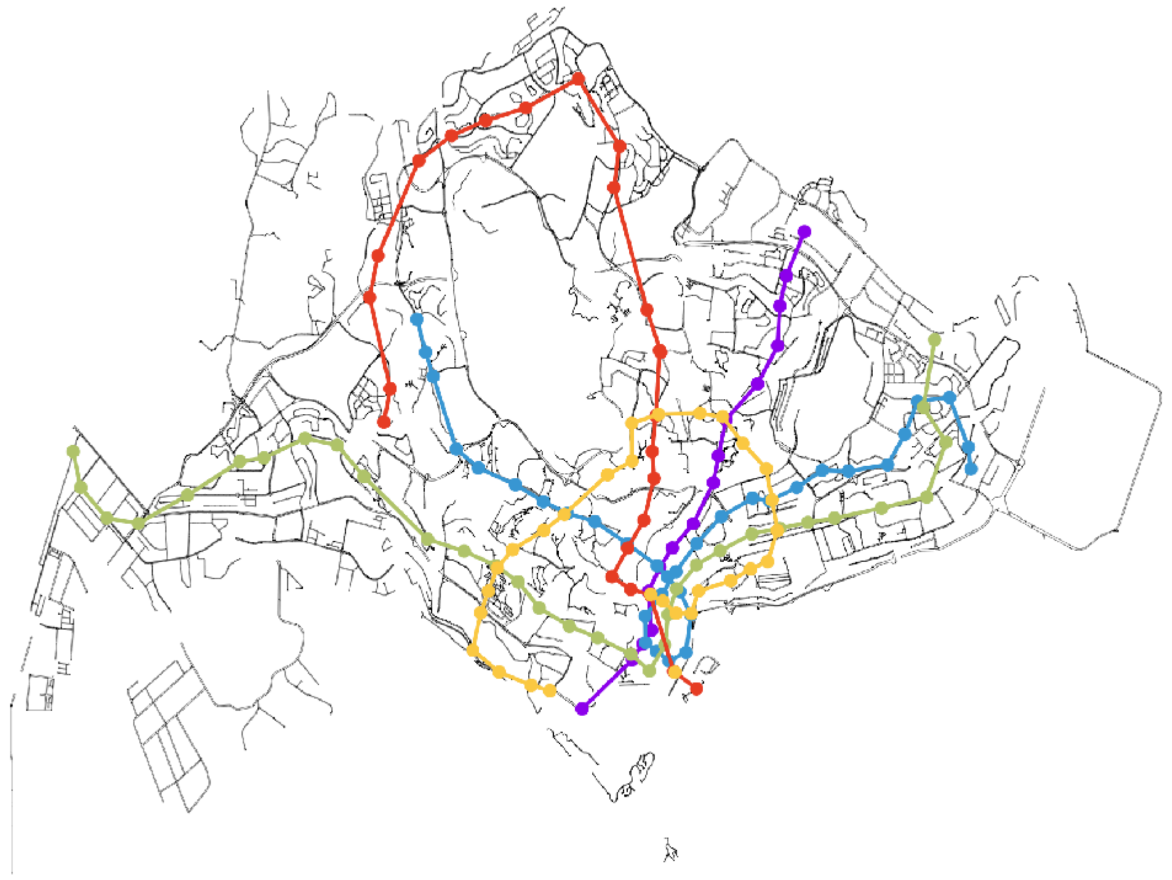
CSCI 357: Algorithmic Game Theory

Lecture 17: Sequential Games

Shikha Singh



Announcements and Logistics

- HW 7 (partner assignment) is due tonight at 11 pm
- HW 8 (short: ~2 questions) will be released tomorrow
  - Due Wed April 20 11pm (a day early)
- Guidelines on project topics will be released on Monday
  - One default project that is more structured
  - Other ideas and links to paper
- I will discuss project options in more detail next week
- Read through it, talk to me and figure out what you want to work on
  - Proposal proposal due Fri April 22 5pm

Questions?

Reminders

- CS TA applications due April 22
  - https://csci.williams.edu/tatutor-application/
- CS pre-registration Info session:
  - Tomorrow during colloquium
  - Come find out about courses for next semester
- CS Thesis applications due April 19
  - https://csci.williams.edu/computer-science-research-application-forms/

Last Time

- Complexity of manipulation strategies in voting
- Fair division of heterogenous divisible good: cake cutting problem
  - Envy free cake cutting for $n = 2$ (Cut and Choose protocol)
  - Envy free cake cutting for $n = 3$ (Selfridge and Conway)
- Today: analyzing sequential (extensive-form) games
  - Subgame perfection equilibrium

Strategic Cake Cutting

- Consider the $n = 2$ cut and choose protocol
- Player's valuation for the cake is their private information
- Are player's incentivized to follow the algorithm?
  - Is it a dominant strategy to "tell the truth" that is, play according to your true valuation
  - Does there exist any valuation profile where cheating gives more utility?
- Suppose player 2 only wants the left $[0, \epsilon]$ of the cake and player 1 value for any piece = size of the piece
  - Is it in player 1's best interest to be truthful, that is, cut the cake in half?

Strategic Cake Cutting: Research

AAAI 2013

Truth, Justice, and Cake Cutting

Yiling Chen
Harvard SEAS
yiling@eecs.harvard.edu

John K. Lai
Harvard SEAS
jklai@post.harvard.edu

David C. Parkes
Harvard SEAS
parkes@eecs.harvard.edu

Ariel D. Procaccia
Harvard SEAS
arielpro@seas.harvard.edu

Abstract

Cake cutting is a common metaphor for the division of a heterogeneous divisible good. There are numerous papers that study the problem of fairly dividing a cake; a small number of them also take into account self-interested agents and consequent strategic issues, but these papers focus on fairness and consider a strikingly weak notion of truthfulness. In this paper we investigate the problem of cutting a cake in a way that is truthful and fair, where for the first time our notion of dominant strategy truthfulness is the ubiquitous one in social choice and computer science. We design both deterministic and randomized cake cutting algorithms that are truthful and fair under different assumptions with respect to the valuation functions of the agents.

SAGT 2018

Truthful Fair Division

Elchanan Mossel* and Omer Tamuz†

November 13, 2018

Abstract

We address the problem of fair division, or cake cutting, with the goal of finding truthful mechanisms. In the case of a general measure space ("cake") and non-atomic, additive individual preference measures - or utilities - we show that there exists a truthful "mechanism" which ensures that each of the k players gets at least $1/k$ of the cake. This mechanism also minimizes risk for truthful players. Furthermore, in the case where there exist at least two different measures we present a different truthful mechanism which ensures that each of the players gets more than $1/k$ of the cake.

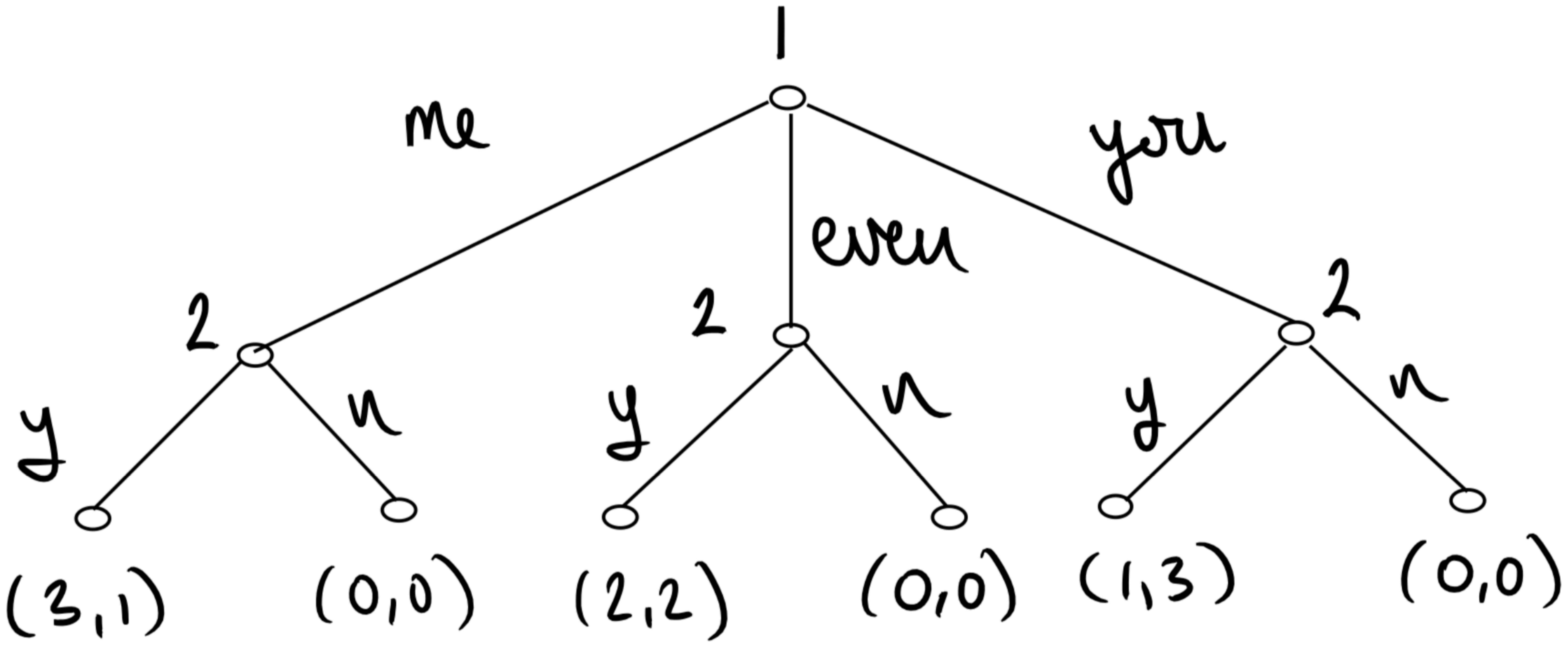
We then turn our attention to partitions of indivisible goods with bounded utilities and a large number of goods. Here we provide similar mechanisms, but with slightly weaker guarantees. These guarantees converge to those obtained in the non-atomic case as the number of goods goes to infinity.

Equilibrium in Sequential Games

- We have so far not talked about incentives or equilibrium in division games
- This is because they are "sequential" in nature
  - One person moves first, second person must respond
- In fact, all the mechanisms/games so far were "one-shot"
  - Everyone picks their whole strategy at once (in the beginning)
  - E.g., in matching/voting/auctions players commit to their entire preference list or bids at the beginning (direct revelation)
- Sequential games introduce new challenges to the analysis of strategic behavior

Bargaining Game

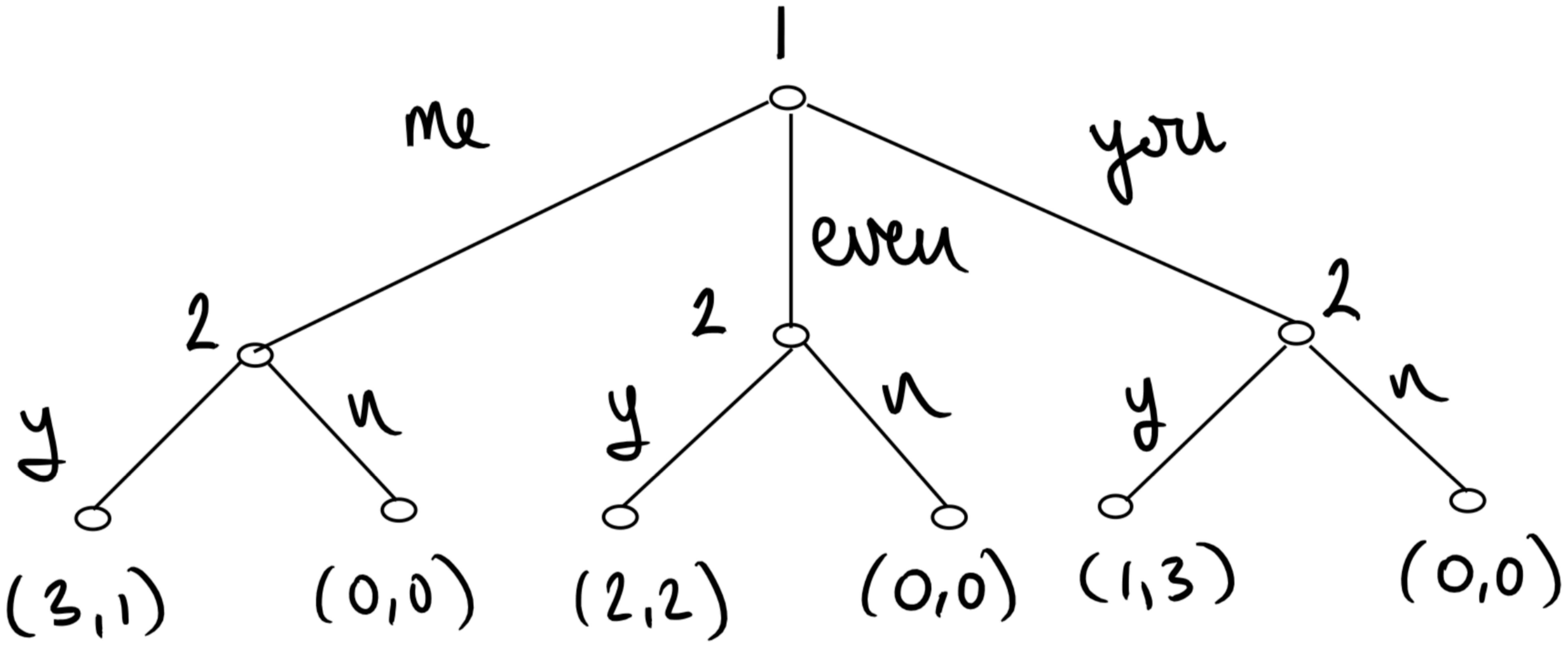
- Two players are bargaining over how to divide $4
- If they do not agree, no one gets the money
- Player 1 goes first and can propose: **me (3, 1)**, **even (2,2)** or **you (1,3)**



Credit: Textbook by Parkes and Seuken

Bargaining Game

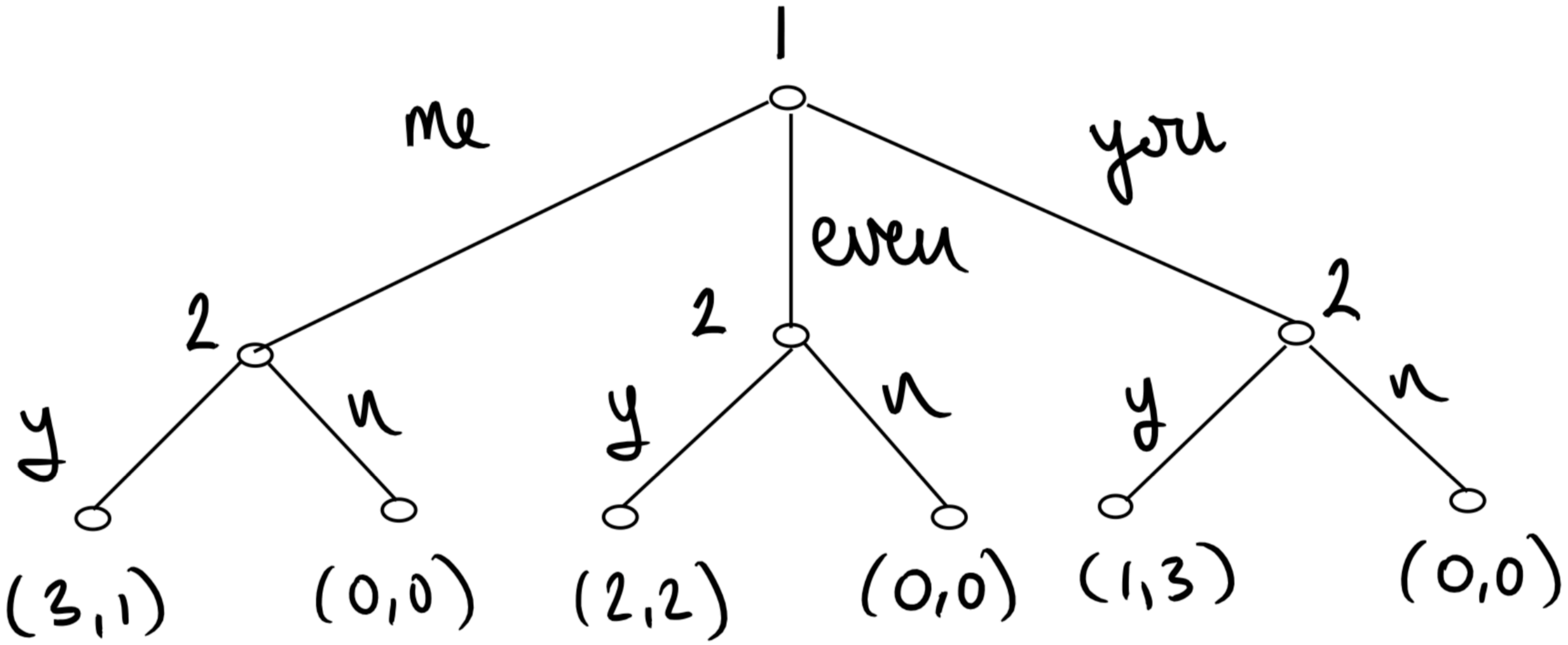
- Seeing this, player 2 can respond by either accept (**y**) or decline (**n**)
- Game tree below shows the utilities of the players at the leaves



