15-780: Graduate AI

Lecture 8. Games

Geoff Gordon (this lecture)
Ziv Bar-Joseph
TAs Geoff Hollinger, Henry Lin

Admin

- *Extension on HW1!*
    - *Until Friday 3PM*
    - *On Friday only, give to Diane Stidle, 4612 Wean Hall*
    - *50% credit until Monday 10:30AM*
    - *No HWs accepted over weekend*

Admin

- *HW2 out today (on website now)*

Admin

- *Poster session for final projects*
  - *5:30PM on Thursday, Dec 13*
- *Final report deadline: beginning of poster session*
  - *This is a **hard** deadline, since course grades are due soon thereafter*

Review

Duality

- *Duality w/ equality constraints*
- *How to express path planning as an LP*
- *Dual of path planning LP*

Optimization in ILPs

- *DFS, with pruning by:*
  - *constraint propagation*
  - *best solution so far*
  - *dual feasible solution*
  - *dual feasible solution for relaxation of ILP with some variables set (branch and bound)*

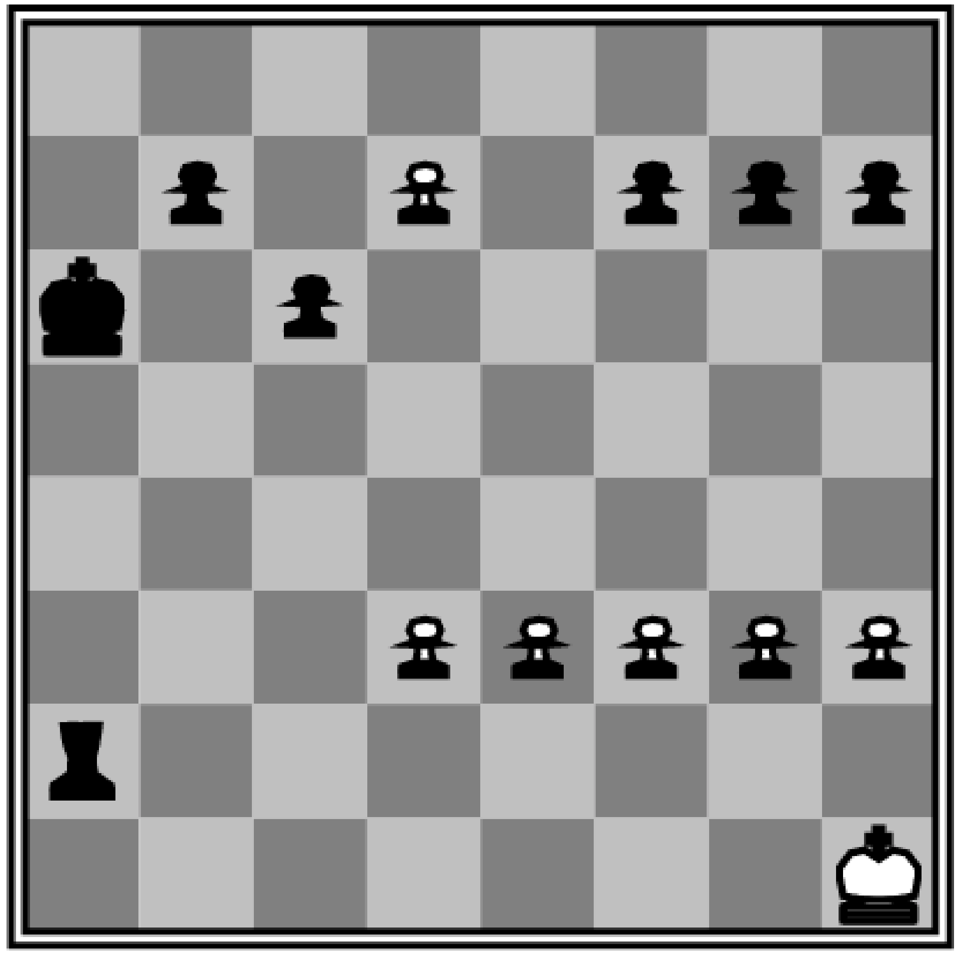
Optimization in ILPs

- *Duality gap*
- *Cutting planes*
- *Branch and cut*

More on optimization

- *Unconstrained optimization: gradient = 0*
- *Equality-constrained optimization*
  - *Lagrange multipliers*
- *Inequality-constrained: either*
  - *nonnegative multipliers, or*
  - *search through bases (for LP: simplex)*

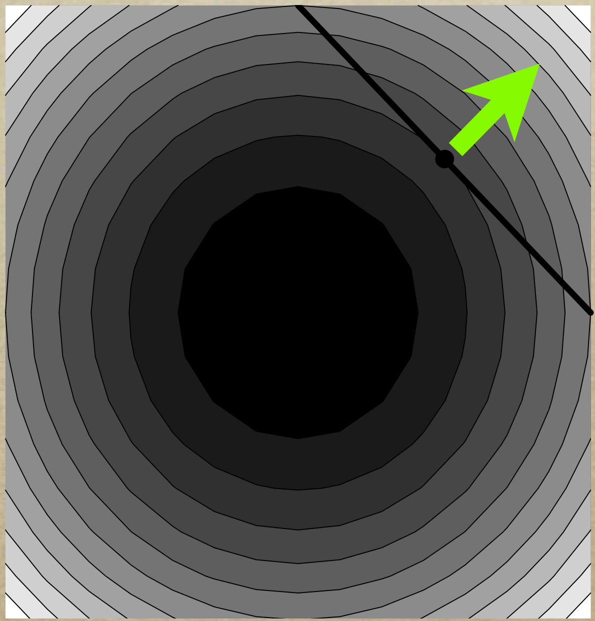
Quiescence



Black to move

Duality as game

- *Yet one more interpretation of duality*
- *Game between minimizer and maximizer*
- $min_{xy}\ x^2 + y^2\ \ s.t.\ \ x + y = 2$



$$min_{xy}\ max_{\lambda}\ \ x^2 + y^2 + \lambda(x + y - 2)$$

Duality as game

- $min_{xy}\ max_{\lambda}\ \ x^2 + y^2 + \lambda(x + y - 2)$
- *Gradients wrt x, y, λ:*
  - $2x + \lambda = 0$
  - $2y + \lambda = 0$
  - $x + y = 2$
- *Same equations as before*

Matrix games

Matrix games

- *Games where each player chooses a single move (simultaneously with other players)*
- *Also called normal form games*
- *Simultaneous moves cause uncertainty: we don't know what other player(s) will do*

Acting in a matrix game

- *One of the simplest kinds of games; we'll get more complicated later in course*
- *But still will make us talk about*
    - *negotiation*
    - *cooperation*
    - *threats, promises, etc.*

Matrix game: prisoner's dilemma

	C	*D*
C	*–1*	*–9*
D	*0*	*–5*

payoff to Row

	C	*D*
C	*–1*	*0*
D	*–9*	*–5*

Payoff to Col

Matrix game: prisoner's dilemma

	C	*D*
C	*–1, –1*	*–9, 0*
D	*0, –9*	*–5, –5*

Can also have n-player games

	H	*T*
H	*0, 0, 1*	*0, 0, 1*
T	*0, 0, 1*	*1, 1, 0*

if Layer plays H

	H	*T*
H	*1, 1, 0*	*0, 0, 1*
T	*0, 0, 1*	*0, 0, 1*

if Layer plays T

Analyzing a game

- *What do we want to know about a game?*
- *Value of a joint action: just read it off of the table*
- *Value of a mixed joint strategy: almost as simple*

Value of a mixed joint strategy

	C	*D*
C	$.6*.3*w$	$.4*.3*x$
D	$.6*.7*y$	$.4*.7*z$

- *Suppose Row plays 30-70, Col plays 60-40*

Payoff of joint strategy

- *Just an average over elements of payoff matrices M_R and M_C*
- *If x and y are strategy vectors like (.3, .7)' then we can write*
    - $x' M_R y$
    - $x' M_C y$

What else?

