15-780: Graduate AI
Lecture 8. Games

Geoff Gordon (this lecture)
Ziv Bar-Joseph
TAs Michael Benisch, Yang Gu

Last time, on Grad AI

Optimization

- *Unconstrained optimization: gradient = 0*
- *Equality-constrained optimization*
  - *Lagrange multipliers*
- *Inequality-constrained: either*
  - *nonnegative multipliers (last W), or*
  - *search through bases (simplex, on M)*

Duality

- *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 and Slater's condition*
- *Cutting planes (how to use, how to find)*
  - *generally, e.g., Gomory*
  - *problem specific, e.g., subtour elimination for TSPs*
- *Branch and cut*

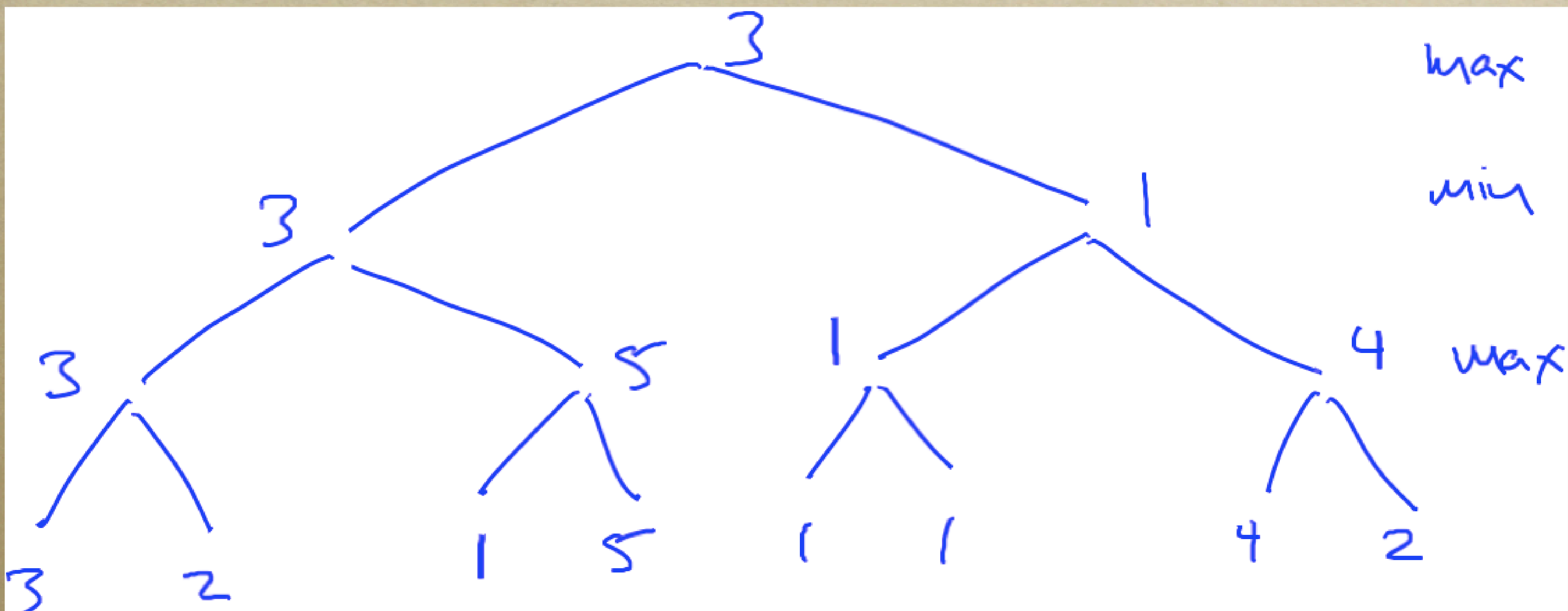
Historical note

- *Gomory's cuts weren't the first poly-time cuts: e.g., Dantzig in 1959*
- *But they were first to guarantee finite termination of cutting plane method for ILPs*
- *Proven by Gomory in 1963*

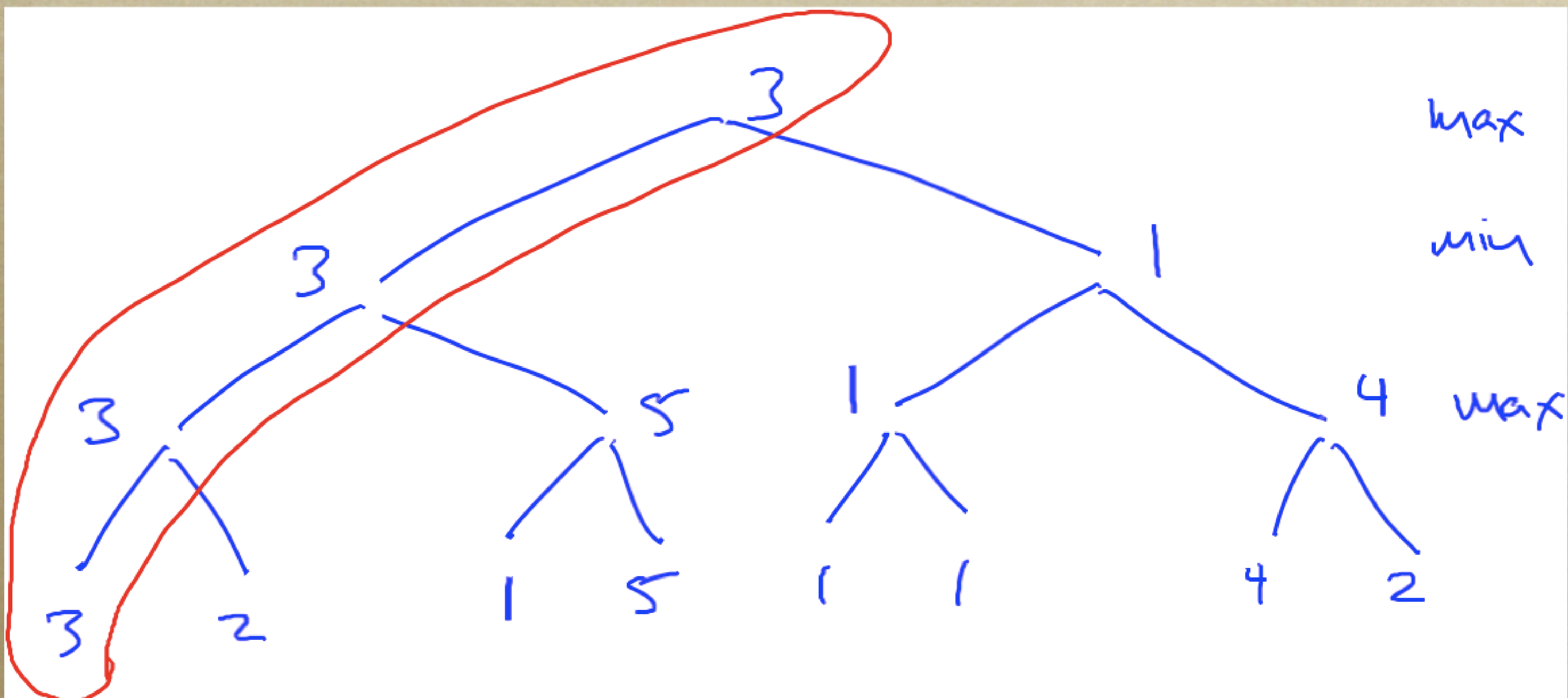
Game search


8
7
6
5
4
3
2
1
a b c d e f g h

Synthetic example



Principal variation



Making it work

- *Minimax is all well and good for small games*
- *But what about bigger ones? 2 answers:*
  - *cutting off search early (big win)*
  - *pruning (smaller win but still useful)*