Credit: Textbook by Parkes and Seuken

Bargaining Game

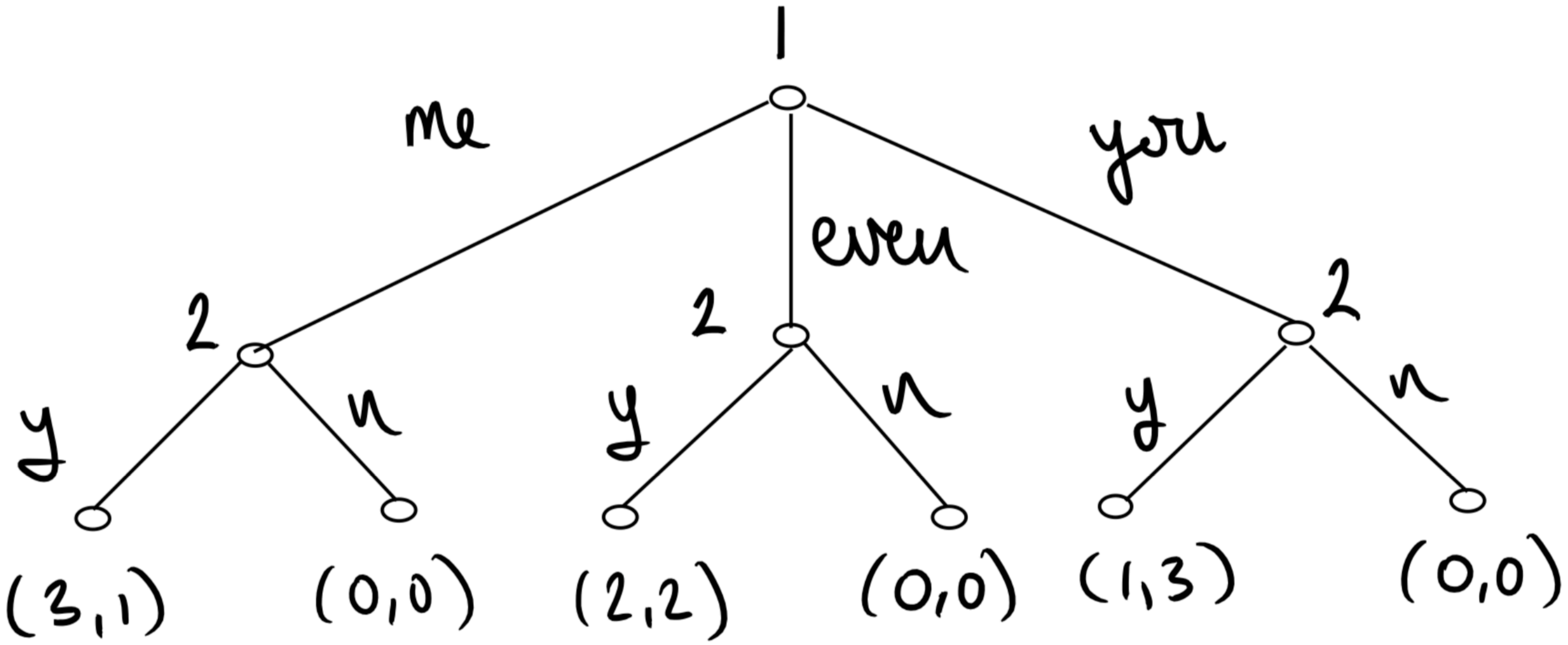
- Strategies in an extensive-form game must specify a complete description of how a player will act
- Player 2 needs an action for all three nodes in the tree: together they form player 2's strategy



Credit: Textbook by Parkes and Seuken

Bargaining Game

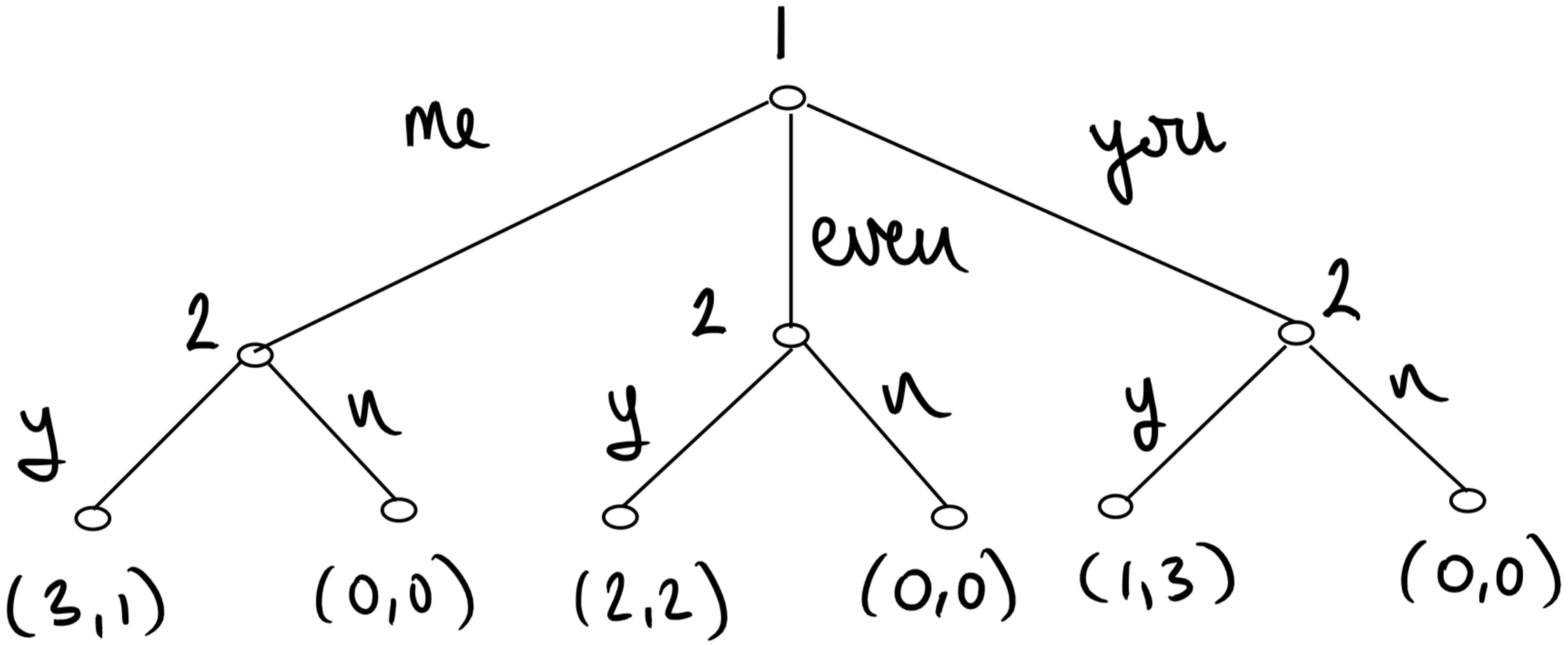
- Player 2's strategy thus needs to specify three actions
- For example, (N, N, Y) represents the action plan to say **no** to **me**, **no** to **even**, and **yes** to **you**
- 2^3 possible strategies



Credit: Textbook by Parkes and Seuken

Strategic-Form

- One can convert an extensive-form game into a strategic (normal-form)
- However, such a representation is far from ideal, and can be confusing



Credit: Textbook by Parkes and Seuken

Strategic-Form

- One can convert an extensive-form game into a strategic (normal-form)
- However, such a representation is far from ideal, and can be confusing
- Strategic-form representation of our bargaining game:
- Can you identify some of the Nash equilibria?

Player 2

		$\langle N,N,N\rangle$	$\langle N,N,Y\rangle$	$\langle N,Y,N\rangle$	$\langle N,Y,Y\rangle$	$\langle Y,N,N\rangle$	$\langle Y,N,Y\rangle$	$\langle Y,Y,N\rangle$	$\langle Y,Y,Y\rangle$
	me	0, 0	0, 0	0, 0	0, 0	3, 1	3, 1	3, 1	3, 1
Player 1	*even*	0, 0	0, 0	2, 2	2, 2	0, 0	0, 0	2, 2	2, 2
	you	0, 0	1, 3	0, 0	1, 3	0, 0	1, 3	0, 0	1, 3

Credit: Textbook by Parkes and Seuken

Nash Equilibrium

- Lots of Nash equilibria of the extensive form game
  - Not meaningful as a predictor of what players will do
- Some of the Nash equilibria are not plausible
  - For example, the Nash equilibrium (you, (N, N, Y)) implies that player 2 would decline $1\$$ or $2\$$ if offered
  - If player 1 did offer it, this would not be rational to decline

Player 2

		$\langle N,N,N\rangle$	$\langle N,N,Y\rangle$	$\langle N,Y,N\rangle$	$\langle N,Y,Y\rangle$	$\langle Y,N,N\rangle$	$\langle Y,N,Y\rangle$	$\langle Y,Y,N\rangle$	$\langle Y,Y,Y\rangle$
	me	0, 0	0, 0	0, 0	0, 0	3, 1	3, 1	3, 1	3, 1
Player 1	*even*	0, 0	0, 0	2, 2	2, 2	0, 0	0, 0	2, 2	2, 2
	you	0, 0	1, 3	0, 0	1, 3	0, 0	1, 3	0, 0	1, 3

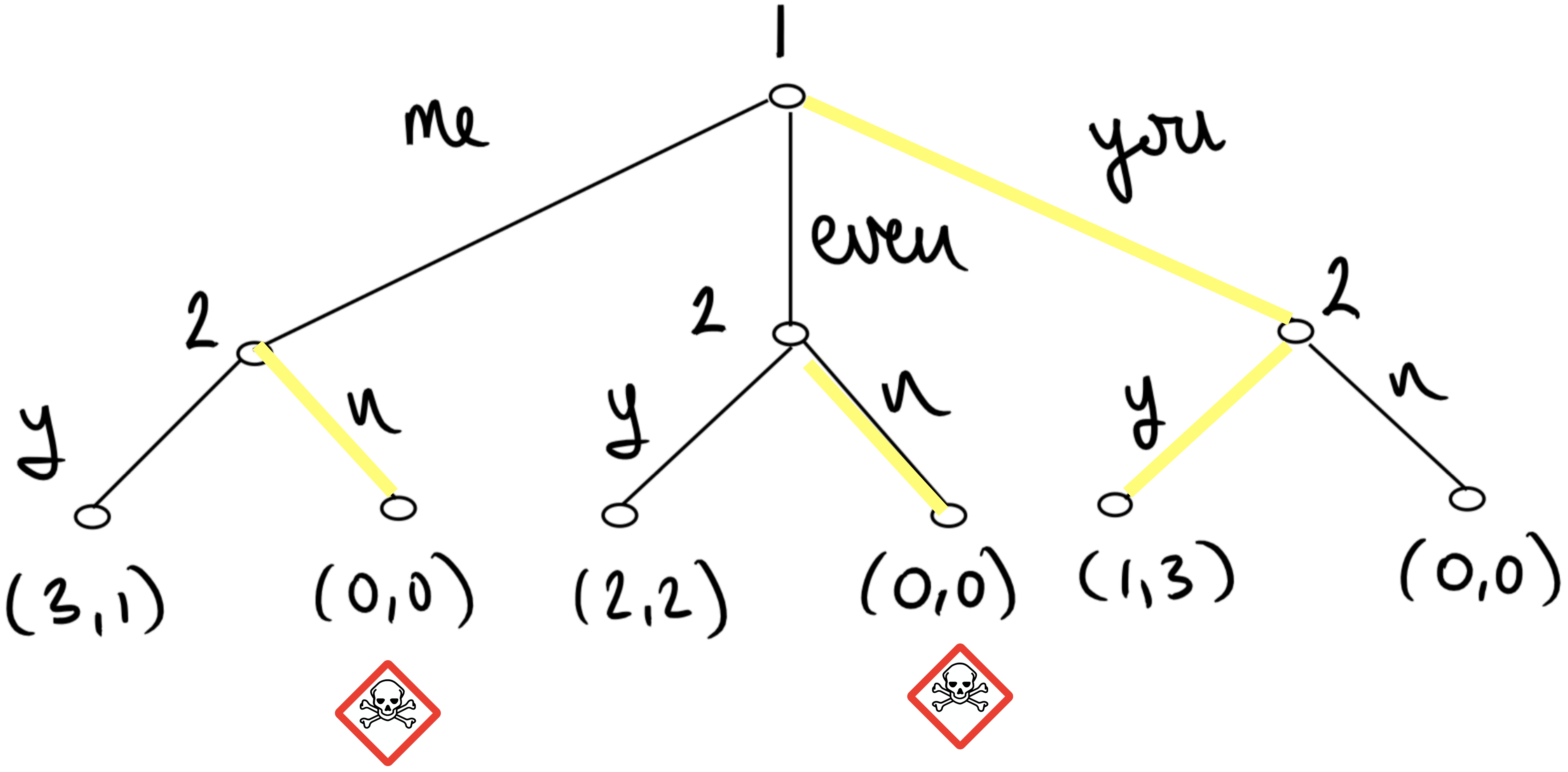
Credit: Textbook by Parkes and Seuken

Empty Threats

- Nash equilibria as a solution concept for extensive-form games is susceptible to **empty threats** or **non-credit threats**
- An empty threat is when Player 2 who will move in a later round threatens to do something irrational
  - The threat is non-credible because it is not in the best interest of Player 1 to carry it out *if it comes to it*
- Player 1's goal is to convince Player 1, who is moving in an earlier round, to take an action that is favorable to Player 1
- Let's see this in the Bargaining game

Empty Threats

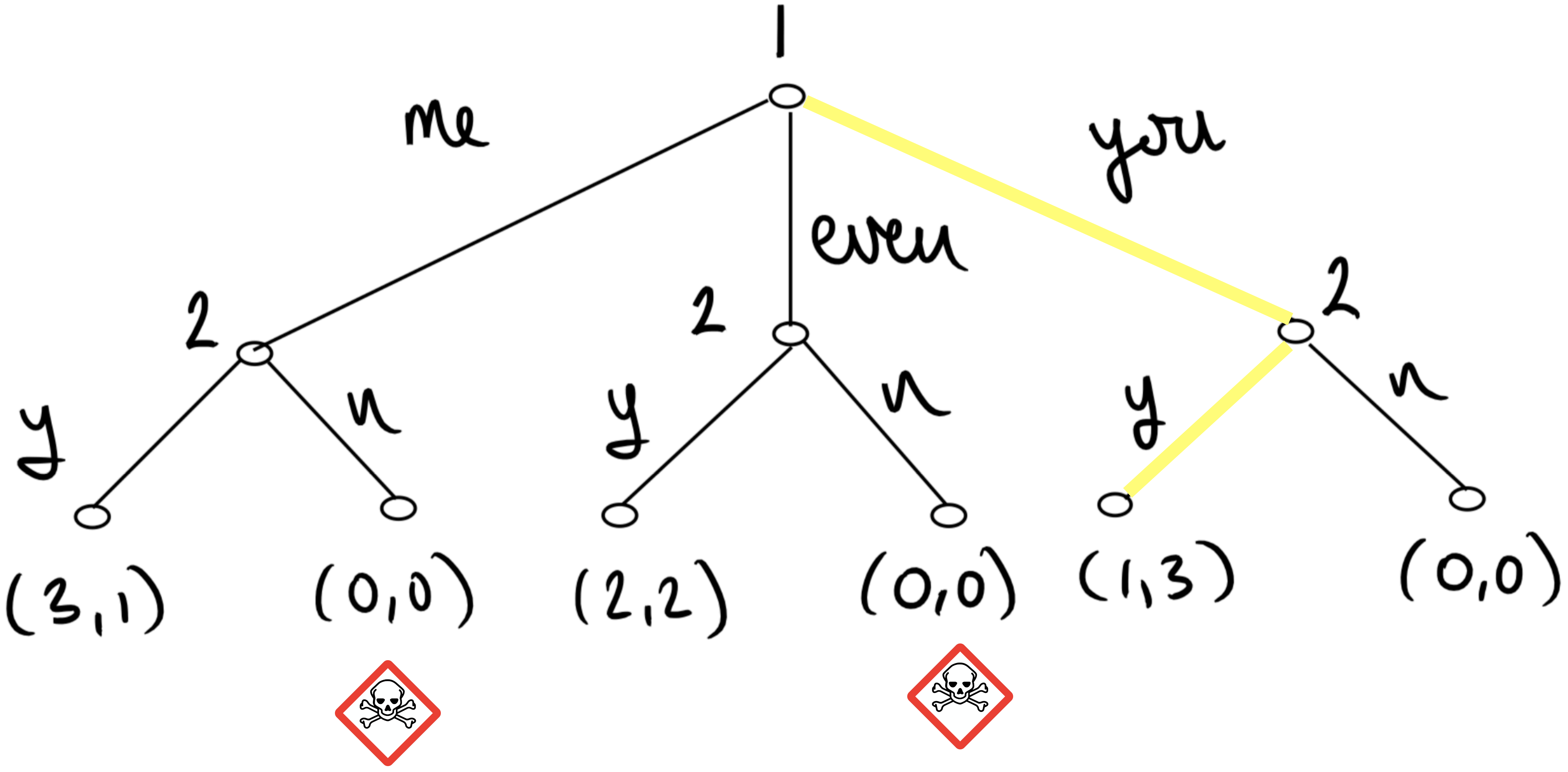
- s_2 playing (N, N, Y) is part of a Nash equilibrium but this is an empty threat
- Aimed at deterring player 1 to pick an option they prefer more