- *Could ask for value of a strategy x under various weaker assumptions about other players' strategies y, z, …*
- *Weakest assumption: other players might do absolutely anything!*
- *How much does a strategy* ***guarantee*** *us in the most paranoid of all possible worlds?*

Paranoia

- *Worst-case value of a row strategy x in 2-player game is*
  - $\min_y x' M_R y$
- *More than two players, min over y, z, …*

Paranoia

- *Paranoid player wants to maximize the worst-case value:*
  - $max_x \, min_y \, x' M_R \, y$
- *Famous theorem of von Neumann: it doesn't matter who chooses first*
  - $max_x \, min_y \, x' M_R \, y = min_y \, max_x \, x' M_R \, y$

Safety value

- *$\min_y \max_x x' M_R y$ is **safety value** or **minimax value** of game*
- *A strategy that guarantees minimax value is a **minimax strategy***
- *Particularly useful in …*

Zero-sum games

Zero-sum game

- *A 2-player matrix game where*
- *(payoff to A) = –(payoff to B) for all combinations of actions*
- *Note: 3-player games are never called zero-sum, even if payoffs add to 0*
- *But if (payoff to A) = 7 – (payoff to B) we sometimes fudge and call it zero-sum*

Zero-sum: matching pennies

	H	*T*
H	*1*	*–1*
T	*–1*	*1*

Minimax

- *In zero-sum games, safety value for Row is negative of safety value for Col*
- *If both players play such strategies, we are in a* ***minimax equilibrium***
  - *no incentive for either player to switch*

Finding minimax

- $\min_x \max_y \; x'My$ *subject to*

  $\mathbf{1}'x = 1$

  $\mathbf{1}'y = 1$

  $x, y \geq 0$

For example

$$\min_x \max_y \quad x_H y_H + x_T y_T - x_H y_T - x_T y_H$$

$$\text{st} \quad x_H + x_T = 1$$

$$y_H + y_T = 1$$

$$x, y \geq 0$$

Finding minimax

- *Eliminate x's equality constraint:*
- $\min_x \max_{y,z}\ \boxed{z(1 - \mathbf{1}'x)} + x'My$ *subject to*