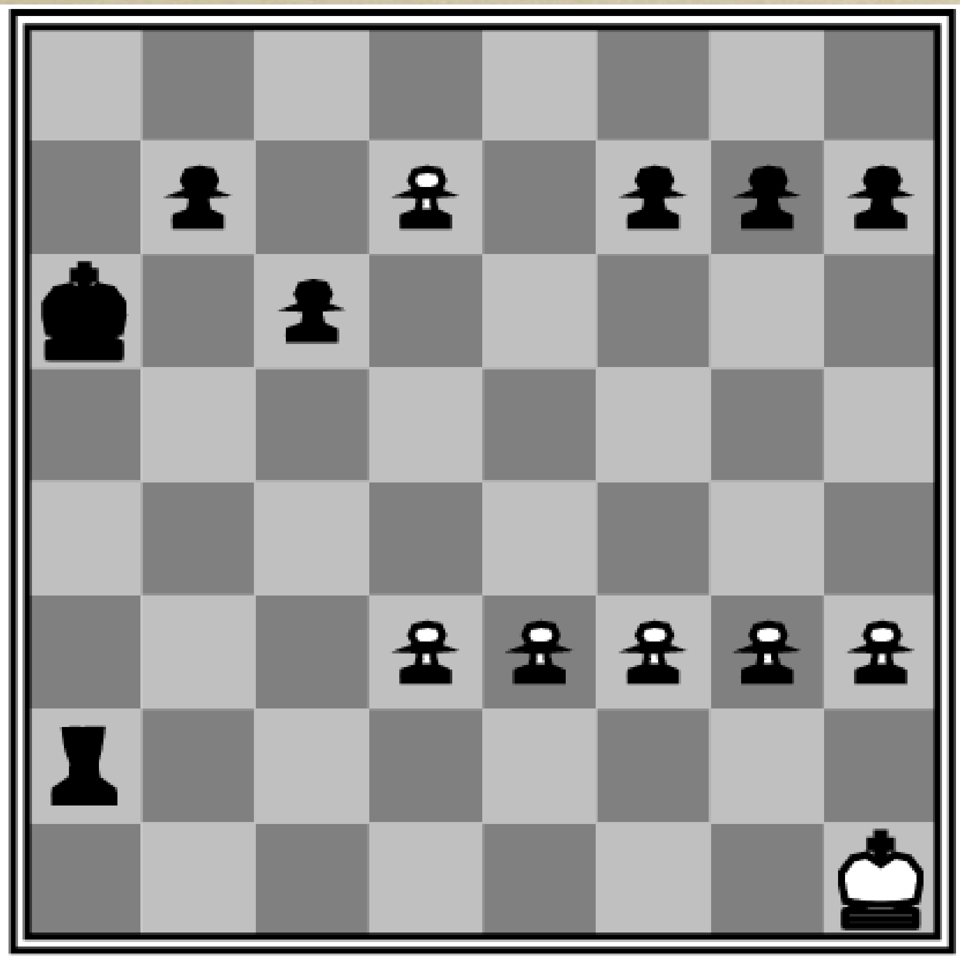
Heuristics

- *Quickly and approximately evaluate a position without search*
- *E.g., Q = 9, R = 5, B = N = 3, P = 1*
- *Build out game tree as far as we can, use heuristic at leaves in lieu of real value*
  - *might want to build it out unevenly (more below)*

Heuristics

- *Deep Blue used: materiel, mobility, king position, center control, open file for rook, paired bishops/rooks, … (> 6000 total features!)*
- *Weights are context dependent, learned from DB of grandmaster games then hand tweaked*