Credit: Textbook by Parkes and Seuken

New Equilibrium

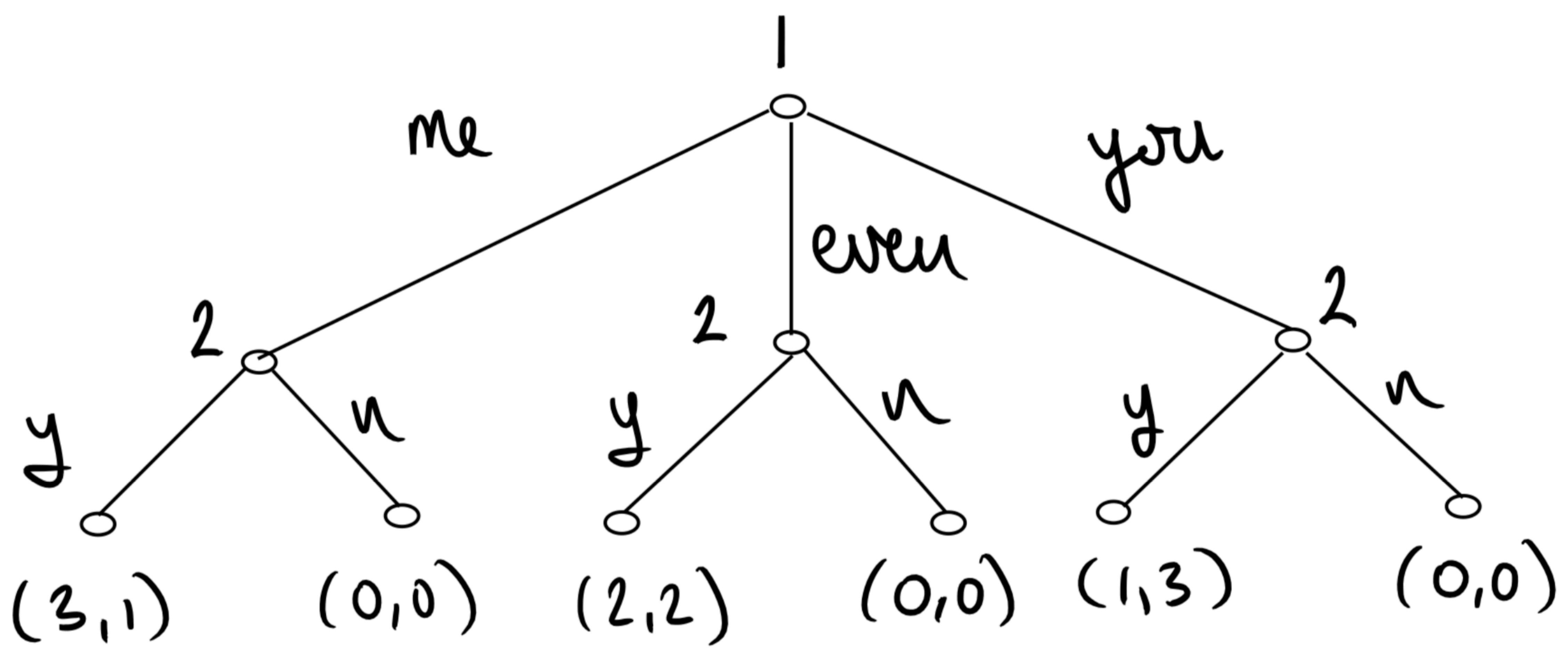
- We need a new equilibrium concept for sequential form games which takes the sequential nature in account and avoids empty threats
- A "refinement" of Nash equilibrium in such games



Credit: Textbook by Parkes and Seuken

Extensive Form Model

- Game tree representation: a path from root to any node is a **history** h
- Terminal histories Z: set of all root to leaf path
- Utility $u_i(h) \in \mathbb{R}$ for each player for each terminal history $h \in Z$
- A player function $P(h) \in N$ for $h \in H \backslash Z$ of non-terminal histories
- Action set $A_i(h)$: set of actions available to player i at non-terminal history h



Credit: Textbook by Parkes and Seuken

Perfect Information

- We will assume the extensive-form games are games of **perfect information**:
  - Each player i knows the complete history h of the game whenever it is i's turn to act $P(h) = i$
  - The structure of the and utilities are common knowledge
- Example of perfect-information sequential game:
  - Chess
- Example of imperfect-information sequential game:
  - Poker
- Extensive-form games with imperfect information are more complicated:
  - Have “information sets” & players’ probabilistic beliefs on histories

Credit: Textbook by Parkes and Seuken

Strategy: Complete Action Plan

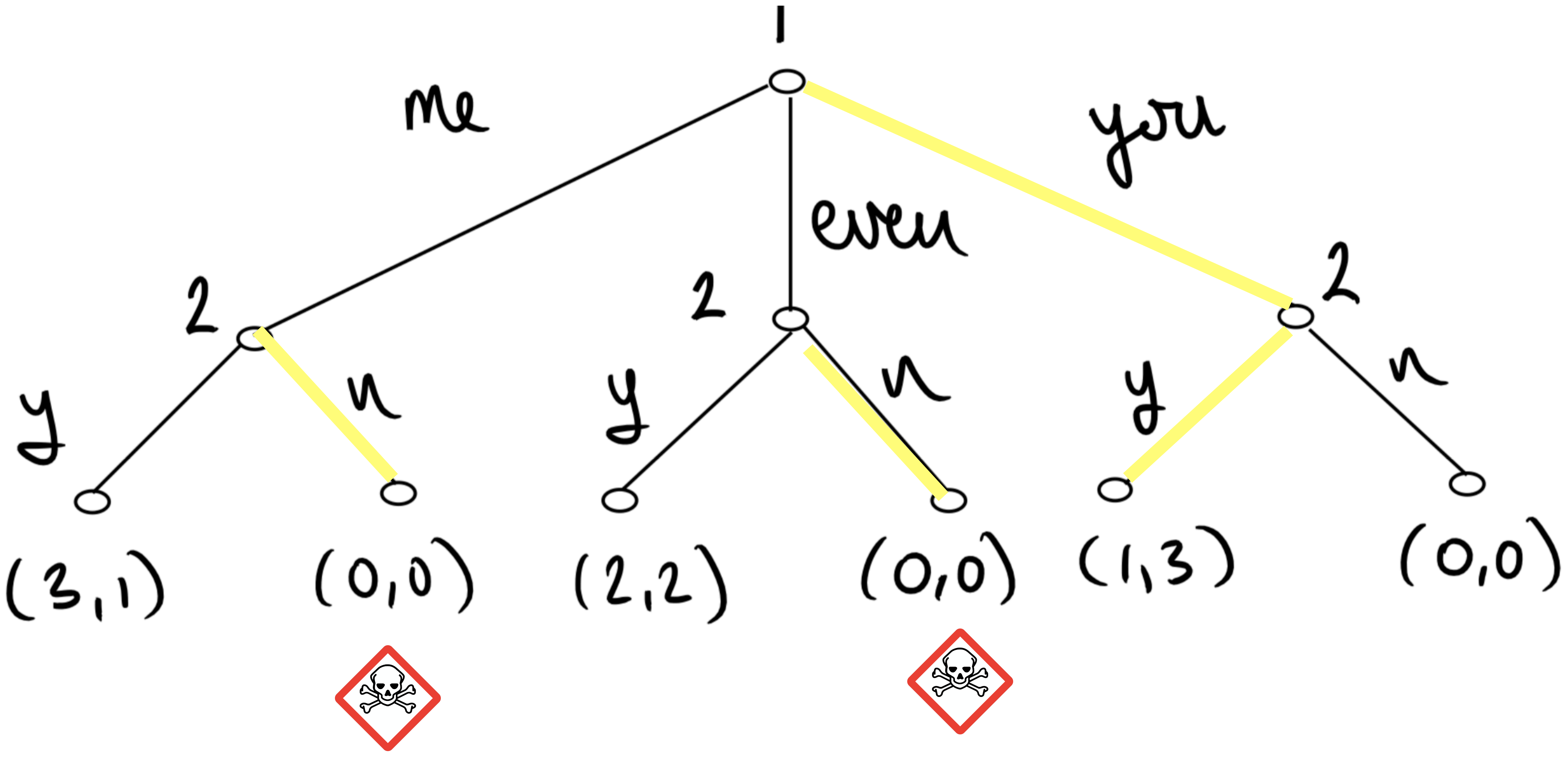
- A strategy must say what action to play at every decision node in the game tree
- **(Definition)** A strategy s_i in extensive-form game defines an action $s_i(h) \in A_i(h)$ for each non-terminal history h for which it is player $i's$ turn, that is, $P(h) = i$
- Fix s_{-i} : other players actions on each decision node they have to play
- **Nash equilibrium**. A strategy profile $s = (s_1, \ldots, s_n)$ is a Nash equilibrium in an extensive-form game iff for all agents i:

$$u_i(s_i, s_{-i}) \geq u(s_i', s_{-i}) \text{ for all strategies } s_i' \text{ of player } i$$

Nash: Bargaining Game

- The Nash equilibrium strategy $(\text{you}, (N, N, Y))$ can be represented as: $s_1(\epsilon) = \text{you}$ and

$$s_2(h) = \begin{cases} N & , \textit{ if } h = (me) \\ N & , \textit{ if } h = (even) \\ Y & , \textit{ if } h = (you) \end{cases}$$



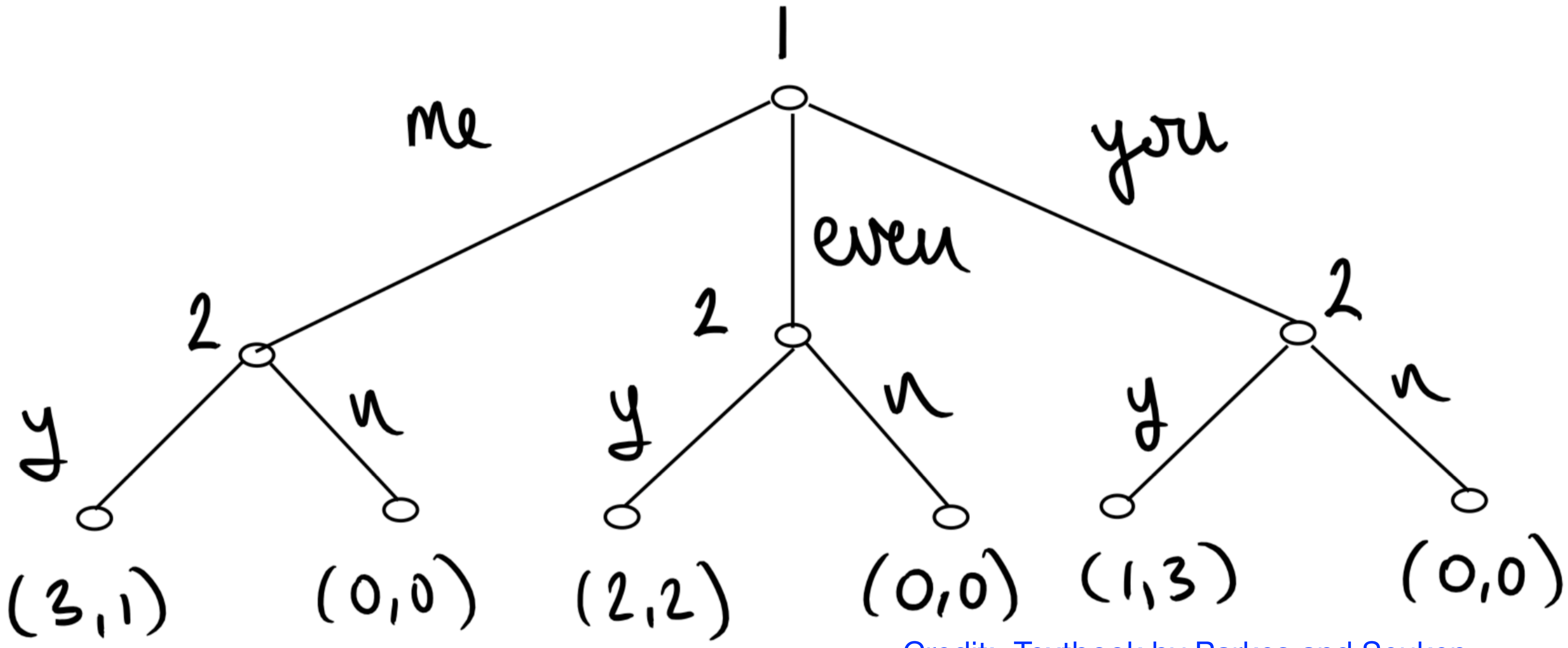
Credit: Textbook by Parkes and Seuken

Subgame-Perfect Equilibrium

(Perfect Information Games)

Subgames

- We define a new solution concept for extensive-form games
- **(Definition).** The subgame starting at history h of an extensive-form game is the extensive-form game rooted at the decision node that corresponds to history h
- Can you identify all subgames in this game?

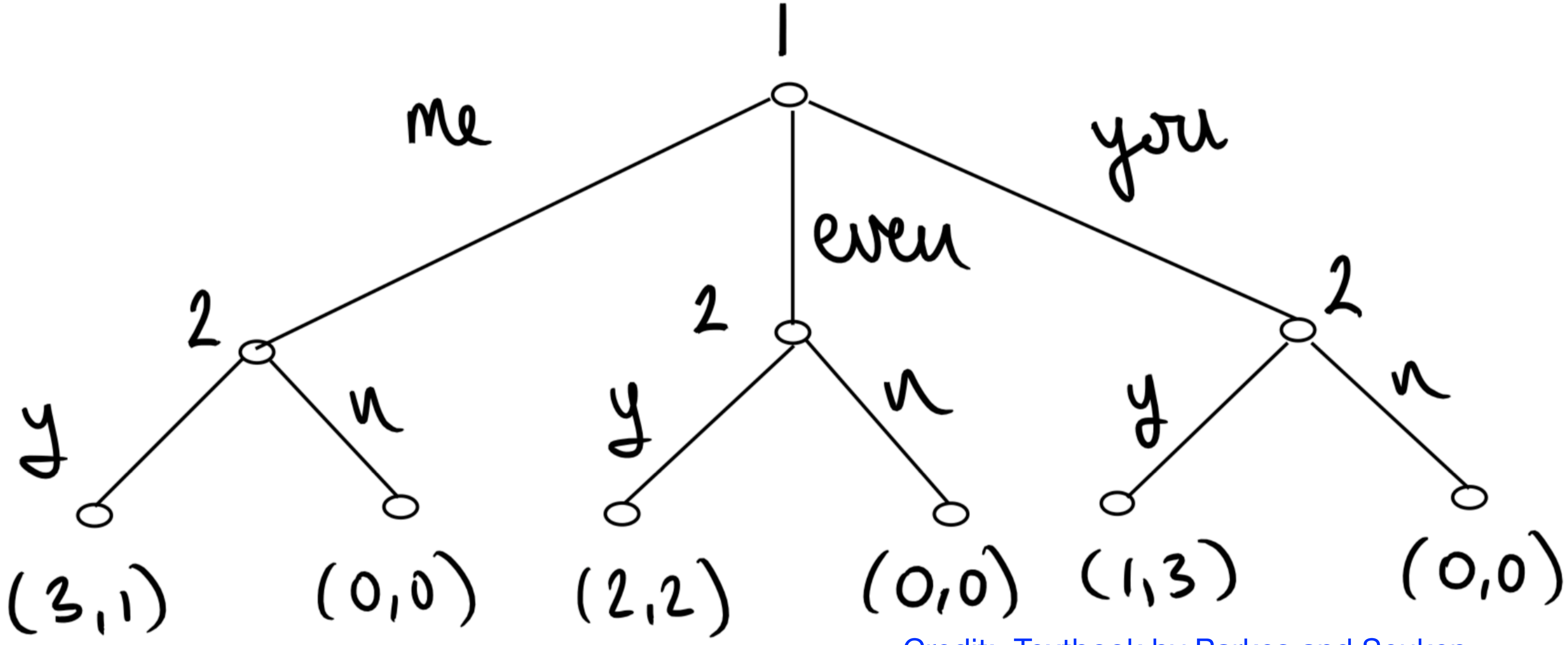


Credit: Textbook by Parkes and Seuken

Subgame-Perfect Equilibrium

- **(Definition)** A strategy profile $s = (s_1, \ldots, s_n)$ is a subgame-perfect equilibrium of an extenstive form game if the strategy profile is a **Nash equilibrium in every subgame** of the game starting at a non-terminal history

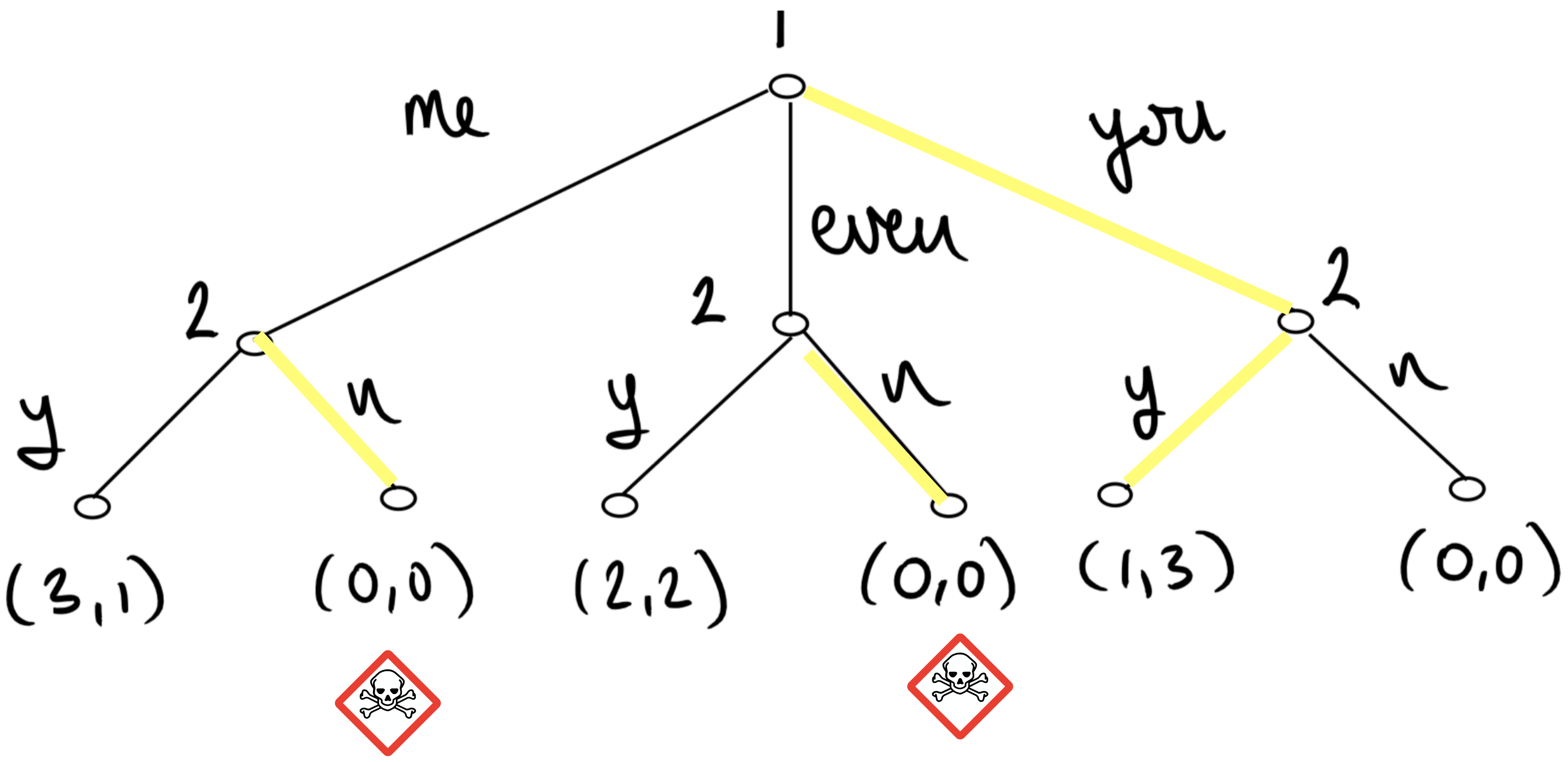
- Enforce that players should play their best responses after each history of the game



