

Graduate AI

Lecture 21:





Game Theory III

Teachers:

Zico Kolter

Ariel Procaccia (this time)

ZERO-SUM GAMES

	
 -1	1
 1	-1

ZERO-SUM GAMES

- **Maximin** (randomized) strategy of player 1 maximizes the worst-case expected payoff
- In the penalty shot game, optimal strategy for both players is playing $(\frac{1}{2}, \frac{1}{2})$
- In the game below, if shooter uses $(p, 1 - p)$:
 - Jump left: $-\frac{p}{2} + 1 - p = 1 - \frac{3}{2}p$
 - Jump right: $p - 1 + p = 2p - 1$
 - Maximize $\min\{1 - \frac{3}{2}p, 2p - 1\}$ over p

$-\frac{1}{2}$	1
1	-1

ZERO-SUM GAMES

- Denote the reward of player 1 from strategies (s_1, s_2) by $R(s_1, s_2)$
- Maximin strategy is computed via LP:

$\max w$

$$\text{s.t. } \forall s_2 \in S, \sum_{s_1 \in S} p(s_1) R(s_1, s_2) \geq w$$

$$\sum_{s_1 \in S} p(s_1) = 1$$

$$\forall s_1 \in S, p(s_1) \geq 0$$



THE MINIMAX THEOREM

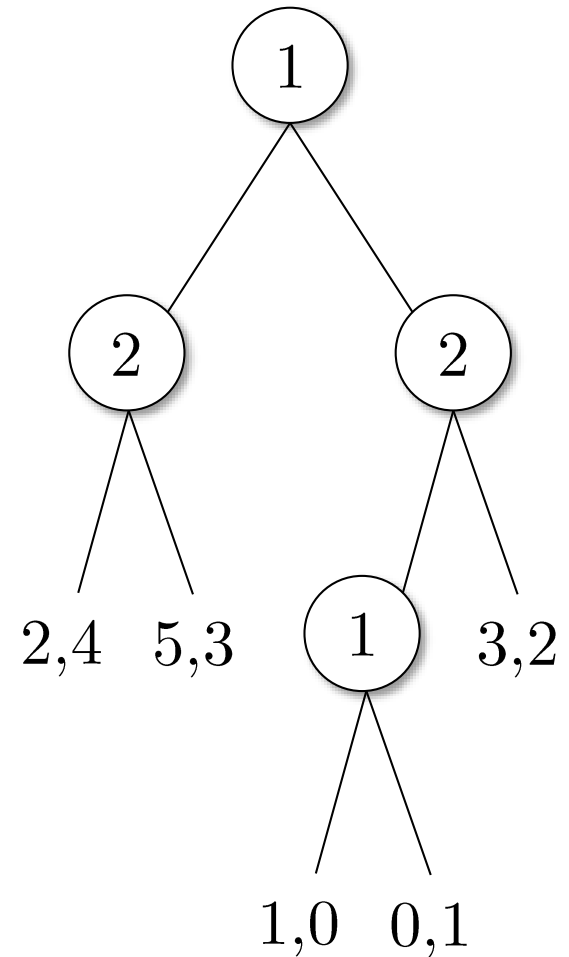
- **Theorem [von Neumann 1928]:**
Every 2-player zero-sum game has a unique value v such that:
 - Player 1 can guarantee value at least v
 - Player 2 can guarantee loss at most v
- **Poll 1:** How many Nash equilibrium **payoffs** do zero-sum games have?
 1. At most one
 2. At least one
 3. Exactly one



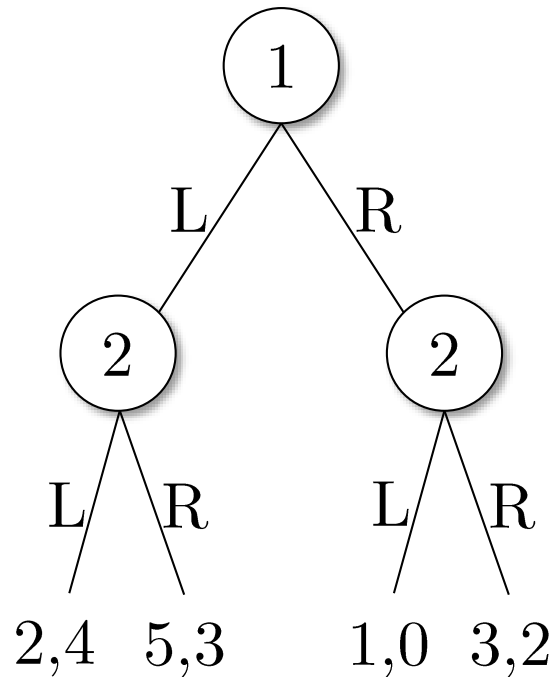
von Neumann

EXTENSIVE-FORM GAMES

- Moves are done sequentially, not simultaneously
- Game forms a tree
- Nodes are labeled by players
- Leaves show payoffs



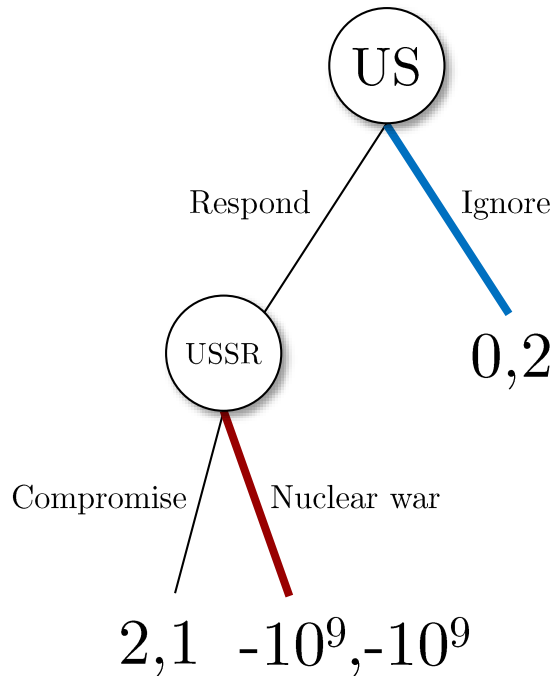
EXTENSIVE VS. NORMAL FORM



	L/L	L/R	R/L	R/R
L	2,4	2,4	5,3	5,3
R	1,0	3,2	1,0	3,2

Problem: Normal-form representation is exponential in the size of the extensive-form representation