  $\mathbf{1}'y = 1$

  $x, y \geq 0$

Finding minimax

- *Gradient wrt x is*
    - $My - \mathbf{1}z$
- $max_{y,\, z}\ z$ *subject to*

  $My - \mathbf{1}z \geq 0$

  $\mathbf{1}'y = 1$

  $y \geq 0$

Interpreting LP

- $\max_{y,\,z}\ z$ *subject to*

  $My \geq \mathbf{1}z$

  $\mathbf{1}'y = 1$

  $y \geq 0$

- *y is a strategy for Col; z is value of this strategy*

For example

$$\max_{y,z} \; z$$

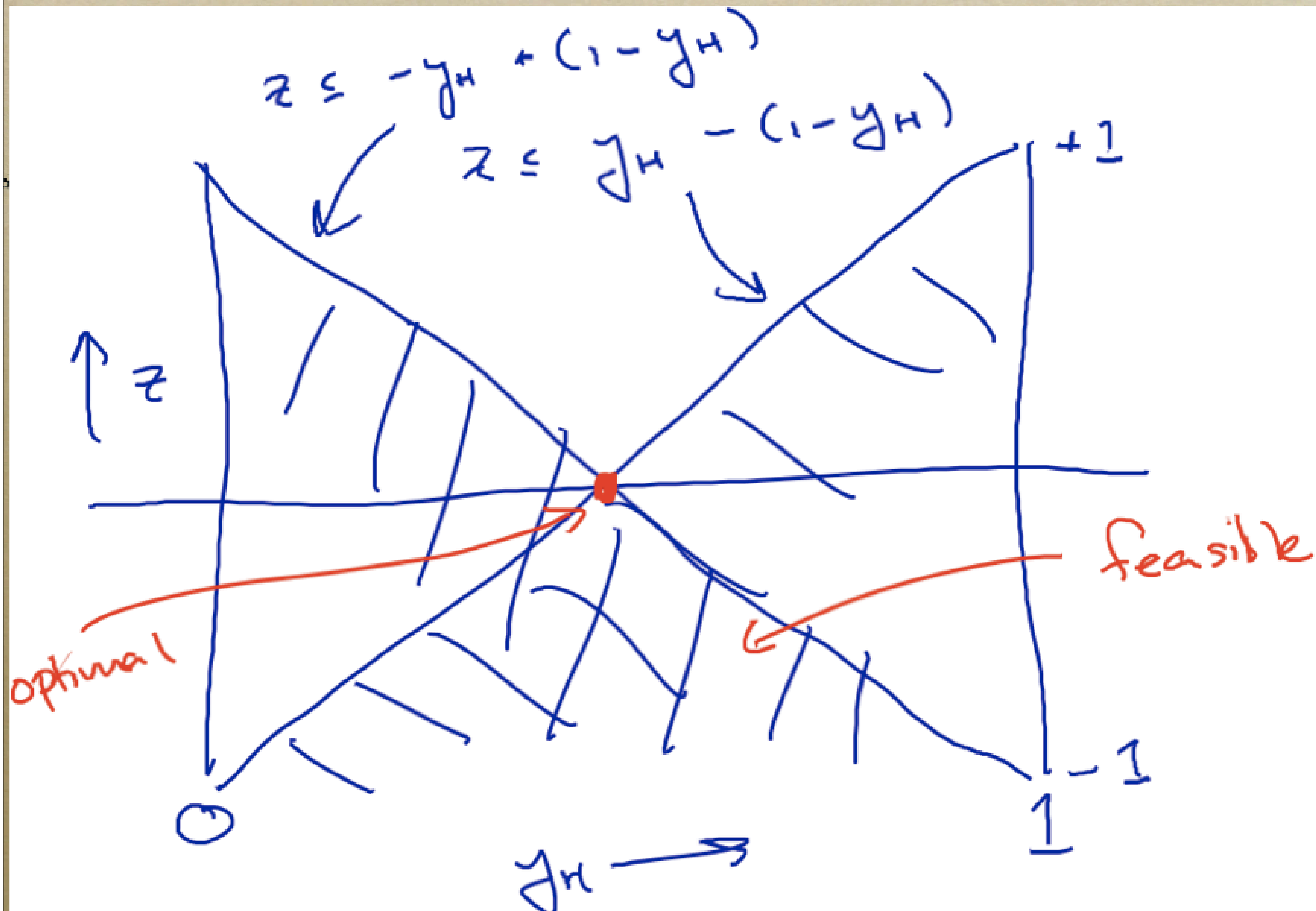
st

$$z \le y_H - y_T$$

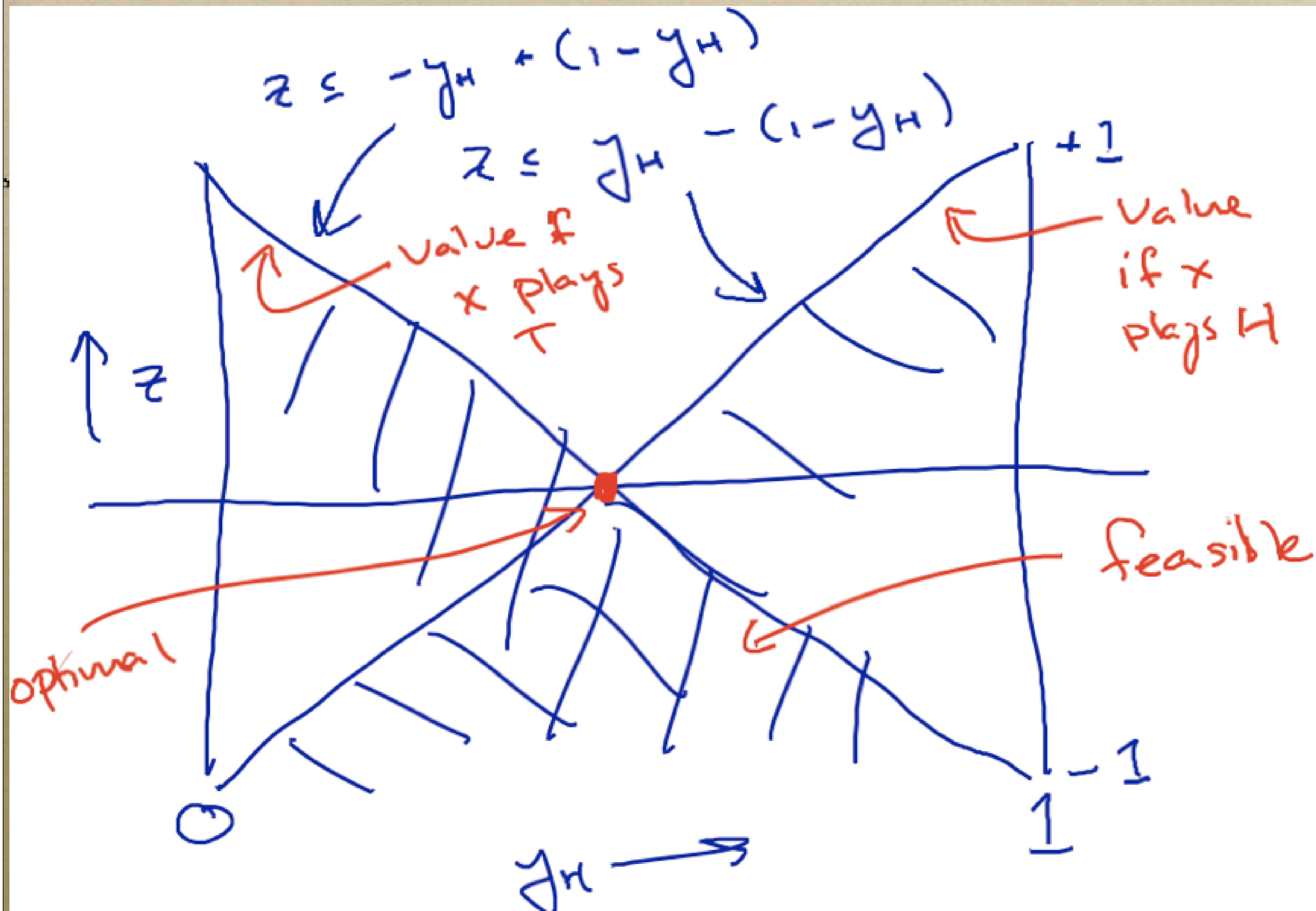
$$z \le -y_H + y_T$$

$$y_H + y_T = 1$$

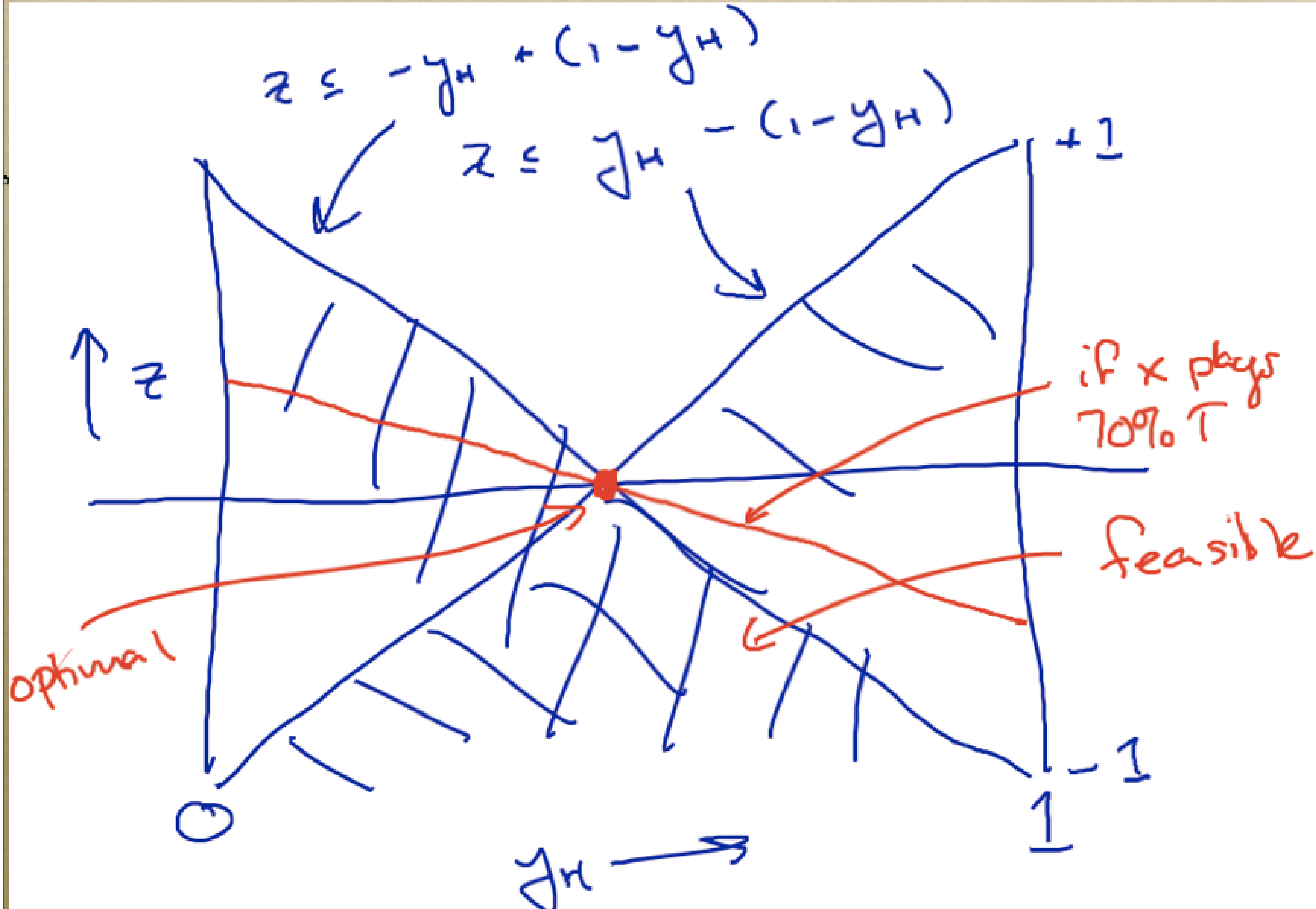
$$y \ge 0$$

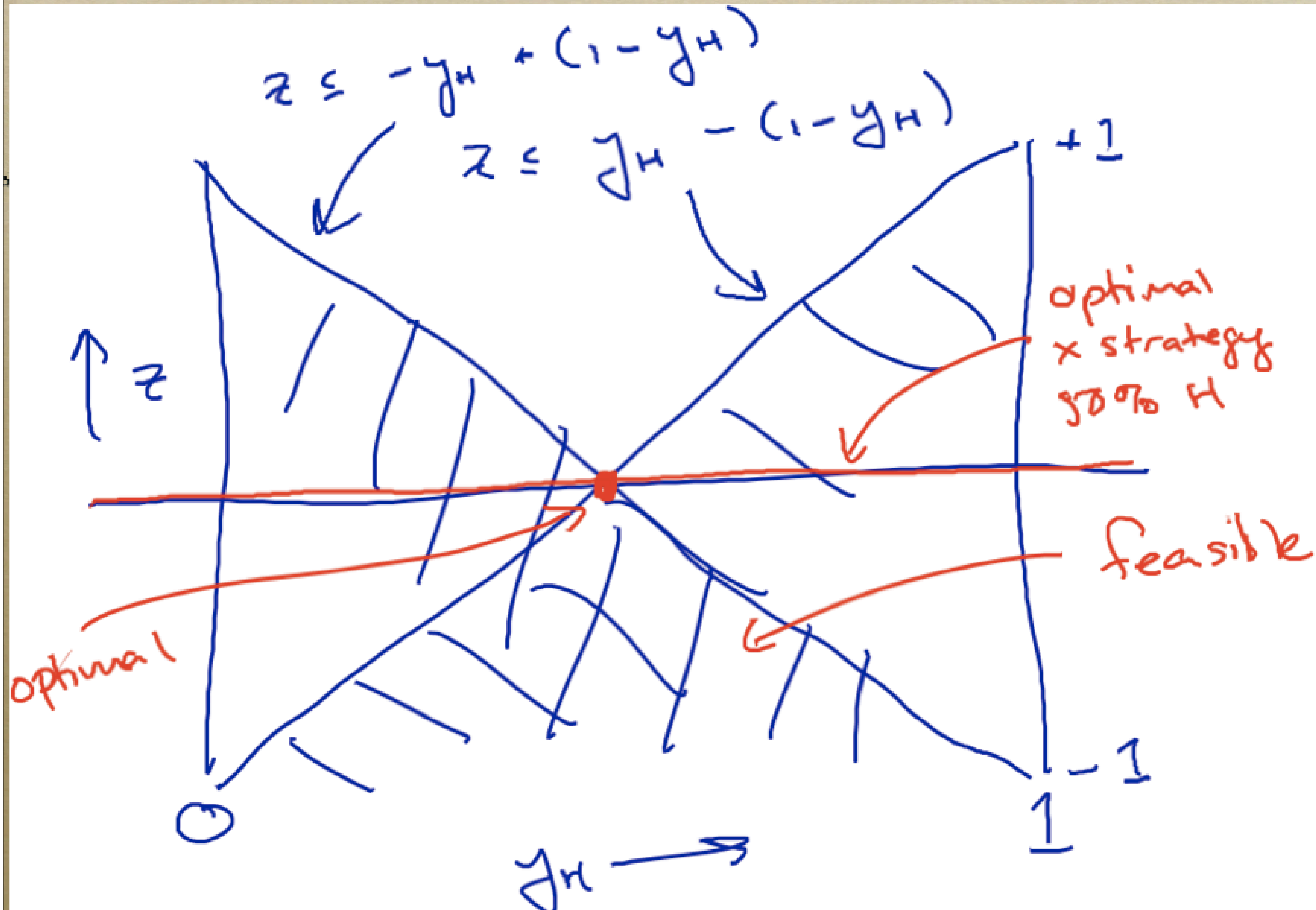


z ≤ -yH + (1 - yH)
z ≤ yH - (1 - yH)
+1
z
feasible
optimal
-1
0
1
yH



z ≤ -yH + (1 - yH)
z ≤ yH - (1 - yH)
+1
value if x plays T
value if x plays H
z
optimal
feasible
0
1
-1
yH


$z \le -y_H + (1 - y_H)$
$z \le y_H - (1 - y_H)$
+1
-1
z
0
1
y_H
if x plays 70% T
feasible
optimal



z ≤ -y_H + (1 - y_H)
z ≤ y_H - (1 - y_H)
+1
optimal
x strategy
50% H
z
feasible
optimal
0
-1
1
y_H

Duality

- *x is dual variable for My* $\geq$ ***1****z*
- *Complementarity: Row can only play strategies where My =* ***1****z*
- *Makes sense: others cost more*
- *Dual of this LP looks the same, so Col can only play strategies where x'M is maximal*

Back to general-sum

- *What if the world isn't really out to get us?*
- *Minimax strategy is unnecessarily pessimistic*

General-sum equilibria

Lunch

	A	*U*
A	*3, 4*	*0, 0*
U	*0, 0*	*4, 3*

A = Ali Baba, U = Union Grill

Pessimism

- *In Lunch, safety value is 12/7 < 2*
- *Could get 3 by suggesting other player's preferred restaurant*
- *Any halfway-rational player will cooperate with this suggestion*

Rationality

- *Trust the other player to look out for his/her own best interests*
- *Stronger assumption than "s/he might do anything"*
- *Results in possibility of higher-than-safety payoff*

Dominated strategies