Credit: Textbook by Parkes and Seuken

Subgame-Perfect Equilibrium

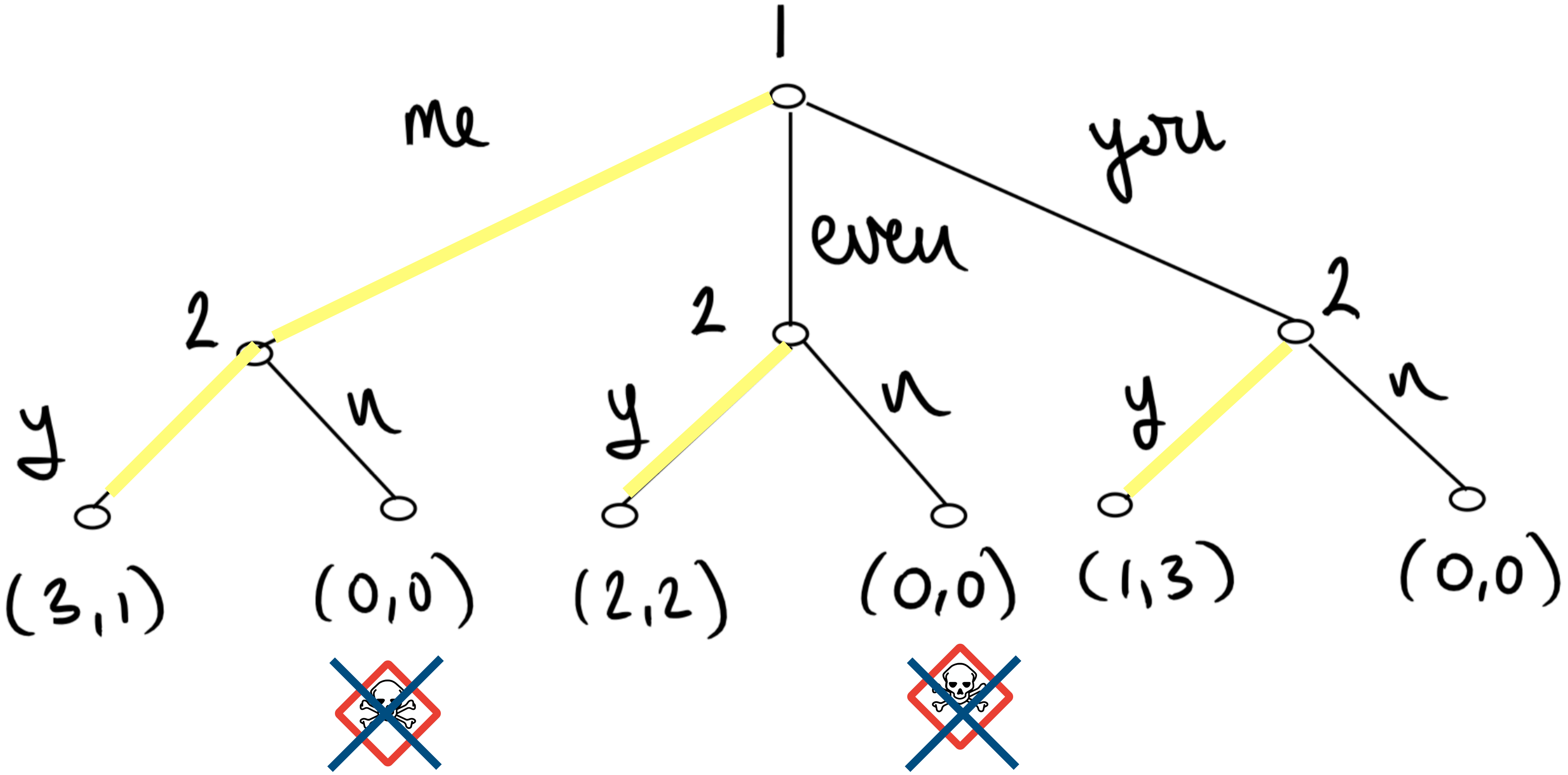
- Is the strategy $(\text{you}, (N, N, Y))$ a subgame perfect equilibrium?



Credit: Textbook by Parkes and Seuken

Subgame-Perfect Equilibrium

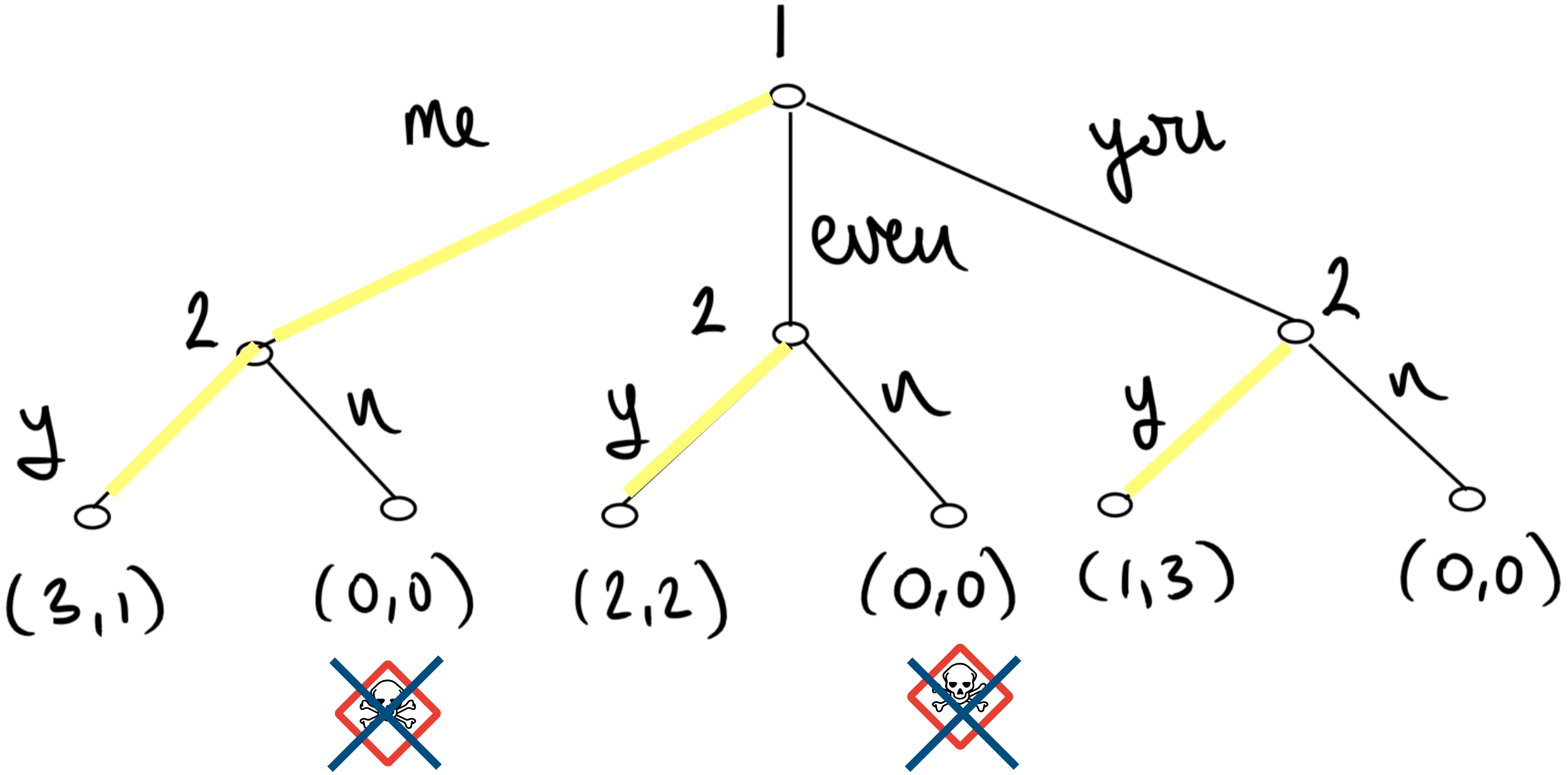
- "Conditioned on reaching" any history where player 2 must act, saying no is never a best response



Credit: Textbook by Parkes and Seuken

Subgame-Perfect Equilibrium

- Given player 2 plays their best response at every node, Player 1 must choose **me**
- This is the unique SPE of this game



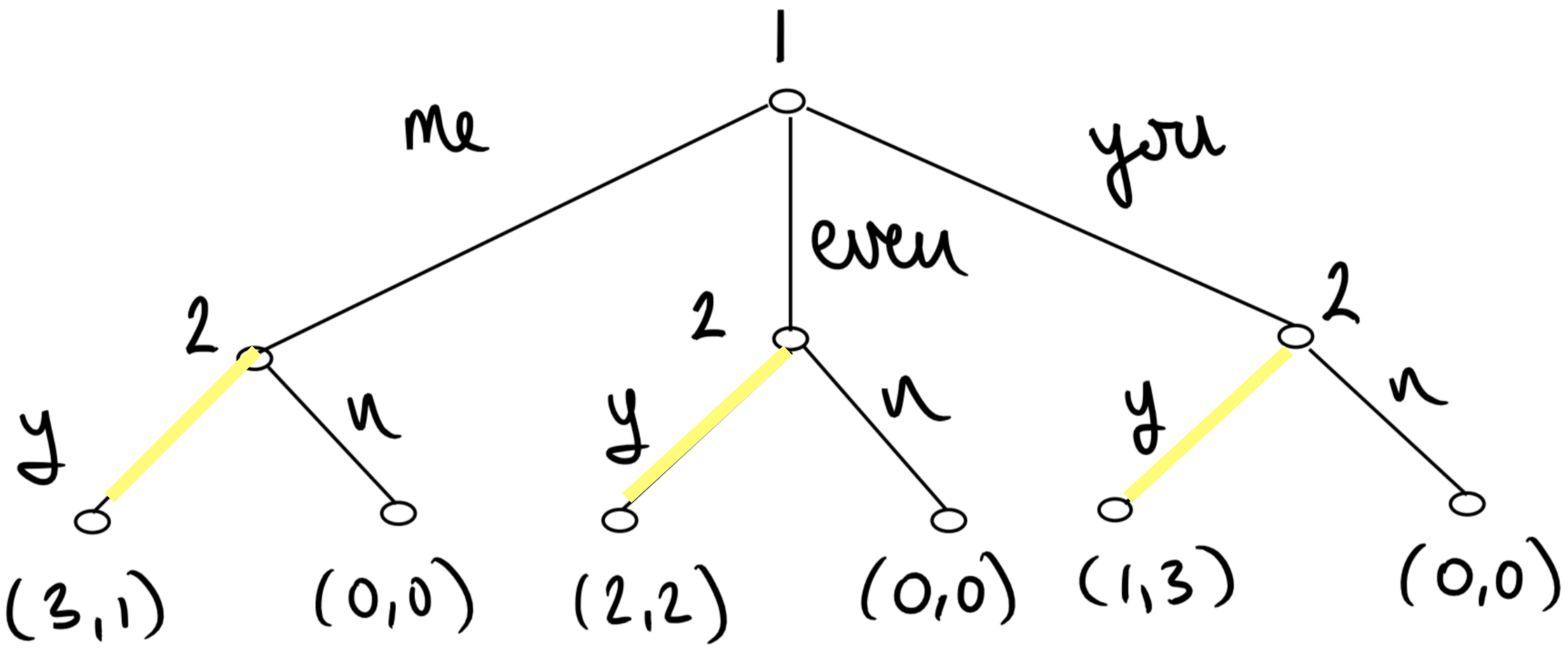
Credit: Textbook by Parkes and Seuken

Backward Induction

- Approach to compute a SPE of an extensive-form game
  - Essentially applies dynamic programming to game trees
- Start at the bottom (say at depth k) and look at the player who acts last P_ℓ
- Conditioned on reaching their decision nodes, figure out $P'_\ell s$ best response
- Fixing the best response of P_ℓ at depth k, we know have a tree of depth $k - 1$
- Continue applying this logic until we reach the root:
  - The resulting strategy profile must be a SPE
  - We need to prove this

Backward Induction

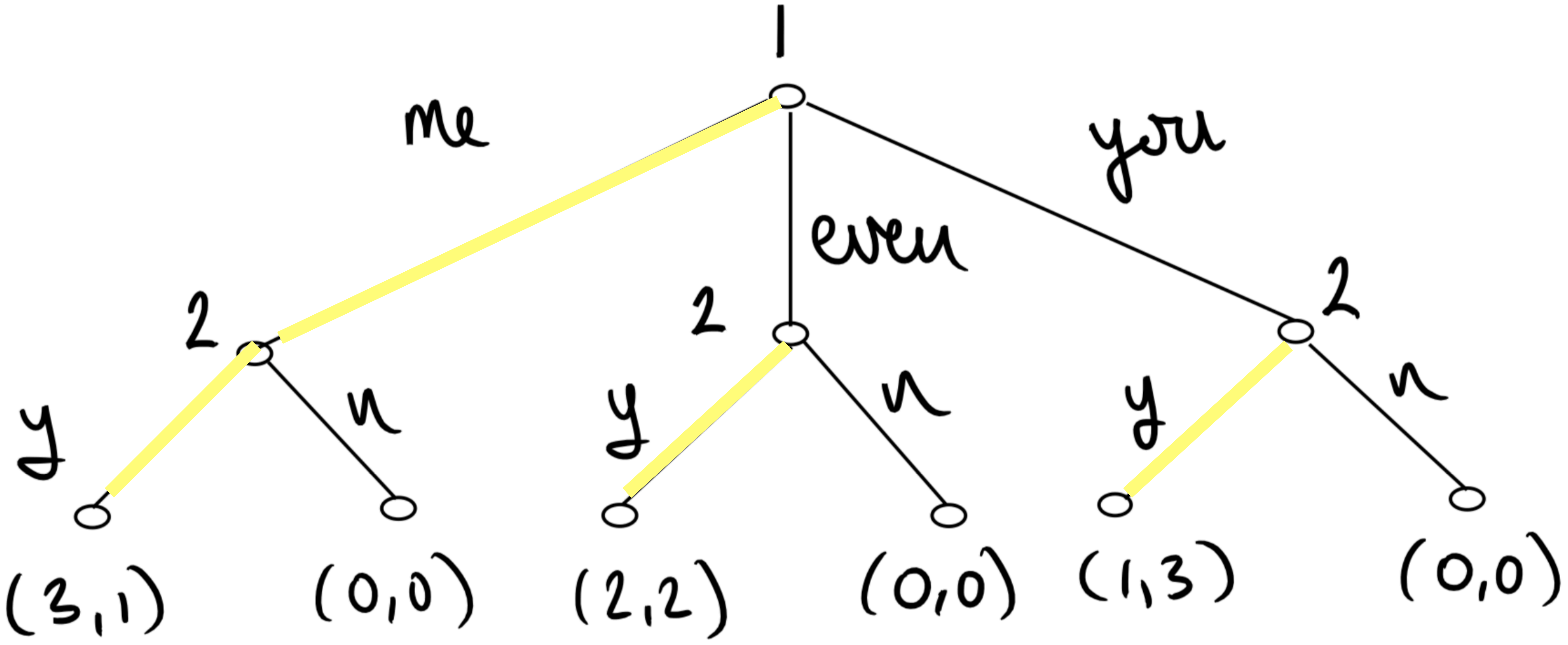
- Backward's induction is essentially "dynamic programming"
- You keep track of the optimal moves as you go up the tree



Credit: Textbook by Parkes and Seuken

Backward Induction

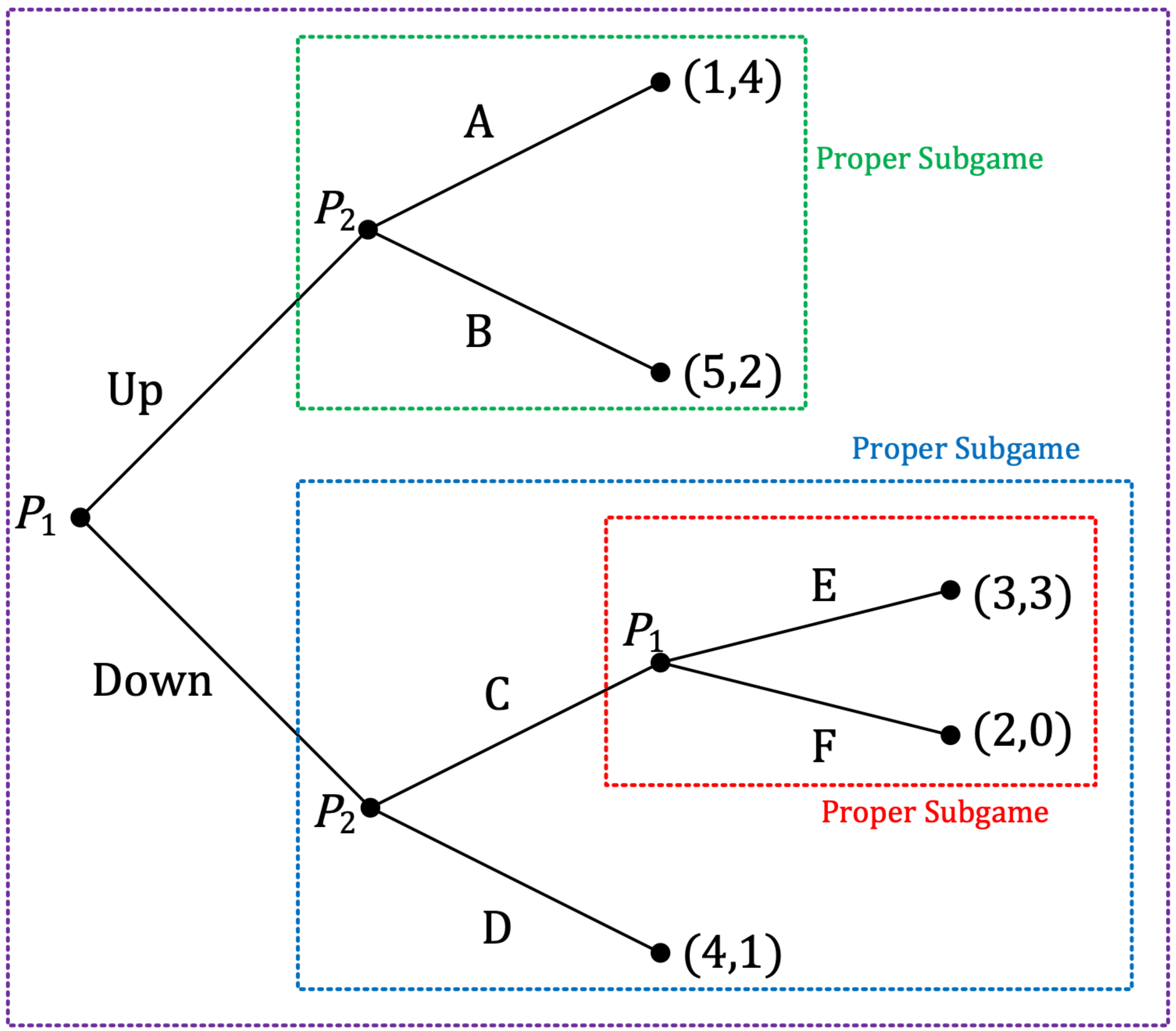
- Backward's induction is essentially "dynamic programming"
- You keep track of the optimal moves as you go up the tree



