (LP).
- LP can be solved in *polynomial time* to produce Nash strategies for P1 and P2.
- Guaranteed to *minimize losses* against the strongest possible opponent.
- “Sequence form” – the LP is linear in the size of the game tree

(Koller, Megiddo, and von Stengel)

Linear Programming

- 2-player, 0-sum game with chance events, mixed strategies, and imperfect information can be formulated as a linear program (LP).
- LP can be solved in
  produce Nash strateg P2.
- Guaranteed to *minimize losses* against the strongest possible opponent.
- "Sequence form" – the LP is lin
  <u>**the size of the game**</u>

(Koller, Megiddo, and v



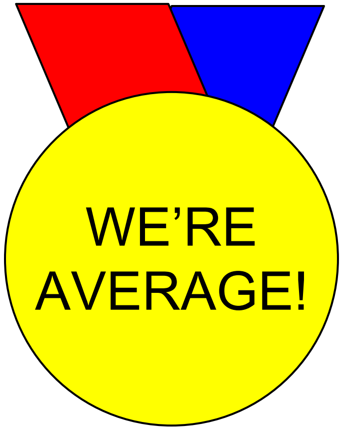
Why Equilibrium?

- In
  - symmetric,
  - two player,
  - zero-sum games,

  playing an equilibrium

  is equivalent to

  having a worst-case performance of tying.
- Given the state of the art
  of modeling of opponents,
  ... not be so bad.

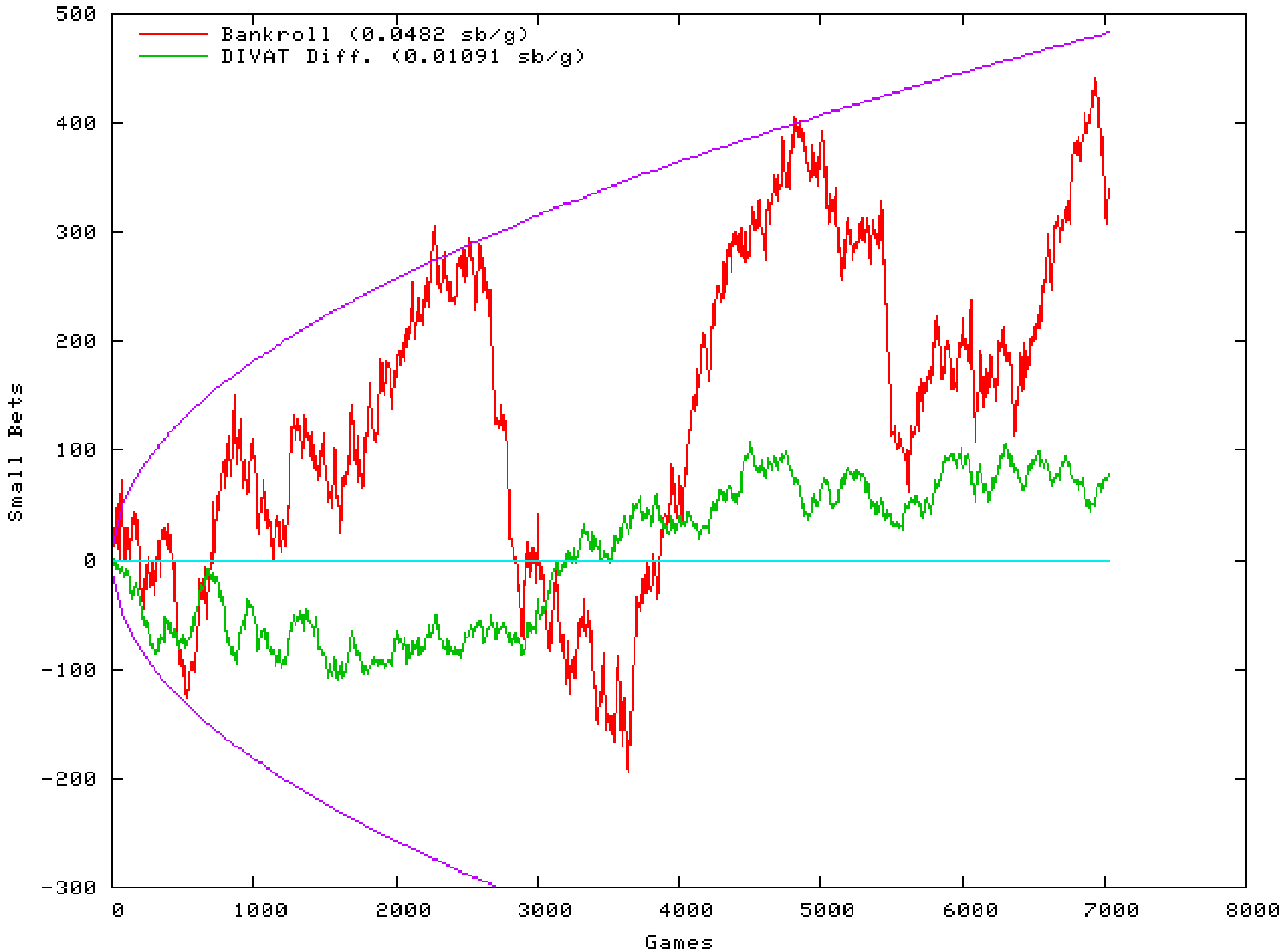


PsOpti (Sparbot)