- *First step towards being rational: if a strategy is bad no matter what the other player does, don't play it!*
- *Such a strategy is* ***(strictly) dominated***
- *Strict = always worse (not just the same)*
- *Weak = sometimes worse, never better*

Eliminating dominated strategies

	C	*D*
C	*-1, -1*	*-9, 0*
D	*0, -9*	*-5, -5*

Prisoner's dilemma

Do we always get a unique answer?

- *No: try Lunch*
- *What can we do instead?*
- *Well, what was special about Row offering to play A?*

	A	*U*
A	*3, 4*	*0, 0*
U	*0, 0*	*4, 3*

Equilibrium

- *If Row says s/he will play A, Col's **best response** is to play A as well*
- *And if Col plays A, then Row's best response is also A*
- *So (A, A) are mutually reinforcing strategies—an **equilibrium***

	A	*U*
A	*3, 4*	*0, 0*
U	*0, 0*	*4, 3*

Equilibrium

- *In addition to assuming players will avoid dominated strategies, could assume they will play an equilibrium*
- *Can rule out some more joint strategies this way*

Nash equilibrium

- *Best-known type of equilibrium*
- *Independent mixed strategy for each player*
- *Each strategy is a best response to others*
  - *puts zero weight on suboptimal actions*
  - *therefore zero weight on dominated actions*

For example

	A	*U*
A	*3, 4*	*0, 0*
U	*0, 0*	*4, 3*

A = Ali Baba, U = Union Grill

Another Nash

		A	*U*
3/7	*A*	*3, 4*	*0, 0*
4/7	*U*	*0, 0*	*4, 3*
		4/7	3/7

Row strategy, Col payoffs

		A	*U*
3/7	*A*	*4*	*0*
4/7	*U*	*0*	*3*
		12/7	*12/7*

Col strategy, Row payoffs

	A	*U*	
A	*3*	*0*	→ *12/7*
U	*0*	*4*	→ *12/7*
	4/7	*3/7*	

Correlated equilibria

Nash at Lunch

- *Nash was still counterintuitive*
    - *Always play U, U or always play A, A*
    - *Or, get bizarrely low payoffs*
- *Any real humans would flip a coin or alternate*
- *Leads to "correlated equilibrium"*

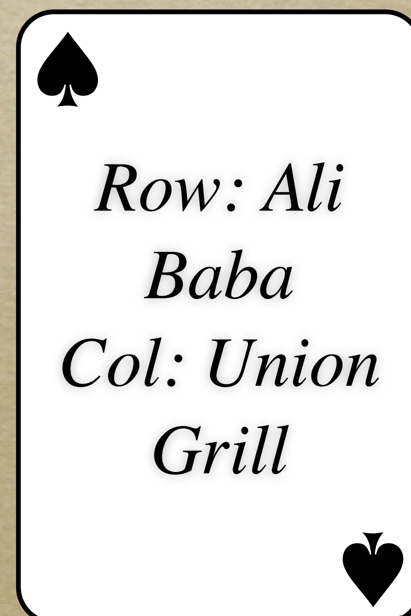
Correlated equilibrium

If there is intelligent life on other planets, in a majority of them, they would have discovered correlated equilibrium before Nash equilibrium.