Credit: Textbook by Parkes and Seuken

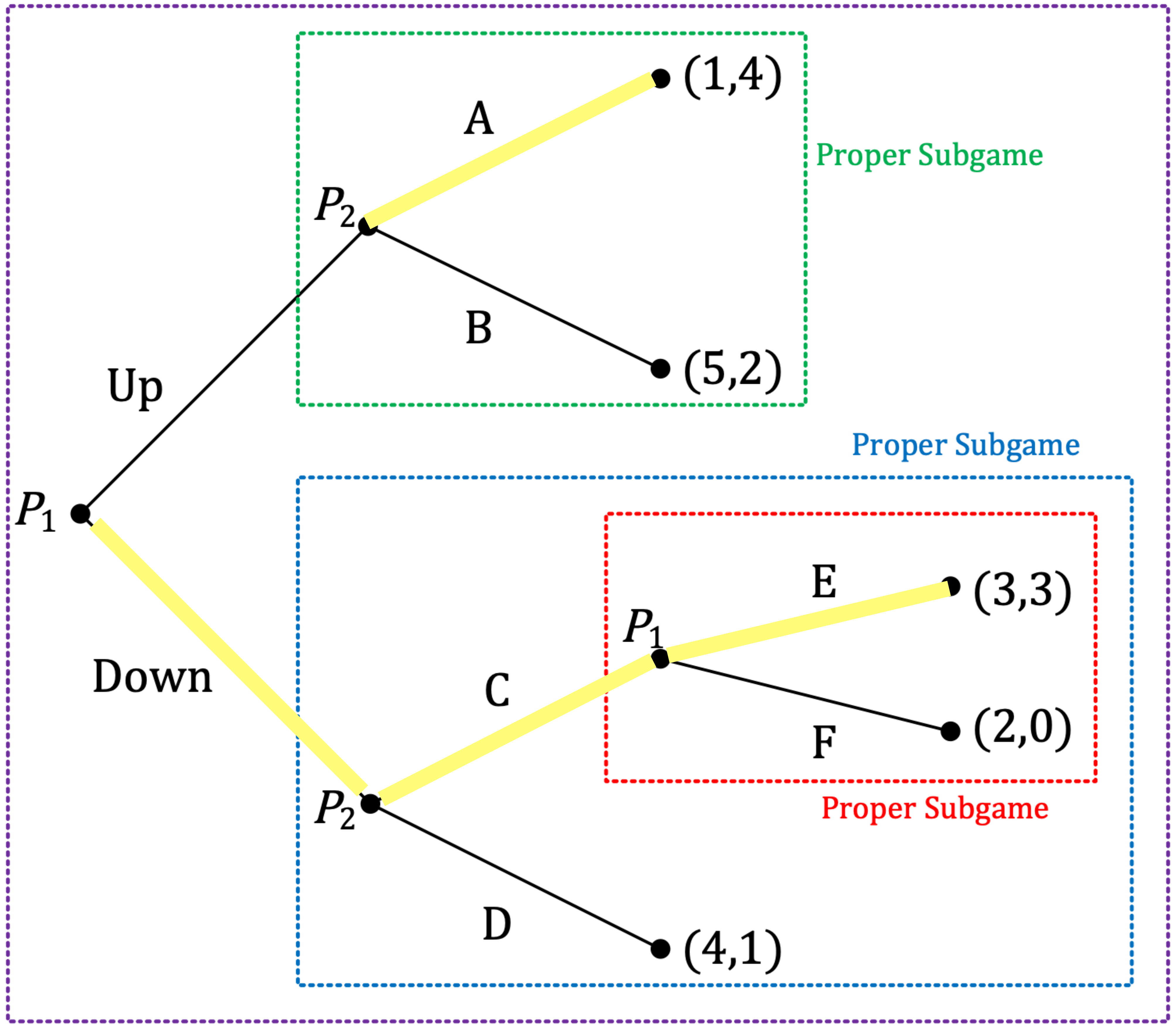
Backward Induction

- **Exercise**. Apply backward induction to find the subgame-perfect Nash equilibrium.



Backward Induction

- **Exercise**. Apply backward induction to find the subgame-perfect Nash equilibrium.



Single Deviation Principle

- Powerful tool to check if a strategy profile forms a SPE
- Assume: finite, extensive-form game s.t. exactly one player moves at each round
- A **single deviation** from strategy s_i is a strategy s_i' that differs from s_i only in the action prescribed at a single history h
- A single deviation is useful if, in the subgame rooted at h, agent i's utility from playing s_i' is strictly better than playing s_i (keeping the strategies s_{-i} fixed)
- **Single-deviation principle**. A strategy s is a subgame-perfect equilibrium in a finite extensive-form game if and only if no player has a useful single deviation.
    - Reduces checking a global optimality problem to checking a local one

Single Deviation Principle Proof

- $(\Rightarrow)$ Suppose s is a SPE, show that there is no single-deviation that benefits a player
  - s is a Nash at every history
  - By definition, no one can benefit by deviating
- $(\Leftarrow)$ Suppose s is not a SPE, show that there exists a history h, a player $P(h)$ such that $P(h)$ deviating from s_i at h (keeping s_{-i} fixed) gives i a better utility
- Since s is a not a SPE, there exists a history h s.t. s is not a Nash in the subgame rooted at h:
  - There exists a player i that has a useful (potentially) multi-step deviation at some subgame rooted at h

Single Deviation Principle Proof

- $(\Leftarrow)$ Suppose s is not a SPE, there is a player i that has a useful (potentially multi-step) deviation starting from some decision node at history h
- Let s_i' be a profitable deviation strategy that is minimally different from s_i
- Let $h*$ be the longest history such that $s_i'(h*) \neq s_i(h*)$:
    - This is the last place where the two disagree (same actions after this point)
    - Such a $h*$ must exist in a finite game
- Consider the subgame rooted at $h*$: action $s_i'(h*)$ is a single deviation from s
- Playing s_i until $h*$ followed by $s_i'(h*)$ must give better utility
    - If not, then s_i' is not minimally different from s_i

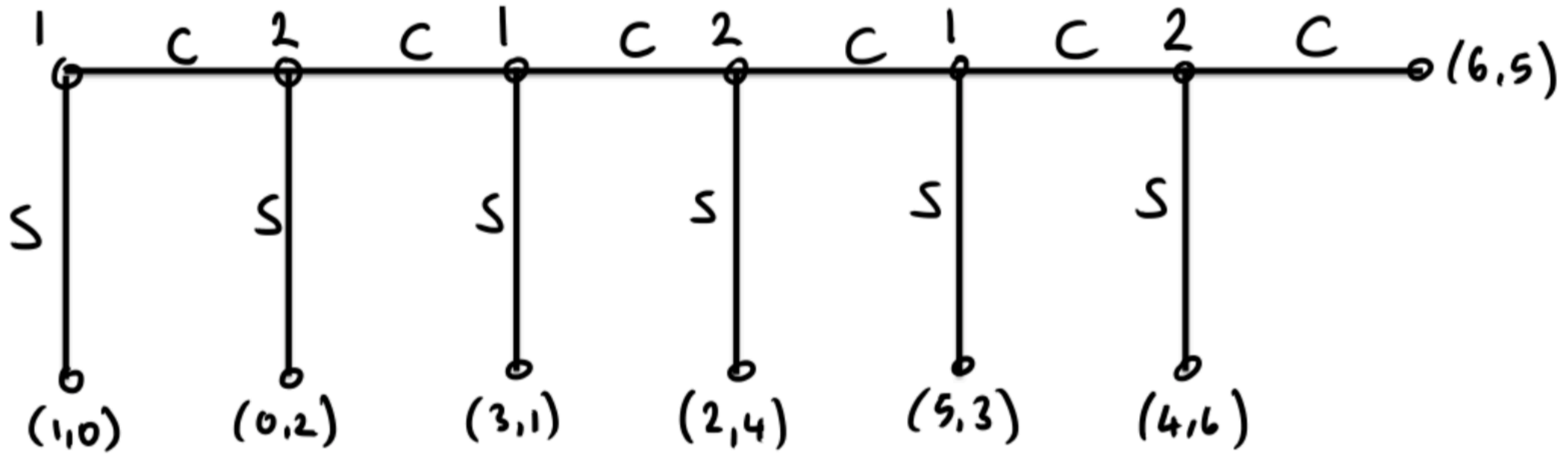
Existence of SPE

- **Theorem**. Any finite extensive-form game has a subgame-perfect equilibrium
- **Proof.**
- A finite extensive-form game has a finite decision tree
    - Each player has a finite number of actions at each history
    - Only a single player moves at any history
- Use backward induction to assign an action to each node in the decision tree that is best response, given the actions assigned to subsequent nodes
- This strategy profile satisfies single-deviation principle by construction and is thus a SPE

SPE: Caveats

Centipede Game

- Two players, each can play S (stop) or C (continue)
- Playing S always stops the game
- Subgame perfect equilibrium?
- Why is this a bit paradoxical?



Credit: Textbook by Parkes and Seuken

Extensive-Form Extensions

- Randomized strategies:
  - modeled through mixed strategies (players randomize over set of strategies in the beginning) or behavioral strategies (randomized over action sets at each history)
- Imperfect information:
  - The history of the game (moves of some players) is not perfectly accessible to other players
- Incomplete information (Bayesian games)
  - Players are not sure of the “type” of the other players
  - Can be turned into imperfect information games if we impose a common prior
- Simultaneous moves in a single turn (e.g. repetition of a one-shot game)

Playing Poker

RESEARCH ARTICLE

COMPUTER SCIENCE

Heads-up limit hold'em poker is solved

Michael Bowling,[1*] Neil Burch,[1] Michael Johanson,[1] Oskari Tammelin[2]

Poker is a family of games that exhibit imperfect information, where players do not have full knowledge of past events. Whereas many perfect-information games have been solved (e.g., Connect Four and checkers), no nontrivial imperfect-information game played competitively by humans has previously been solved. Here, we announce that heads-up limit Texas hold'em is now essentially weakly solved. Furthermore, this computation formally proves the common wisdom that the dealer in the game holds a substantial advantage. This result was enabled by a new algorithm, CFR$^+$, which is capable of solving extensive-form games orders of magnitude larger than previously possible.

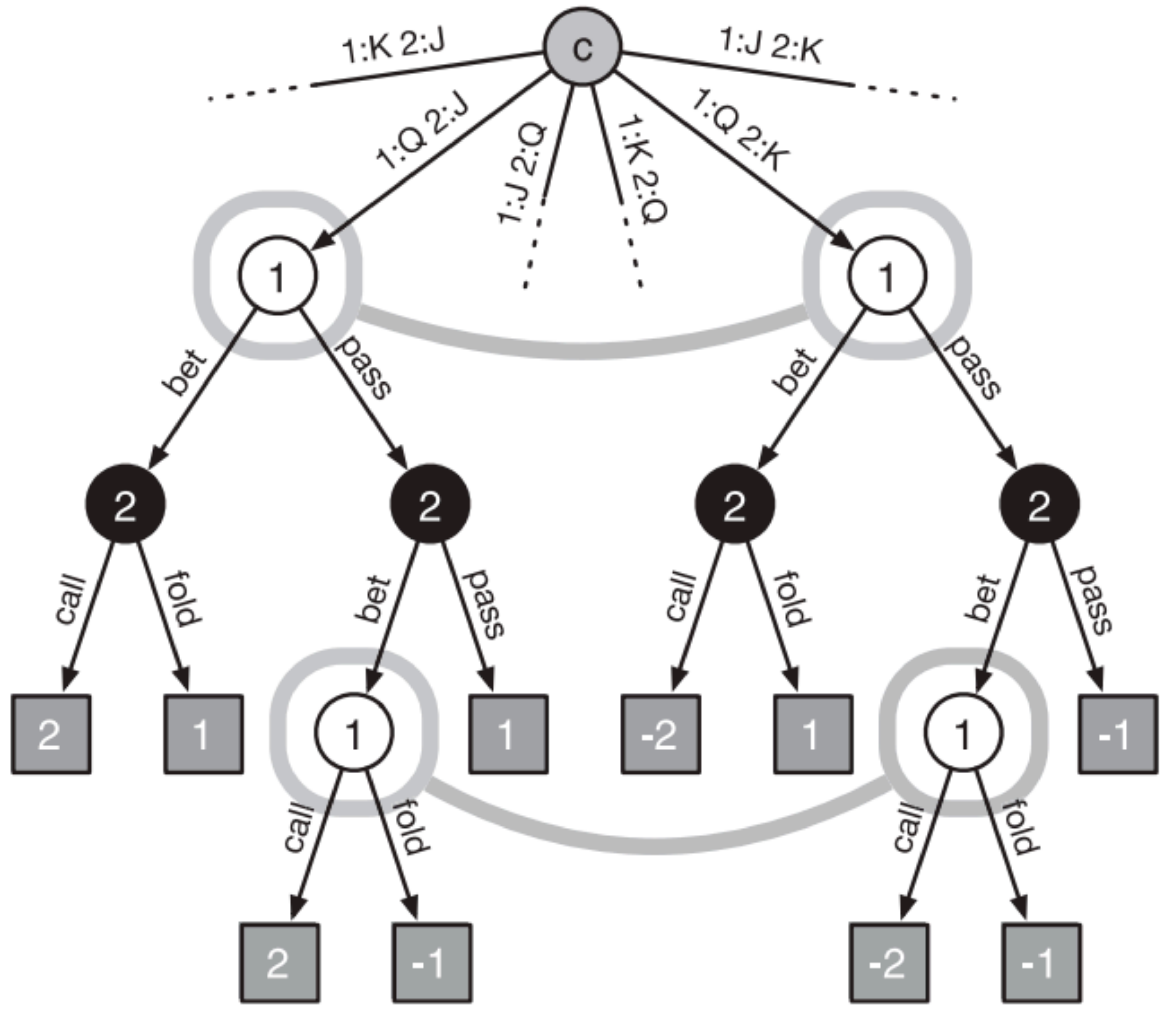


Fig. 1. Portion of the extensive-form game representation of three-card Kuhn poker (*16*). Player 1 is dealt a queen (Q), and the opponent is given either the jack (J) or king (K). Game states are circles labeled by the player acting at each state ("c" refers to chance, which randomly chooses the initial deal). The arrows show the events the acting player can choose from, labeled with their in-game meaning. The leaves are square vertices labeled with the associated utility for player 1 (player 2's utility is the negation of player 1's). The states connected by thick gray lines are part of the same information set; that is, player 1 cannot distinguish between the states in each pair because they each represent a different unobserved card being dealt to the opponent. Player 2's states are also in information sets, containing other states not pictured in this diagram.

Repeated Games

Prisoner's Dilemma

- One shot game:
    - (D, D) is the unique DSE
- What if the players played it over and over again?
- Would cooperation emerge?

