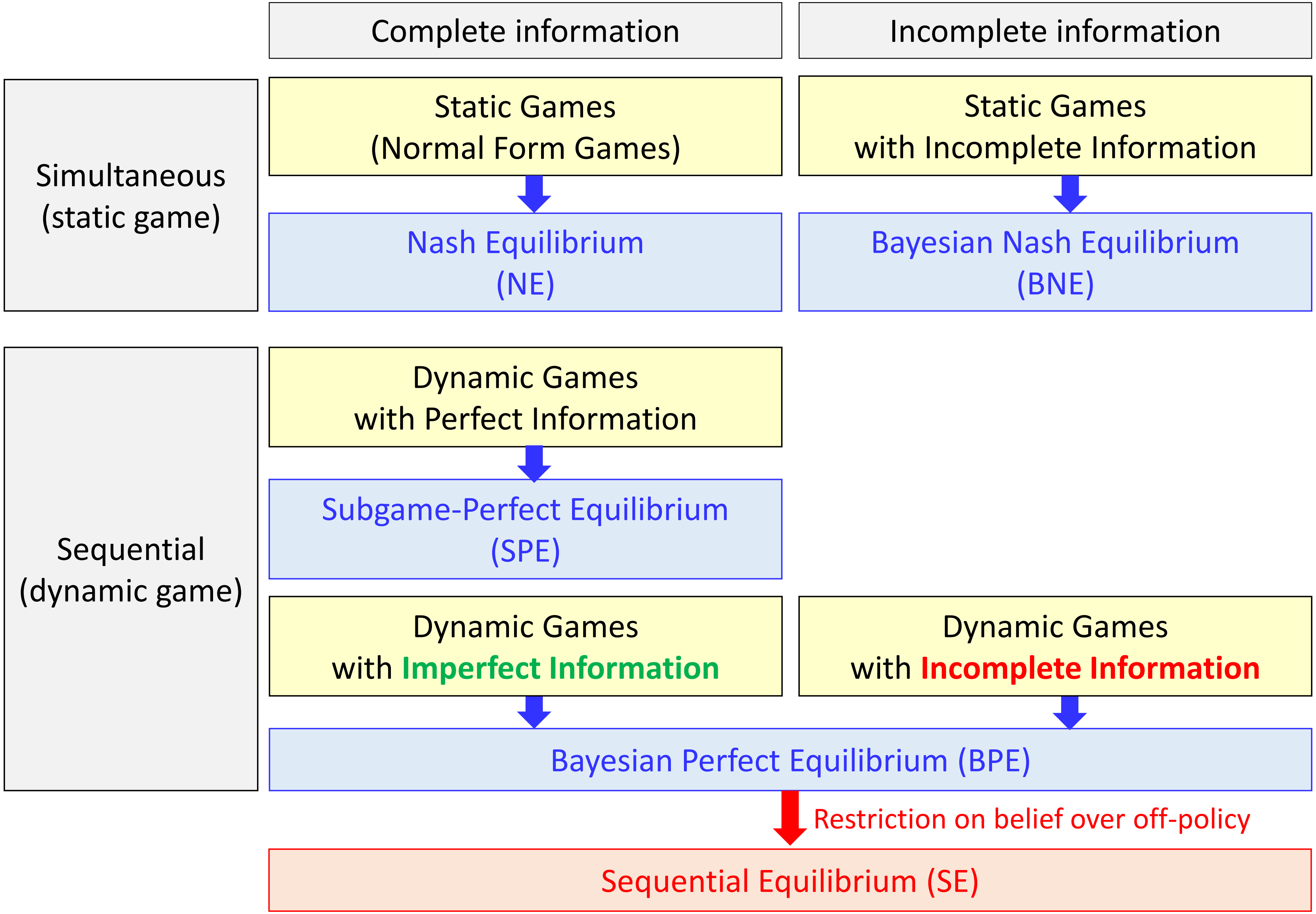
2. Sequential Games with Complete Information

Introduction



	Complete information	Incomplete information
Simultaneous (static game)	Static Games (Normal Form Games) ↓ Nash Equilibrium (NE)	Static Games with Incomplete Information ↓ Bayesian Nash Equilibrium (BNE)
Sequential (dynamic game)	Dynamic Games with Perfect Information ↓ Subgame-Perfect Equilibrium (SPE)	
	Dynamic Games with **Imperfect Information** ↓ Bayesian Perfect Equilibrium (BPE)	Dynamic Games with **Incomplete Information**

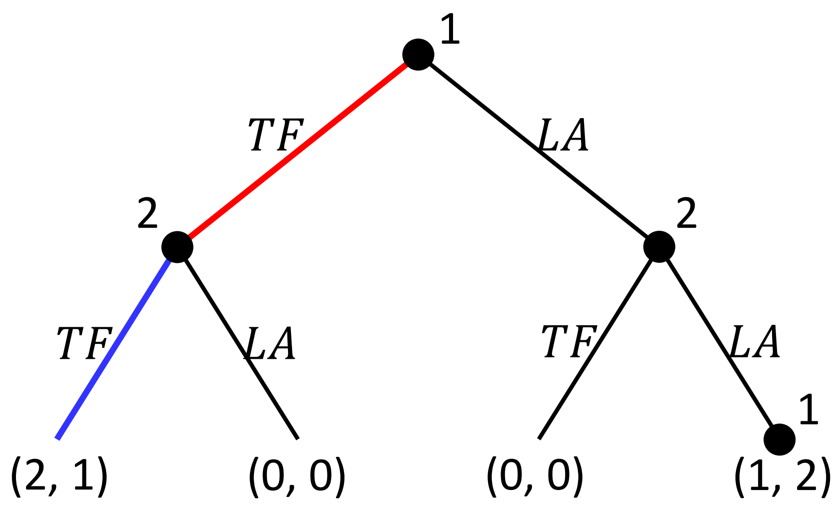
Restriction on belief over off-policy

Sequential Equilibrium (SE)

Perfect information extensive-form game

Motivations

		Player 2 (Husband)	
		TF	LA
Player 1 (wife)	TF	2, 1	0, 0
	LA	0, 0	1, 2



- Assumes that
  - ✓ Wife finishes work at 3:00 p.m. while husband finish work at 5:00 p.m.
- Wife can go to a movie theater for TF and call husband to come to the theater
  - ✓ If husband accept to come, they will get (2,1), otherwise they will get (0,0)
- The fundamental difference of this game with comparing to the previous simultaneous version is that
  - ✓ when husband moves he know what wife have done! (she choses what she like to see)
  - ✓ Furthermore, wife knows, by common knowledge of rationality, that husband will choose to follow her decision because it is his best response to do so (no choice….)

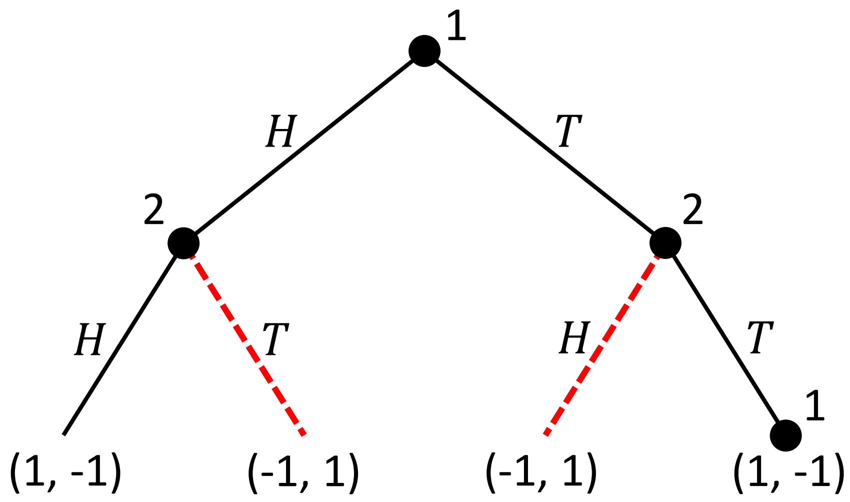
Motivations

- Is moving first always better?

- Is moving first always better?

No,

	H	T
H	$1, -1$	$-1, 1$
T	$-1, 1$	$1, -1$



- Moving first, in some case, reveals player's strategy given that other player can see first mover's action

Motivations

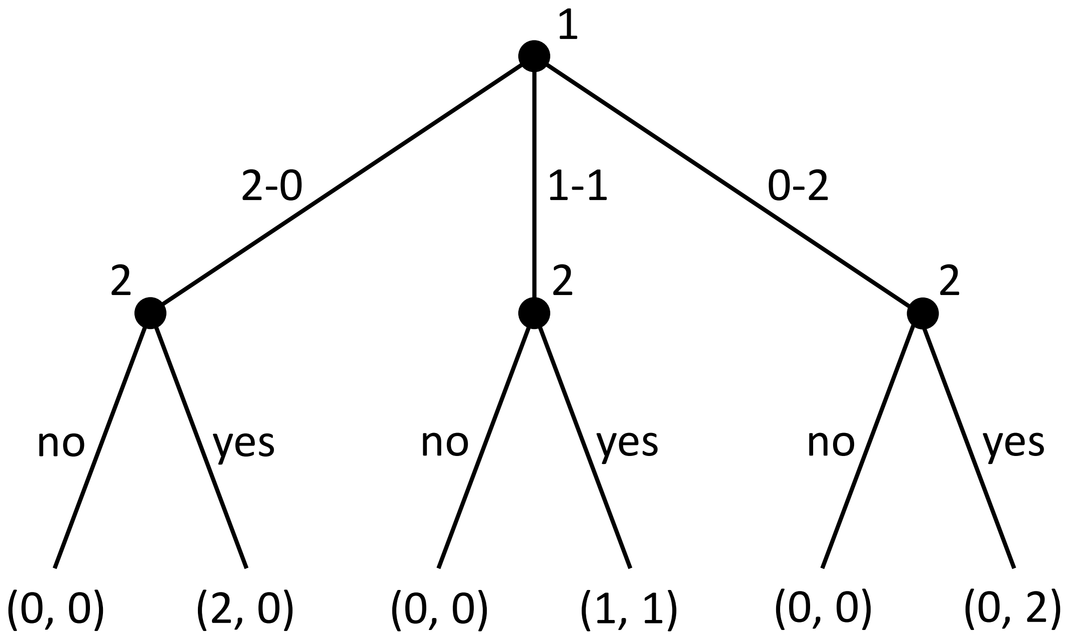
- The normal form game representation does not incorporate any notion of sequence, or time, of the actions of the players
  - Used for simultaneous game
  - Preserves game-theoretic properties
- The extensive form is an alternative representation that makes the temporal structure explicit
  - ➢ More precisely, it is not the chronological order of play that matter, but what players know when they make their choices
- Two variants:
  - perfect information extensive-form games
  - imperfect-information extensive-form games
- We will restrict our discussion to finite games, that is, to games represented as finite trees

Motivations

- **Perfect Information**
  - All players know the game structure (complete information).
  - Each player, when making any decision, is perfectly informed of all the events that have previously occurred.

- **Imperfect Information**
  - All players know the game structure (complete information).
  - Each player, when making any decision, may not be perfectly informed about some (or all) of the events that have already occurred.

Perfect-information game



The sharing games

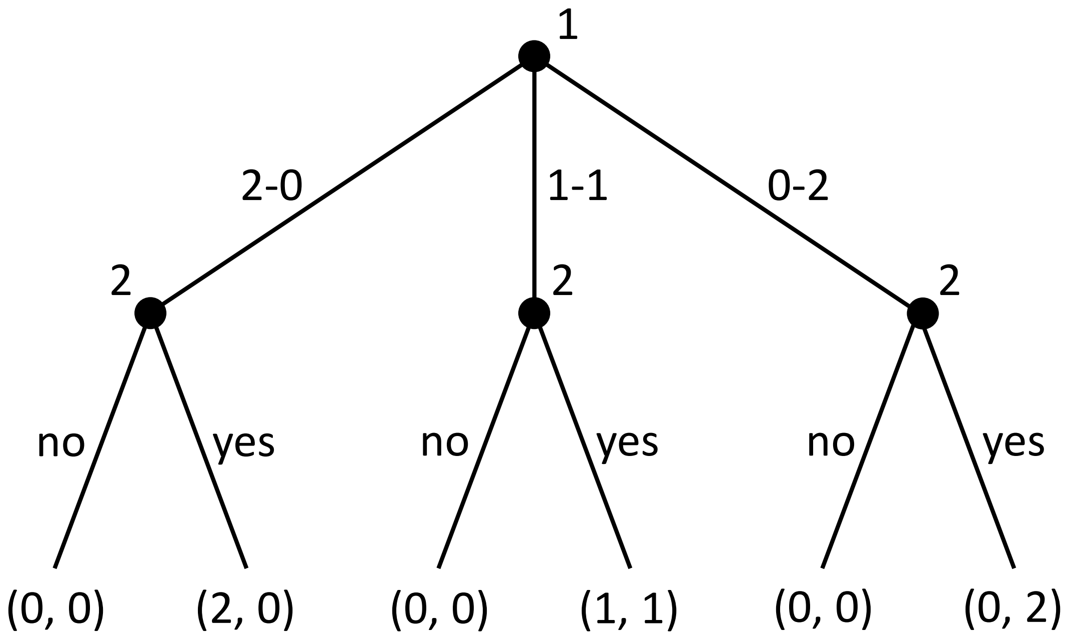
- A perfect-information game in extensive form (or, simply, a perfect information game) is a tree in the sense of graph theory
  - Each node represents the choice of one of the players
  - Each edge represents a possible action
  - Leaves represent final outcomes over which each player has a utility function

Definition (Perfect-information game)

A (finite) perfect-information game (in extensive form) is a tuple $G = (N, A, H, Z, \chi, \rho, \sigma, u)$:

- N is a set of n players;
- A is a (single) set of actions;
- H is a set of nonterminal choice nodes;
- Z is a set of terminal nodes, disjoint from H;
- $\chi: H \longmapsto A$ is the **action function**, which assigns to each choice node a set of possible actions
- $\rho: H \longmapsto N$ is the **player function**, which assigns to each nonterminal node a player $i \in N$ who choose an action at that time
- $\sigma: H \times A \longmapsto H \cup Z$ is the **successor function**, which maps a choice node and an action to a new choice node or terminal node such that for all $h_1, h_2 \in H$ and $a_1, a_2 \in A$, if $\sigma(h_1, a_1) = \sigma(h_2, a_2)$ then $h_1 = h_2$ and $a_1 = a_2$; and
- $u = (u_1, \ldots, u_n)$, where $u_i: Z \longmapsto \mathbb{R}$ is a real-valued **utility function** for player i on the terminal node Z