—Roger Myerson

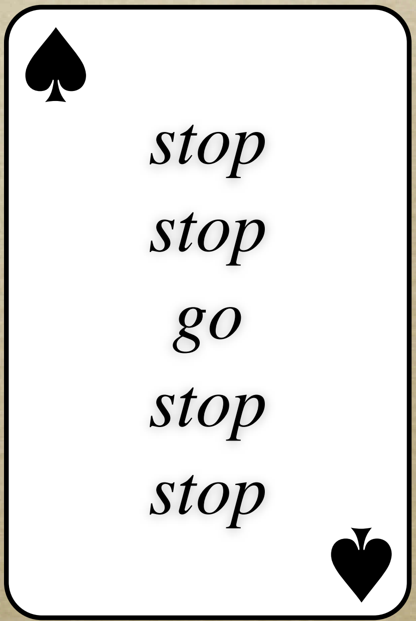
Moderator

- *A moderator has a big deck of cards*
- *Each card has written on it a recommended action for each player*
- *Moderator draws a card, whispers actions to corresponding players*
  - *actions may be correlated*
  - *only find out your own*
  - *may infer others*



Correlated equilibrium

- *Since players can have correlated actions, an equilibrium with a moderator is called a* ***correlated equilibrium***
- *Example: 5-way stoplight*
- *All NE are CE*
- *At least as many CE as NE in every game (often strictly more)*



Finding correlated equilibrium

	A	*U*
A	*3, 4*	*0, 0*
U	*0, 0*	*4, 3*

	A	*U*
A	*a*	*b*
U	*c*	*d*

Finding correlated equilibrium

	A	*U*
A	*a*	*b*
U	*c*	*d*

- *P(Row is recommended to play A) = a + b*
- *P(Col recommended A | Row recommended A) = a / (a + b)*
- *Rationality: when I'm recommended to play A, I don't want to play U instead*

Rationality constraint

Rpayoff(A, A) P(col A | row A)

Rpay(U, A) P(A | A)

$$4\frac{a}{a+b} + 0\frac{b}{a+b} \geq 0\frac{a}{a+b} + 3\frac{b}{a+b} \qquad \text{if } a+b>0$$

Rpay(A, U) P(U | A)

Rpay(U, U) P(U | A)

	A	U
A	a	b
U	c	d

	A	U
A	4,3	0,0
U	0,0	3,4

Rationality constraint is linear

$$4\frac{a}{a+b} + 0\frac{b}{a+b} \geq 0\frac{a}{a+b} + 3\frac{b}{a+b} \qquad \text{if } a+b>0$$

$$4a + 0b \geq 0a + 3b$$

All rationality constraints

	A	U
A	a	b
U	c	d

	A	U
A	4,3	0
U	0	3,4

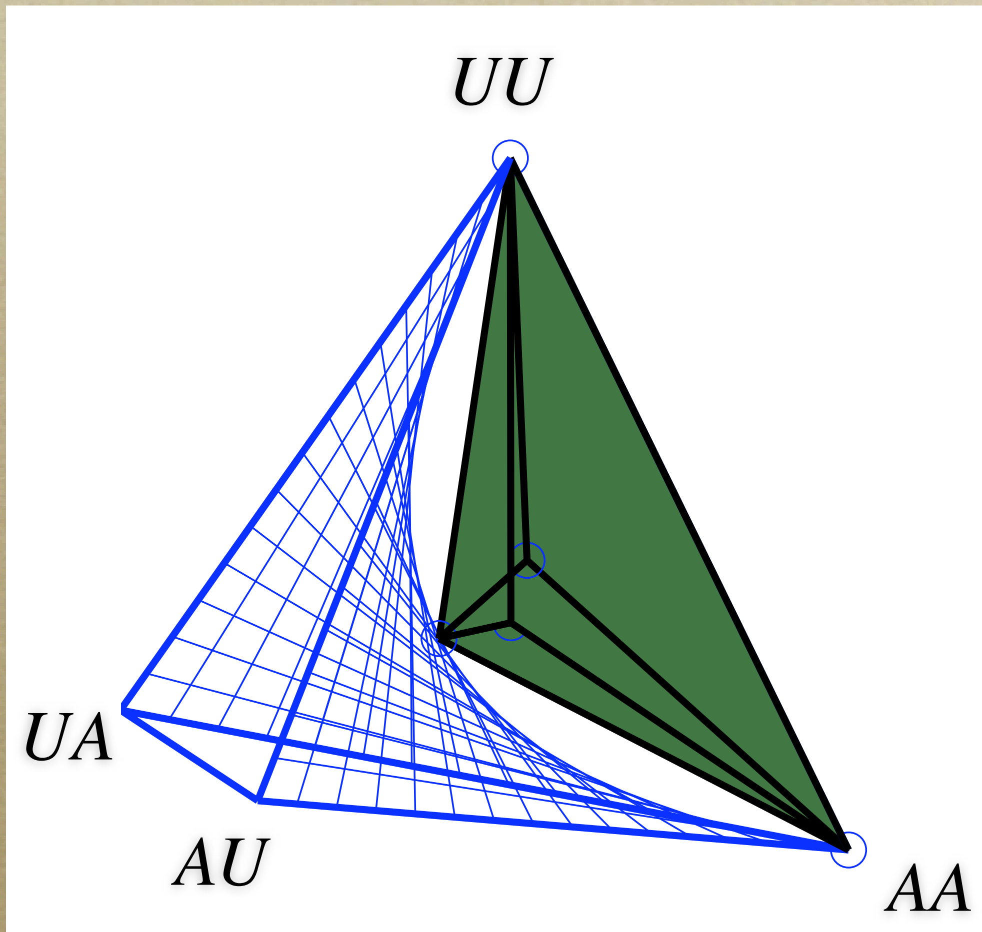
Row recommendation A $4a + 0b \geq 0a + 3b$

Row recommendation U $0c + 3d \geq 4c + 0d$

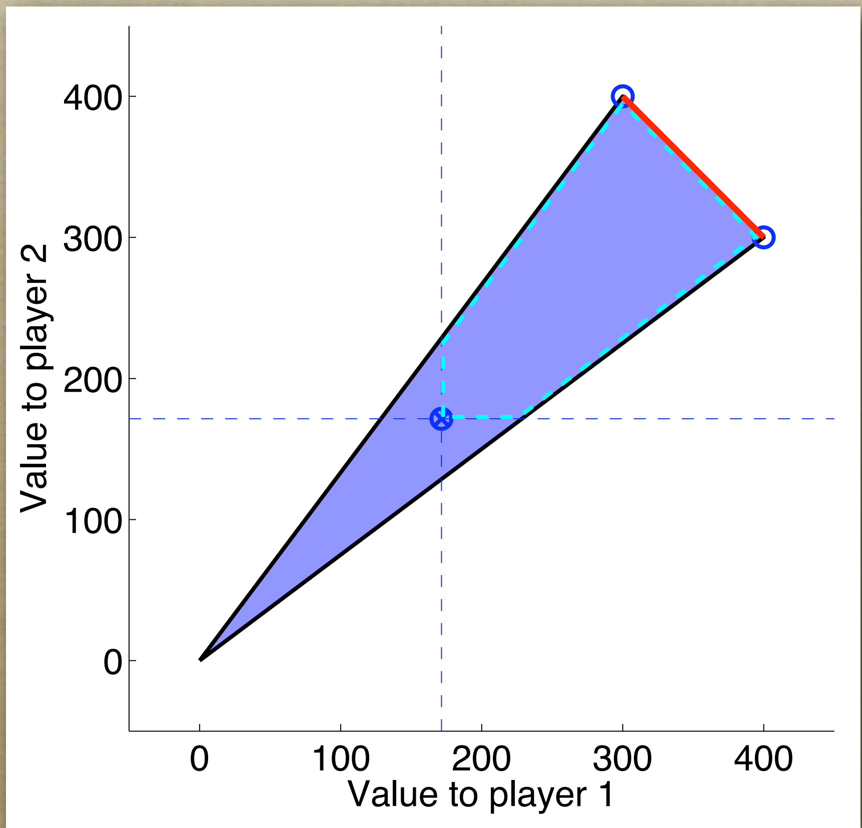
Col recommendation A $3a + 0c \geq 0a + 4c$

Col recommendation U $0b + 4d \geq 3b + 0d$

Correlated equilibrium



Correlated equilibrium payoffs



Realism?