EXTENSIVE VS. NORMAL FORM

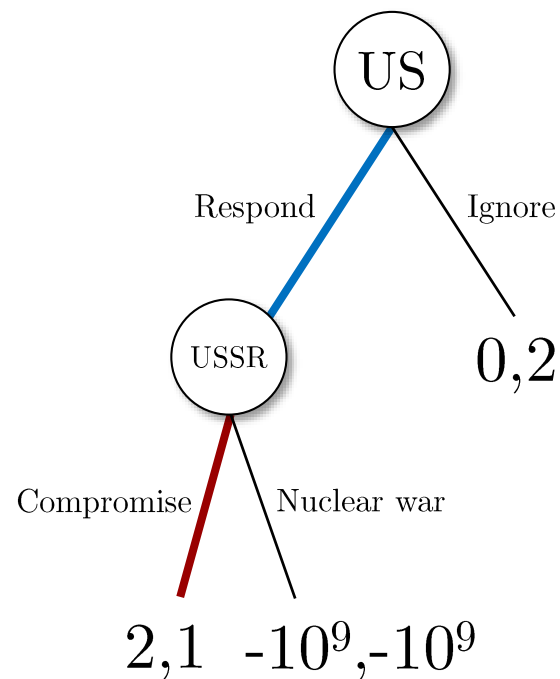


	Compromise	Nuclear war
Respond	2,1	$-10^9, -10^9$
Ignore	0,2	0,2

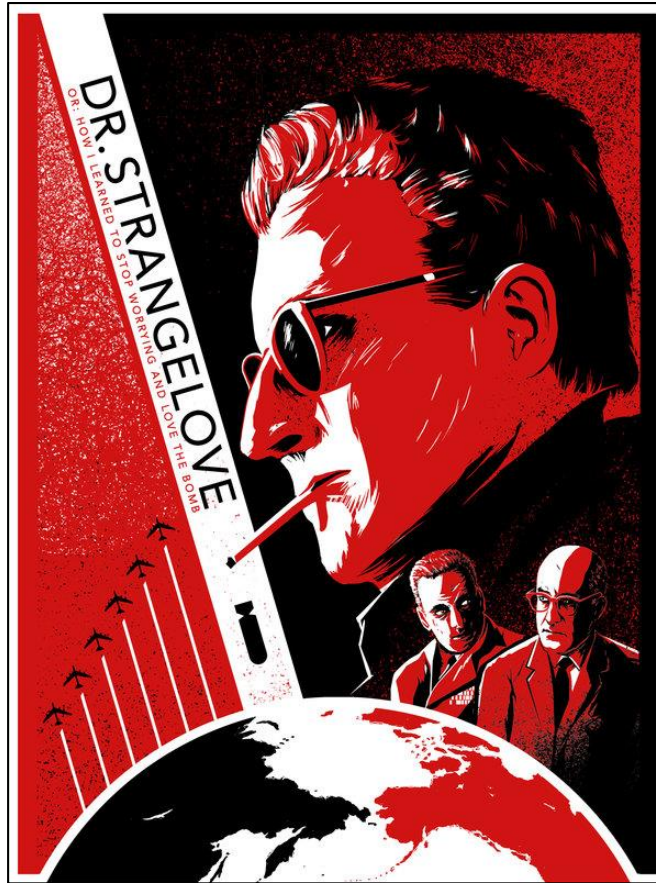
Problem: (ignore, nuclear war) is a Nash equilibrium, but threat isn't credible!

SUBGAME-PERFECT EQUILIBRIUM

- Each subtree forms a subgame
- A set of strategies is a **subgame-perfect equilibrium** if it is a Nash equilibrium in each subgame
- A player may be able to improve his equilibrium payoff by eliminating strategies!