Pure strategies



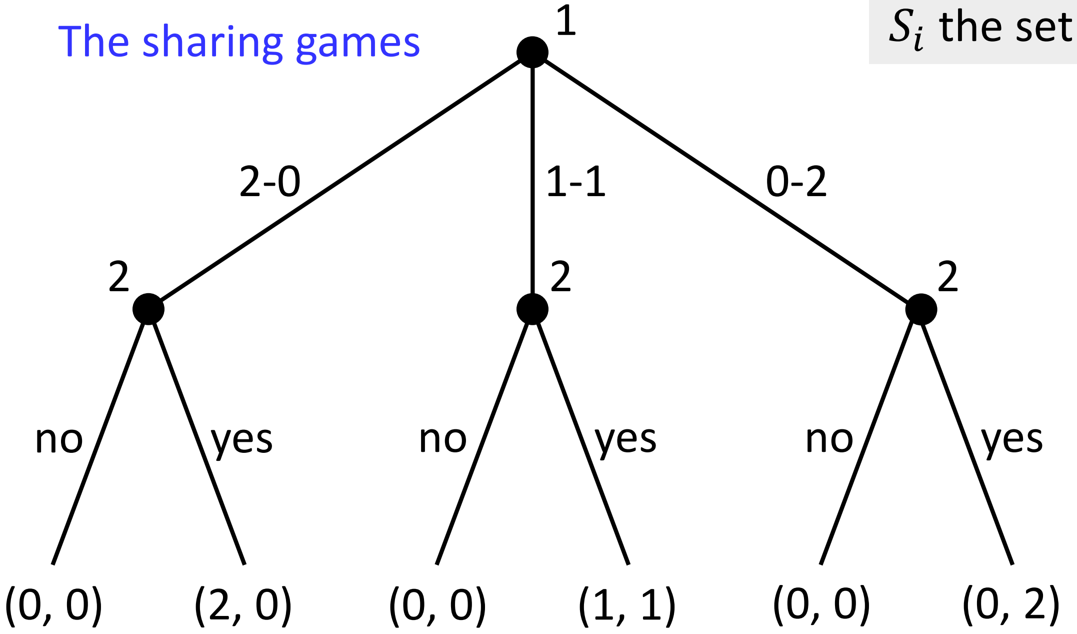
The sharing games

- A pure strategy for a player in a perfect-information game is a ***complete specification*** of which deterministic action to take at every node belonging to that player

Definition (Pure strategy in a perfect information game)

Let $G = (N, A, H, Z, \chi, \rho, \sigma, u)$ be a perfect-information extensive-form game. Then the pure strategies of player i consists of the Cartesian product $\prod_{h \in H, \rho(h)=i} \chi(h)$

- A pure strategy for a player in a perfect-information game is a complete specification of which deterministic action to take at every node belonging to that player
- An agent's strategy requires a decision at each choice node, regardless of whether or not it is possible to reach that node given the other choice nodes



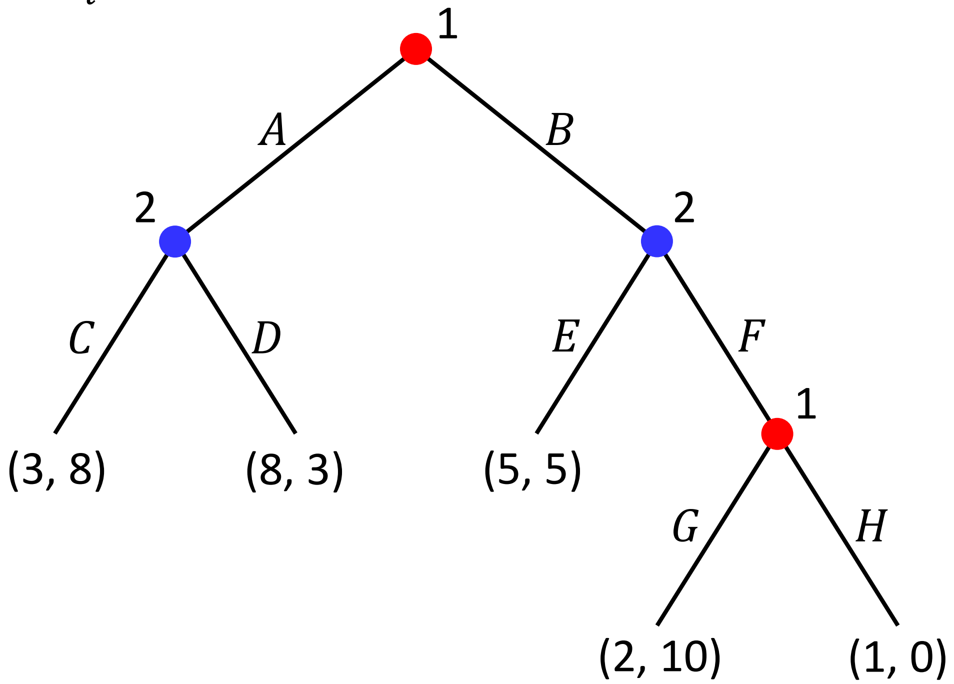
S_i the set of all pure-strategies mappings $s_i \in S_i$

$$S_1 = \{2-0, 1-1, 0-2\}$$

$$S_2 = \{ (yes, yes, yes), (yes, yes, no), (yes, no, yes), (yes, no, no), (no, yes, yes), (no, yes, no), (no, no, yes), (no, no, no) \}$$

Pure strategies example

- In summary, a pure strategy for player i is a mapping $s_i\colon H_i \to A_i$ that assigns an action $s_i(h_i) \in A_i(h_i)$ for every node $h_i \in H_i$ for player i. We denote by S_i the set of all pure-strategy mappings $s_i \in S_i$.



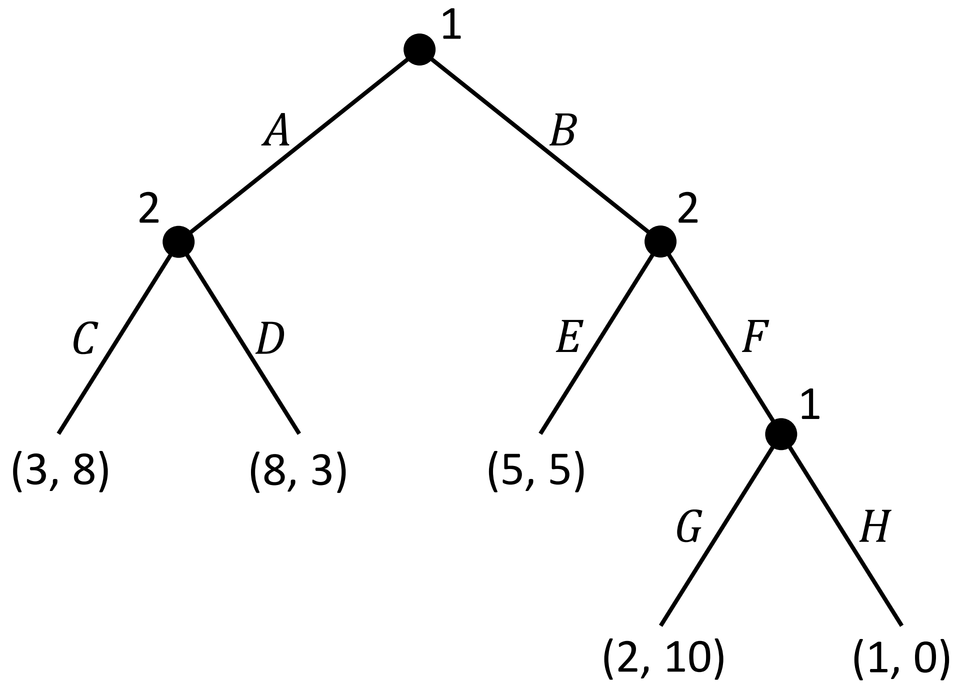
- In order to define a complete strategy for this game, each of the players must choose an action at each of his two choice node:
  - $S_1 = \{(A,G),(A,H),(B,G),(B,G)\}$
  - $S_2 = \{(C,E),(C,F),(D,E),(D,F)\}$
- It is important to note that we have to include the strategies (A,G) and (A,H) even though the choice between G and H is not available conditional on taking A

Mixed strategies definition

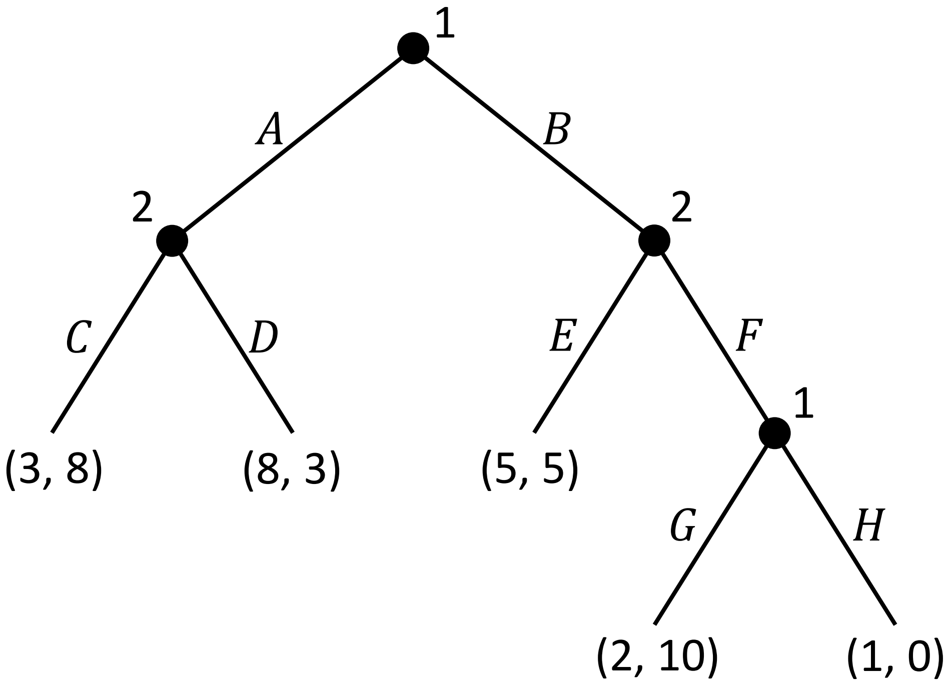
Definition (Mixed strategy in a perfect information game)

A mixed strategy for player i is a probability distribution over his pure strategies $s_i \in S_i$.

- When a mixed strategy is used, the player selects a plan randomly before the game is played and then follows a particular pure strategy
- A mixed strategy for player 2 is a probability distribution over his pure strategies $S_2 = \{(C,E),(C,F),(D,E),(D,F)\}$, for example, $\{0.15(C,E), 0.35(C,F), 0.15(D,E), 0.35(D,F)\}$
- After one of pure strategies is selected form mixed strategy (probability distribution), the player is choosing a pure plan of action



Behavior strategy motivation



- Is " If player 1 chooses A then I will play C, while if he plays B then I will mix and play E with probability 1/3" possible?

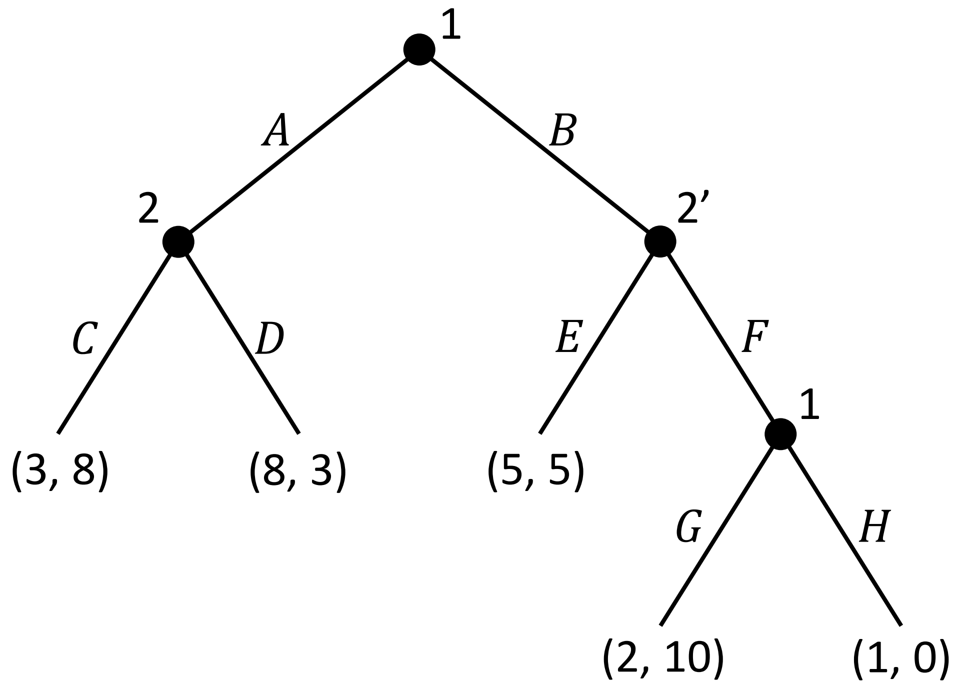
- That is, a player can randomize their action at each node they encounter as the game unfolds

Definition (Behavioral strategy)

A behavioral strategy specifies for each decision node $h_i \in H_i$ an independent probability distribution over $A_i(h_i)$, action available for agent i at node h_i, and is denoted by b_i: $H_i \rightarrow \Delta A(h_i)$ where $b_i(a_i(h_i))$ is the probability that player i plays action $a_i(h_i) \in A_i(h_i)$ in node h_i.