Quiescence

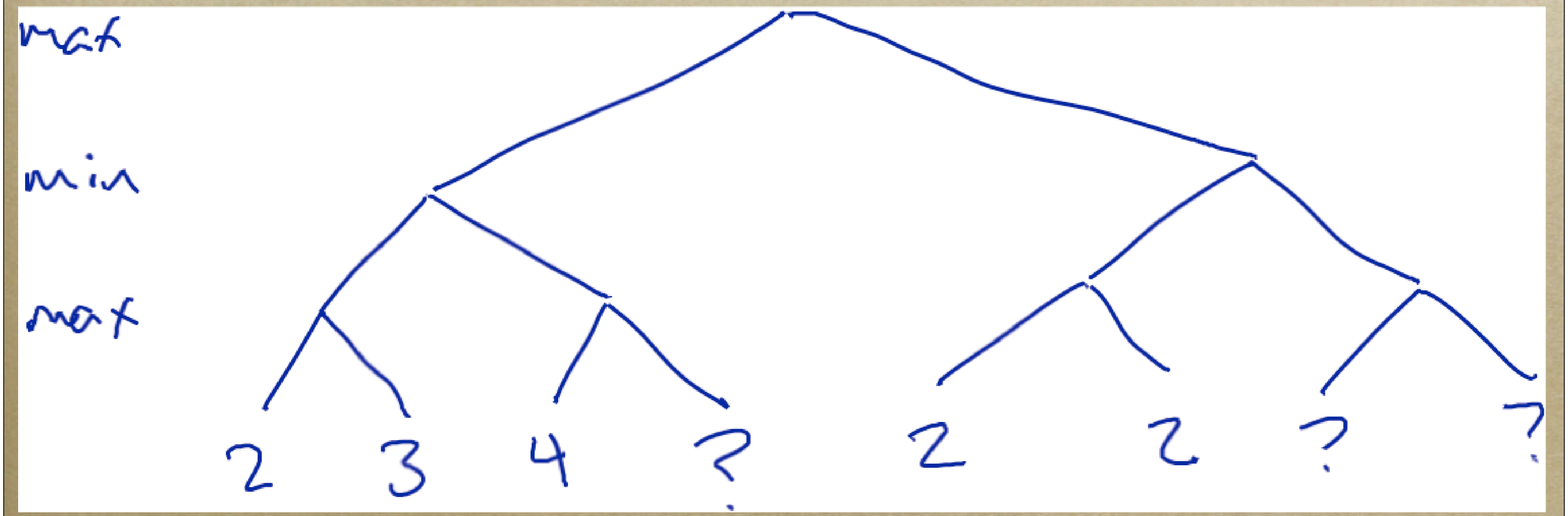


Black to move

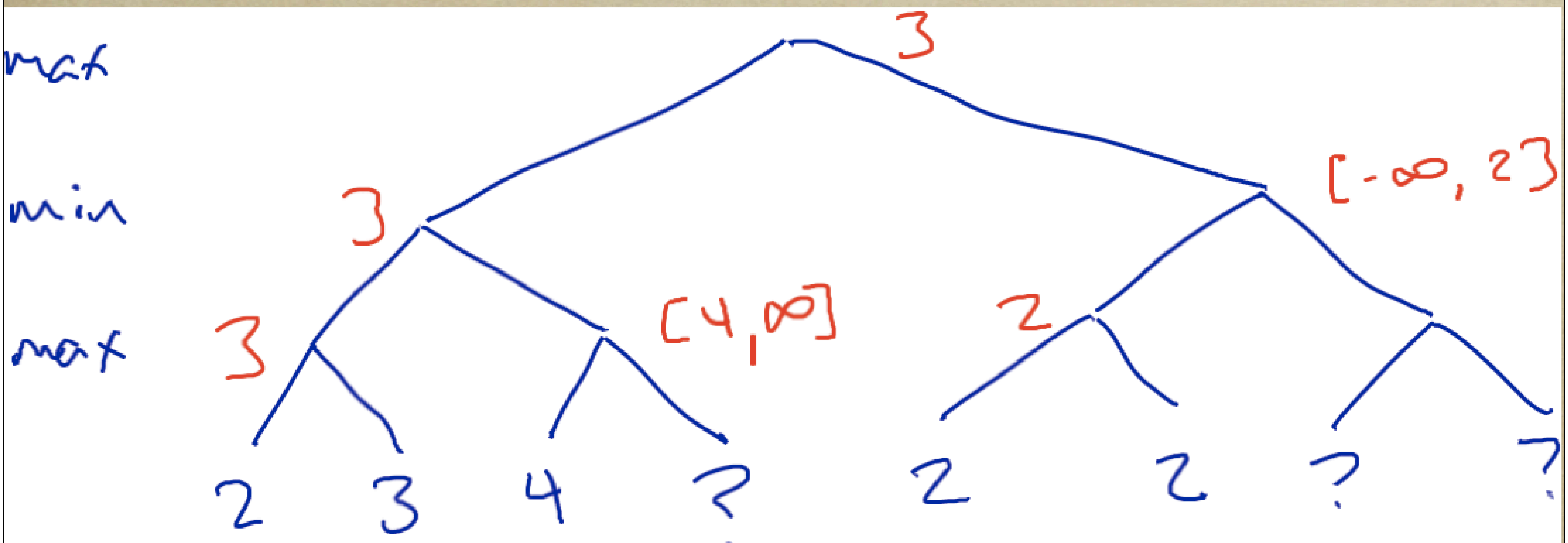
Pruning

- *Idea: don't bother looking at parts of the tree we can prove are irrelevant*

Pruning example



Pruning example



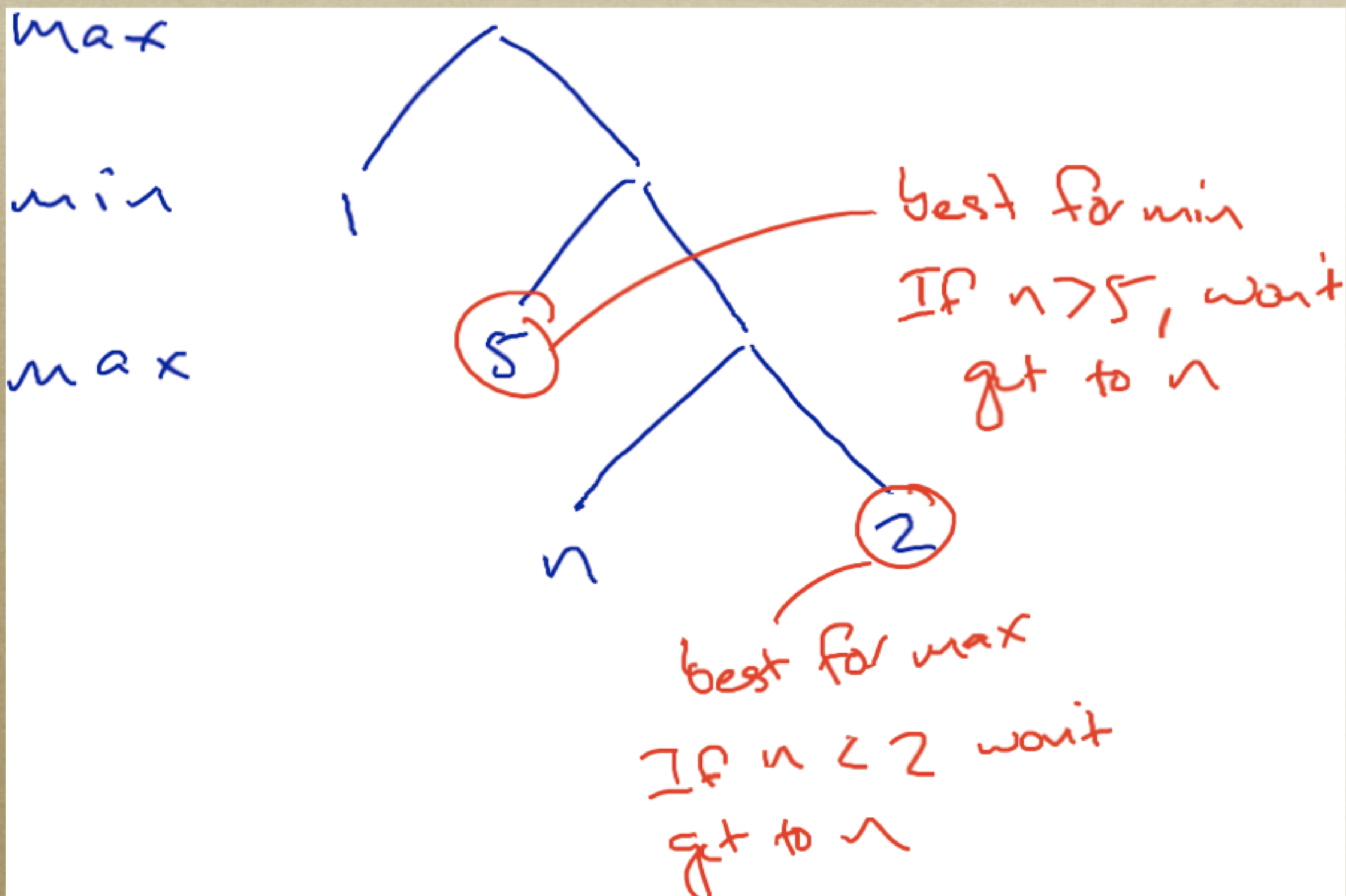
Alpha-beta pruning

- *Do a DFS through game tree*
- *At each node n on stack, keep bounds*
  - α(*n*)*: value of best deviation so far for MAX along path to n*
  - β(*n*)*: value of best deviation so far for MIN along path to n*

Alpha-beta pruning

- *Deviation = way of leaving the path to n*
- *So, to get α,*
  - *take all MAX nodes on path to n*
  - *look at all their children that we've finished evaluating*
  - *best (highest) of these children is α*
- *Lowest of children of MIN nodes is β*