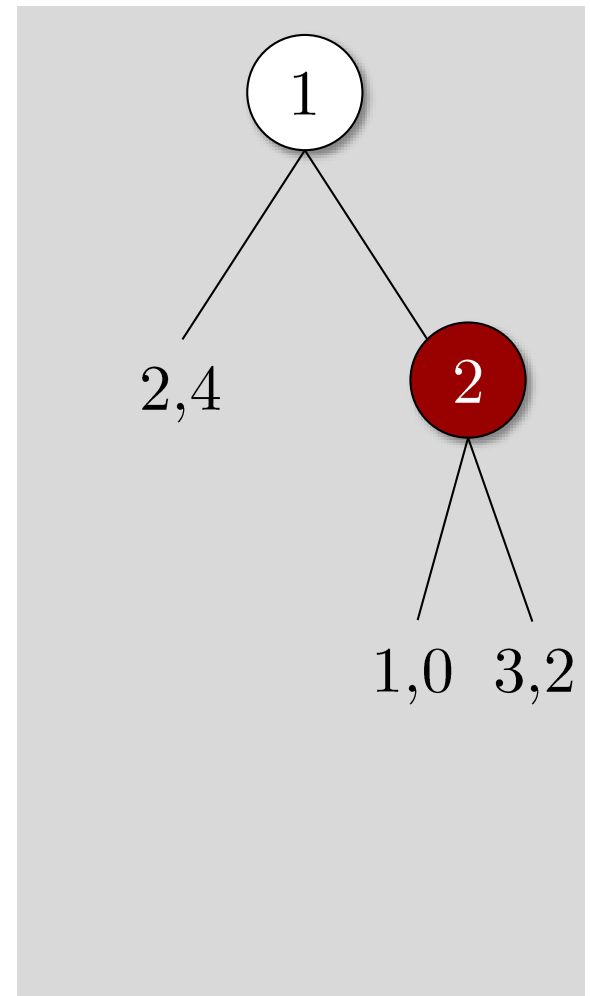
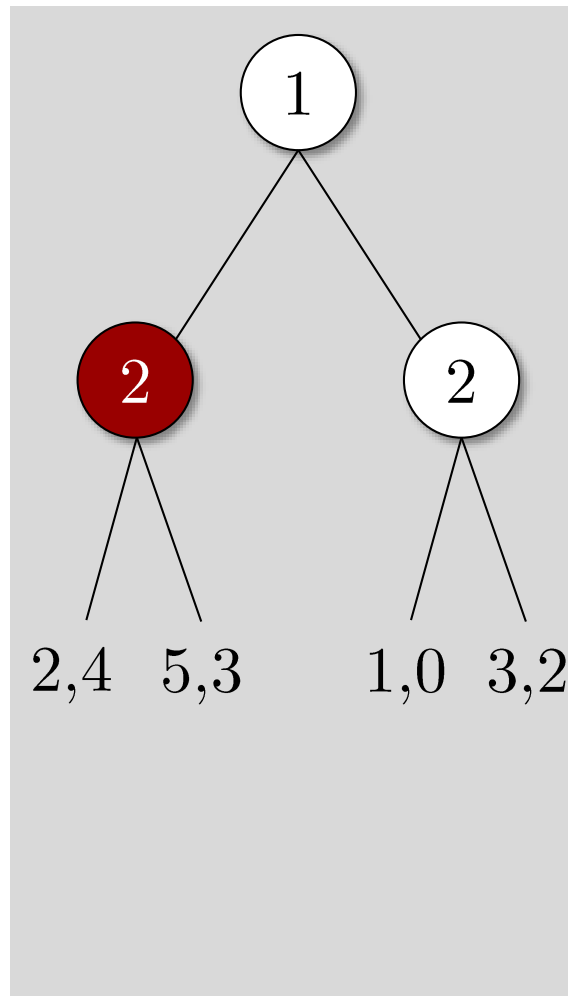
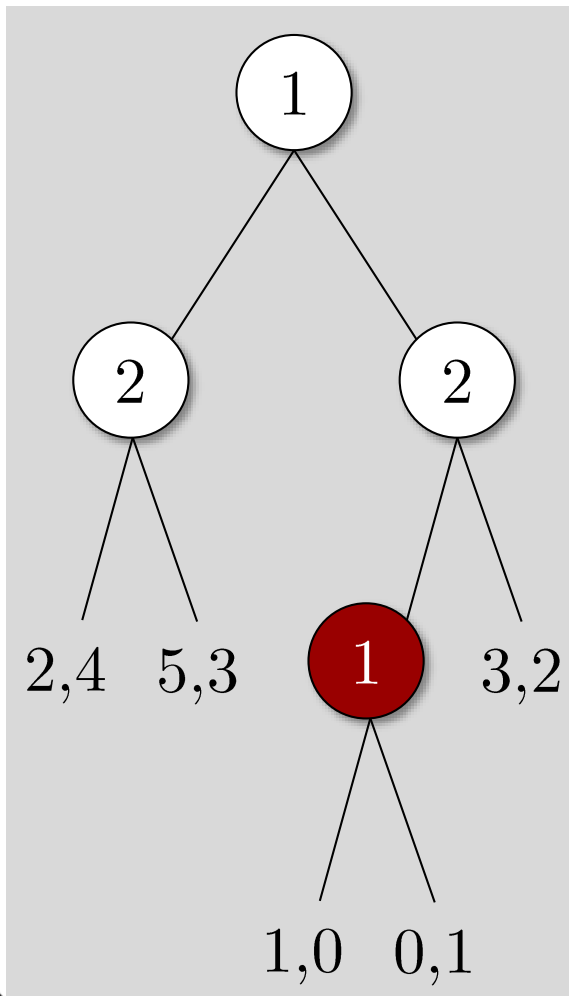