- A behavioral strategy is more in tune with the dynamic nature of the extensive-form game.
- When using such a strategy, a player mixes among his actions whenever he is called to play

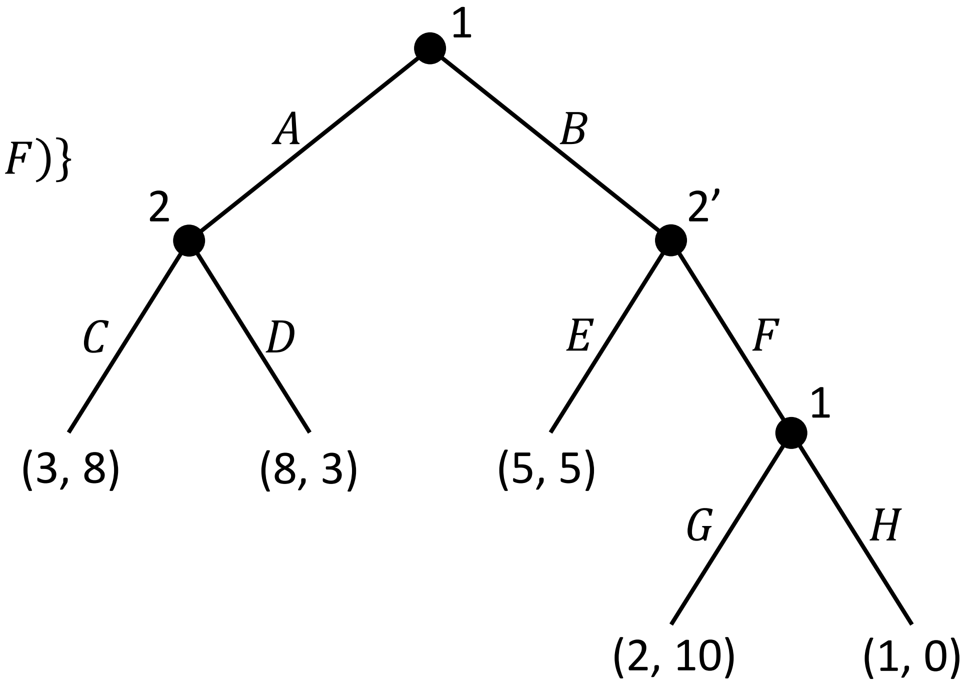
A behavioral strategy
-at node 2 : $\{0.5C, 0.5D\}$
-at node 2': $\{0.3E, 0.7F\}$



Behavior strategy

A mixed strategy for player 2,
$\{0.15(C,E), 0.35(C,F), 0.15(D,E), 0.35(D,F)\}$

A behavioral strategy
-at node 2 : $\{0.5C, 0.5D\}$
-at node 2': $\{0.3E, 0.7F\}$



Definition (Perfect recall game)

A game of perfect recall is one in which no player ever forgets information that he previously knew

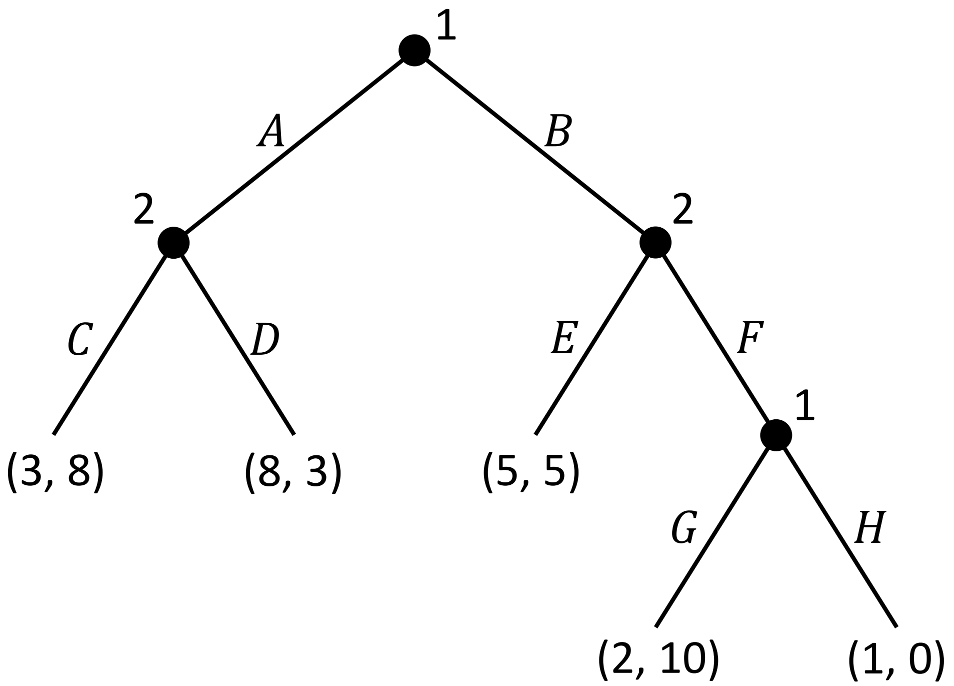
- For the class of perfect-recall games, mixed and behavioral strategies are equivalent, in the sense that given strategies of i's opponents, the same distribution over outcomes can be generated by either a mixed or behavioral strategy of player i.

Nash equilibria

Theorem

Every (finite) perfect-information game in extensive form has a pure-strategy Nash equilibrium (PSNE).

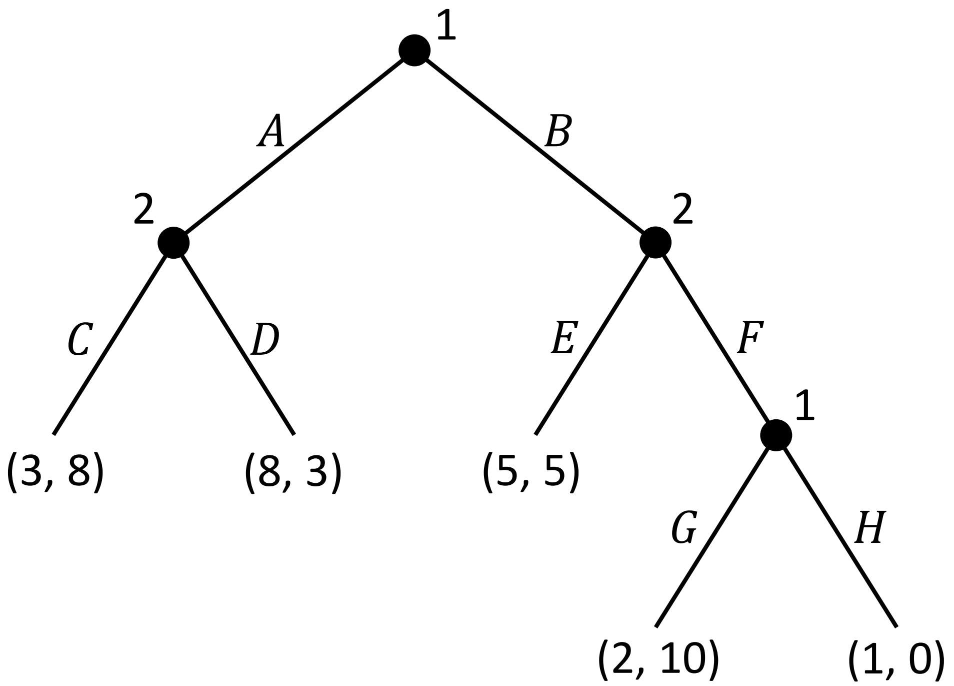
Induced normal form



	CE	CF	DE	DF
AG	$3, 8$	$3, 8$	$8, 3$	$8, 3$
AH	$3, 8$	$3, 8$	$8, 3$	$8, 3$
BG	$5, 5$	$2, 10$	$5, 5$	$2, 10$
BH	$5, 5$	$1, 0$	$5, 5$	$1, 0$

- Every perfect-information game there exists a corresponding normal-from game as bellow:
- The temporal structure of the extensive form presentation can result in a certain redundancy within the normal form game
  - E.g., we write down 16 payoff pairs instead of 5 in normal form game
  - It can result in an exponential blowup of the game representation
- While we can write any extensive-from game as a normal form game, we can't do the reverse
  - For example, matching pennies cannot be written as perfect information extensive form

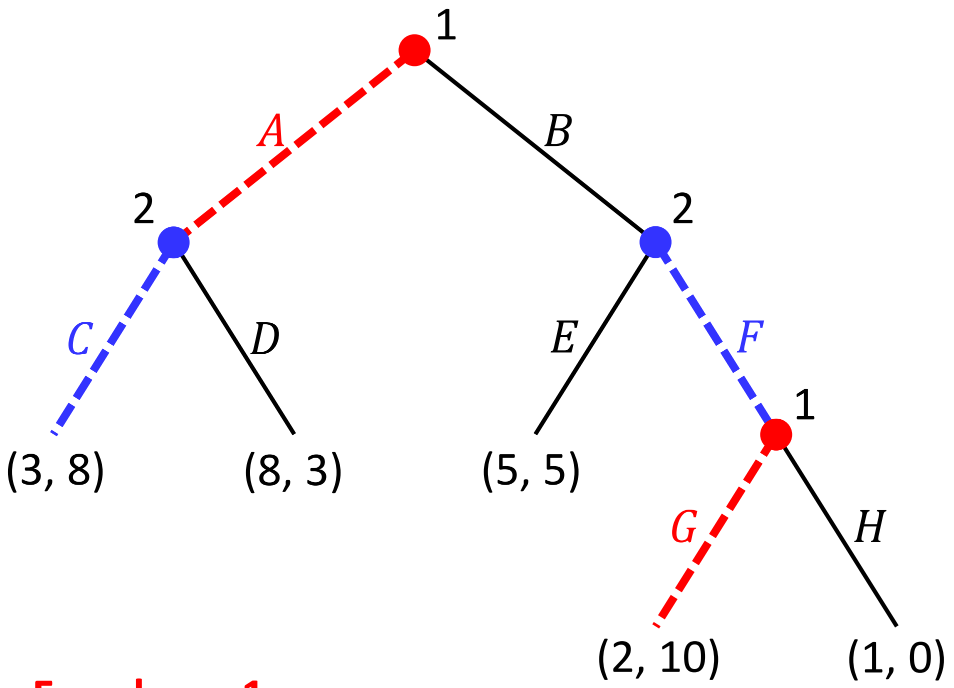
Nash equilibria



	CE	CF	DE	DF
AG	$3,8$	$\boxed{3,8}$	$8,3$	$8,3$
AH	$3,8$	$\boxed{3,8}$	$8,3$	$8,3$
BG	$5,5$	$2,10$	$5,5$	$2,10$
BH	$\boxed{5,5}$	$1,0$	$5,5$	$1,0$

- What are the (three) pure-strategy equilibria?
  - $(A,G),(C,F)$
  - $(A,H),(C,F)$
  - $(B\ H),(C,E)$

Nash equilibria



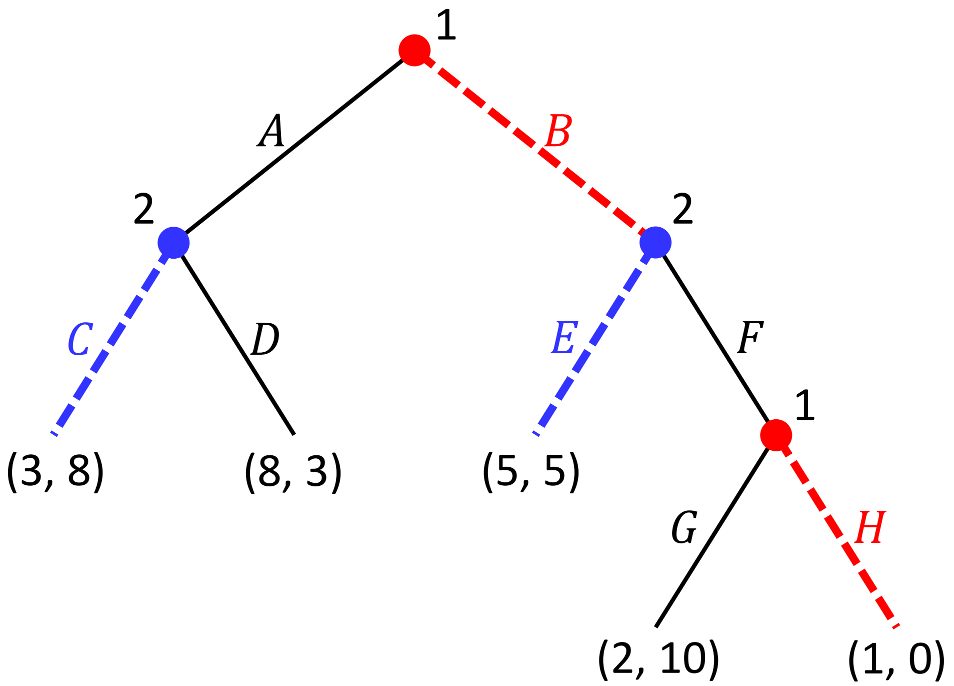
	CE	CF	DE	DF
AG	$3,8$	$3,8$	$8,3$	$8,3$
AH	$3,8$	$3,8$	$8,3$	$8,3$
BG	$5,5$	$2,10$	$5,5$	$2,10$
BH	$5,5$	$1,0$	$5,5$	$1,0$

- **For player 1:**
  - If player 1 plays B rather than A at the first node, he will get a payoff 2 instead of 3; thus, there is no incentive to change the action
  - If player 1 plays H rather than G at the second node, there is no change in his payoff

- **For player 2:**
  - If player 2 plays D rather than C at the first node, he will get a payoff 3 instead of 8; thus, there is no incentive to change the action
  - If player 2 plays E rather than F at the second node, there is no change in his payoff

Thus, $\{(A, G), (C, F)\}$ is an equilibrium

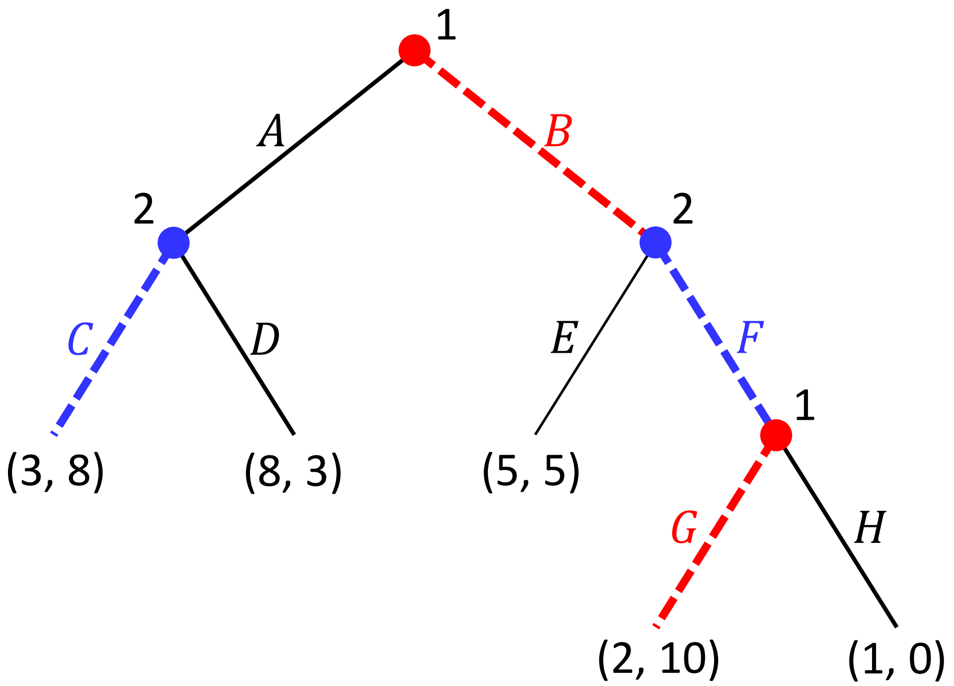
Nash equilibria



	CE	CF	DE	DF
AG	$3, 8$	$3, 8$	$8, 3$	$8, 3$
AH	$3, 8$	$3, 8$	$8, 3$	$8, 3$
BG	$5, 5$	$2, 10$	$5, 5$	$2, 10$
BH	$\boxed{5, 5}$	$1, 0$	$5, 5$	$1, 0$

- What will happen if player 1 choose to paly BG instead BH?