Example of alpha and beta



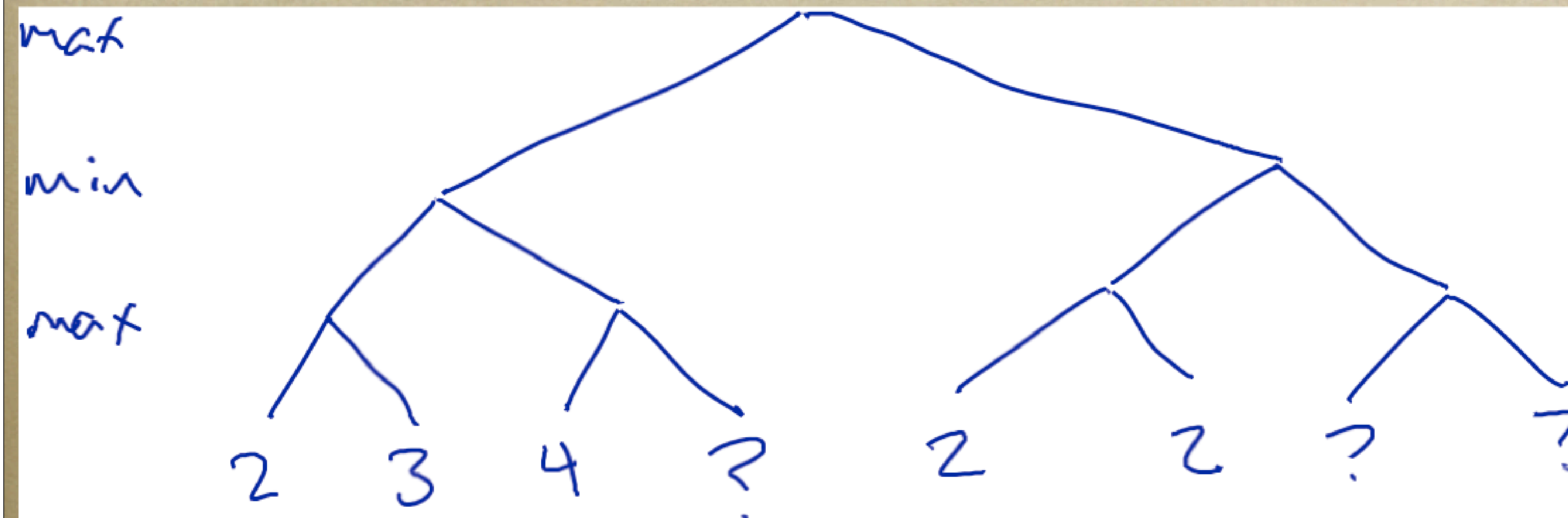
Alpha-beta pruning

- *At max node:*
  - *receive* α *and* β *values from parent*
  - *expand children one by one*
  - *update* α *as we go*
  - *if* α *ever gets higher than* β, *stop*
  - *won't ever reach this node (return* α*)*

Alpha-beta pruning

- *At min node:*
  - *receive* α *and* β *values from parent*
  - *expand children one by one*
  - *update* β *as we go*
  - *if* β *ever gets lower than* α, *stop*
  - *won't ever reach this node (return* β*)*

Example



How much do we save?

- *Original tree: b^d nodes*
  - *b = branching factor*
  - *d = depth*
- *If we expand children in random order, pruning will touch $b^{3d/4}$ nodes*
- *Lower bound (best node first): $b^{d/2}$*
- *Can often get close to lower bound w/* ***move ordering heuristics***

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*

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 $x' M_R y$ and $x' M_C y$*

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?*

Safety value

- *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, …*
- *Best worst-case value is **safety value** or **minimax value** of game*
  - $\max_x \min_y x' M_R y$

What else?

- *If the world really is out to get us, the safety value is the end of the story*
- *This is the case 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 (famous theorem of Nash)*
- *A strategy that guarantees minimax value is a **minimax strategy***
- *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:*
- $min_x\ max_{y,z}\ \ z(1 - \mathbf{1}'x) + x'My$ *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*