- *Often more realistic than Nash*
- *Moderators are often available*
- *Sometimes have to be kind of clever*
- *E.g., can simulate a moderator if we can talk (may need crypto, though)*
- *Or, can use private function of public randomness (e.g., headline of NY Times)*

How good is equilibrium?

- *Does an equilibrium tell you how to play?*
- *Sadly, no.*
  - *while CE included reasonable answer, also included lots of others*
- *To get further, we'll need additional assumptions*

Bargaining

Bargaining

- *In the standard model of a matrix game, players can't communicate*
- *To allow for bargaining, we will extend the model with **cheap talk***

Cheap talk

- *Players get a chance to talk to one another before picking their actions*
- *They cay say whatever they want—lie, threaten, cajole, or even be honest*
  - *“cheap” because no guarantees*
- *What will happen?*

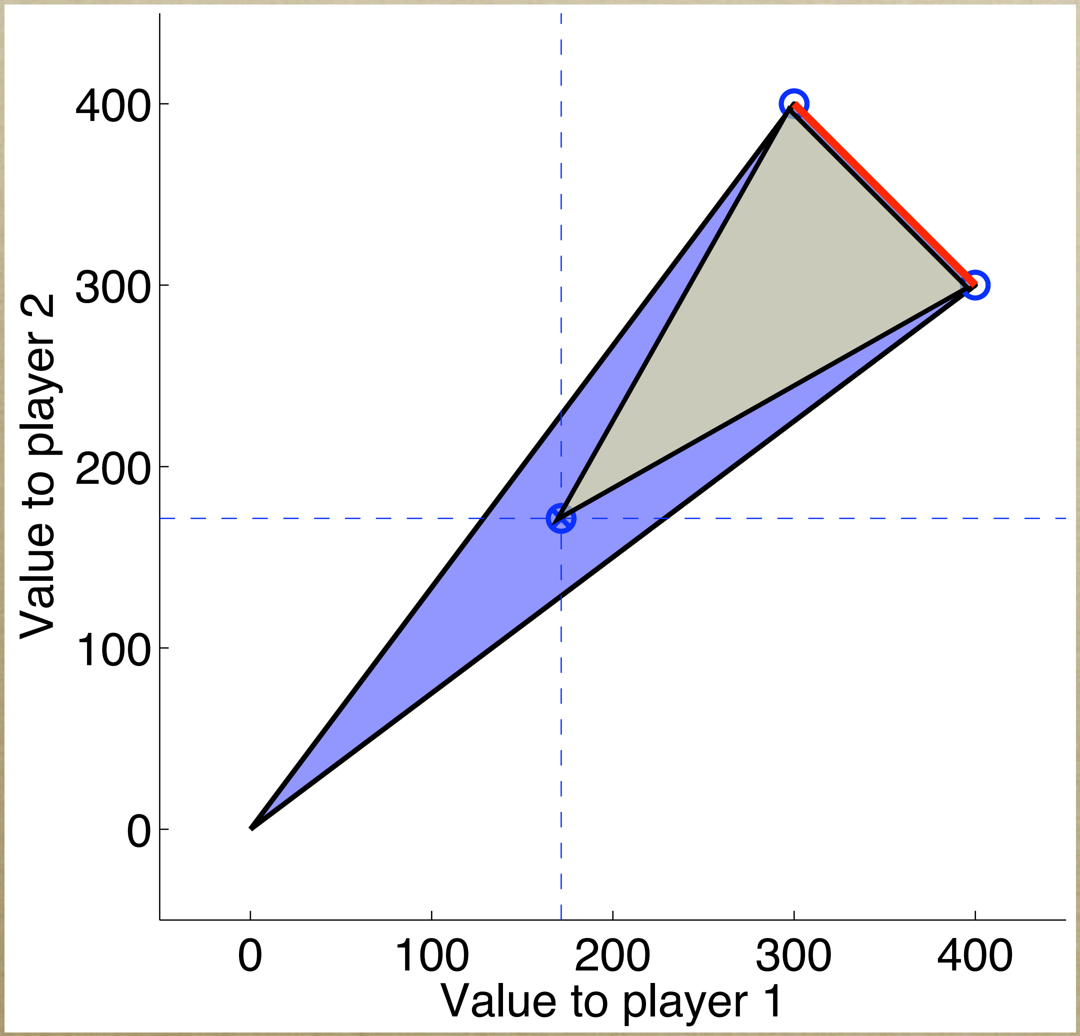
Coordination

- *Certainly the players will try to* ***coordinate***
- *That is, they will try to agree on an equilibrium*
    - *agreeing on a non-equilibrium will lead to deviation*
- *But which one?*

Which one?

- *In Lunch, there are 3 Nash equilibria*
  - *and 5 corner CE + combinations*
- *Players could agree on any one, or agree to randomize among them*
  - *e.g., each simultaneously say a binary number, XOR together, use result to pick equilibrium*

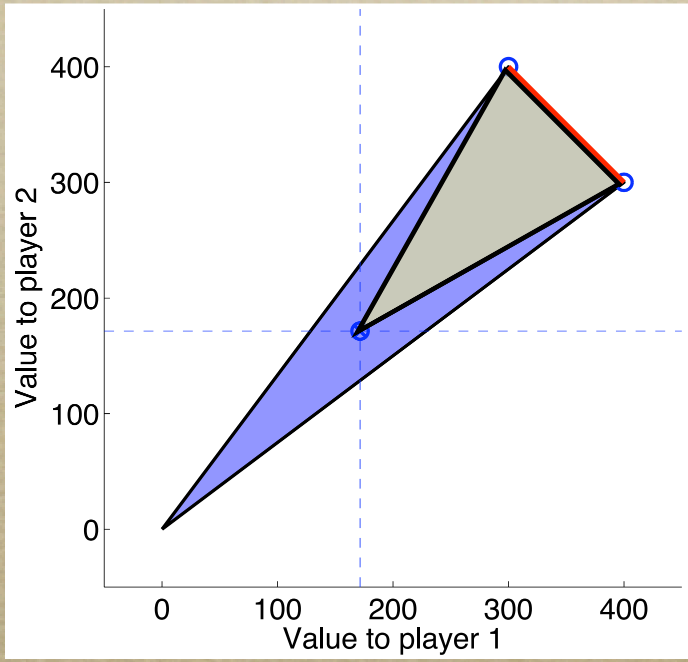
Which one?