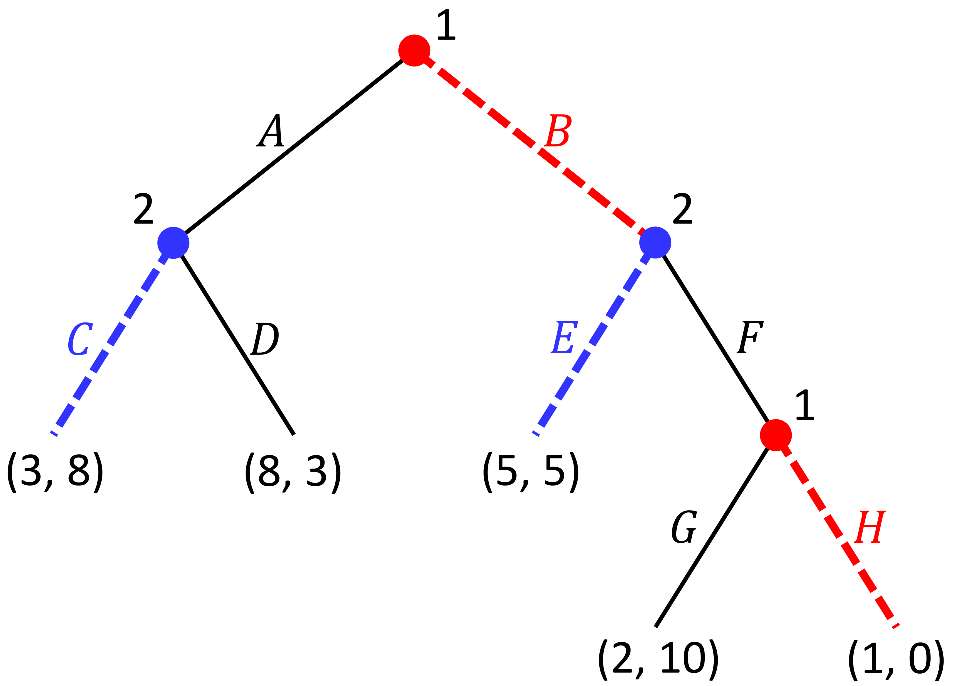
Nash equilibria



	CE	CF	DE	DF
AG	3, 8	3, 8	8, 3	8, 3
AH	3, 8	3, 8	8, 3	8, 3
BG	5, 5	2, 10	5, 5	2, 10
BH	5, 5	1, 0	5, 5	1, 0

- What will happen if player 1 choose to paly BG instead BH?
    - ➢ Player 2 will choose to play CF instead CE to get a payoff 10 instead 5
    - ➢ As a result, player 1 will get a payoff 2 instead of 5 (bad for player 1)

Nash equilibria



	CE	CF	DE	DF
AG	$3, 8$	$3, 8$	$8, 3$	$8, 3$
AH	$3, 8$	$3, 8$	$8, 3$	$8, 3$
BG	$5, 5$	$2, 10$	$5, 5$	$2, 10$
BH	$\boxed{5, 5}$	$1, 0$	$5, 5$	$1, 0$

- That is, player 1 is threatening player 2 to choose E by playing H by committing to choose an action that is harmful to player 2 in this second decision node
- If player 2 choose to play F, then would player 1 really follow through on his threat and play H?
  - This action is not reasonable because choosing H instead of G will reduce his payoff given that player 1 reaches that decision node

We need to define an equilibrium **refinement concept** that does not suffer from this issue

Sequential rationality

- We will insist that a player use strategies that are optimal at every node (stage) in the game tree
- We call this principle *sequential rationality*, because it implies that players are playing rationally at every strategy in the sequence of the game, whether it is on or off the equilibrium path of play

Definition (Sequentially rational)

Given strategies $s_{-i} \in \prod A_{-i}$ of i's opponents, we say that s_i is sequentially rational if and only if player i is playing a best response to s_{-i} in each of his decision node.

Subgame perfection

Definition (Subgame of G rooted at h)

Given a perfect-information extensive-form game G, the subgame of G rooted at node h is the restriction of G to the descendants of h.

Definition (Subgames of G)

The set of subgames of G consists of all of subgames of G rooted at some node in G

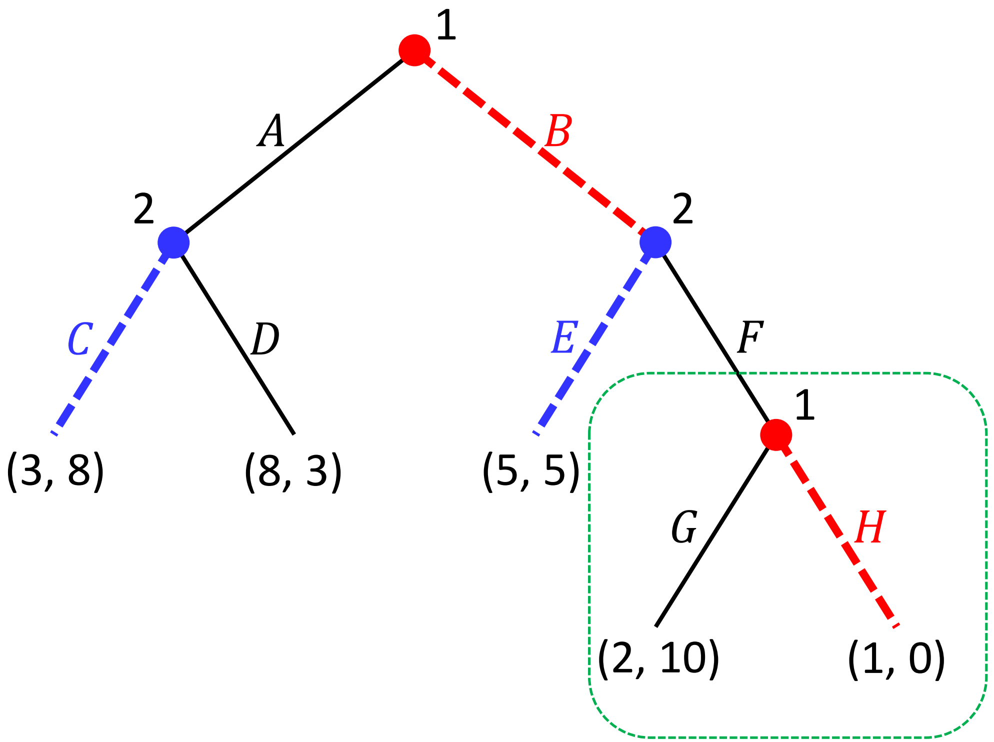
Subgame perfection

Definition (Subgame-perfect equilibrium)

The subgame-perfect equilibria (SPE) of a game G are all strategy profiles s such that for any subgame G' of G, the restriction of s to G' is a Nash equilibrium of G'.

- SPE is a refinement of the Nash equilibrium in perfect-information games in extensive form
- Since G is its own subgame, every SPE is also Nash equilibrium
- SPE is a stronger concept than Nash equilibrium
  (i.e., every SPE is a NE, but not every NE is a SPE)
- Every perfect-information extensive-form game has at least one subgame-perfect equilibrium
- The concept of SPE rules out unwanted Nah equilibria with "nonedible threats"
- SPE requires not only that a Nash equilibrium profiles of strategies be combination of best responses on the equilibrium path, which is a necessary condition of a Nash equilibrium, but also that the profile of strategies consist of mutual best Reponses off the equilibrium path

Subgame perfection

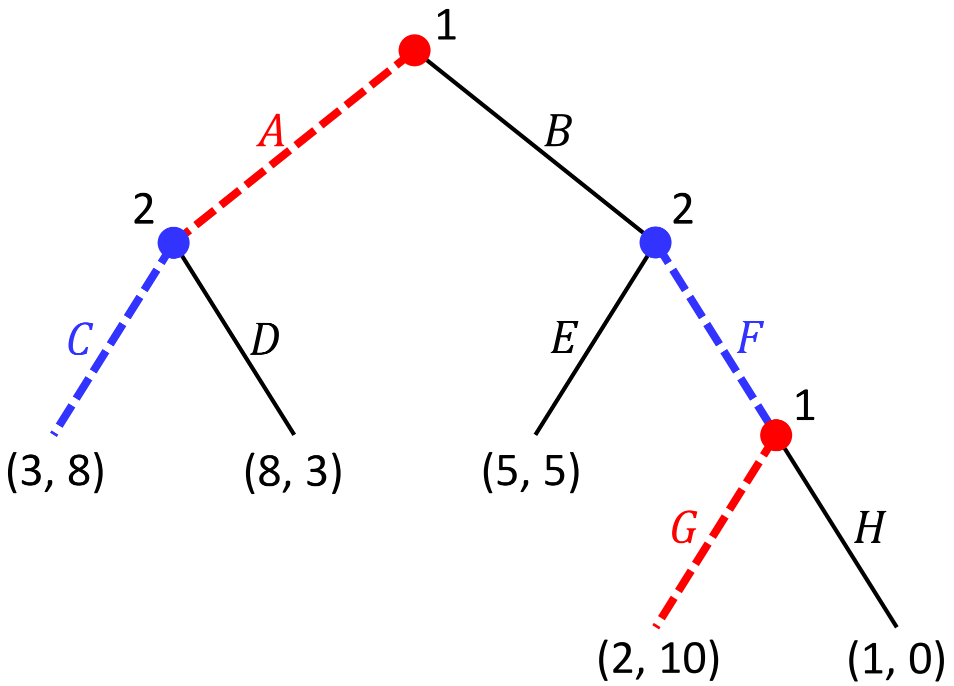


	CE	CF	DE	DF
AG	3, 8	3, 8	8, 3	8, 3
AH	3, 8	3, 8	8, 3	8, 3
BG	5, 5	2, 10	5, 5	2, 10
BH	$\boxed{5, 5}$	1, 0	5, 5	1, 0

Subgame rooted player 1's second decision node

- The unique Nash equilibrium for this subgame is for paler 1 to paly G
- Thus, the action H, the restriction of the strategies (B, H) to this subgame, is not optimal in this subgame
  - ➢ (B, H) cannot be part of a subgame perfect equilibria

Subgame perfection example



	CE	CF	DE	DF
AG	$3, 8$	$\boxed{3, 8}$	$8, 3$	$8, 3$
AH	$3, 8$	$3, 8$	$8, 3$	$8, 3$
BG	$5, 5$	$2, 10$	$5, 5$	$2, 10$
BH	$5, 5$	$1, 0$	$5, 5$	$1, 0$

- The only SPE for this game is $\{(A, G), (C, F)\}$

Computing subgame perfect equilibria

- The principle of backward induction:
  - One identifies the equilibria in the “bottom-most” subgame tress, and assumes that those equilibria will be played as one backs up and considers increasingly large tree
  - Guarantee a subgame perfect equilibrium and simple computation
  - Can be implemented as a single depth-first traversal of the game tree and thus requires time linear in the size of the game representation
- The Nash equilibrium for a general sum game:
  - Finding Nash equilibria of general games require time exponential in the size of the normal form.
  - In addition, the induced normal form of an extensive-form game is exponentially larger than the original representation.