	C	D
C	4, 4	0, 5
D	5, 0	2, 2

	C	D
C	a, a	b, c
D	c, b	d, d

$c > a > d > b$

Split or Steal

- Players can choose split or steal the prize money
  - If both steal, no one gets any money
  - If one splits, other steals: the thief gets all the money
  - If both split: they share the only in half
- Weakly dominant action?
  - Steal weakly dominates Split for both players
- In both the video game and game show, the game is multi-stage and current decisions have future consequences
- Cooperation is often seen in all these situations
- https://www.youtube.com/watch?v=S0qjK3TWZE8&ab_channel=spinout3

	Split	Steal
Split	1/2, 1/2	0, 1
Steal	1, 0	0, 0

Motivation: Incentives in P2P

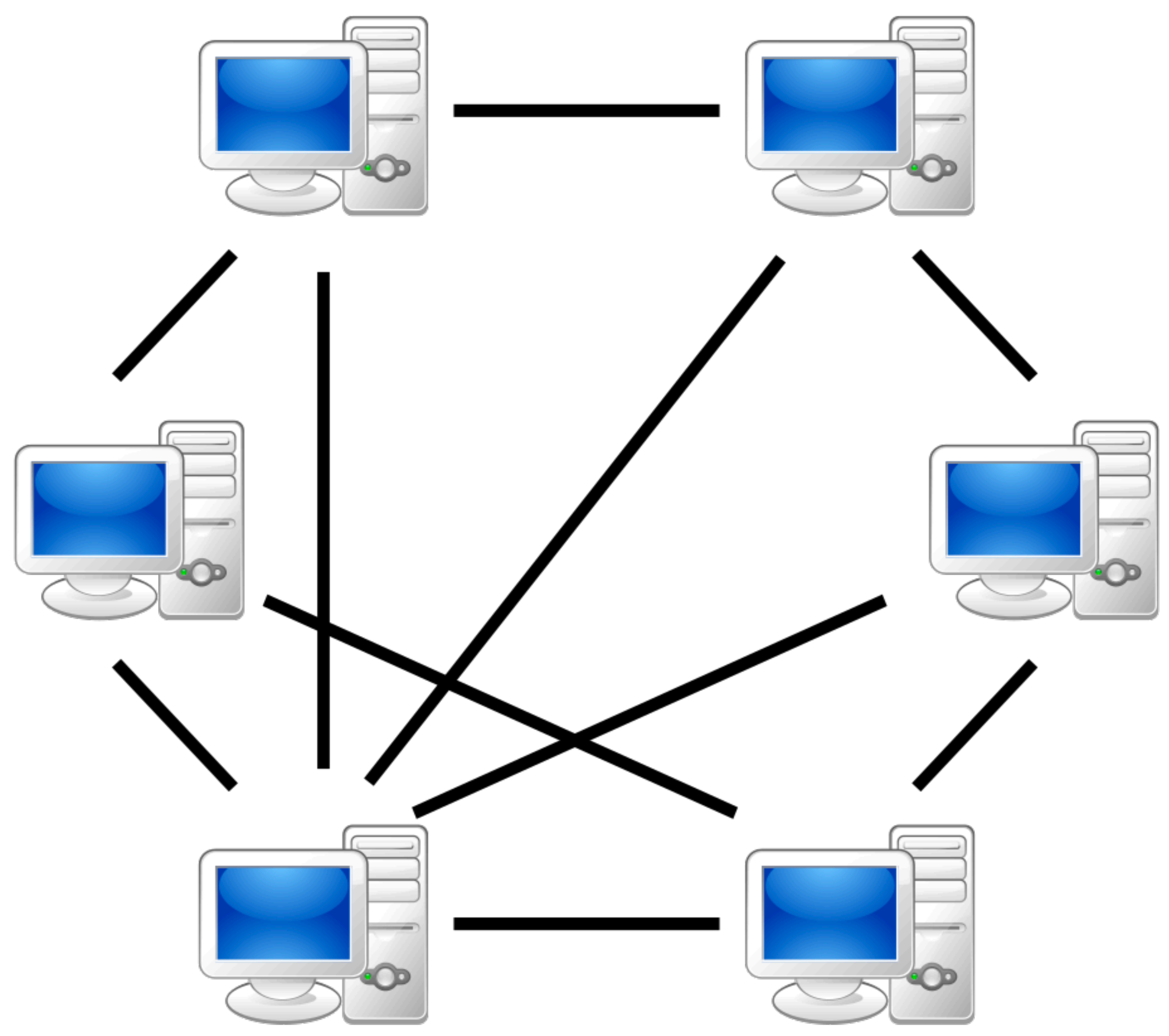
- P2P systems provide an intriguing case study of how a system evolves in response to incentive issues
  - Incentive properties vary widely across different protocols
- Peer-to-peer file sharing:
  - Way to distribute a file between users where they upload and download from each other
- P2P is now fundamental to blockchain platforms, such as Bitcoin and Ethereum
- **AGT view**: **do peers in a P2P system to have an incentive to cooperate**?
  - For file sharing, do users have incentive to upload while downloading from peers?

Failure of Centralization

- In the days of early internet, file sharing was done in an ad hoc way
- Napster (1999): provided a centralized, searchable directory listing which users have copies of various files (e.g. mp3s)
    - Matchmaker (matched up people who want file to people who have the file)
    - File transfer was then done directly between users
- Lawsuits against Napster for copyright infringement (2000s)
    - By RIAA, Metallica, etc
- After Napster failed to comply, it was shut down in 2001
- Napster's rise (25 million users) pointed to the demand for such systems but its failure motivated **decentralized designs**

Benefits of P2P

- Client-server model: server provider is associated with the server machines, users device is a client machine
    - These platforms need to make use of millions of distributed servers in order to cache content on machines close to users to provide low latency and maintaining this infrastructure
- In contrast, P2P systems there is no distinction between client and servers: each computer acts as both and is called a peer
- Main advantage: can scale well to large numbers of users while keeping the costs low for the initial uploader of the content
    - Provide robustness by avoiding a single point of failure
- Disadvantages: no control over content and who will download it, for how long the files will be available, etc



Decentralized: Gnutella

- First decentralized P2P network of its kind
- Design highlighted various incentive issues inherent in P2P networks
- Functionality of Gnutella rests on users conforming to the following behavior:
    - Upon receiving a file request, either upload the file to the requester (if the user has the file) or forward to other peers
- Problem with the design. Users were not given any incentive to actually behave in this way
- **Free-riding in Gnutella**:
    - A user is a free-rider who downloads but never uploads
  - A study by researchers showed that free-riding was the dominant behavior in Gnutella: 2/3rd of the users were free riders
  - In follow up study in 2005, free riding had climbed to 85% leading to the extinction of the system

File Sharing Game

- Consider two players: Aamir and Beth
- Aamir has a file that Beth wants and vice versa
- They simultaneously and independently decide whether or not to upload the requested file
- For each player, the benefit of receiving the file is 3 and the cost of uploading is 1 (bandwidth charges, opportunity costs, etc)
- Four possibilities of a file transfer game in normal form

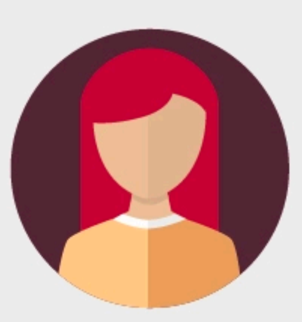
	Upload	Don't Upload
Upload	$2, 2$	$-1, 3$
Don't Upload	$3, -1$	$0, 0$

File Sharing Game

- Consider two players: Aamir and Beth
- Aamir has a file that Beth wants and vice versa
- They simultaneously and independently decide whether or not to upload the requested file
- For each player, the benefit of receiving the file is 3 and the cost of uploading is 1 (bandwidth charges, opportunity costs, etc)

Aamir

Beth



	Upload	Don't Upload
Upload	$2, 2$	$-1, 3$
Don't Upload	$3, -1$	$0, 0$

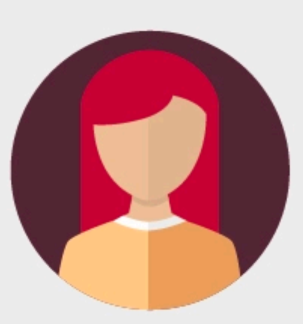
Prisoner's Dilemma

- Our payoff matrix is just a variant of the prisoner's dilemma game from Lecture 2
- Each player has a strictly dominant strategy to defect
    - In this case, to not upload
- When Aamir and Beth play their dominant strategy neither uploads and each gets a payoff of zero
- Prisoner's dilemma summarizes the essential conflict between individual good and the collective good

	Upload	Don't Upload
Upload	$2, 2$	$-1, 3$
Don't Upload	$3, -1$	$0, 0$

Aamir

Beth



	C	D
C	a, a	b, c
D	c, b	d, d

$$c > a > d > b$$

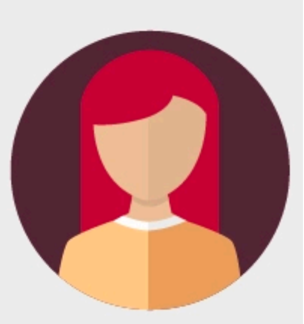
Repeated Prisoner's Dilemma

- In real life examples of Prisoner's dilemma players do seem to cooperate: how can we explain this?
- **Intuition**: in real-life settings, the short term gain from defecting is outweighed by its long-term costs
- The model of repeated game try to capture this aspect:
  - How cooperation develops in long-term play
- **Idea**: Suppose we repeated Prisoner's dilemma n times
  - What behavior to be expect to see?

	Upload	Don't Upload
Upload	$2, 2$	$-1, 3$
Don't Upload	$3, -1$	$0, 0$

Aamir

Beth



	C	D
C	a, a	b, c
D	c, b	d, d

$c > a > d > b$

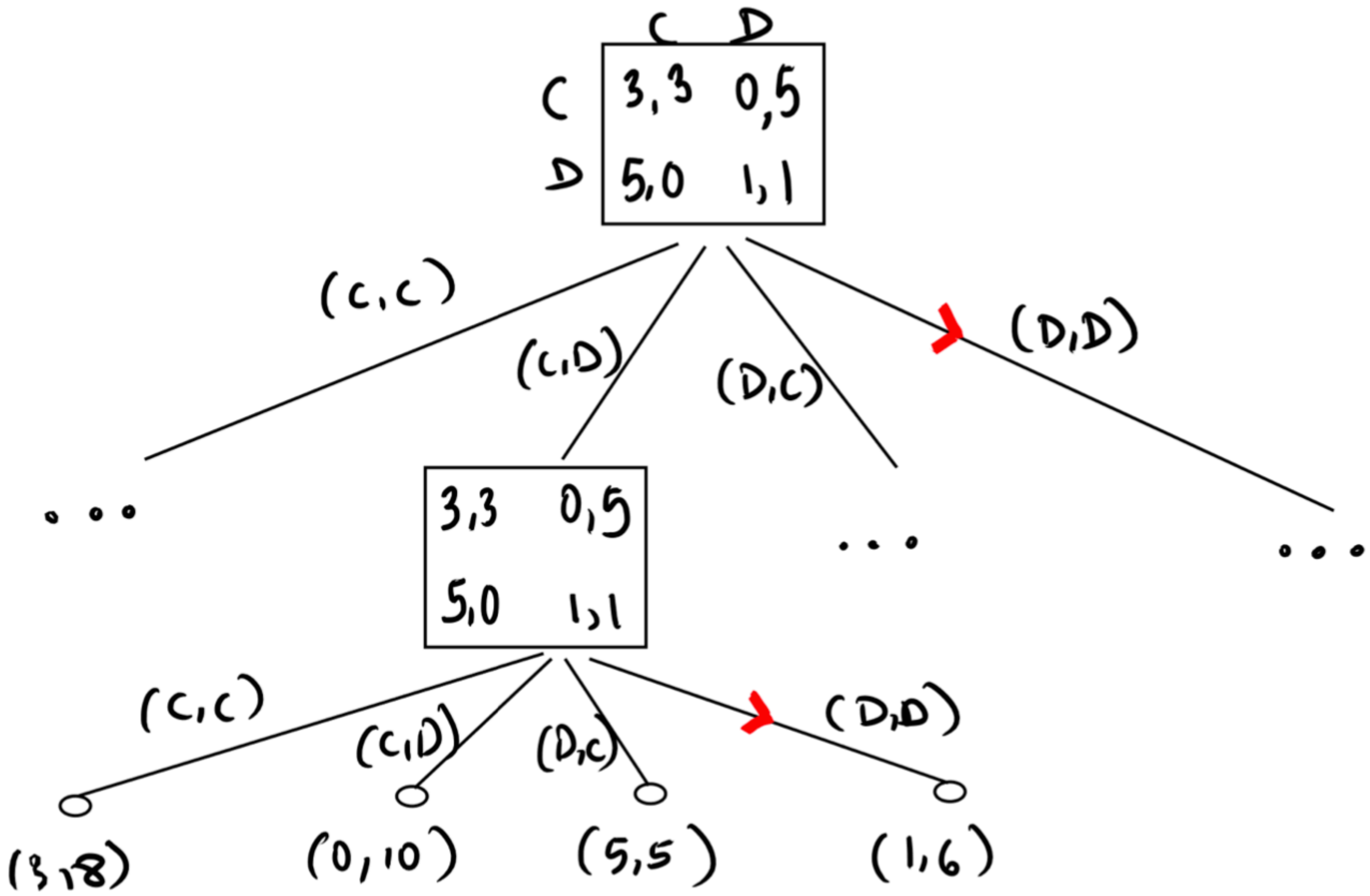
Finitely Repeated Games

- Consider the following finitely-repeated Prisoner's dilemma
- **Finitely-repeated Prisoner's Dilemma**:
  - Aamir and Beth play the one-shot simultaneously move Prisoner's dilemma game n times (for some known $n \geq 1$)
  - The total utility of each is the sum of utilities across the n rounds
- Note that this is a sequential game in the sense that the action chosen by players in round i can depend on the history
- As we grow n, does cooperation emerge?
- Use backward induction: best response in round n?
  - Best response in any round?
- Unique SPE: **defect in each round** 😔

Finitely Repeated Games

- Consider a normal-form game (**"stage game"**) $G = (N, \tilde{A}, \tilde{u})$ with action set $\tilde{A}_i$ for player i and utility $\tilde{u}_i(a)$ for player i on outcome profile $a = (a_1, \ldots, a_n)$
- **Finitely repeated game.** In a finitely-repeated game G^T, the stage game G is played by the same players for $T \geq 1$ rounds, such that each player has perfect information about the history of actions in all previous rounds
- To represent a repeated game in **extensive-form**, we allow simultaneous moves in each round
  - The history of the game is now a sequence of action profiles (instead of a sequence of individual actions)
  - The utility of each player is the sum of utilities along the history

Example: Repeated PD



Credit: Textbook by Parkes and Seuken

Finitely Repeated Games

- We can generalize the trend we saw for repeated Prisoner's dilemma to any normal-form game
- **Theorem**. If the stage game has a unique Nash equilibrium, then the only subgame-perfect equilibrium strategy in a finitely-repeated game is to play the stage game Nash equilibrium strategy after each possible history.
- **Proof**. By backward induction, from the final period
  - Only best response in final round is to play the unique Nash equilibrium
  - In second-last round, the actions of the players does not effect the payoff in the last round, so best response is to play the unique Nash in this round and so on

Takeaways

- The model of finitely-repeated games fails to capture the emergence of cooperation in real-world settings
- What is possible missing from the model?
  - Element of **uncertainty about the future** (when does the game really come to an end?)
  - How players might value short term payment differently from payment that is long in the future
    - "**Discount**" future payments
- The model of **infinitely repeated games** captures this intuition
- In fact, there is a SPE of prisoner's dilemma where (C, C) is sustained in every period: this shift happens when we move to infinitely repeated games

Infinitely Repeated Games

- **Infinitely-repeated Prisoner's dilemma**:
  - Aamir and Beth play the one-shot simultaneously move Prisoner's dilemma game
  - With probability δ (where $0 < \delta < 1$) the game continues for another round
  - With probability $1 - \delta$ the game ends at this round
- In the literature, this δ is often represented as players "**discounting the future payoffs**"
- Suppose players get utility a, a', a'' from three rounds, then their "discounted utility" is $a + \delta a' + \delta^2 a''$
- You can think of it as **expected utility** or **discounted utility**

Infinitely Repeated Games

- **Example**. Suppose the sequence of play in a three-round Prisoner's dilemma is $(C, C), (C, C)$ and (D, C)
- Suppose $\delta = 0.9$
- What is player 1's "discounted" utility
  - With probability 1 player 1 gets 2 in round 1
  - With probability 0.9 player 1 gets 2 in round 2
  - With probability 0.9^2 player 1 gets 3 in round 3
- Overall expected/discounted utility is thus $2 + 0.9(2) + 0.81(3) = 6.23$

	C	D
C	2, 2	−1, 3
D	3, −1	0, 0

Infinitely Repeated Games

- **Infinitely-repeated Prisoner's dilemma**:
  - Aamir and Beth play the one-shot simultaneously move Prisoner's dilemma game
  - With probability δ (where $0 < \delta < 1$) the game continues for another round
  - With probability $1 - \delta$ the game ends at this round
- Both players now want to maximize their total expected utility or total discounted utility defined as $u_i(h) = \sum_{k=0}^{\infty} \delta^k \tilde{u}_i(a^{(k)})$ where $a^{(k)}$ is the action profile chosen in the kth round
- Intuition if δ is sufficiently large, then cooperation should emerge
  - Note that $\delta = 0$, we are back to the one-shot game

When δ is sufficiently large: players are interpreted as "**patient**"

Infinitely Repeated Games

- Caution about infinitely-repeated games:
  - Can potentially have a **lot of equilibria**
  - Specific equilibria thus reduces predictive power
  - Known problem with the repeated-game framework and a topic of ongoing research
- Often the focus is on what type of strategies can be sustained by subgame-perfect and Nash equilibria
- First, we will discuss what type of **symmetric strategies** can be sustained in equilibrium
- Second, we will discuss what type of strategies seem to do well in an **asymmetric environment** based on **empirical study**

Trigger Strategies

- In repeated games, a **trigger strategy** essentially threatens the opponent with a “worse” punishment if they deviate from an implicit agreed upon action profile
- The most extreme (and unforgiving) trigger strategy is the **grim trigger strategy** that punishes forever after a single deviation
- Suppose Beth plays the following **grim trigger strategy**:
  - Start by cooperating, but if opponent has ever defected in the past, then defect
  - Otherwise, cooperate
- If Beth plays such a strategy, what is Aamir’s best response in a Nash or subgame perfect equilibrium?

Grim Trigger Strategy

- First let us reason about the Nash equilibria (before we think about subgame-perfect)
- If Aamir ever defects, Beth punishes him forevermore
- What is Aamir's best response?
- Suppose we are in stage i
  - If Aamir ever defected in the past, then Beth's future behavior is fixed so Aamir might as well defect forever
  - If Aamir has cooperated in the past $i-1$ stages for $i \geq 1$, should he continue to cooperate?