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

    $\mathbf{1}'y = 1$

    $y \geq 0$

For example

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

st

$$z \le y_H - y_T$$
$$z \le -y_H + y_T$$

$$y_H + y_T = 1$$

$$x, y \ge 0$$

Interpreting LP

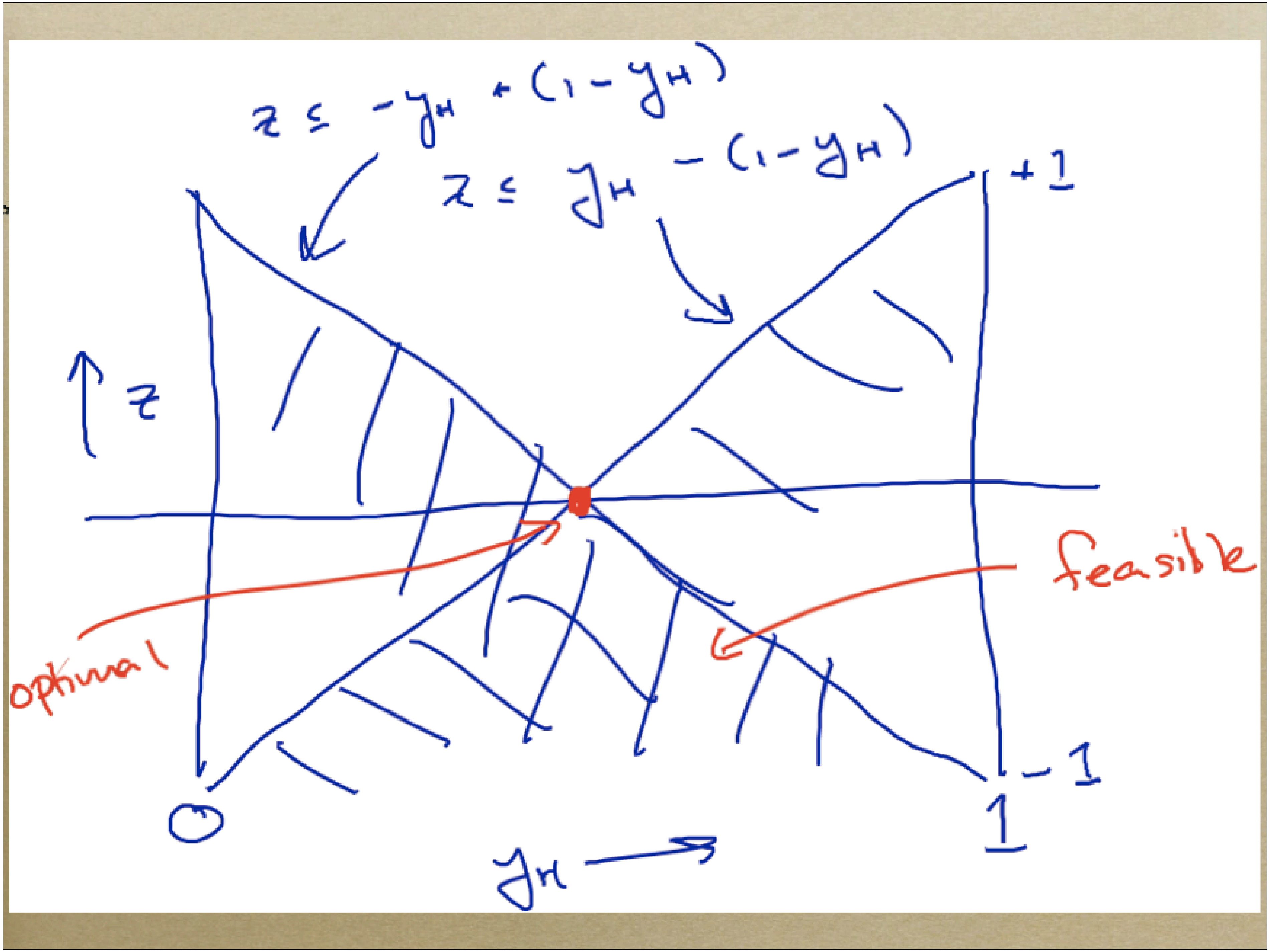
- $max_{y,\,z}$ z *subject to*

  $My \geq \mathbf{1}z$

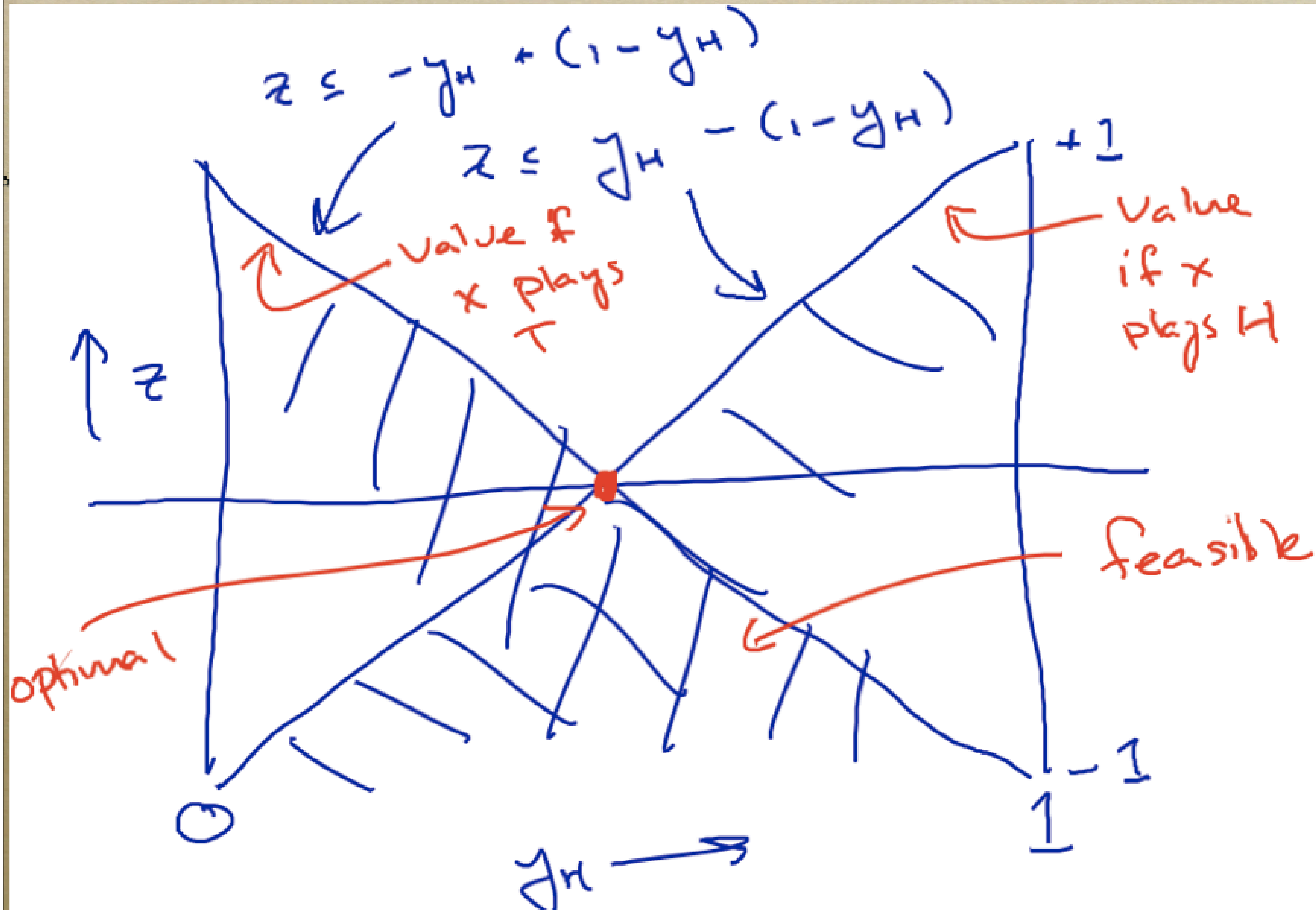
  $\mathbf{1}'y = 1$

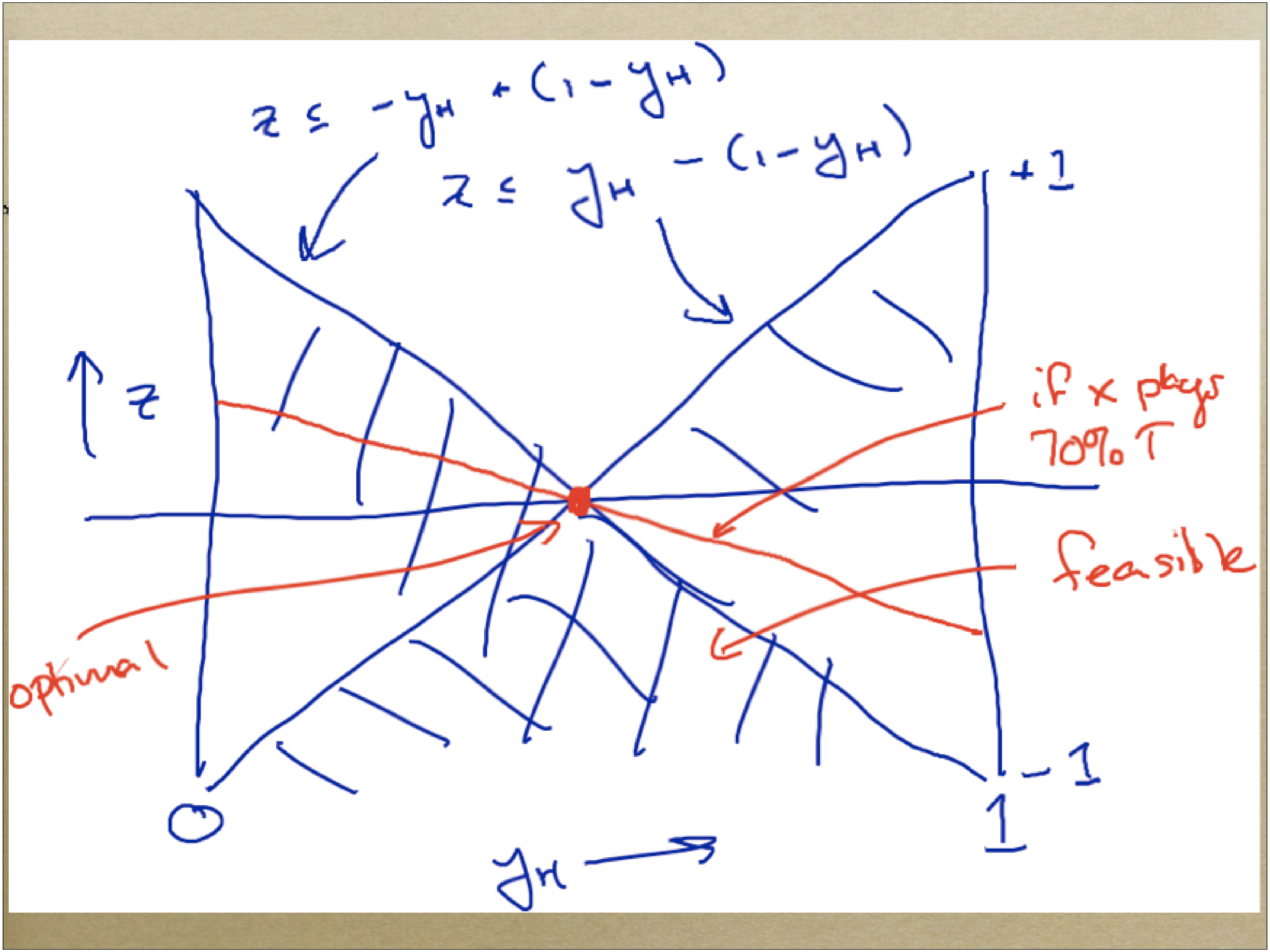
  $y \geq 0$

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

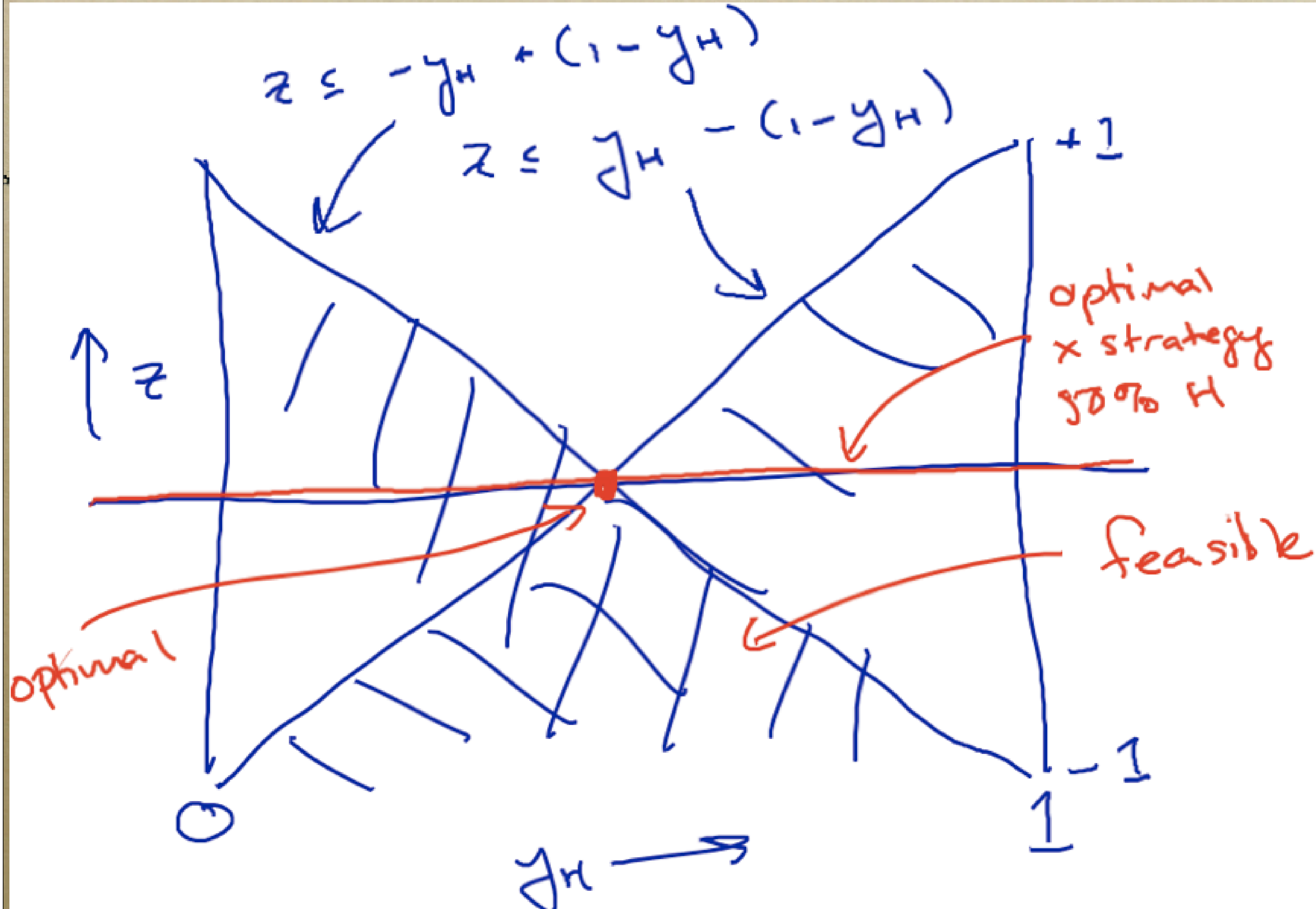


z ≤ -yH + (1 - yH)
z ≤ yH - (1 - yH)
+1
-1
z
0
1
yH
feasible
optimal


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



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



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

Duality

- *x is dual variable for My* $\geq \mathbf{1}z$
- *Complementarity: Row can only play strategies where My* $= \mathbf{1}z$
- *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 less-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

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) is a mutually reinforcing pair of strategies—an* ***equilibrium***