Computing subgame perfect equilibria

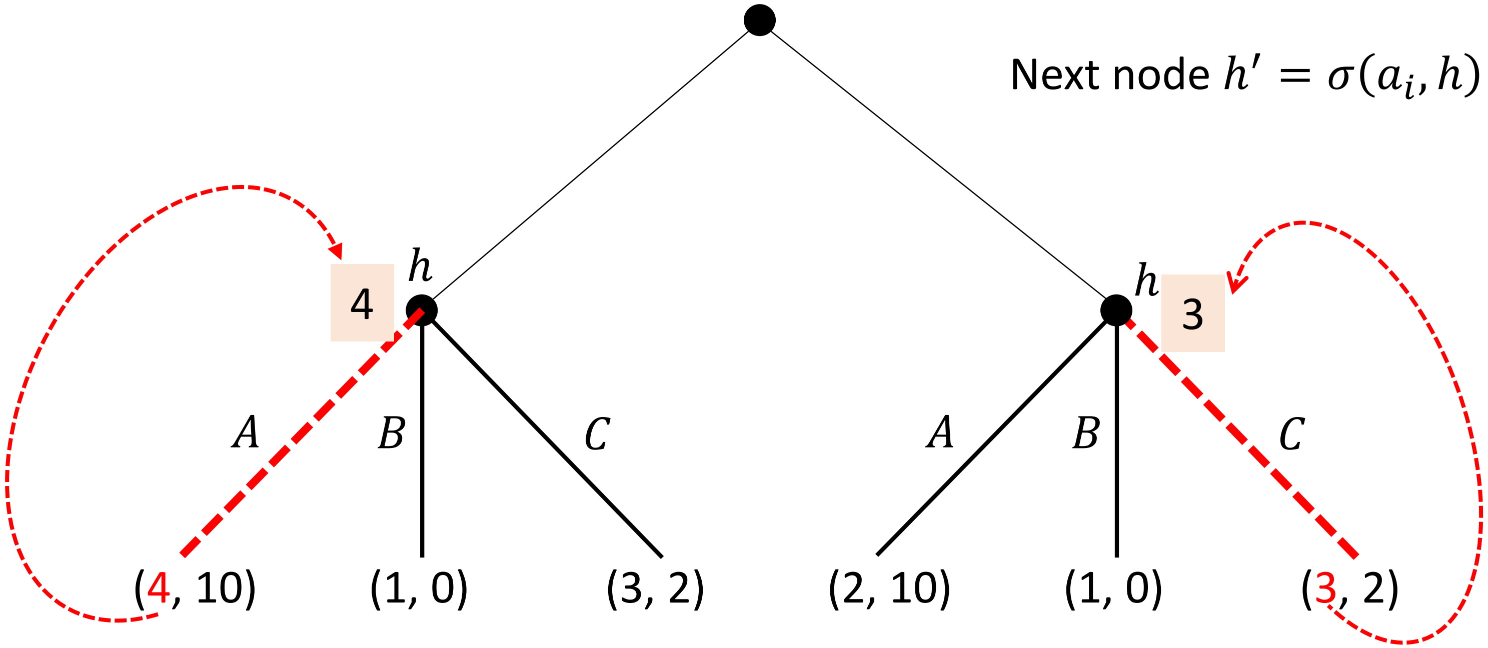
- Procedure for finding the value of a sample (subgame-perfect) Nash equilibrium of a perfect-information extensive-form game

- Every time a given player i has the opportunity to act at a given node $h \in H$ $(i.e., \rho(h) = i)$:

$$a_i^* = \underset{a_i \in \chi(h)}{\operatorname{argmax}}\, u_i(\sigma(a_i, h))$$

Available action $= \chi(h) = \{\mathrm{A, B, C}\}$

Next node $h' = \sigma(a_i, h)$



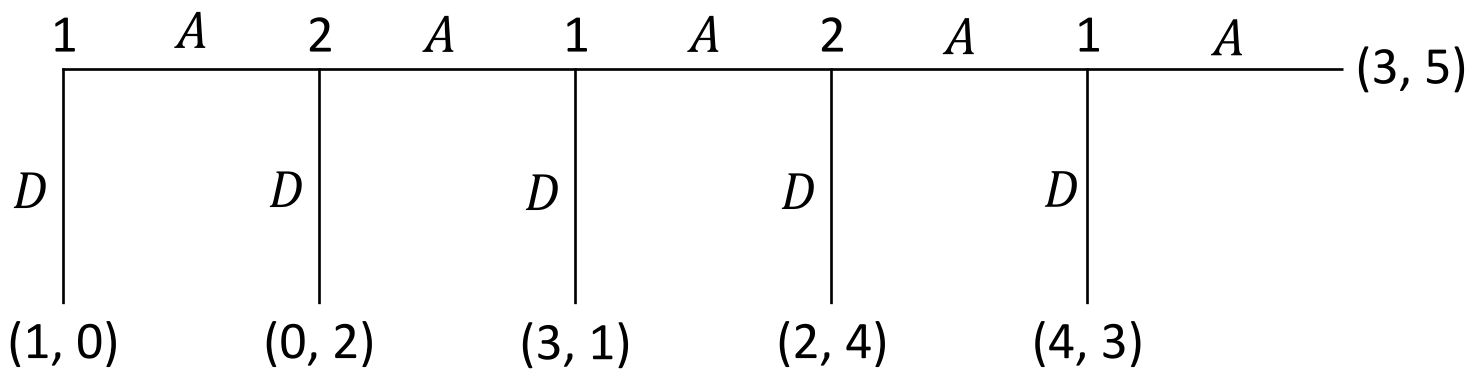
Backward induction algorithm

- Procedure for finding the value of a sample (subgame-perfect) Nash equilibrium of a perfect-information extensive-form game

function $\textsc{BackwardInduction}$ (node h) **returns** $u(h)$
if $h \in Z$ **then**
  return $u(h)$
$best\_util \leftarrow -\infty$
forall $a \in \chi(h)$ **do**
  $util\_at\_child \leftarrow \textsc{BackwardInduction}(\sigma(h, a))$
  if $util\_at\_child_{\rho(h)} > best\_util_{\rho(h)}$ **then**
    $best\_util \leftarrow util\_at\_child$
return $best\_util$

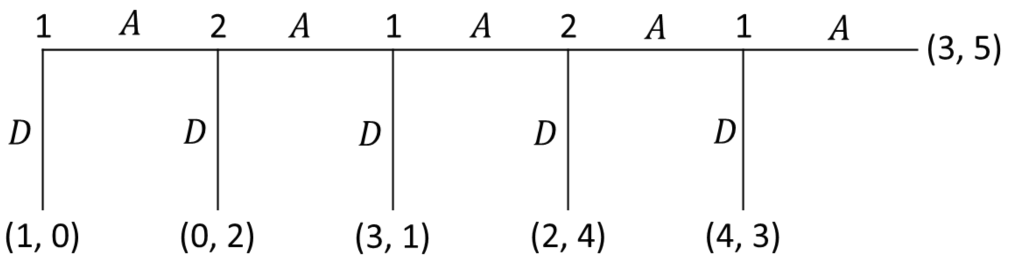
- *util_at_child* is a vector denoting the utility of each player
- The procedure doesn't return an equilibrium strategy, but rather labels each node with a vector of real numbers
  - This labeling can be seen as an extension of the game's utility function to the non-terminal nodes H
  - The equilibrium strategies: take the best action at each node.

SPE example: Centipede Game



Let’s play this game as a fun

SPE example: Centipede Game



- What happens when we use this procedure on Centipede?
    - In the only equilibrium, player 1 goes down in the first move.
    - However, this outcome is Pareto-dominated by all but one other outcome

- Two considerations:
    - Practical: human subjects don't go down right away
    - Theoretical: what should you do as player 2 if player 1 doesn't go down?
        - SPE analysis says to go down. However, that same analysis says that P1 would already have done down. How do you update your beliefs upon observation of a measure zero event?
        - But if player 1 knows that you will do something else, it is rational for him not to go down anymore… a paradox
        - There's a whole literature on this question

SPE example: Stackelberg Duopoly

- Two identical firms, players 1 and 2, produce some good
- Firm i produce quantity q_i
- Cost for production is $c_i(q_i) = c_i q_i$
- Price is given by $d = a - b(q_1 + q_2)$
- The profit of company i given its opponent chooses quantity q_j is

$$u_i(q_i, q_j) = (a - bq_i - bq_j)q_i - c_i q_i = -bq_i^2 + (a - c_i)q_i - bq_j q_i$$

- The best-response function for each firm is given by the first-order condition

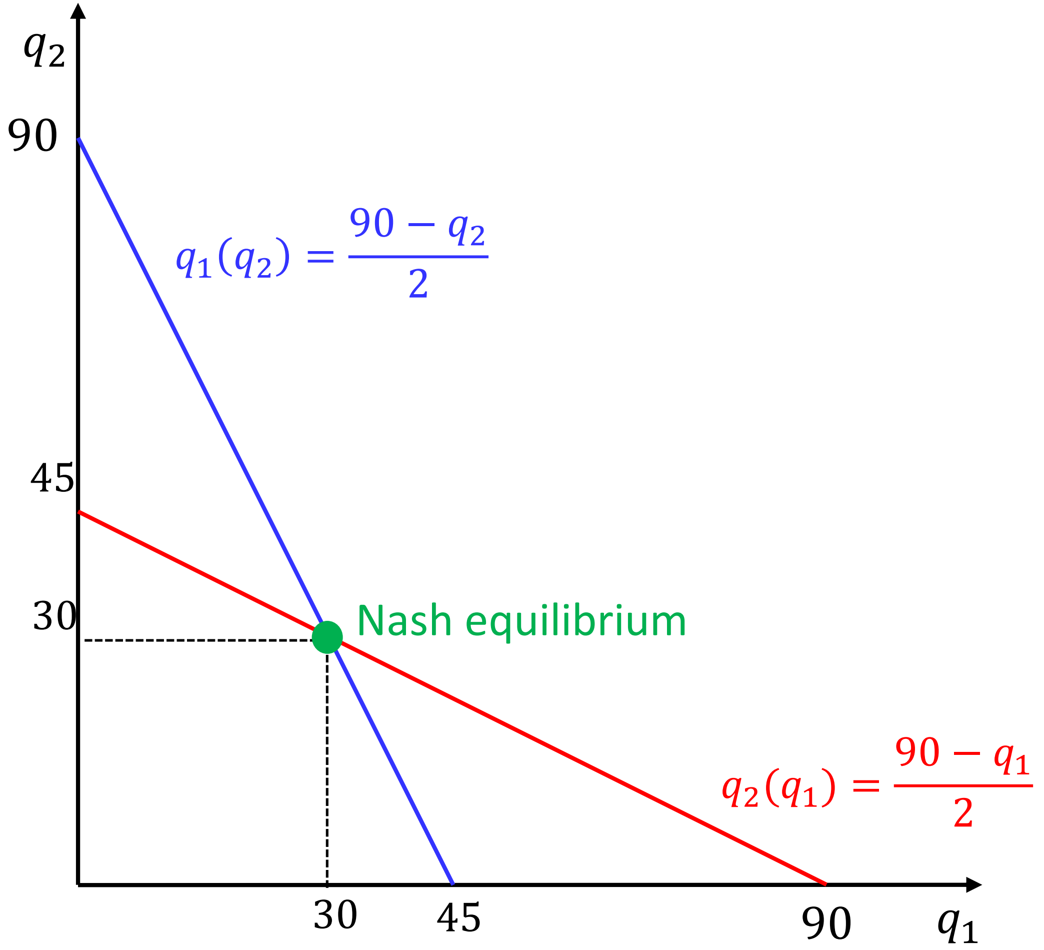
$$BR_i(q_j) = \frac{a - bq_j - c_i}{2b}$$

SPE example: Stackelberg Duopoly

- In case there are two firms, we have two best-response equations:

$$q_1 = \frac{a - bq_2 - c_1}{2b} \quad \text{and} \quad q_2 = \frac{a - bq_1 - c_2}{2b}$$

$a = 100, b = 1, c_1 = c_2 = 10$



SPE example: Stackelberg Duopoly

- Now, assume player 1 will choose q_1 first and player 2 will observe the choice made by player 1 before it makes its choice of q_2

- Assume player 1 choose q_1, then player 2 will best respond to it

$$q_2(q_1) = \frac{90 - q_1}{2}$$

- Assuming common knowledge of rationality, what should player 1 do?
  - ✓ Player 1 know exactly how a rational player 2 would respond to tits choice of q_1
  - ✓ Thus, player 1 will replace the fixed q_2 in its profit function with the best response of firm 2 and choose q_1 to solve

$$\max_{q_1}[100 - q_1 - q_2]q_1 - 10q_1$$
$$\Rightarrow \max_{q_1}\left[100 - q_1 - \left(\frac{90 - q_1}{2}\right)\right]q_1 - 10q_1$$
$$\Rightarrow 100 - 2q_1 - 45 + q_1 - 10 = 0$$
$$\Rightarrow q_1 = 45, q_2 = 22.5$$

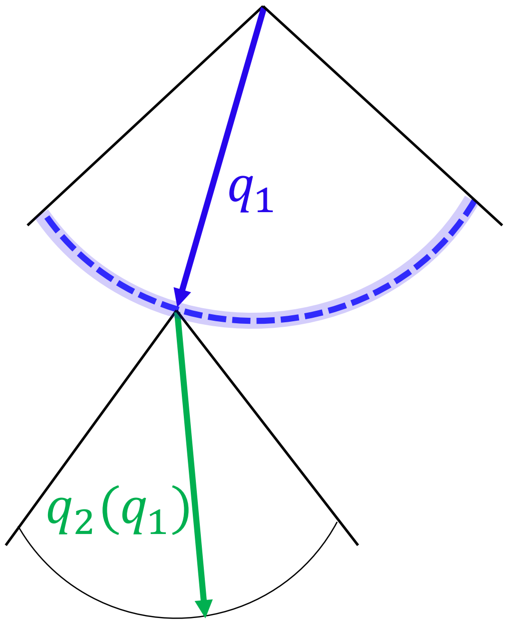
- $u_1(45,22.5) = (100 - 45 - 22.5) \times 45 - 10 \times 45 = 1012.5 > 900\ (\mathrm{NE})$
- $u_2(45,22.5) = (100 - 22.5 - 45) \times 22.5 - 10 \times 22.5 = 506.25 < 900(\mathrm{NE})$

First-mover advantage

SPE example: Stackelberg Duopoly

- If a profile of strategies survives backward induction then this profile is also a subgame-perfect equilibrium (SPE), and in particular Nash equilibrium (NE)

- Be careful to specify the strategies of player 2:
  - Player 2 has a continuum of information sets, each being a particular choice of q_1
  - We must specify q_2 for every information set, each of which correspond to an action q_1 chosen by player 1