DOOMSDAY MACHINE

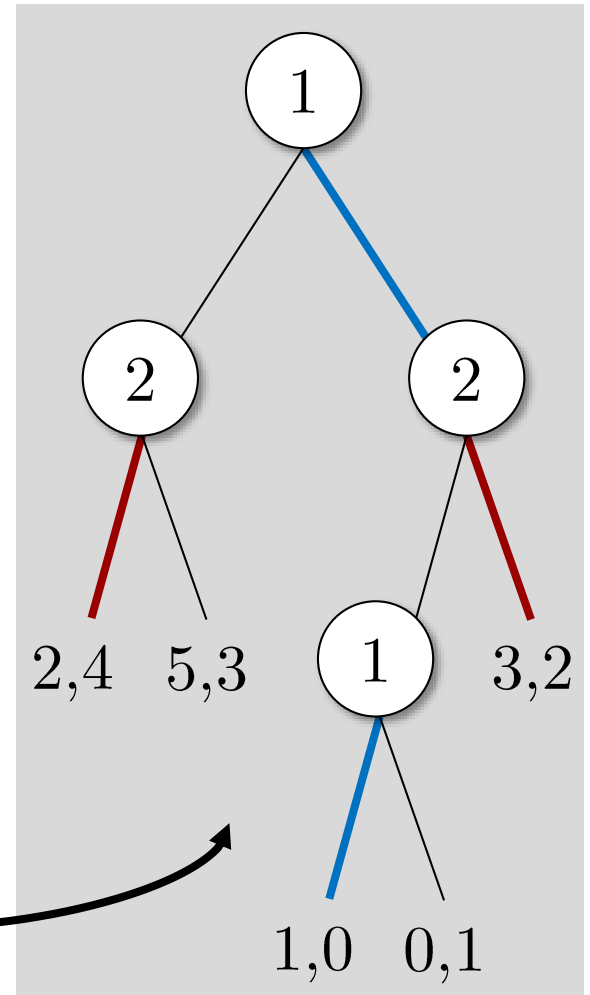
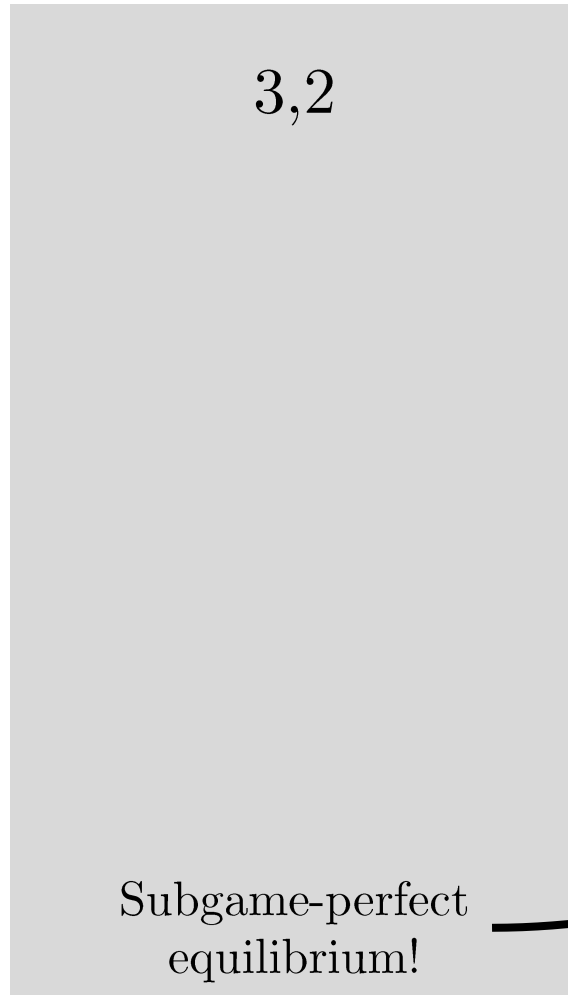
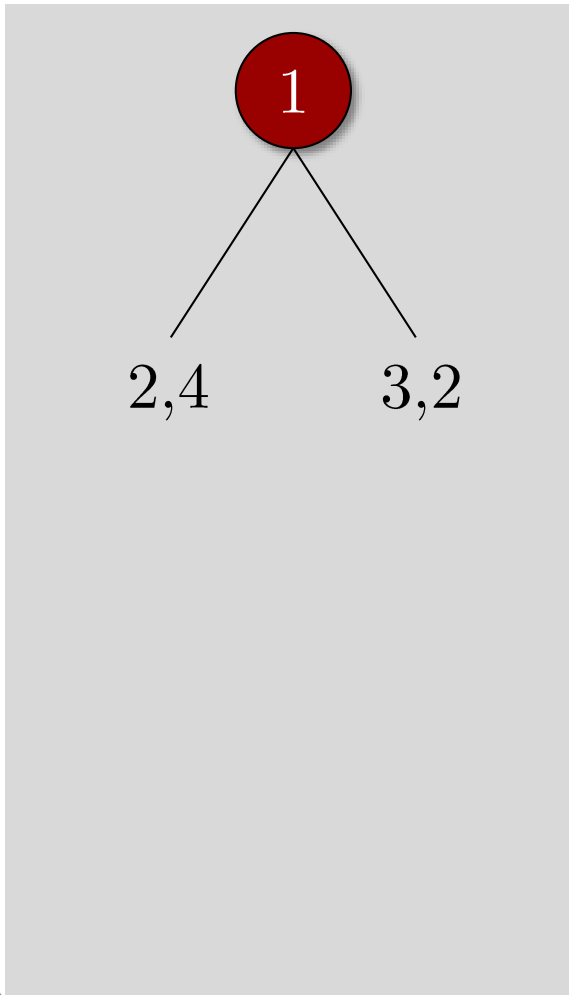


<https://youtu.be/2yfXgu37iyI>

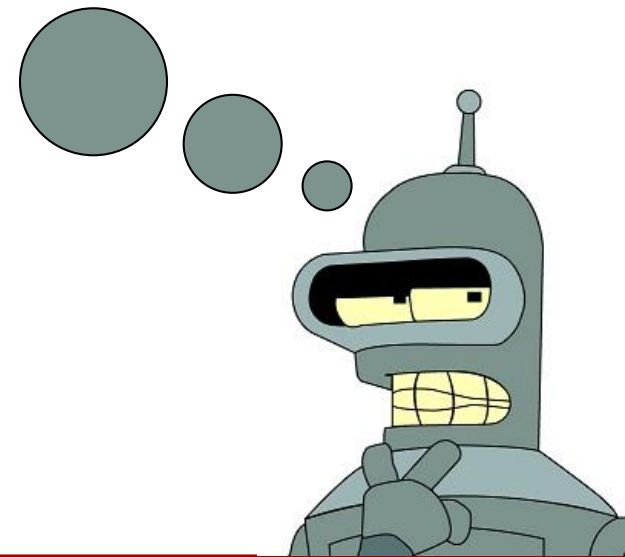
BACKWARD INDUCTION



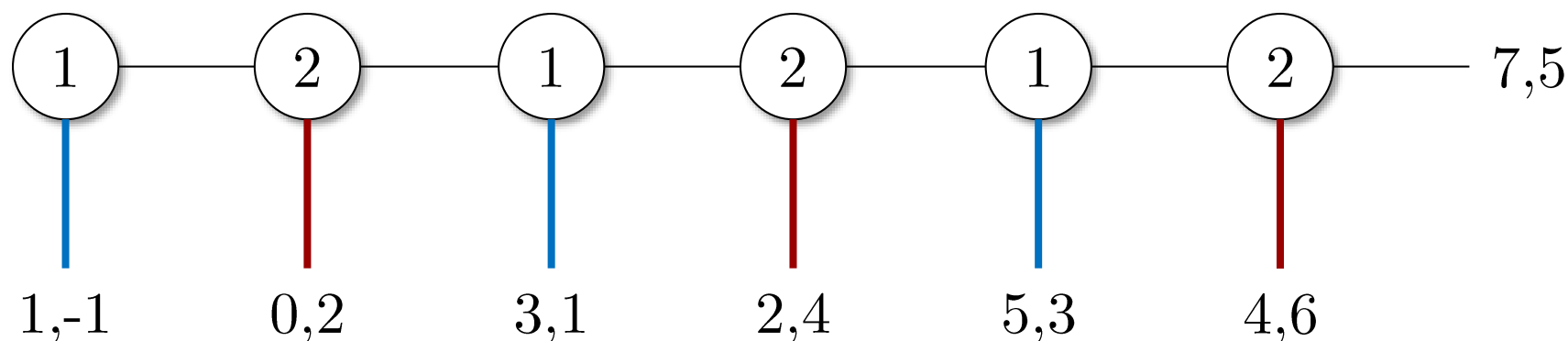
BACKWARD INDUCTION



Extensive-form games
can be represented as
normal-form games.
How come they
always have a pure
equilibrium?