	C	D
C	$2,\ 2$	$-1,\ 3$
D	$3,\ -1$	$0,\ 0$

	stage i	stages $\geq i+1$
C	2	$\geq 2\delta$
D	3	≤ 0

Grim Trigger Strategy

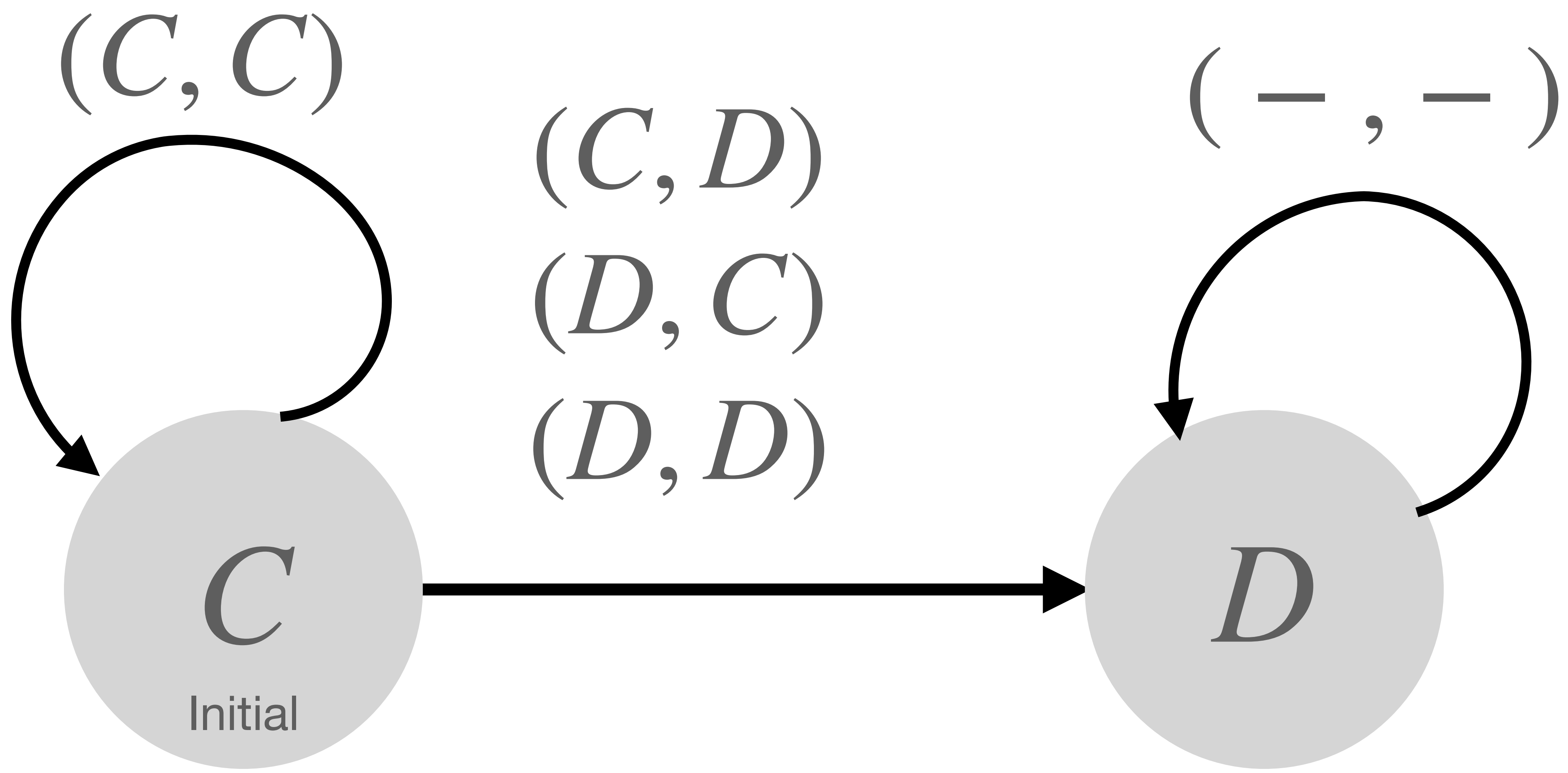
- Thus, the strategy pair (Grim, Grim) is a Nash equilibria if $\delta \geq 1/2$
- Notice that this generates the outcome (C, C) is each stage of the game
- Is this also a subgame perfect equilibria?
- Intuitively, for a Nash equilibria to be subgame perfect, the **threats must be credible**
- Challenge in analyzing subgame perfect:
    - Need to account for behavior on histories that may never be reached under equilibrium
- Matters a lot how we exactly we define the strategies followed by the players

	C	D
C	$2, 2$	$-1, 3$
D	$3, -1$	$0, 0$

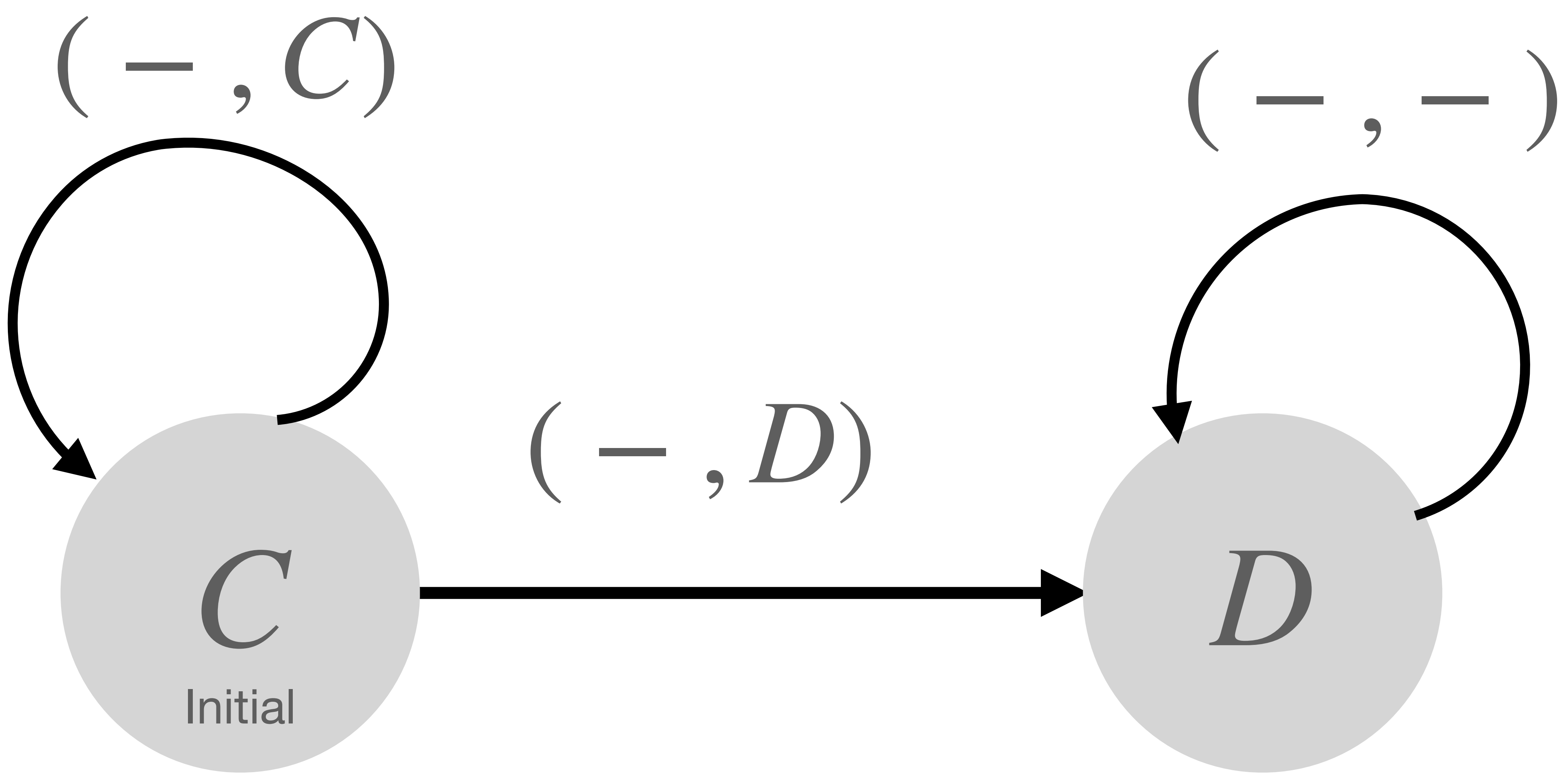
	stage i	stages $\geq i+1$
C	2	$\geq 2\delta$
D	3	≤ 0

Automaton Strategies

- Consider the following two automaton strategies for grim trigger
- Grim Trigger (left) is a subgame-perfect equilibrium for $\delta > 1/2$
- But the modified Grim Trigger (right) is **NOT** a subgame-perfect equilibrium for any $\delta \in (0,1)$



Grim Trigger



Modified Grim Trigger

Single Deviation Principle

- Single-deviation principle holds in an infinitely repeated game with discounting
- **Theorem**. A strategy profile is a subgame-perfect equilibrium in an infinitely-repeated game with discounting if, and only if, there is no useful single deviation.
- We will use this result without proof
- To reason about single deviation, its is useful to draw out the "reduced" decision tree for each strategy
  - (Game tree on board for Grim trigger)
  - The actions $(C, D), (D, C), (D, D)$ all have the same outcome and can be "reduced" for analysis

	C	D
C	$2, 2$	$-1, 3$
D	$3, -1$	$0, 0$

Grim in SPE

- **Lemma**. Grim Trigger strategy is a symmetric subgame-perfect Nash for all $\delta > 1/2$.
- **Proof**. We only need to consider two types of histories at stage $i \geq 1$:
- **Case 1: Cooperation history** at stage i. Histories in which D has never been played by any player in stages $i-1$
- **Case 2: Defection history** at stage i. Histories where at least one player has played D in stages $i-1$.
- For case 1, suppose both players continue with (C, C) prescribed by Grim trigger, then their payoff is
  - $\geq 2 + 2\delta$

	C	D
C	2, 2	−1, 3
D	3, − 1	0, 0

Grim in SPE

- **Lemma**. Grim Trigger strategy is a symmetric subgame-perfect Nash for all $\delta > 1/2$.
- **Proof**. We only need to consider two types of histories at stage $i \geq 1$:
- **Case 1: Cooperation history** at stage i. Histories in which D has never been played by any player in stages $i-1$
- For case 1, suppose both players continue with (C, C) prescribed by Grim trigger, then their payoff is
  - $\geq 2 + 2\delta$
- If either player deviates to D in stage i, their payoff is
  - At most $3 + 0$ (for all future stages)
- Thus for $\delta > 1/2$, this deviation is not beneficial for Case 1

	C	D
C	$2,\ 2$	$-1,\ 3$
D	$3,\ -1$	$0,\ 0$

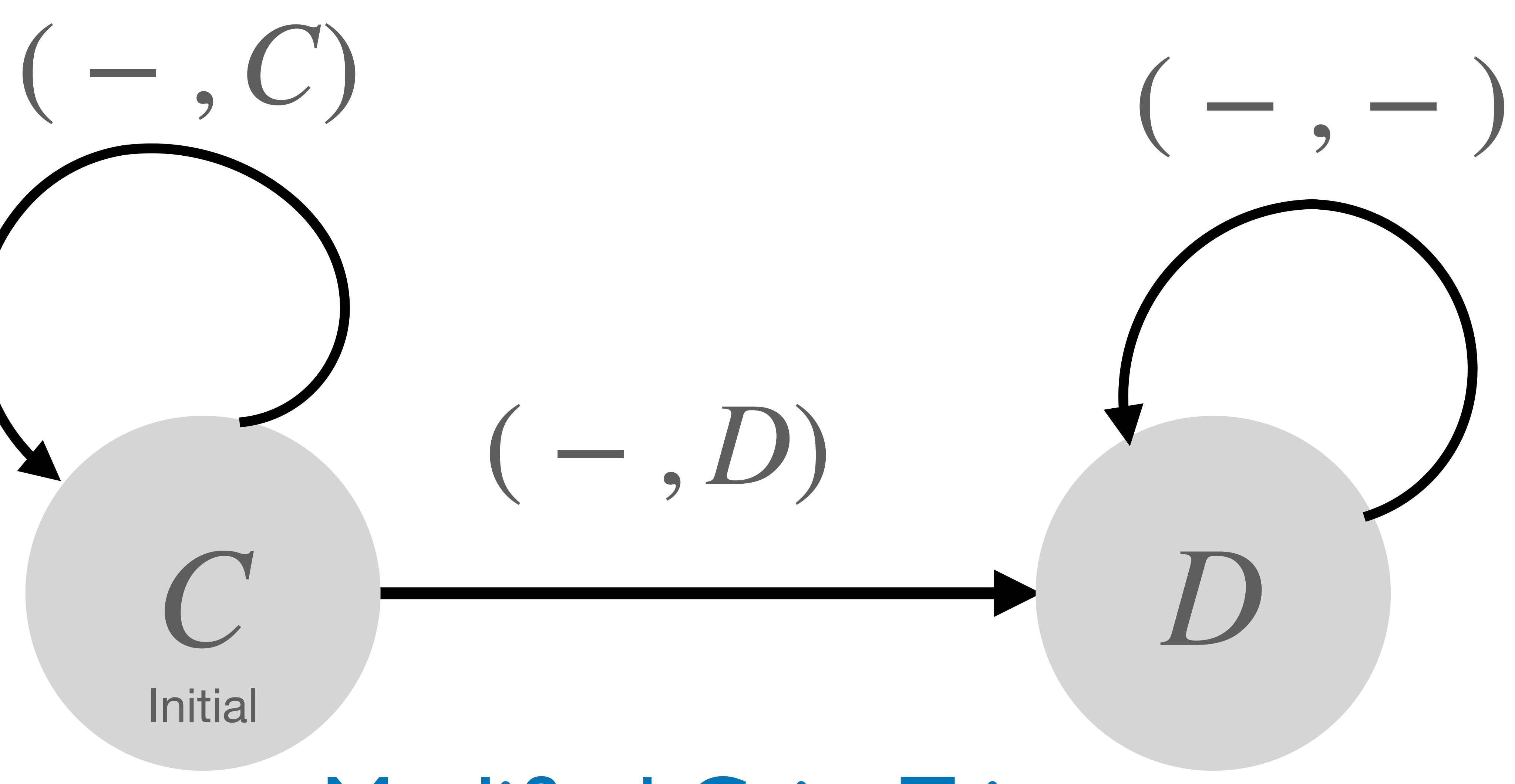
Grim in SPE

- **Lemma**. Grim Trigger strategy is a symmetric subgame-perfect Nash for all $\delta > 1/2$.
- **Proof**. We only need to consider two types of histories at stage $i \geq 1$:
- **Case 2: Defection history** at stage i. Histories where at least one player has played D in stages $i - 1$.
- If both players follow grim trigger strategy (D, D) in stage i
  - Their payoff is 0
- If either player deviates to C, their payoff becomes
  - $-1 + 0$
- Thus single-deviation is not useful for this case as well

	C	D
C	$2, 2$	$-1, 3$
D	$3, -1$	$0, 0$

Modified Grim is not SPE

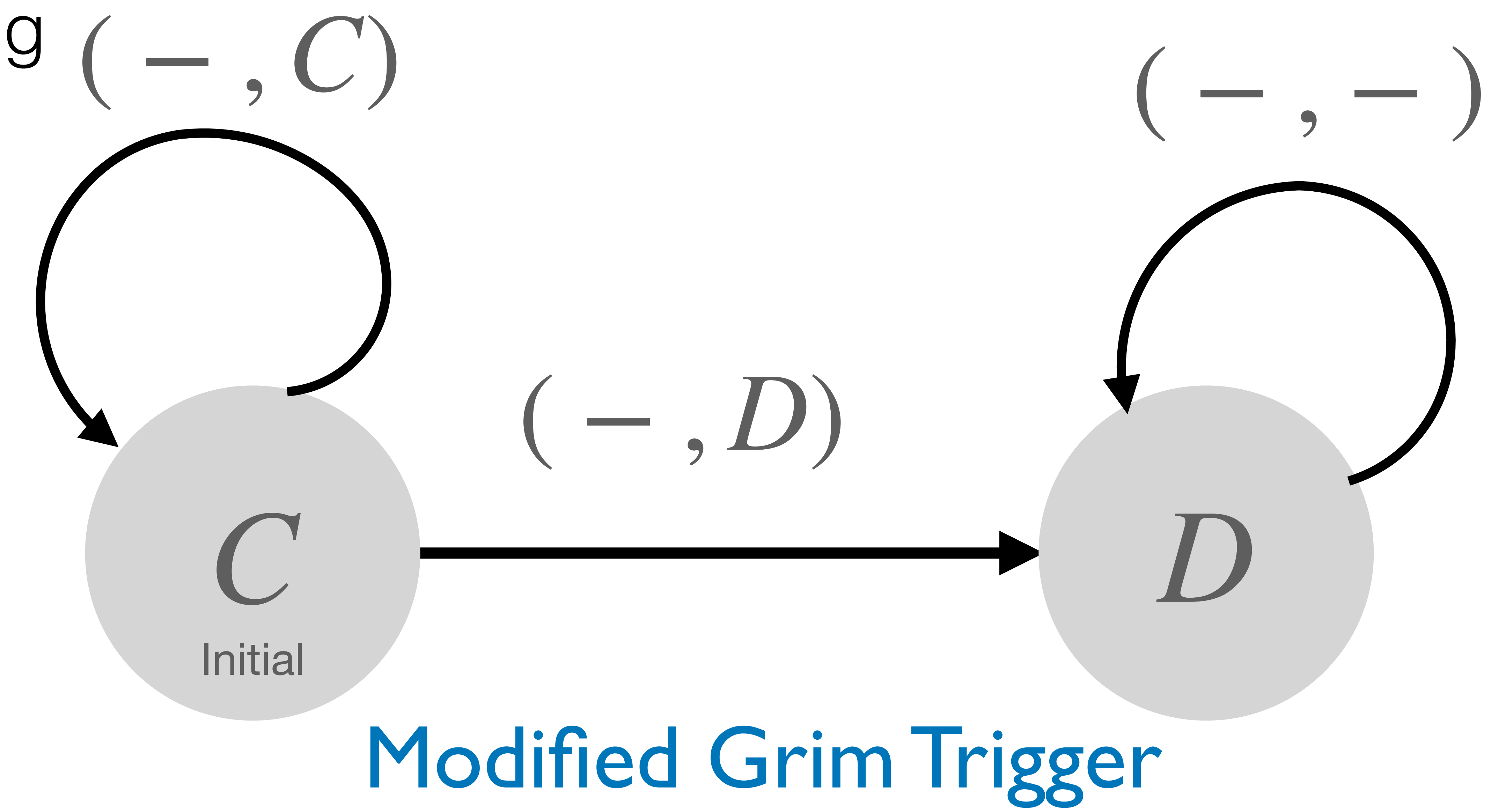
- **Lemma**. Modified Grim Trigger strategy is not a symmetric subgame-perfect equilibrium for any $\delta \in (0,1)$
- **Proof**. The difference now is that the action profiles (C, D) and (D, C) diverge in the game tree
- To show this is not a subgame-perfect equilibrium we look at the root of the tree (stage 0)
- Consider the subgame following outcome (C, D)
- Suppose player 1 adheres to modified grim and plays D in next stage
- **Claim**. It is not optimal for player 2 to play C according to modified grim