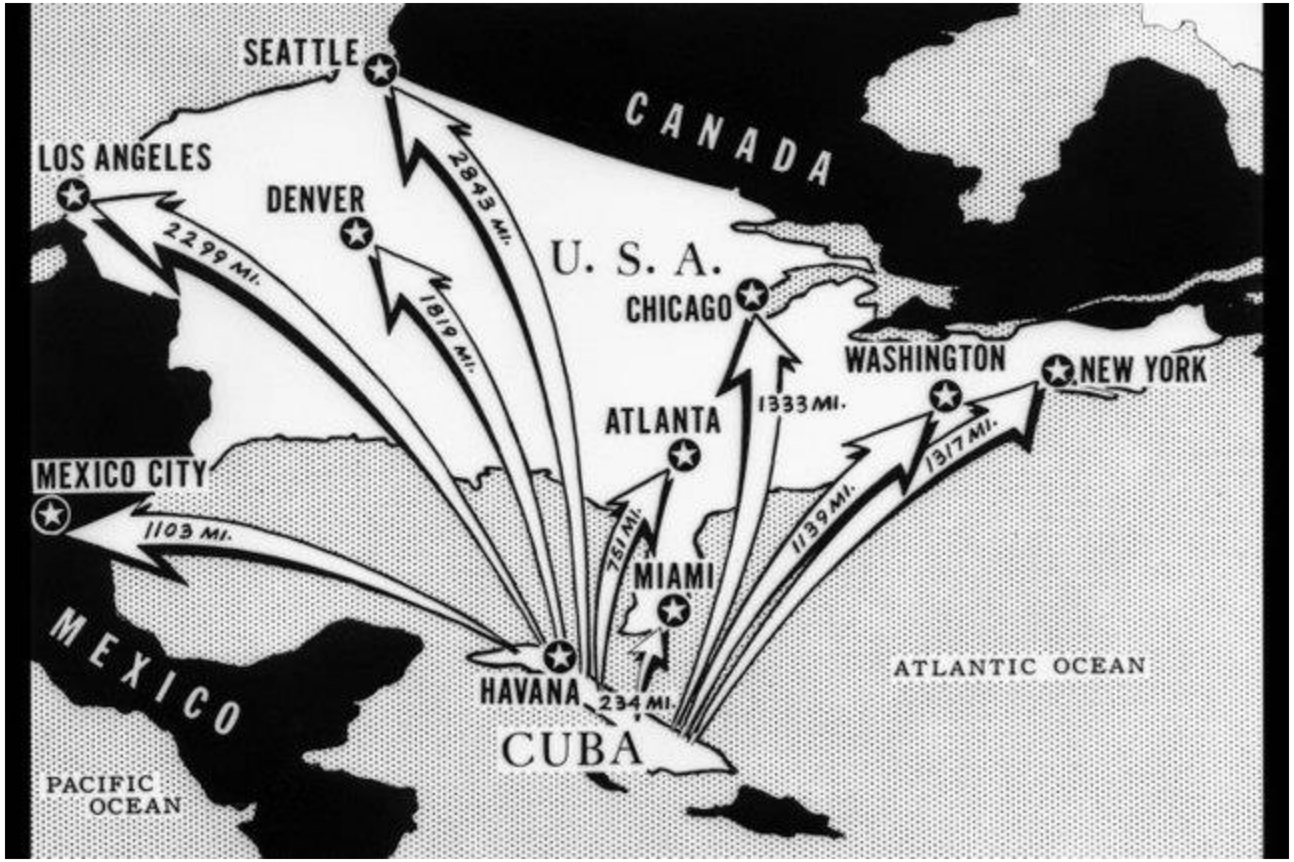
$$\text{SPE} = \left(q_1, q_2(q_1)\right) = \left(45, \frac{90-q_1}{2}\right)$$

$$\cancel{(45{,}22.5)}$$

$$u_1(q_1, q_2) = (100 - q_1 - q_2)q_1 - 10q_1$$
$$u_2(q_1, q_2) = (100 - q_1 - q_2)q_2 - 10q_2$$

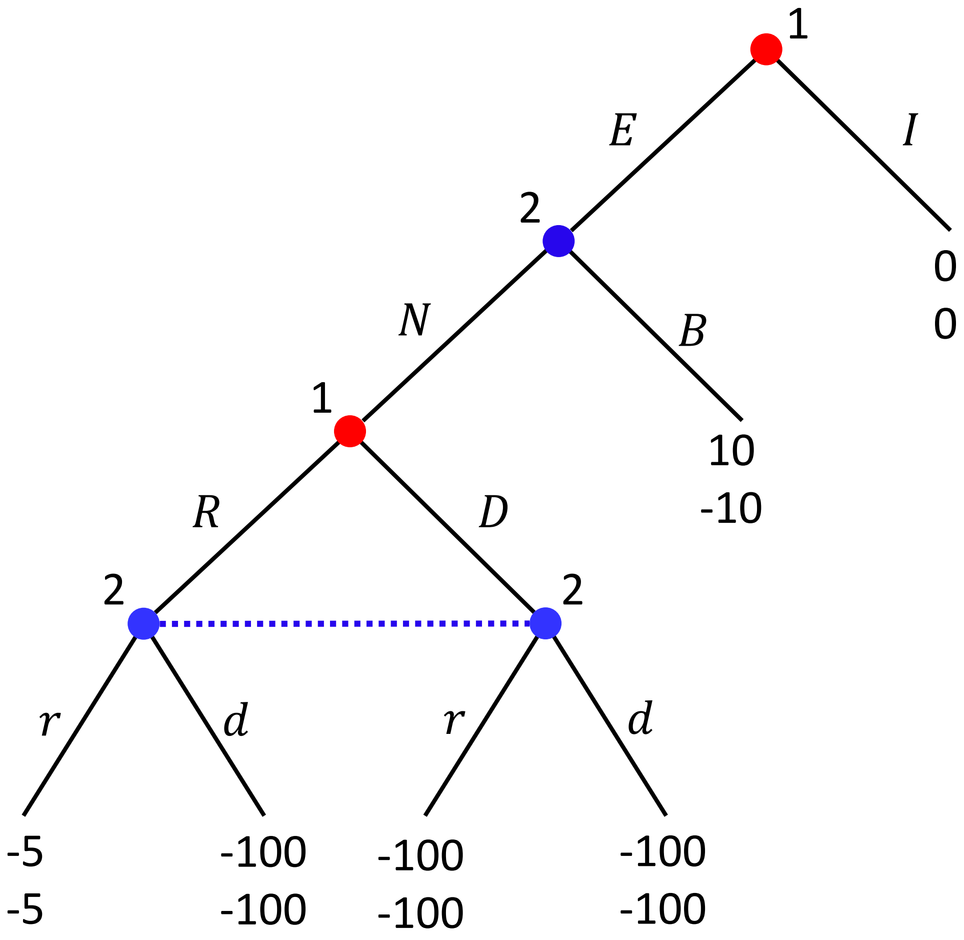
SPE example: Mutually Assured Destruction

Cuban missile crisis of 1962



SPE example: Mutually Assured Destruction

Cuban missile crisis of 1962



Player 1 (U.S.):

- Ignore incident (I)
- Escalate situation (E)

Player 2 (USSR.):

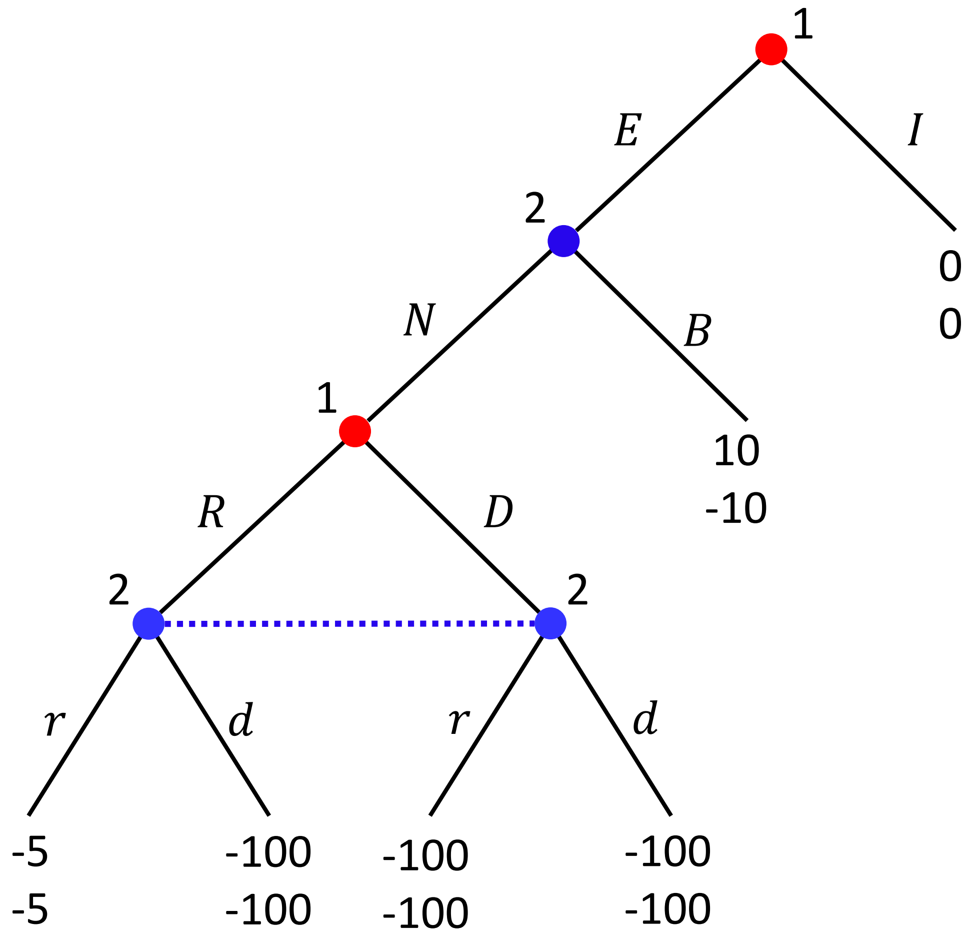
- Nuclear confrontation(N)
- Back down (B)

Player 1 & Player 2 (War game):

- Retreat (R for PL1 and r for PL2)
- Doomsday (D for PL1 and d for PL2)

SPE example: Mutually Assured Destruction

- Convert the extensive form game into a normal form game



	Br	Bd	Nr	Nd
IR	0, 0	0, 0	0, 0	0, 0
ID	0, 0	0, 0	0, 0	0, 0
ER	10, -10	10, -10	-5, -5	-100, -100
ED	10, -10	10, -10	-100, -100	-100,-100

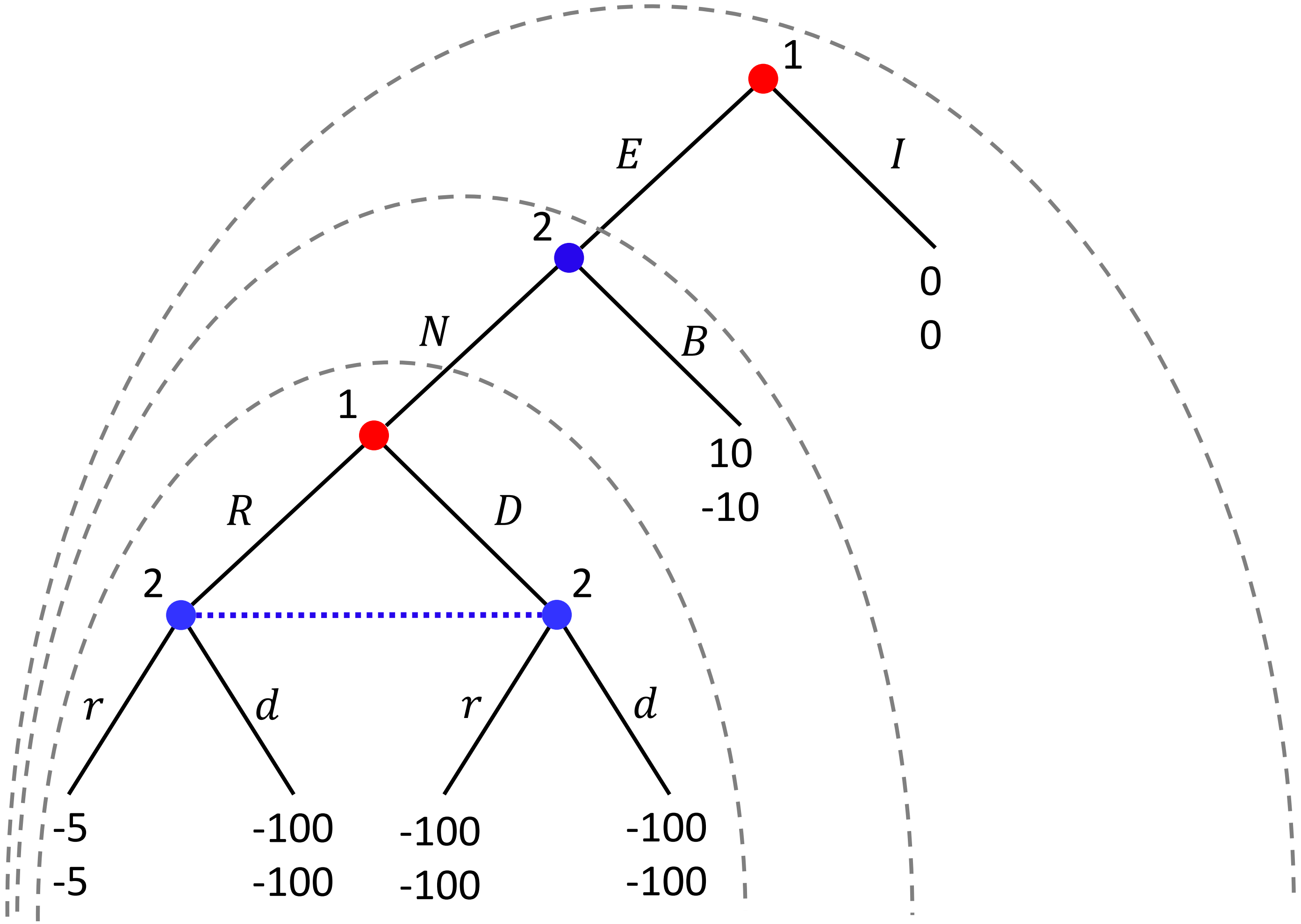
Are they subgame perfect equilibrium?