- Abstract game tree of size 10^7
- Bluffing, slow play, etc.
  fall out from the mathematics.
- Best 2-player program to date !
- Has held its own against 2 world-class humans
- Won the AAAI'06 poker-bot competitions

IJCAI'03

PsOpti2 vs. “theCount”

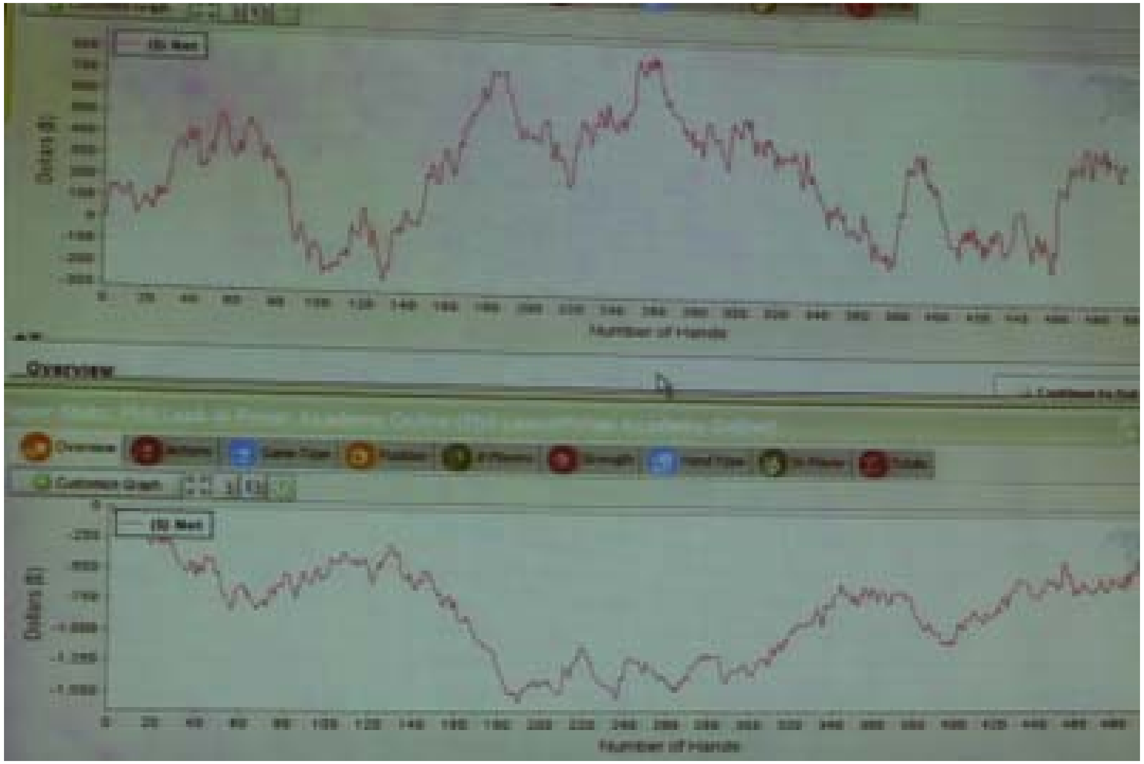


PsOpti's Weaknesses

- The equilibrium strategy for the highly abstract game is far from perfect.
- No opponent modelling.
  - Nash equilibrium not the best strategy:
    - Non-adaptive
    - Defensive
  - Even the best humans have weaknesses that should be exploited

http://www.poker-academy.com

Man-Machine Poker Match! (2007)



- A graph for each half of the duplicate match plotted in Poker Academy Prospector

http://games.cs.ualberta.ca/poker/man-machine/

Results

AAAI 2007

- 4 sessions; each 500-hard ***duplicate matches***

1. Ali won $390; Phil lost $465.
   - -$75 → DRAW
2. Phil: $1570; Ali: –$2495
   - -$925 → Polaris WON!
3. Ali: –$625; Phil: +$1455
   - +830 → Polaris LOST!
4. Ali: +$4605; Phil: +$110
   - +$570 → Polaris LOST!

- Total: 1-2-1
  ... but only $395 over 2000 hands!

Man vs Machine Poker

- Comparable with top human players (2007)
- Attracted international media attention

“We won, not by a significant amount, and the bots are closing in.”
– Phil Laak

“I really am happy it's over. I'm surprised we won ... It's already so good it will be tough to beat in future.”
– Ali Eslami

BBC
USA TODAY
The Sydney Morning Herald
smh.com.au
San Francisco Chronicle
The New York Times
CTV
CBS NEWS
The Washington Post

55