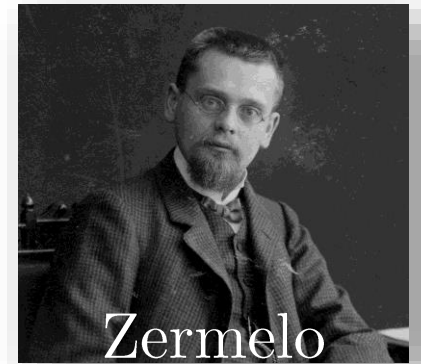
EXAMPLE: CENTIPEDE GAME



Even subgame-perfect equilibrium can lead to strange outcomes!

CHECKERS IS SOLVED

- **Zermelo's Theorem [1913]:** Either white can force a win, or black can force a win, or both sides can force a draw
- **Proof:** Backward induction ■
- Schaeffer solved the game in 2007, after 18 years of computation: It's a tie!
- Checkers game tree has 10^{20} nodes; chess has 10^{40} ; go has 10^{170}



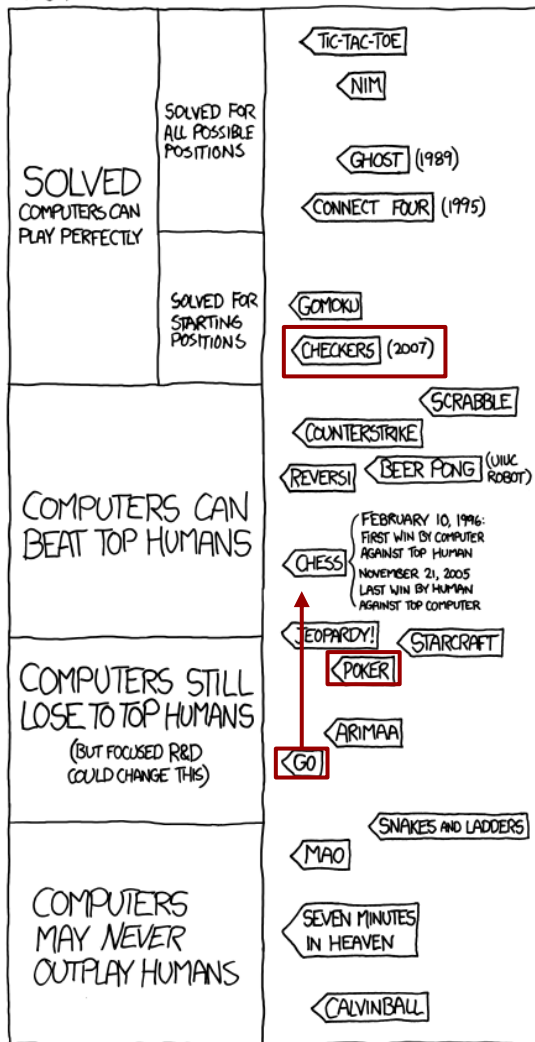
ALPHA GO

- In 2016, AlphaGo beat Lee Sedol, one of the strongest players in the history of go, in a 5-game match
- A milestone that experts thought was a decade away
- Combination of tree search techniques and deep reinforcement learning