- There are six pure strategy Nash equilibria:

$$\text{NE}s = \{(IR, Nr), (IR, Nd), (ID, Nr), (ID, Nd), (ED, Br), (ED, Bd)\}$$

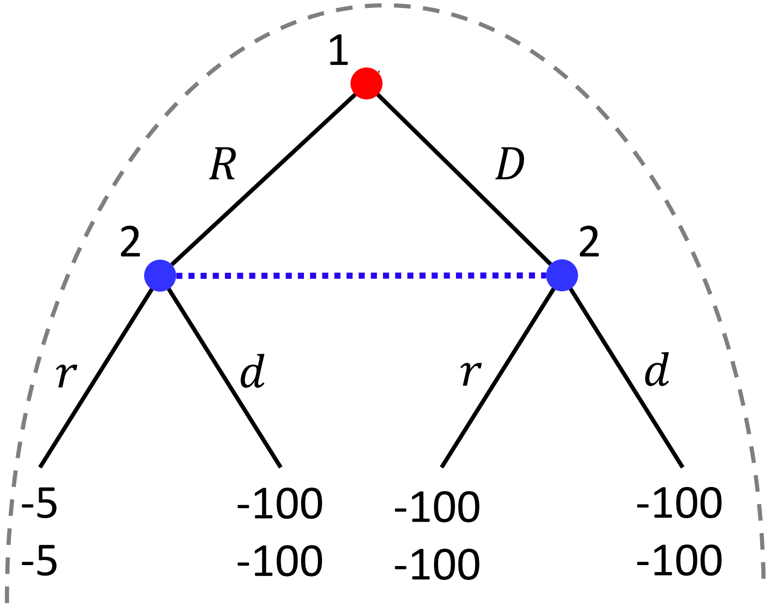
SPE example: Mutually Assured Destruction

- Find the subgame-perfect equilibria using backward induction



SPE example: Mutually Assured Destruction

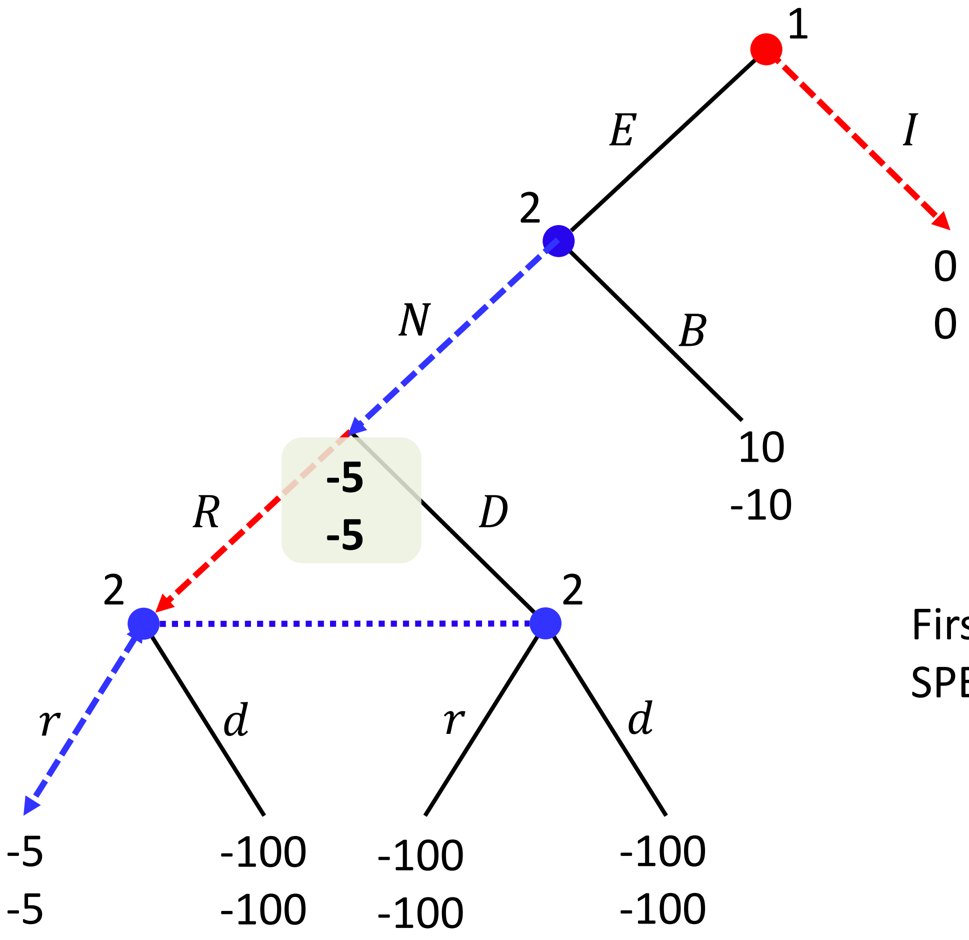
- Find the subgame-perfect equilibria using backward induction



	r	d
R	$\mathbf{-5, -5}$	$-100, -100$
D	$-100, -100$	$\mathbf{-100, -100}$

SPE example: Mutually Assured Destruction

- Find the subgame-perfect equilibria using backward induction

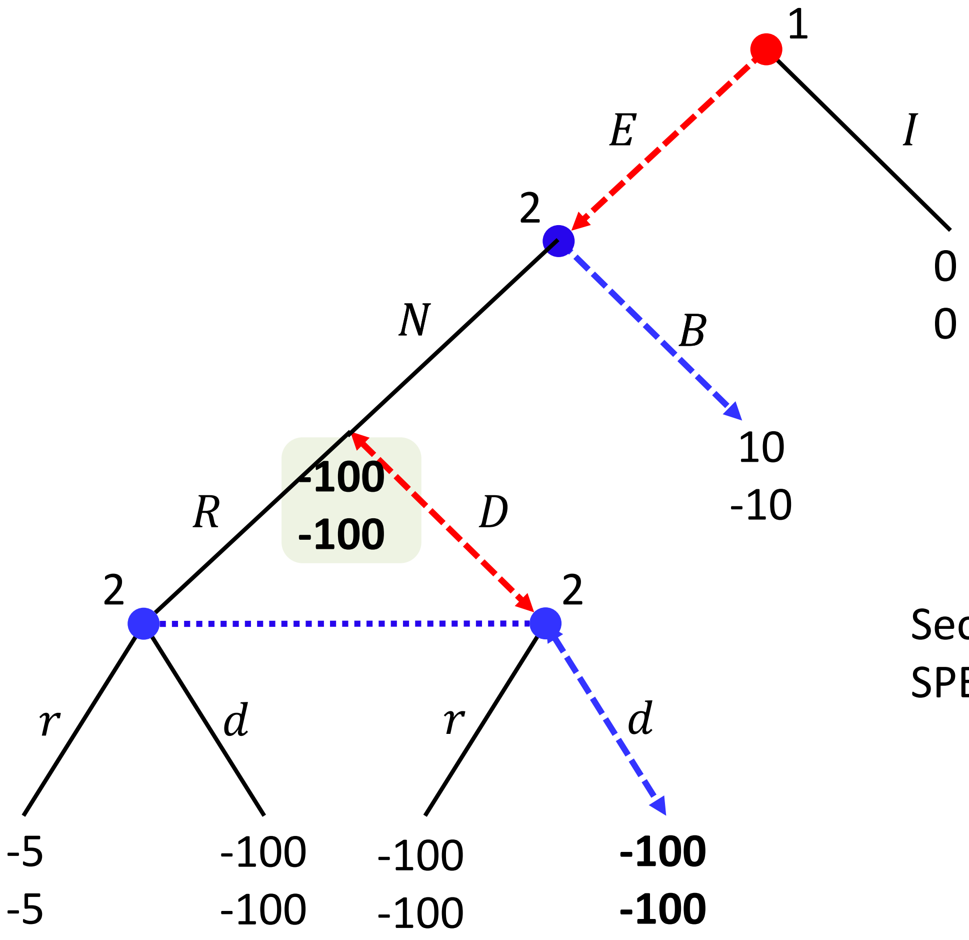


First subgame perfect equilibrium is SPE1 =(IR, Nr)

Case 1: NE1=(R, r) is used at the last subgame

SPE example: Mutually Assured Destruction

- Find the subgame-perfect equilibria using backward induction



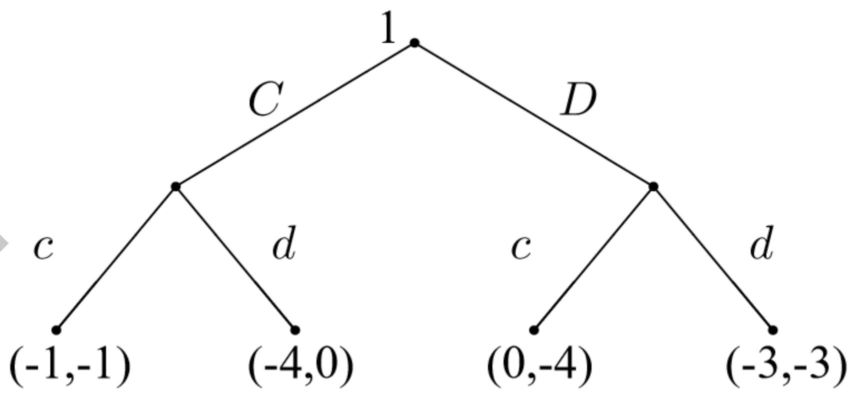
Second subgame perfect equilibrium is SPE2 =(ED, Bd)

Case 2: NE2=(D, d) is used at the last subgame

Imperfect information extensive-form game

Can we represent prisoner's dilemma game into extensive form?

Player 1 \ Player 2	c	d
C	-1, -1	-4, 0
D	0, -4	-3, -3



Motivations

- Up to this point, in our discussion of extensive-form games we have allowed players to specify the action that they would take at every choice node of the game.

- This implies that players know the node they are in and all the prior choices, including those of other agents (perfect information game)

- We may want to model agents needing to act with partial or no knowledge of the actions taken by others, or even with limited memory of their own past actions.

- This is possible using imperfect information extensive-form games.
  - each player's choice nodes are partitioned into **information sets**
  - if two choice nodes are in the same information set then the agent cannot distinguish between them.

Formal definition

Definition (Imperfect-information game)

An imperfect-information game (in extensive form) is a tuple $(N, A, H, Z, \chi, \rho, u, I)$, where:

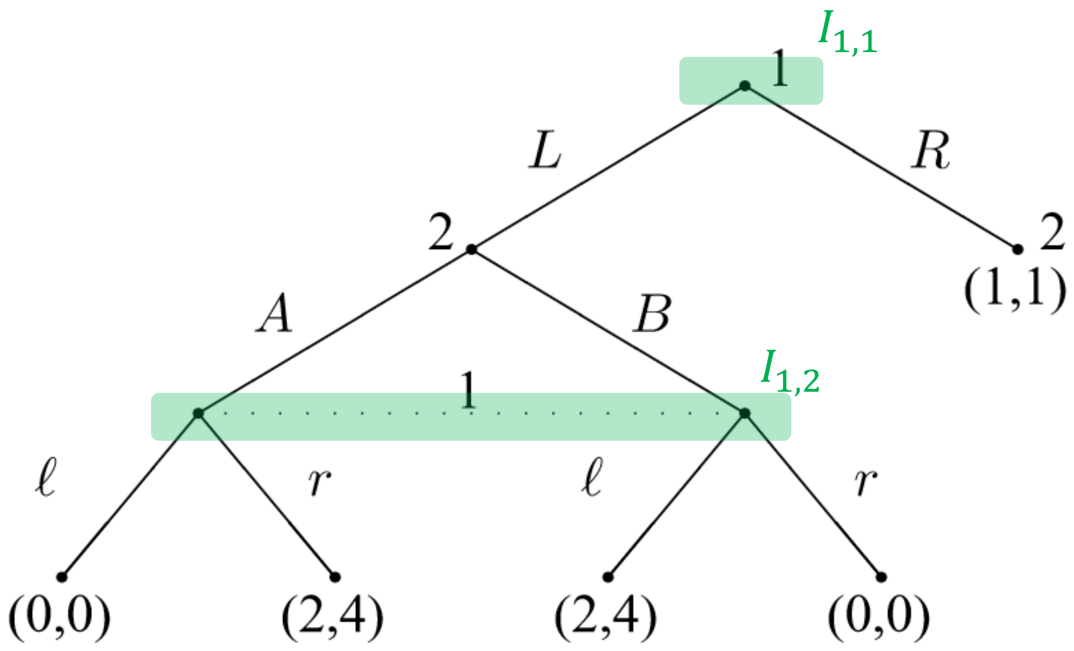
- $(N, A, H, Z, \chi, \rho, u)$ is a perfect-information extensive-form game; and
- $I = (I_1, \dots, I_n)$, where $I_i = \left(I_{i,1}, \dots, I_{i,k_i}\right)$ is a set of equivalence classes on (i.e., a partition of) $\{h \in H : \rho(h) = i\}$ with the property that $\chi(h) = \chi(h')$ and $\rho(h) = \rho(h')$ whenever there exists a j for which $h \in I_{i,j}$ and $h' \in I_{i,j}$.

- in order for the choice nodes to be truly indistinguishable, we require that the set of actions at each choice node in an information set be the same (otherwise, the player would be able to distinguish the nodes)

$$\chi(h) = \chi(h') \text{ and } \rho(h) = \rho(h')$$

- Thus, if $I_{i,j} \in I_i$ is an equivalence class, we can unambiguously use the notation $\chi\left(I_{i,j}\right)$ to denote the set of actions available to player i at any node in information set $I_{i,j}$

Example



An imperfect-information game

- player 1 has two information sets:

$$I_1 = (I_{1,1}, I_{1,2})$$

- The information set $I_{1,2}$ has the same set of possible actions

$$\chi(I_{1,2}) = \{l, r\}$$

- We can regard player 1 as not knowing whether player 2 chose A or B when he makes his choice between l and r

Definition (Pure strategies for imperfect-information game)

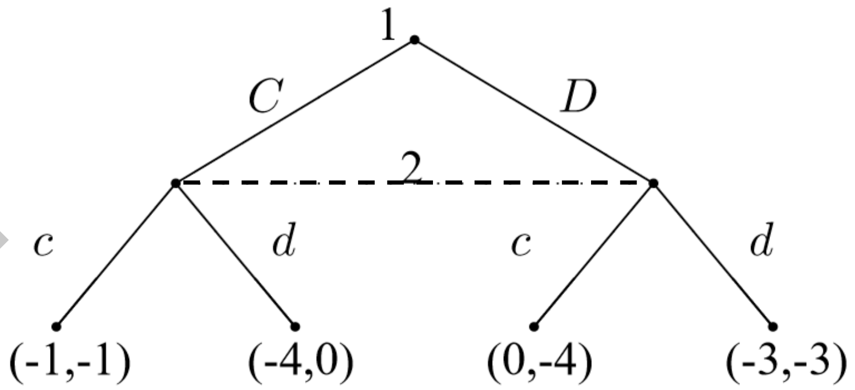
Let $G = (N, A, H, Z, \chi, \rho, u, I)$ be an imperfect-information game. Then the pure strategies of player i consist of the Cartesian product $\prod_{I_{i,j}} \chi(I_{i,j})$

- In other words, a pure strategy for an agent in an imperfect-information game selects one of the available actions in each information set of that agent
- Thus perfect-information games can be thought of as a special case of imperfect-information games, in which every equivalence class of each partition is a singleton.

Normal-form game

- Consider the Prisoner's Dilemma game, represented as the following imperfect information game:

		Player 2: c	Player 2: d
Player 1	C	-1, -1	-4, 0
	D	0, -4	-3, -3



- Recall that perfect-information games were not expressive enough to capture the Prisoner's Dilemma game and many other ones

- In contrast, as is obvious from this example, any normal-form game can be trivially transformed into an equivalent imperfect-information game.