400
300
200
100
0
Value to player 2
0
100
200
300
400
Value to player 1

Pareto dominance

- *Not all equilibria are created equal*
- *For any in brown triangle's interior, there is one on red line that's better for **both** players*
- *Red line = **Pareto dominant***



Beyond Pareto

- *We still haven't achieved our goal of actually predicting what will happen*
- *We've narrowed it down a lot: Pareto-dominant equilibria*
- *Further narrowing is the subject of much argument among game theorists*

So let's try it

	A	*U*
A	*3, 4*	*0, 0*
U	*0, 0*	*4, 3*

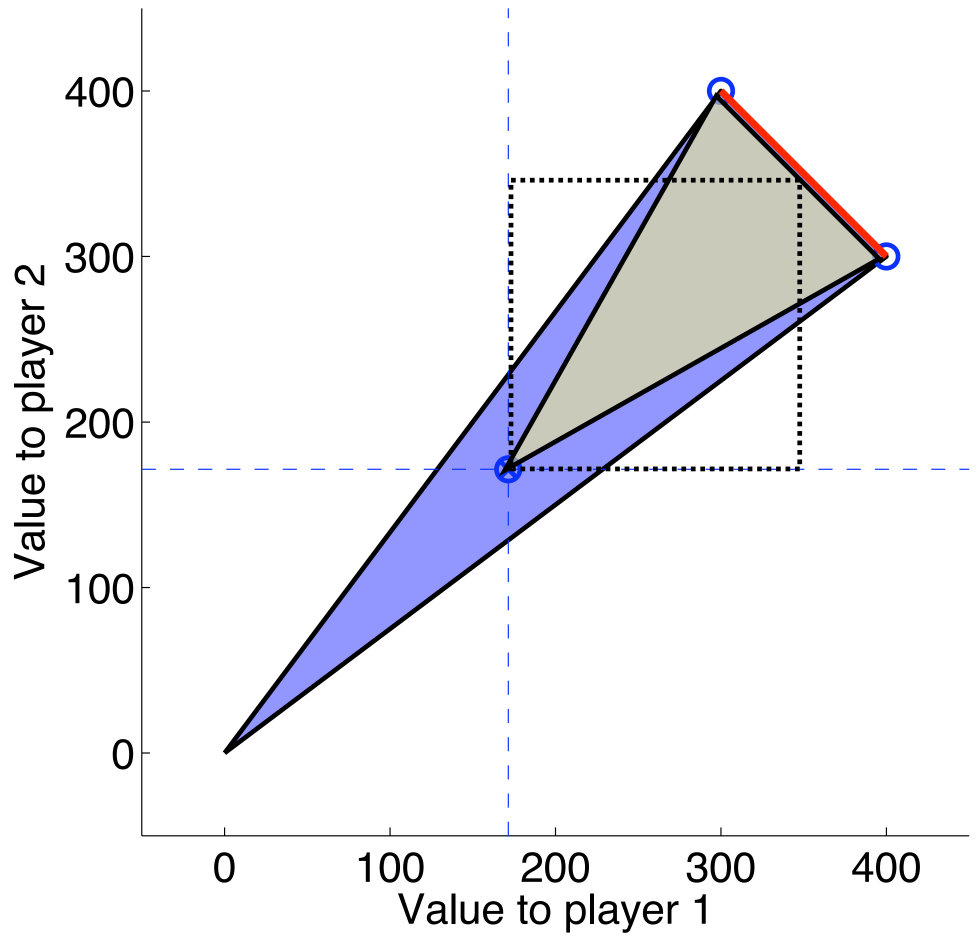
A = Ali Baba, U = Union Grill

Nash bargaining solution

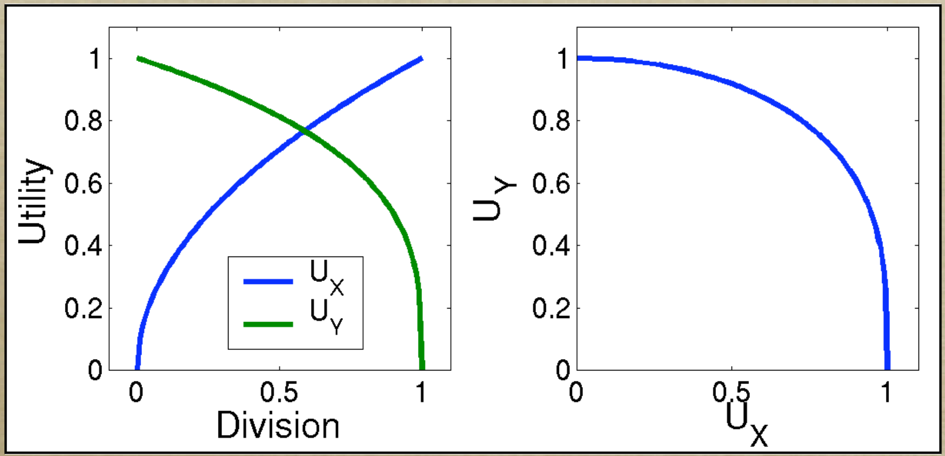
- *Nash built model of bargaining process*
- *Rubinstein later made the model more detailed and implementable*
- *Model includes offers, threats, and impatience to reach an agreement*
- *In this model, we finally have a unique answer to "what will happen?"*

Nash bargaining solution

- *Predicts players will agree on the point on Pareto frontier that maximizes product of extra utility*
- *Invariant to axis rescaling, player exchanging*



Rubinstein's game

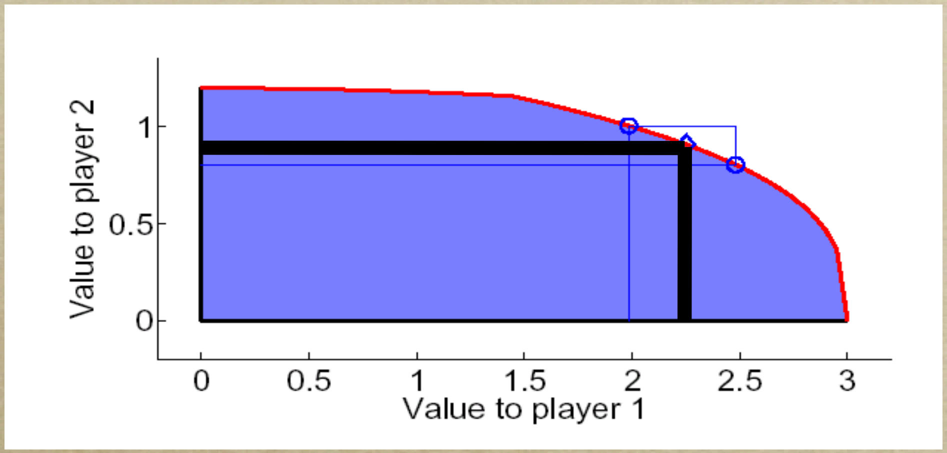


- *Two players split a pie*
- *Each has concave, increasing utility for a share in [0,1]*

Rubinstein's game

- *Bargain by alternating offers:*
  - *Alice offers 60-40*
  - *Bob says no, how about 30-70*
  - *Alice says no, wants 55-45*
  - *Bob says OK*
- *Alice gets* $\gamma^2 U_A(0.55)$*, Bob:* $\gamma^2 U_B(0.45)$
- *In case of disagreement, no pie for anyone*