DIFFICULTY OF VARIOUS GAMES FOR COMPUTERS

EASY

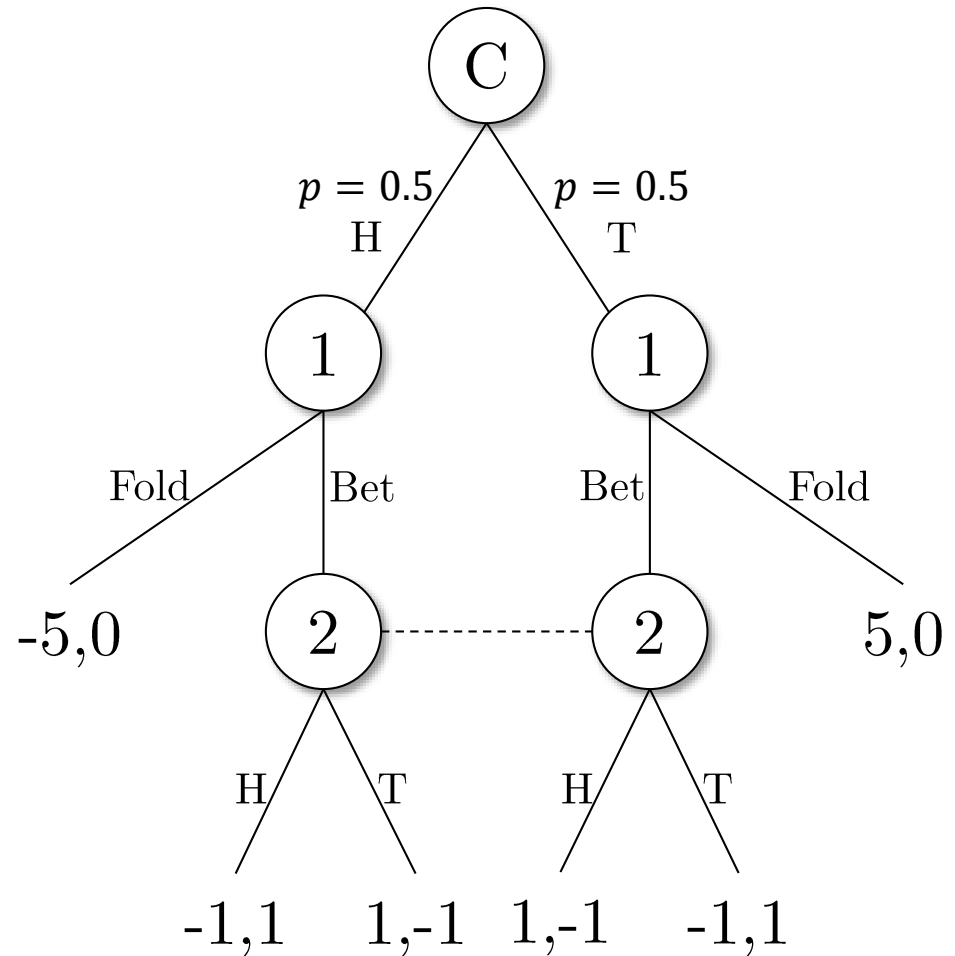


HARD



IMPERFECT INFORMATION GAMES

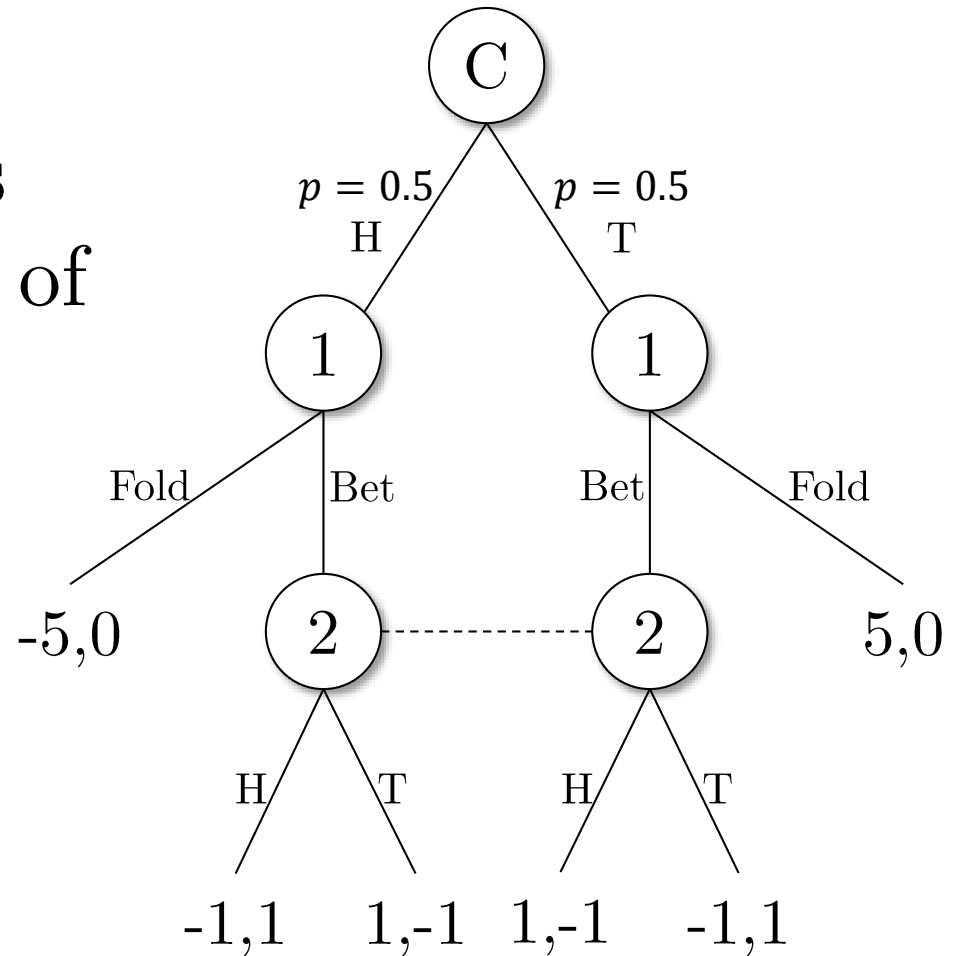
- A **chance node** chooses between several actions according to a known probability distribution
- An **information set** is a set of nodes that a player may be in, given the available information
- A strategy must be identical for all nodes in an information set



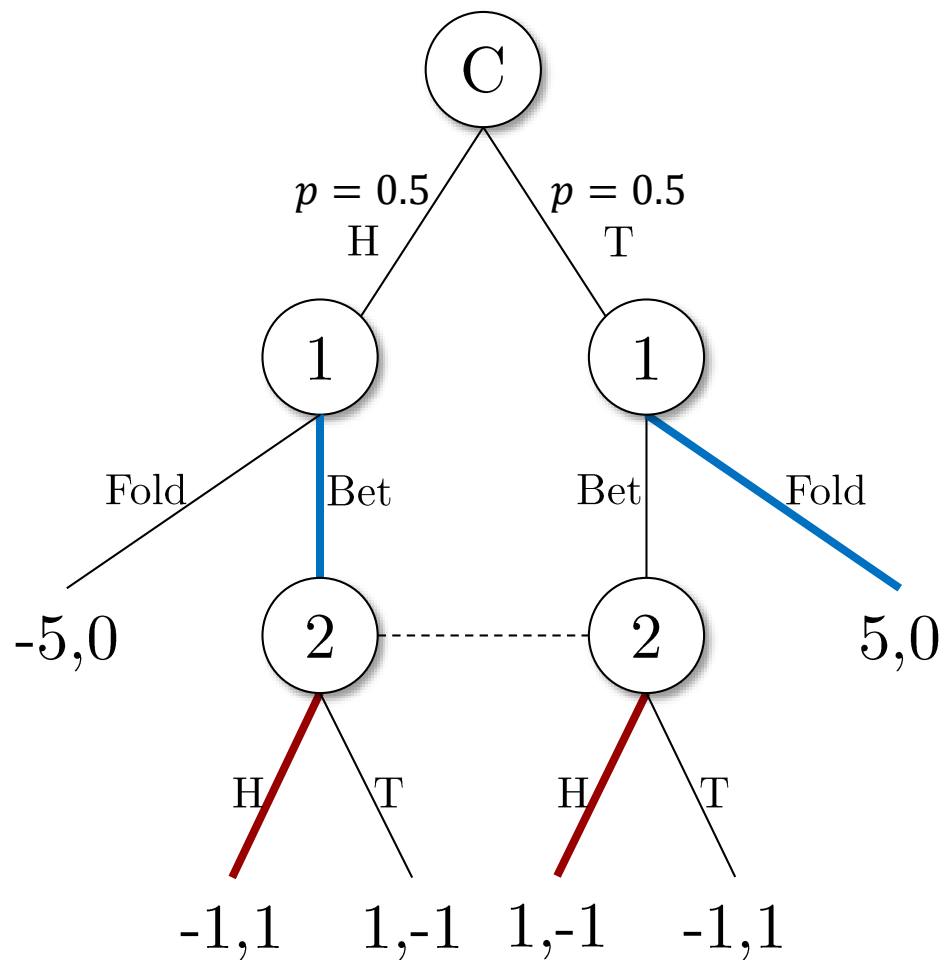
EXAMPLE: SPACESHIP GAME

- **Poll 2:** In Nash equilibrium, what is the expected payoff of player 1?