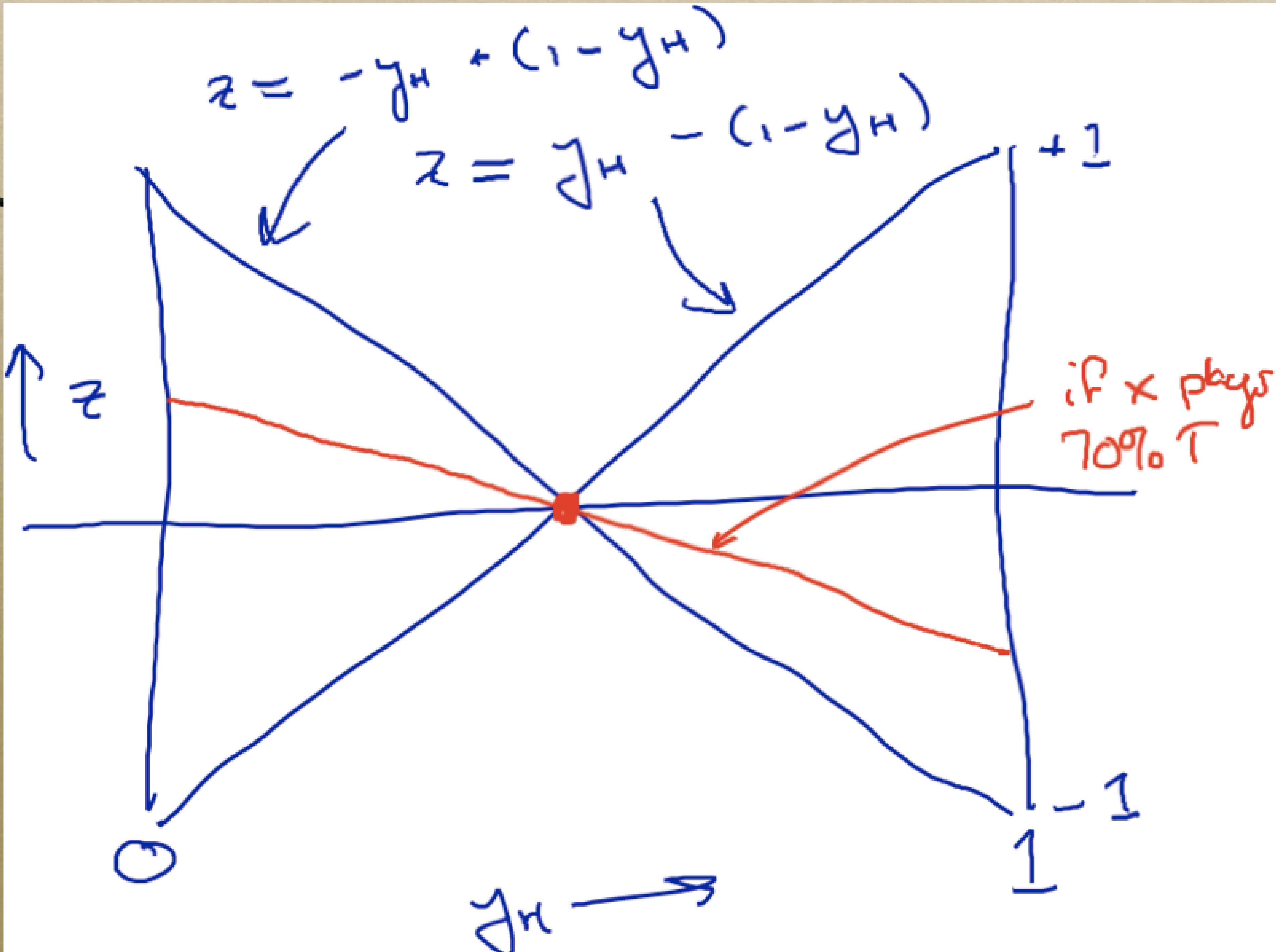
	A	U
A	3, 4	0, 0
U	0, 0	4, 3

Finding equilibria

- *The idea of equilibrium allows us to rule out some more joint strategies beyond what dominance gave us*
- *The particular type of equilibrium we are about to describe is due to Nash*
    - *his name keeps coming up…*

Finding equilibria

- *In a Nash equilibrium, we have a (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*



z = -y_H + (1 - y_H)
z = y_H - (1 - y_H)
+1
z
if x plays 70% T
0
1
-1
y_H

How good is equilibrium?

- *Does an equilibrium tell you how to play?*
- *Sadly, no.*
- *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 two ways:*
  - *first,* ***cheap talk***
  - *second, a* ***moderator***

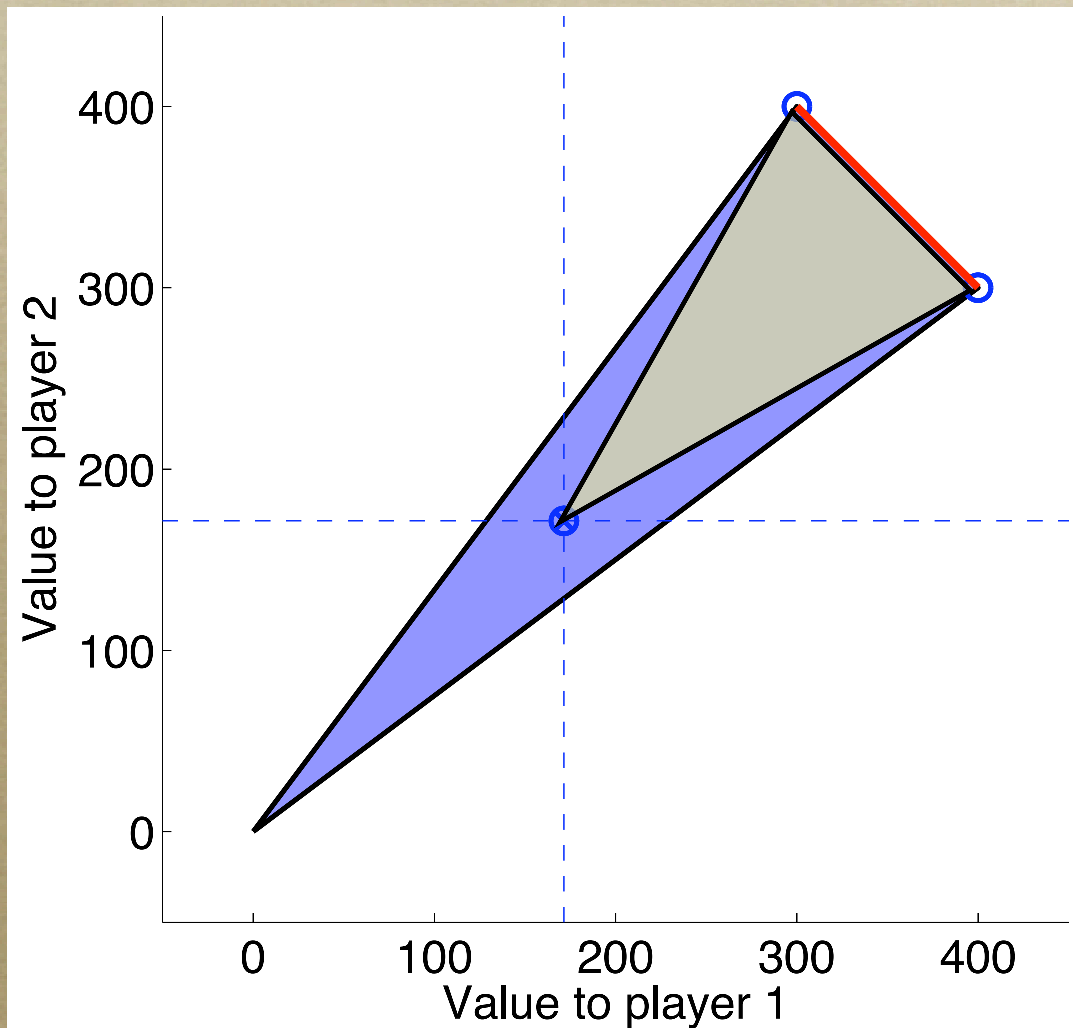
Cheap talk

- *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*
- *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?*

Pareto dominance

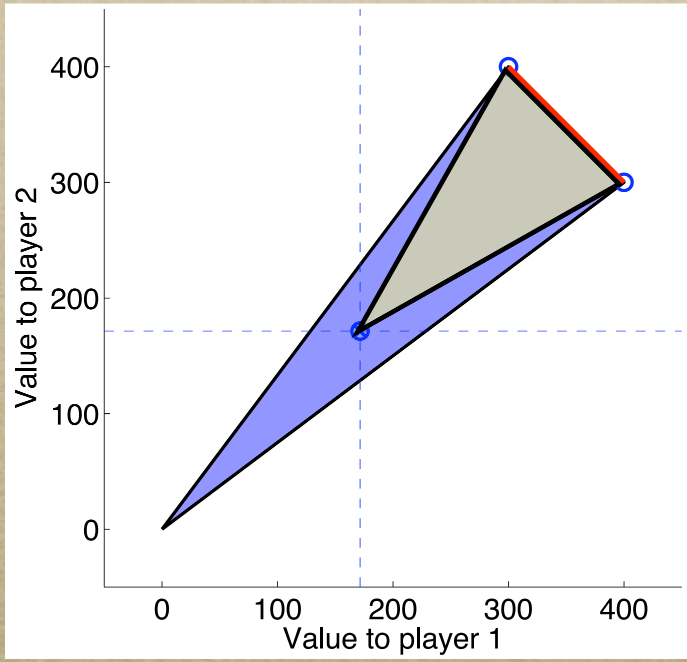


Pareto dominance

- *In Lunch, there are 3 Nash equilibria*
- *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*