Theorem

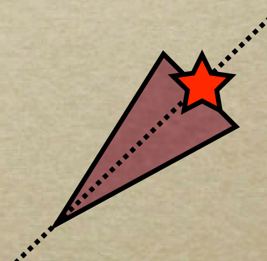


- *In this model, we can finally predict what "rational" players will do*
- *Will arrive (near)* ***Nash bargaining point****, which maximizes product of extra utilities*

$$(U_1 - min_1)\,(U_2 - min_2)$$

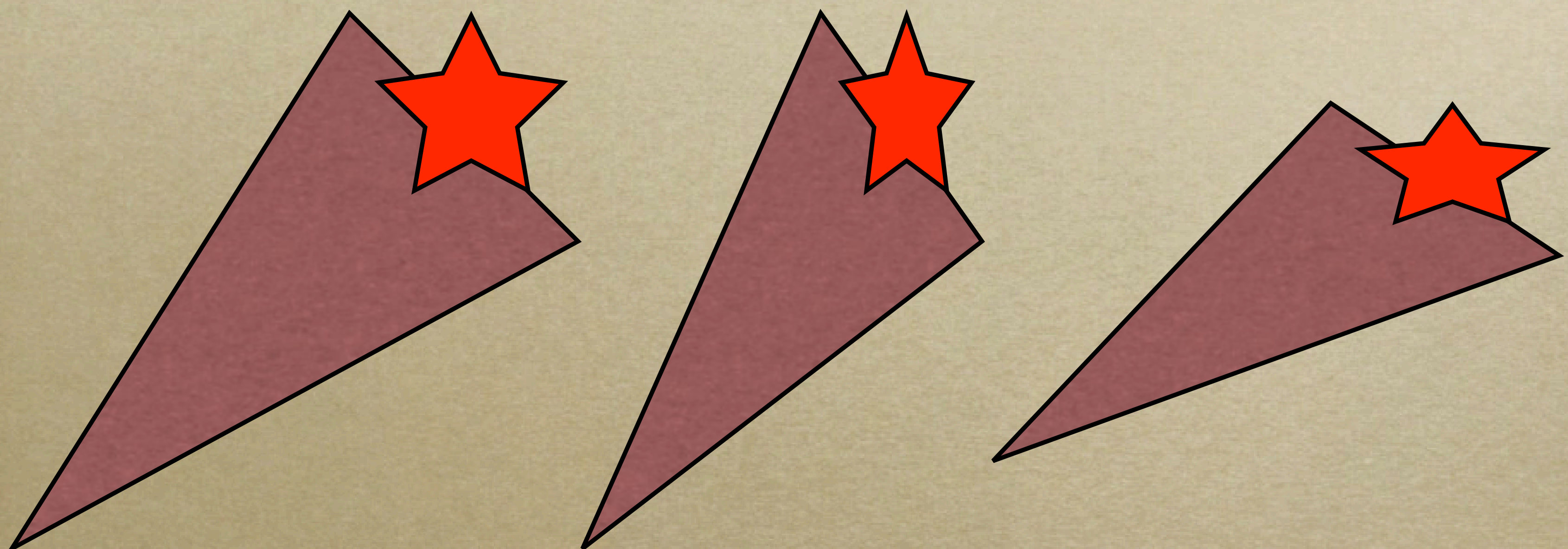
Theorem

- *NBP is **unique** outcome that is*
  - ***optimal** (on Pareto frontier)*
  - ***symmetric** (utilities are equal if possible outcomes are symmetric)*

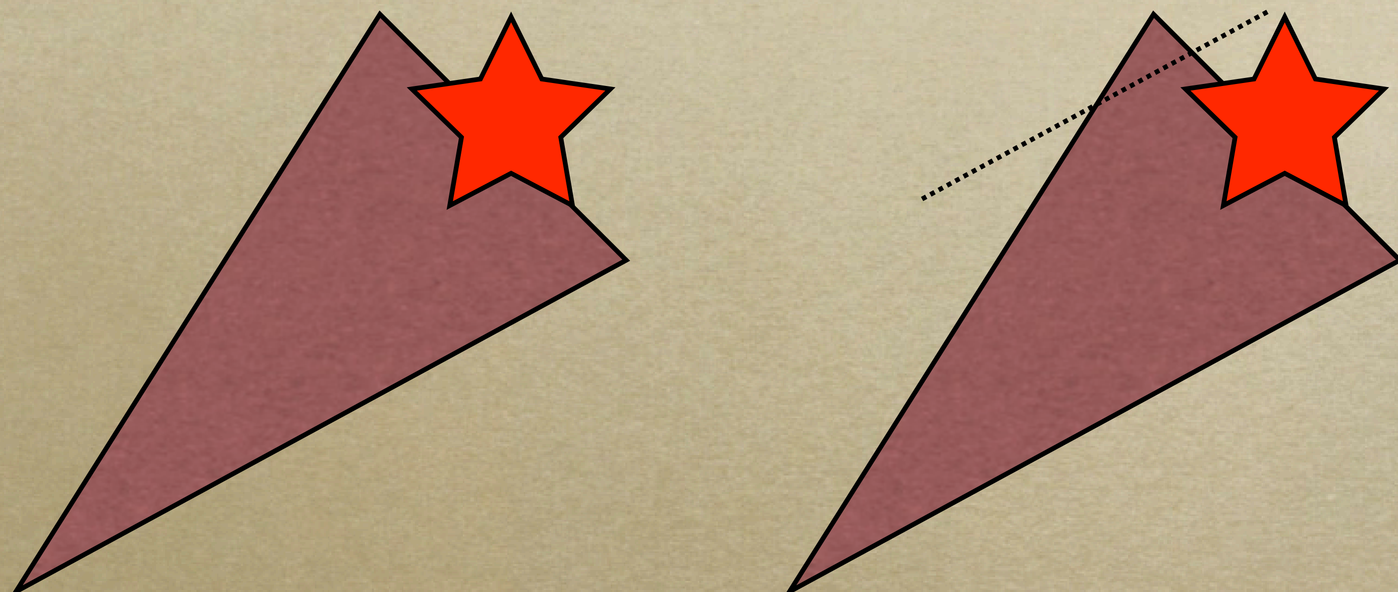


  - ***scale-invariant***
  - ***independent of irrelevant alternatives***

Scale invariance

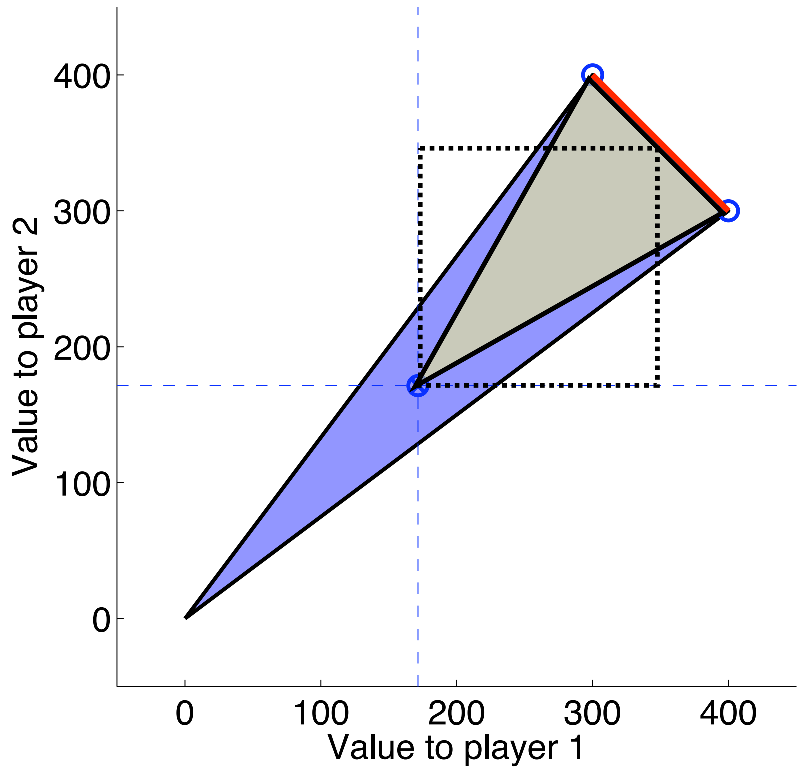


Independence of irrelevant alternatives



Lunch with Rubinstein

- *Use Rubinstein's game to predict outcome of Lunch*
- *Offer = "let's play this equilibrium"*
- *Arrive at "rational" solution*



Bargaining over time

Bargaining over time

- *If we're playing more than once, life gets really interesting*
- *Threats, promises, punishment, trust, concessions, …*

A political game

	C	*W*	*O*
C	*–1, 5*	*0, 0*	*–5, –3*
W	*0, 0*	*0, 0*	*–5, –3*
O	*–3, –10*	*–3, –10*	*–8, –13*

A political game