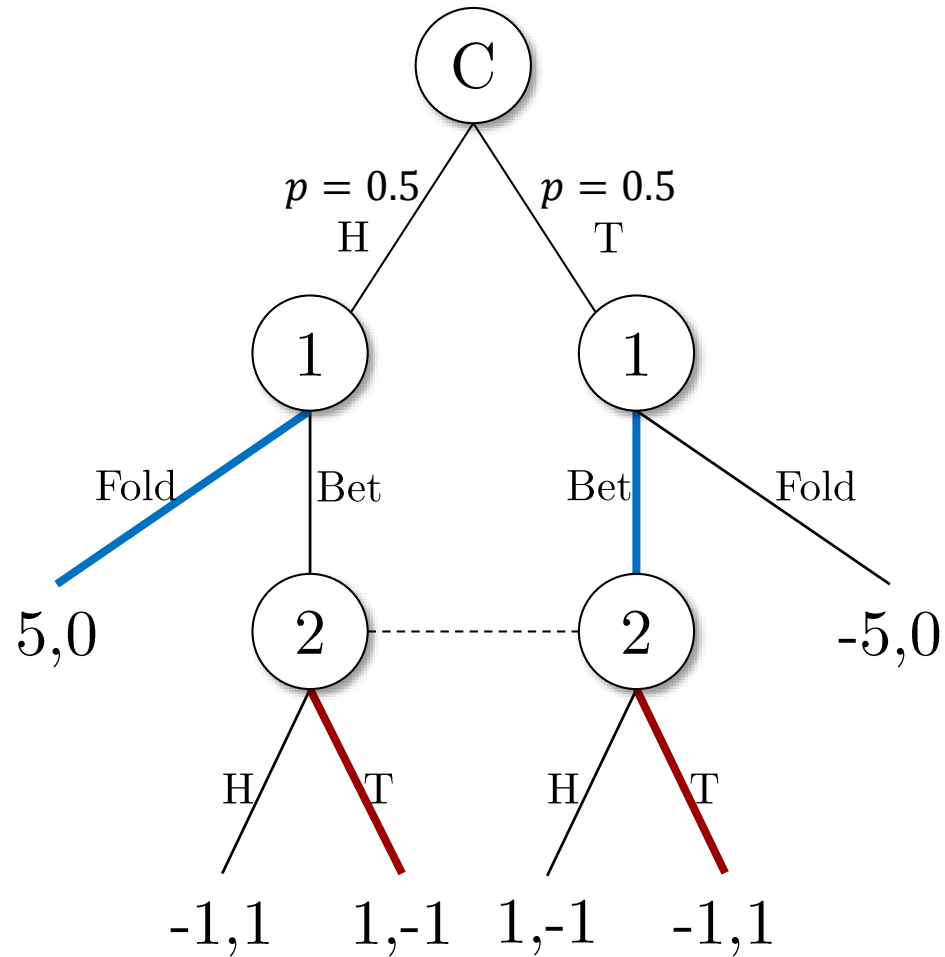
1. 0.5
2. 1
3. 1.5
4. 2
5. 2.5



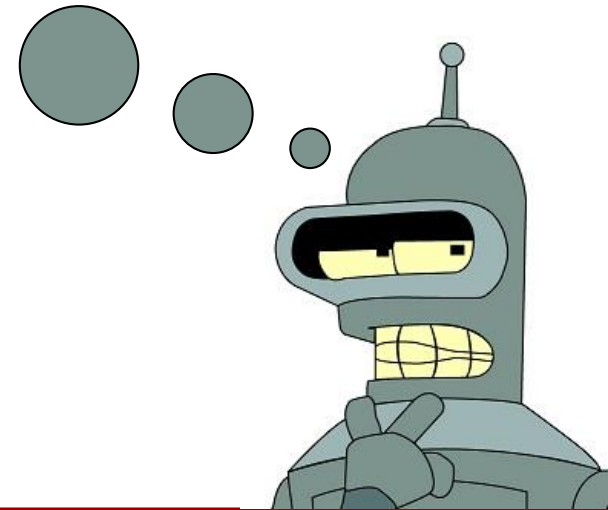
EXAMPLE: SPACESHIP GAME



EXAMPLE: SPACESHIP GAME



Impossible to
compute the optimal
strategy of a subgame
in isolation, unlike
perfect info games!