Modified Grim Trigger

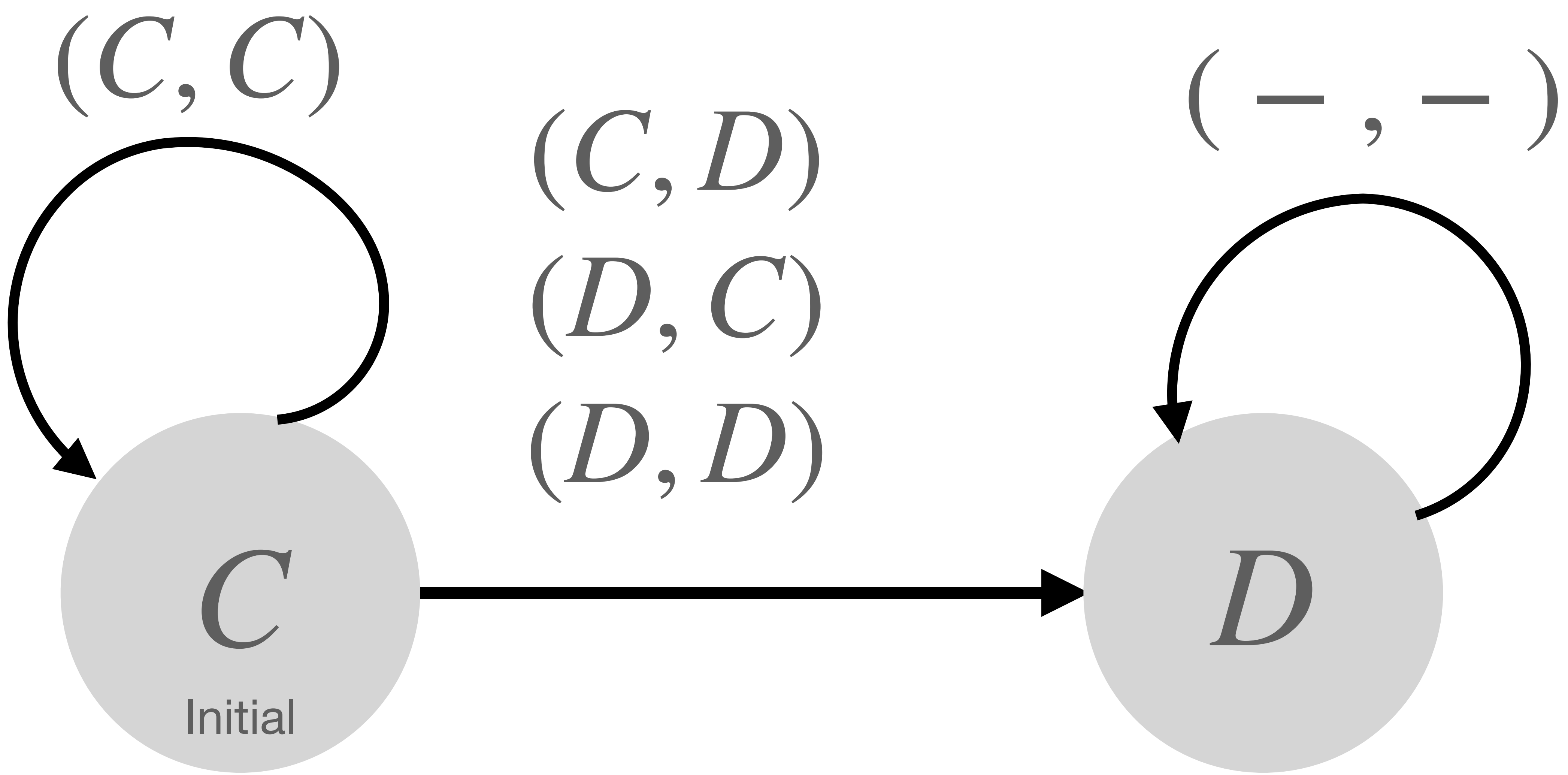
Modified Grim is not SPE

- **Lemma**. Modified Grim Trigger strategy is not a symmetric subgame-perfect equilibrium for any $\delta \in (0,1)$
- **Proof**. Consider the subgame following outcome (C, D)
- Suppose player 1 adheres to modified grim and plays D in next stage
- **Claim**. It is not optimal for player 2 to play C according to modified grim
- If player 2 adheres to modified grim, then outcome in next stage is (D, C) followed by (D, D) forever
  - Payoff -1
- If player 2 deviates to (D, D) gets at least 0

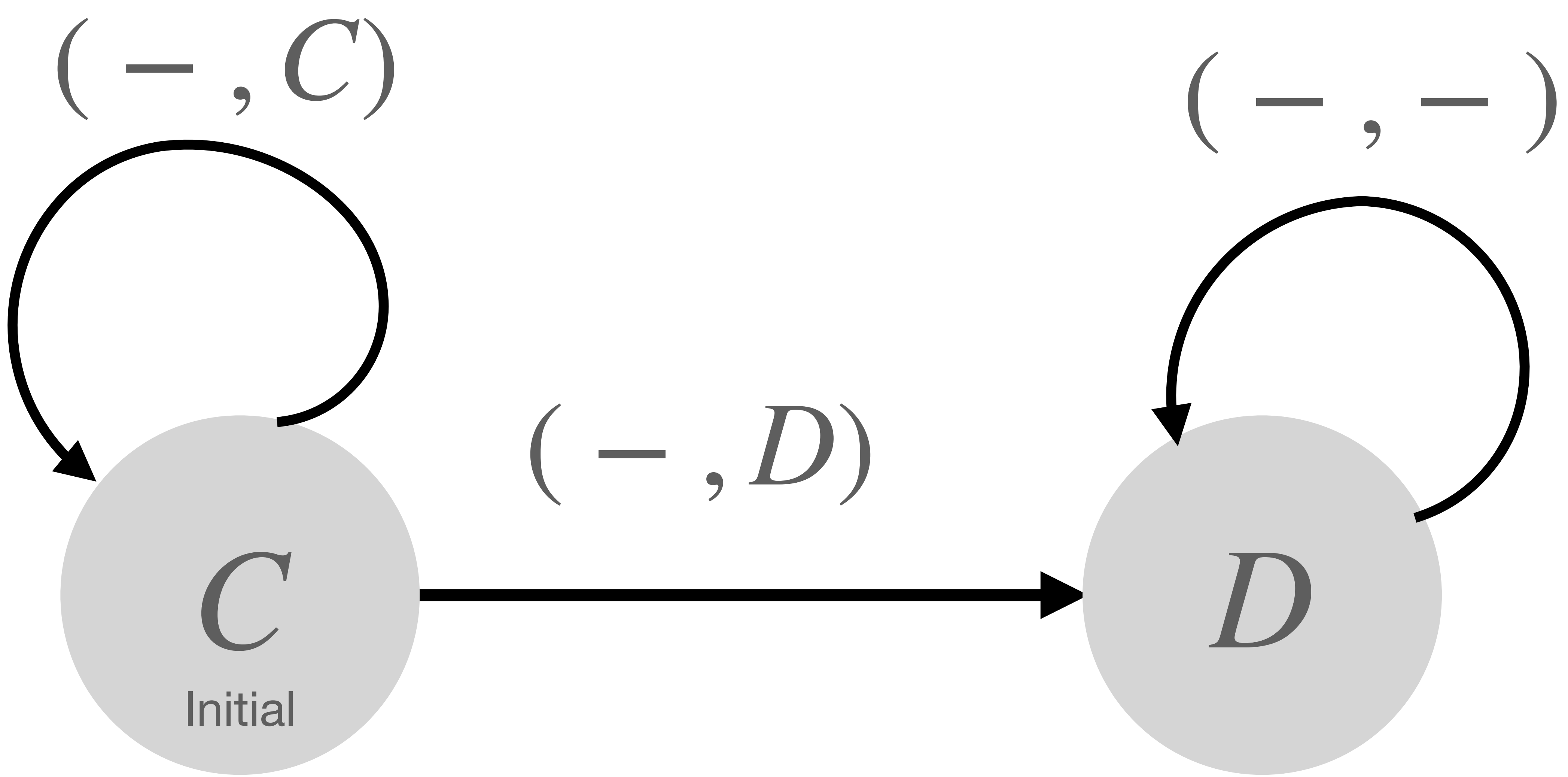


Takeaway: Nash vs SPE

- Even though "on the equilibrium path" no player should deviate to D is opponent has not deviated to D in the past
- SPE requires that a threat be credible even on histories that may never be played in equilibrium
- SPE is fragile wrt slight changes in how strategy is defined in repeated Prisoner's dilemma



Grim Trigger



Modified Grim Trigger

Tit-for-Tat

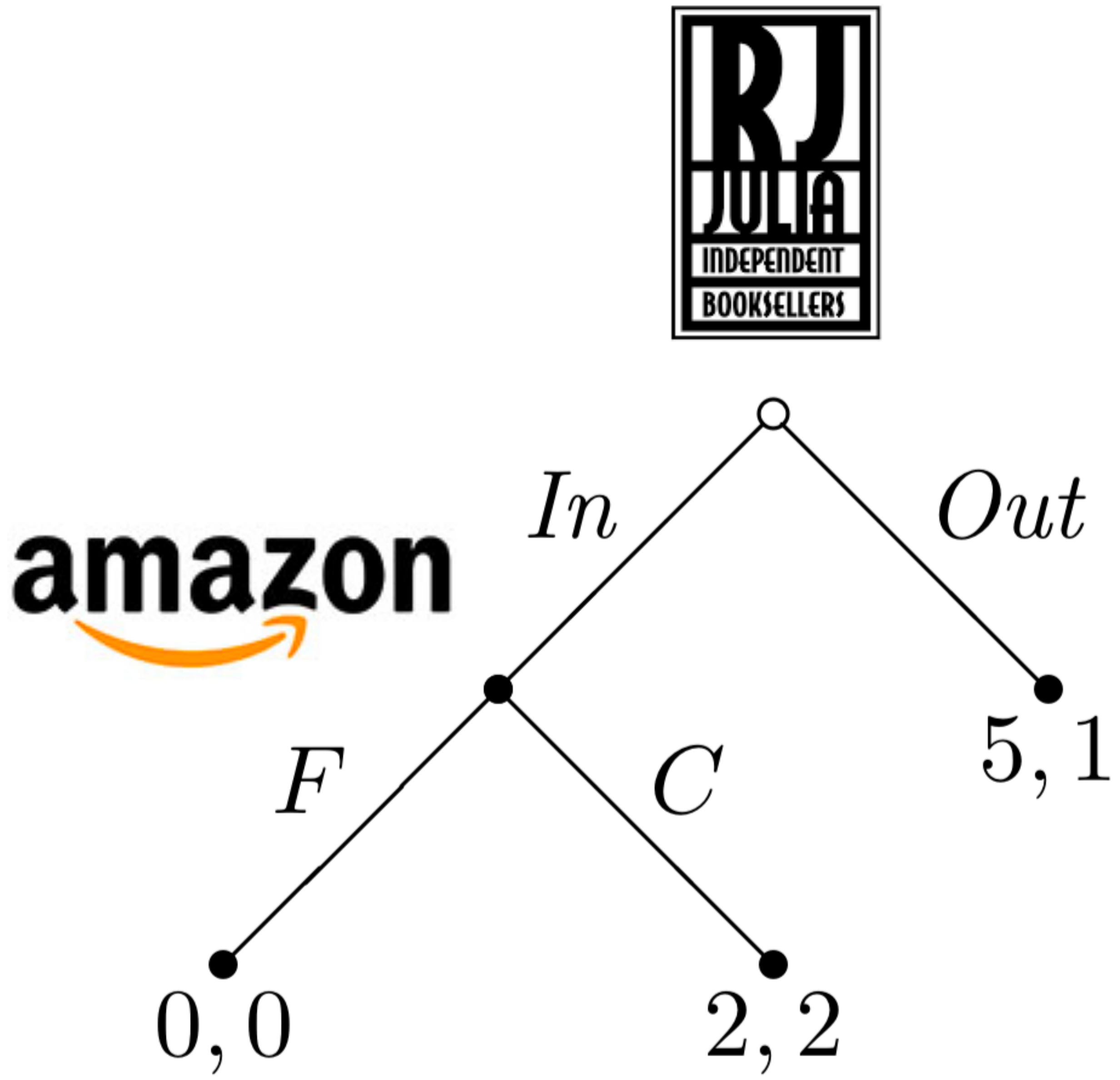
- Grim trigger strategy is pretty extreme (holding a grudge in perpetuity for a single defection)
- Even though Grim is at equilibrium with itself, a more robust strategy (for asymmetric environments) should involve some forgiveness
- **Tit-for-tat strategy**:
  - Start by cooperating
  - Do in stage i whatever the opponent does in stage $i - 1$
- Thus, tit-for-tat starts optimistically, punishes immediately and forgives quickly
- Turns out to be a good strategy in repeated prisoner's dilemma
  - Also perhaps in life?

Tit-for-Tat

- We show that TfT is a symmetric Nash equilibrium
- Suppose Beth is playing TfT, and we consider if Aamir can benefit by deviating in some stage i
  - Notice that stage i's payoff only depends on stage $i-1$
  - Deviating to D will lead to 3 now but at most 0 in next round
  - Cooperating now will lead to 2 now and 2δ in the next round
- Thus for $\delta > 1/2$, Aamir has no incentive to deviate
- In **Homework 8**, you are asked to show that
  - TfT is not a SPE for any $\delta < 1$
  - Modifying TfT to become a conditional cooperator (always return to C after D) is a SPE for sufficiently large δ

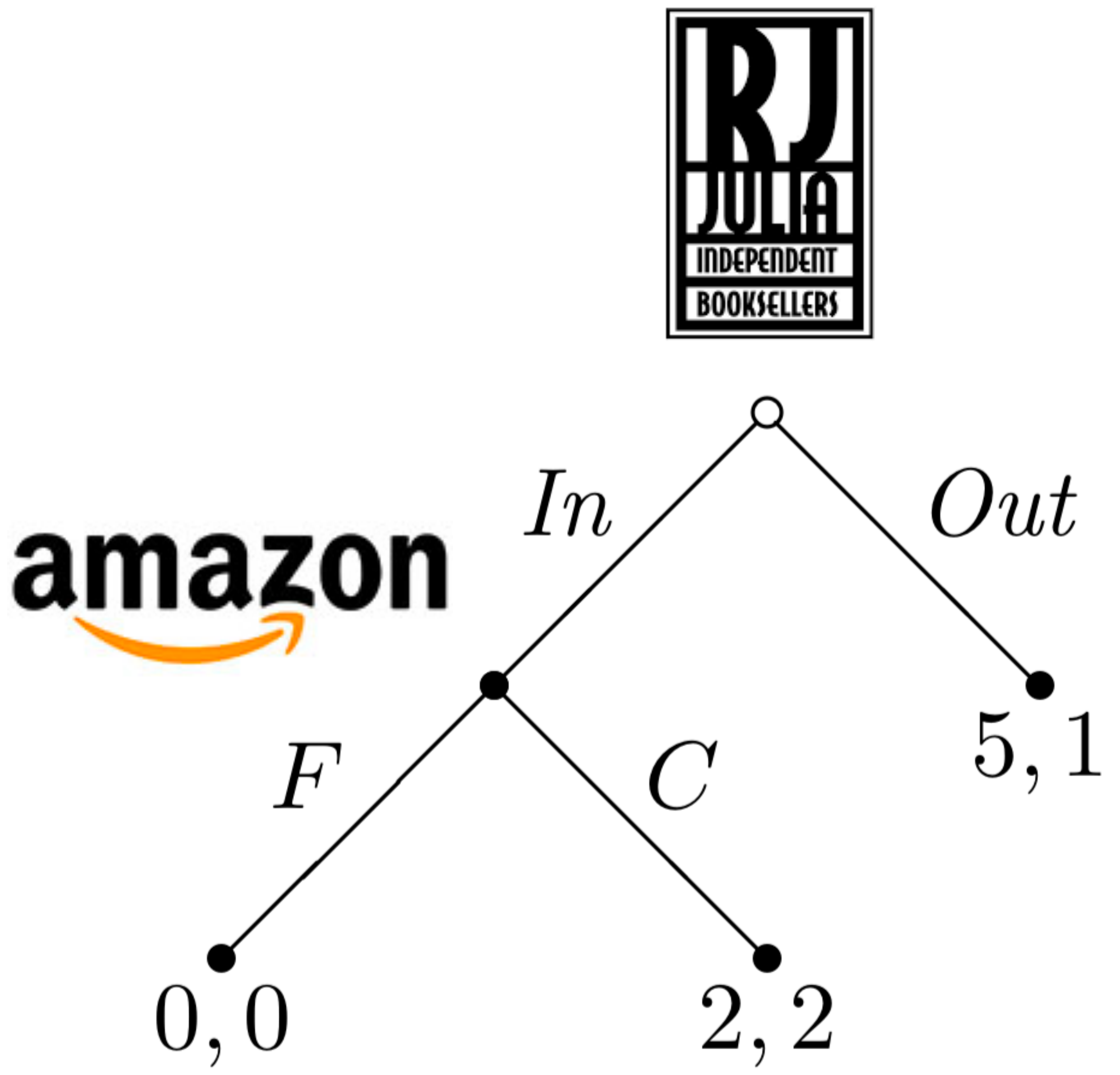
Chain Store Game

- Consider a game between a monopolist "chain store" and an local business "entrant"
  - Say amazon and RJ Julias (independent book store)
- If entrant enters, the monopolist can either begin a price war (fight) or share the market (cooperate)



Chain Store Game

- Backward induction says, if entrant enters market, it is in best interest of chain store to cooperate
- Given best response of chain store, entrant should always enter



Chain Store Game

- Now suppose we repeat this process in k different cities with different local business as entrants
- Say on day i, business in city i decides to enter or not for $1 \leq i \leq k$
- Backward induction says, regardless of how we got to a particular history h (chain store could have chosen to "fight" in these histories), each player should play best response at h
- Critiques of backward induction
  - An entrant that has observed a chain store play "fight" may not believe that it will play best response "cooperate"
  - A chain store may want to deter future entrants by playing "fight" (deterrence theory)