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

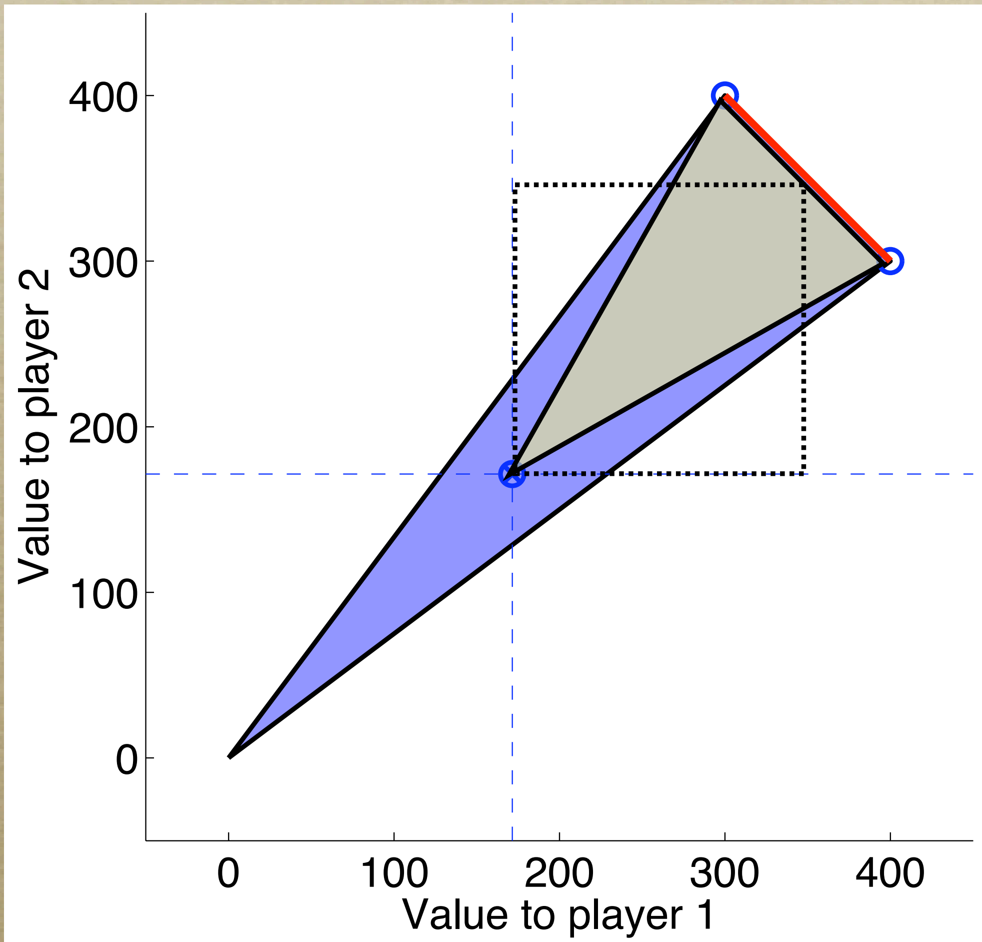
- *We still haven't achieved our goal of actually predicting what will happen*
- *We've narrowed it down a lot: Pareto-dominant equilibria*
- *Further narrowing is the subject of much argument among game theorists*

Nash bargaining solution

- *Nash designed a model of the bargaining process (there's that name again…)*
- *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*



Moderator

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

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)*

Realism?

- *Moderators are often available*
- *Sometimes have to be kind of clever*
- *E.g., can simulate a moderator using cheap talk and some crypto*

Correlated equilibrium

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

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

Correlated equilibrium

- *Probability that Row is recommended to play A = a + b*
- *Given recommendation for A, probability that Col also plays A = a / (a + b)*
- *Rationality: when I'm recommended to play A, I don't want to play B instead*

Correlated equilibrium

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

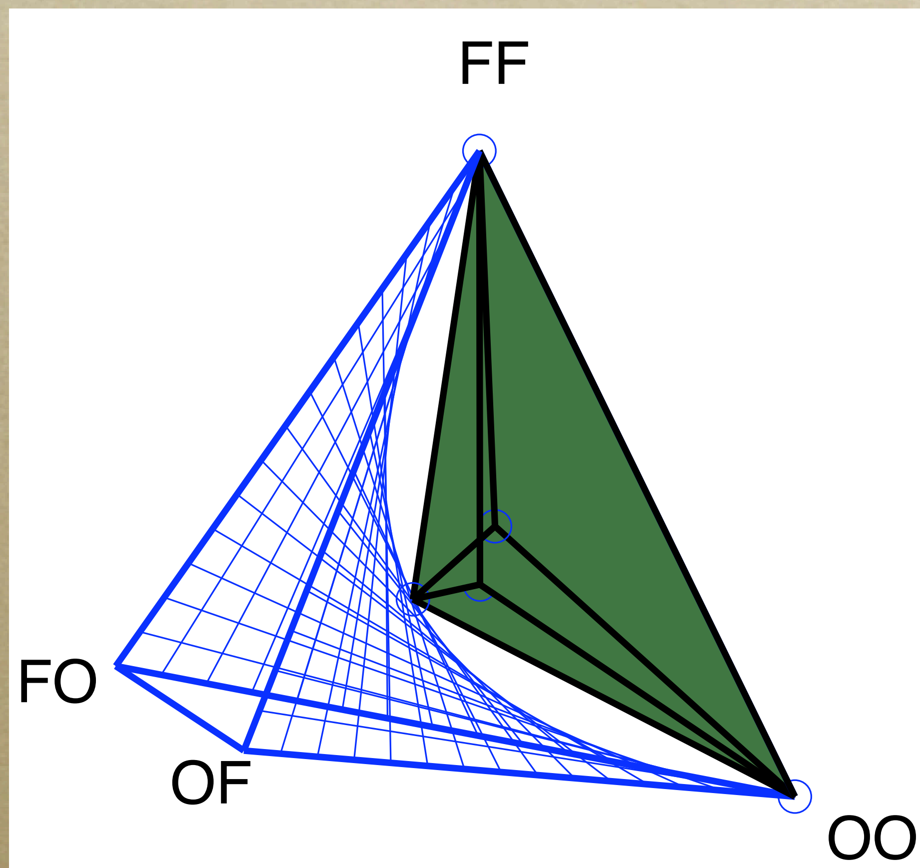
$$4a + 0b \geq 0a + 3b$$

$$0c + 3d \geq 4c + 0d$$

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

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

Correlated equilibrium



Bargaining

- *Can use Nash bargaining model to select among CE*
- *Same results hold: unique answer on Pareto frontier (but now Pareto frontier might be better)*