SOLVING IMPERFECT INFO GAMES

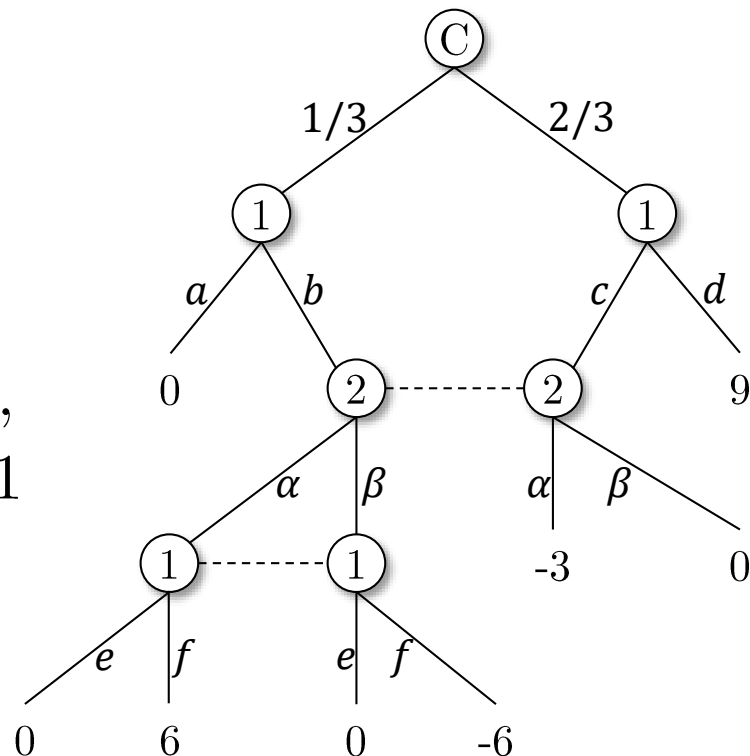
- Focus on **zero-sum** games (such as poker)
- We just saw that linear programming solves normal-form, zero-sum games in polynomial time
- But size of the normal-form game is exponential in the extensive-form representation!
- Work directly on extensive-form game



SOLVING IMPERFECT INFO GAMES

- Player 1 constraints are linear:
 - $p_a + p_b = 1$
 - $p_c + p_d = 1$
 - $p_e + p_f = 1$
 - $\forall x, p_x \geq 0$
- Fix a strategy q_α, q_β for player 2, then the best response of player 1 is:

$$\max_p 2p_b q_\alpha p_f - 2p_b q_\beta p_f - 2p_c q_\alpha + 6p_d$$
 which leads to a nonconvex problem!



SEQUENCE FORM

- **Insight:** last action taken by a player is the same for all nodes in an information set
 - **Perfect recall:** A player never forgets something he knew in the past
 - This is a restriction on the structure of the game
- Introduce scaled probability variables p'_x
- Information set constraint: $\sum_{x \in A_I} p'_x = p'_y$, where A_I is the set of actions in information set I , and y is the last action before reaching I
- To recover probabilities, set $p_x = p'_x / p'_y$

SEQUENCE FORM

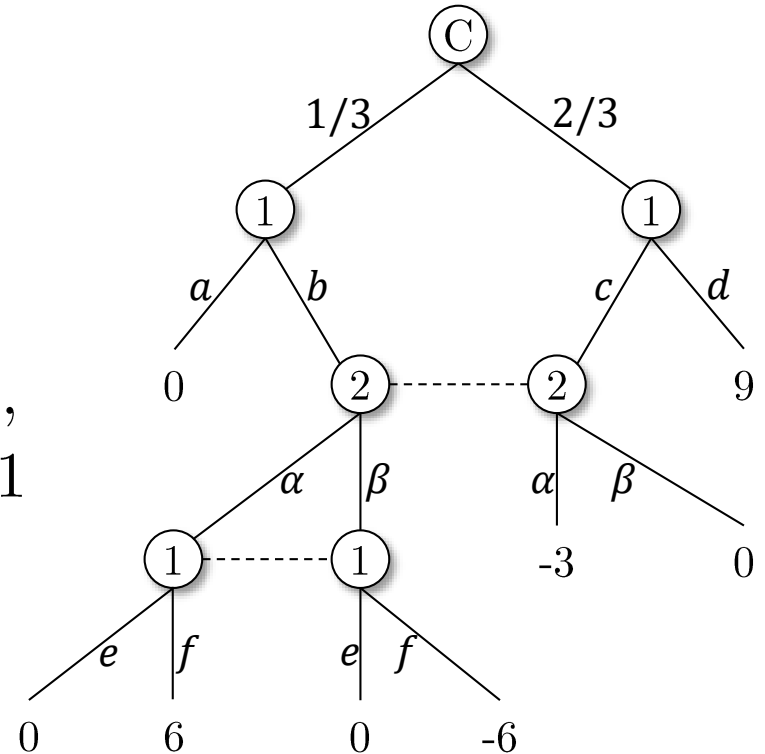
- Player 1 constraints are linear:

- $p'_a + p'_b = 1$
- $p'_c + p'_d = 1$
- $p'_e + p'_f = p'_b$
- $\forall x, p'_x \geq 0$

- Fix a strategy q_α, q_β for player 2, then the best response of player 1 is:

$$\max_p 2q_\alpha p'_f - 2q_\beta p'_f - 2p'_c q_\alpha + 6p'_d$$

which is linear!



SEQUENCE FORM

- We showed how to compute a best response for a fixed opponent strategy
- **Fact:** Using LP duality, we can compute best responses for both players simultaneously
- **Fact:** This gives a method for computing optimal strategies
- Used to compute optimal strategies for Rhode Island Hold'em poker, which has roughly 10^8 nodes [Gilpin and Sandholm 2007]
- But No Limit Texas Hold'em has 10^{167} nodes





BRAINS VS. ARTIFICIAL INTELLIGENCE

Be sure to tweet @WinBigRivers and @SCSatCMU using #BrainsvsAI



JANUARY 11-30 | 11AM-7PM

WE ARE UPPING THE ANTE!
120,000 HANDS NO-LIMIT HOLD 'EM

Each hand starts with each player having 200 big blinds.
One big blind is \$100, and one small blind is \$50.

Hands Dealt: 120,000/120,000

BRAINS : (\$1,766,250)

LIBRATUS : \$1,766,250

DONG KIM : (\$85,649)

JASON LES : (\$880,087)

LIBRATUS : \$85,649

LIBRATUS : \$880,087

JIMMY CHOU : (\$522,857)

DANIEL MCAULAY : (\$277,657)

LIBRATUS : \$522,857

LIBRATUS : \$277,657

January 11-30, 2017, at Rivers Casino, Pittsburgh
The first time a computer program has defeated top human pros
in a heads-up, no-limit poker game



SUMMARY

- Terminology:
 - Extensive-form game
 - Subgame perfect equilibrium
 - Imperfect information, information set
 - Perfect recall
- Algorithms:
 - Solving zero-sum games via LP
 - Sequence form-based approach to solving imperfect information games