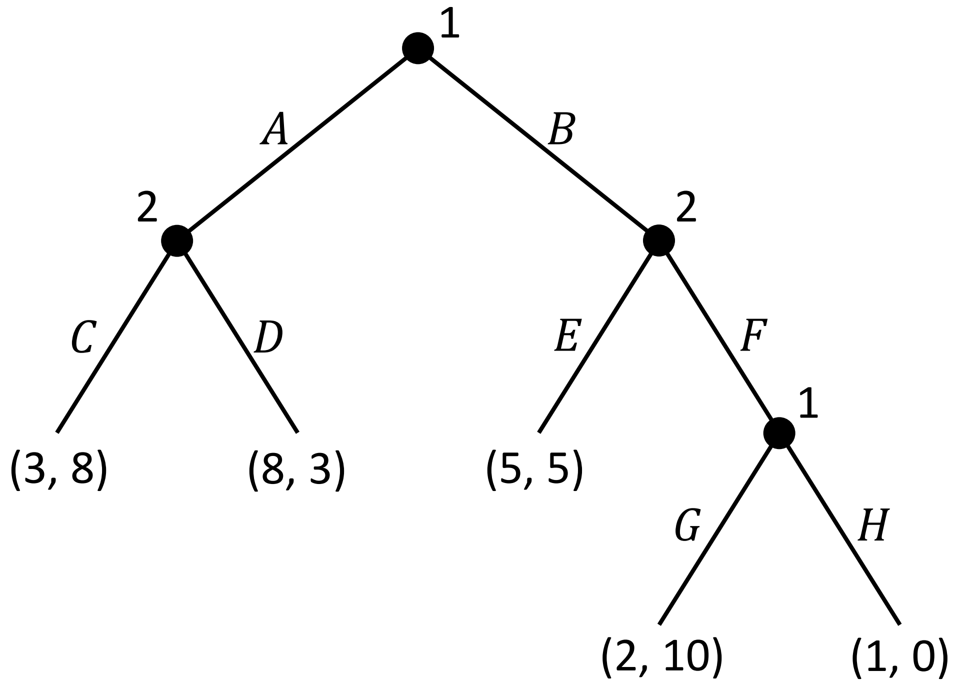
Induced normal-form game

- Same as before: enumerate pure strategies for all agents
- Mixed strategies are just mixtures over the pure strategies as before.
- Nash equilibria are also preserved.
- Note that we've now defined both mapping from normal form games (NF) to Imperfect information extensive form games (IIEF) and a mapping from IIEF to NF.
    - what happens if we apply each mapping in turn?
    - we might not end up with the same game, but we do get one with the same strategy spaces and equilibria.

Randomized strategies

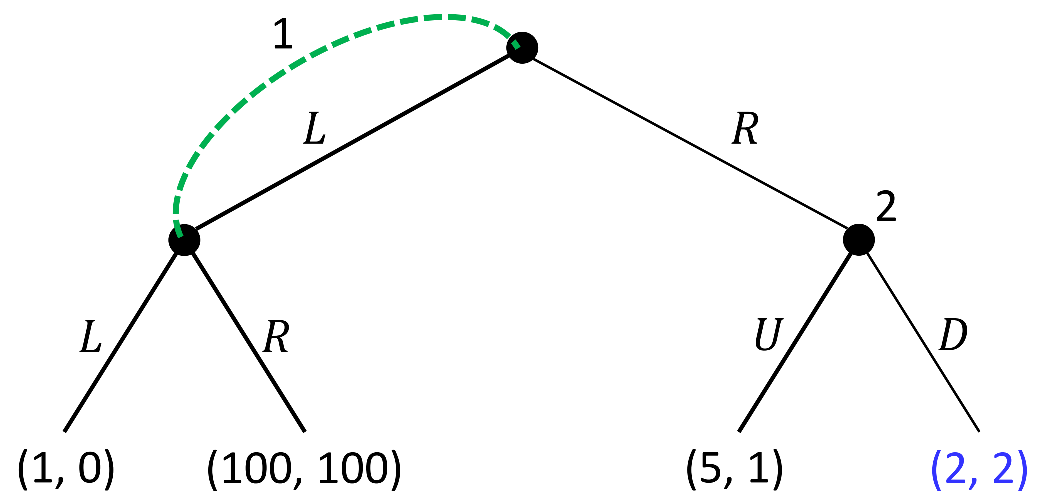
- It turns out there are two meaningfully different kinds of randomized strategies in imperfect information extensive form games
  - Mixed strategy: randomize over pure strategies (i.e., distribution over vectors)
  - Behavioral strategy: **independent coin toss** every time an information set is encountered (vector of distribution)

Randomized strategies example



- Set of all pure strategy
  - $S_1 = \{(A, G), (A, H), (B, G), (B, G)\}$
  - $S_2 = \{(C, E), (C, F), (D, E), (D, F)\}$
- Give an example of a mixed strategy
  - $\{0.15(A, G), 0.35(A, H), 0.15(B, G), 0.35(B, H)\}$
- Give an example of a behavioral strategy:
  - A with probability 0.5 and G with probability 0.3
- In this game every behavioral strategy corresponds to a mixed strategy...

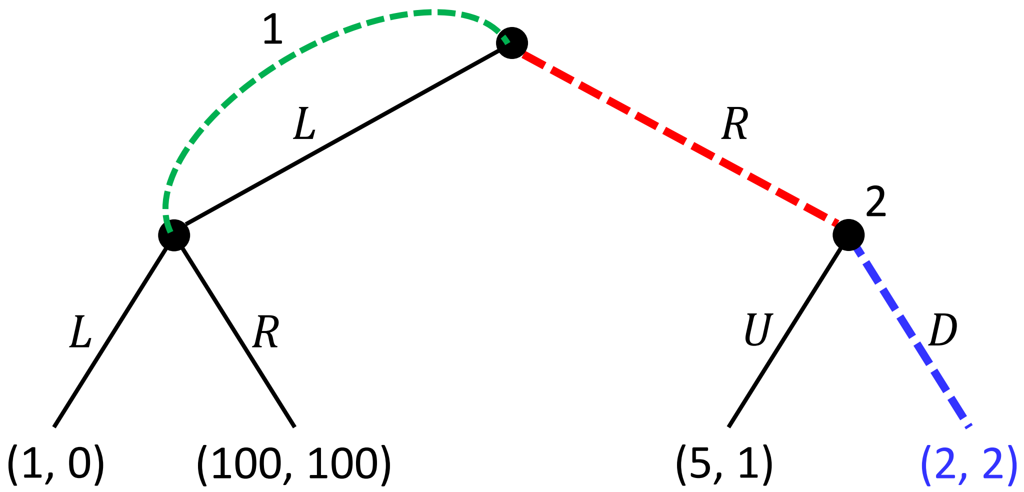
Consider the following game with imperfect recall



- the space of pure strategies
  - Agent 1: $\{L, R\}$,
  - Agent 2: $\{U, D\}$

Games of imperfect recall

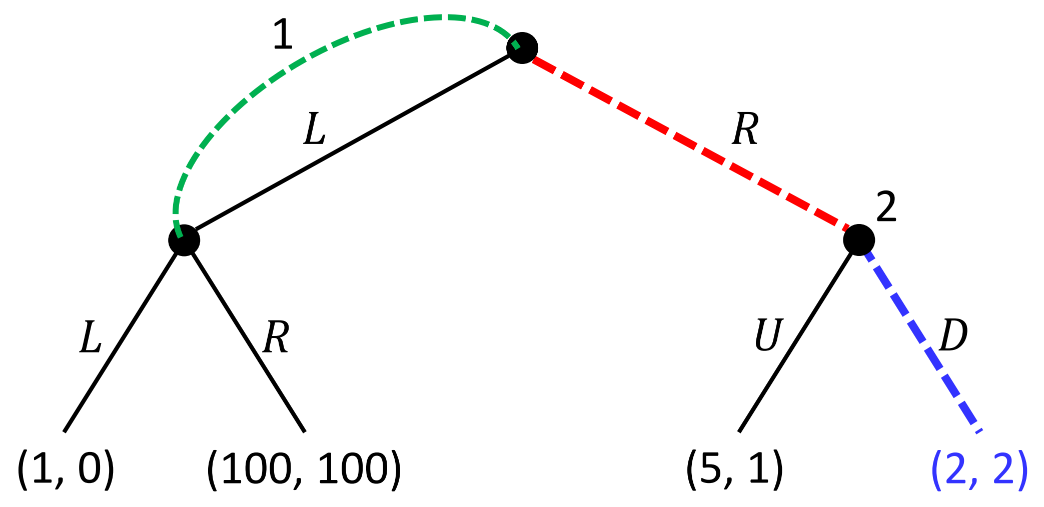
Consider the following game with imperfect recall



- the space of pure strategies
  - Agent 1: $\{L, R\}$,
  - Agent 2: $\{U, D\}$

- What is the **mixed strategy equilibrium** ?
  - ➢ Agent 1 decides probabilistically whether to play L or R in his information set
  - ➢ Once he decides, he plays that pure strategy consistently.
    - ✓ R is a strictly dominant strategy for agent 1,
    - ✓ D is agent 2's strict best response
  - ➢ thus (R,D) is the unique Nash equilibrium.
  - ➢ thus the payoff of 100 is irrelevant in the game

Consider the following game with imperfect recall



- the space of pure strategies
  - Agent 1: $\{L, R\}$,
  - Agent 2: $\{U, D\}$

- What is an equilibrium in **behavioral strategies**?

- With behavioral strategies agent 1 gets to randomize afresh each time he finds himself in the information set

  - Again, D strongly dominant for player 2
  - If 1 uses the behavioral strategy $(p, 1-p)$, his expected utility is $1 \times p^2 + 100 \times p(1-p) + 2 \times (1-p) = -99p^2 + 98p + 2$
  - Maximum at $p = 98/198$
  - Thus equilibrium is $\big((98/198, 100/198), (0{,}1)\big)$

Perfect recall

- In a sequential game, perfect recall refers to the assumption that, at every opportunity to act,
  - each Player remembers what he did in prior moves,
  - each player remembers everything that he knew before.
- Effectively, the assumption is one that players never forget information once it is acquired.

Theorem

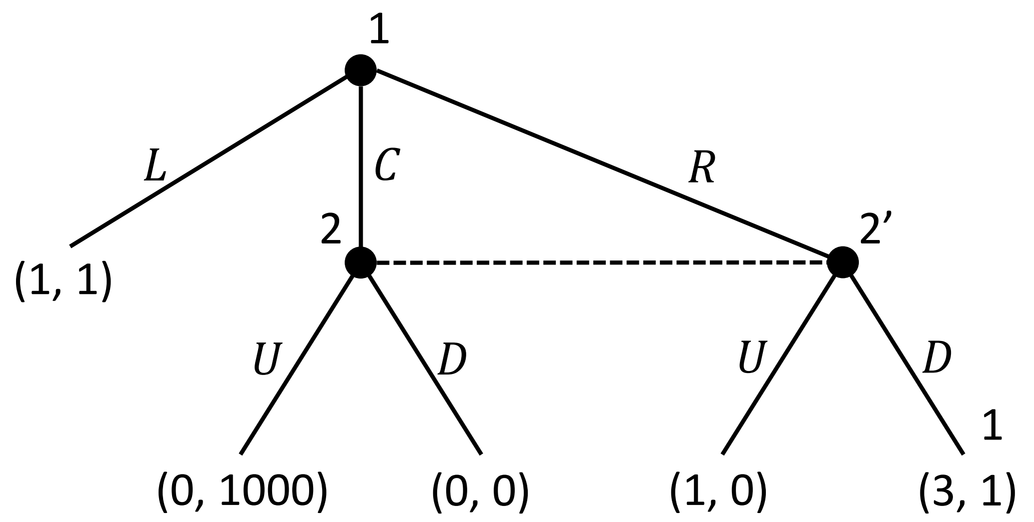
In a game of perfect recall, any mixed strategy of given agent can be replaced by an equivalent behavioral strategy, and any behavioral strategy can be replaced by an equivalent mixed strategy. Here two strategies are equivalent in the sense that they induce the same probabilities on outcomes, for any fixed strategy profile (mixed or behavioral) of the remaining agents

- Perfect information game is a perfect recall game
  - We can find behavioral strategies to find the Nash equilibrium

- In general, imperfect-information games, mixed and behavioral strategies yield noncom parable sets of equilibria

Sequential equilibrium

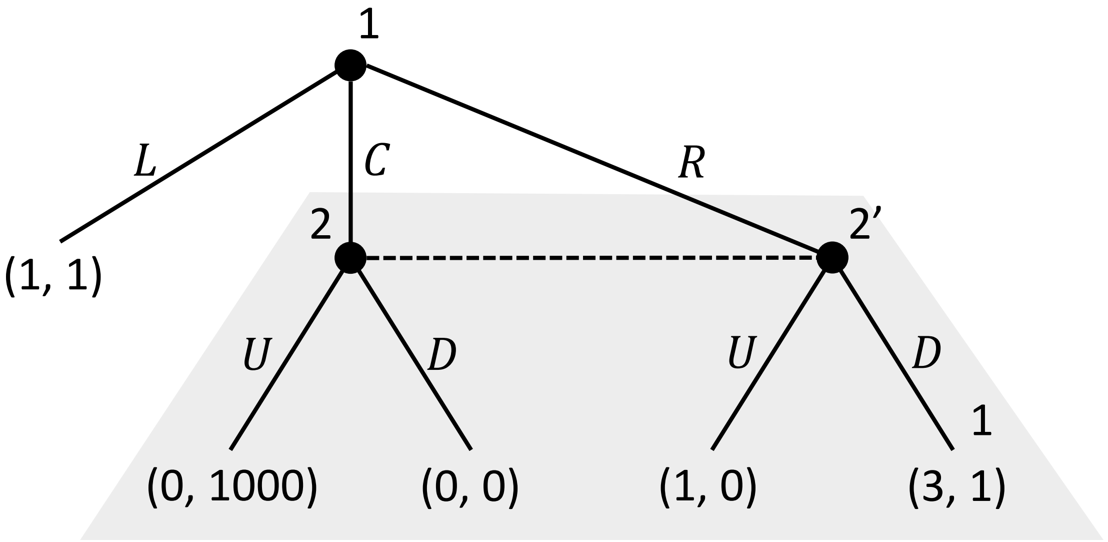
Consider the following imperfect information game in a extensive form game



- In a subgame-perfect equilibrium, we require that the strategy of each agent be a best response in every subgame
  - ✓ It cannot be applied to imperfect information game because a subgame cannot be well defined

Sequential equilibrium

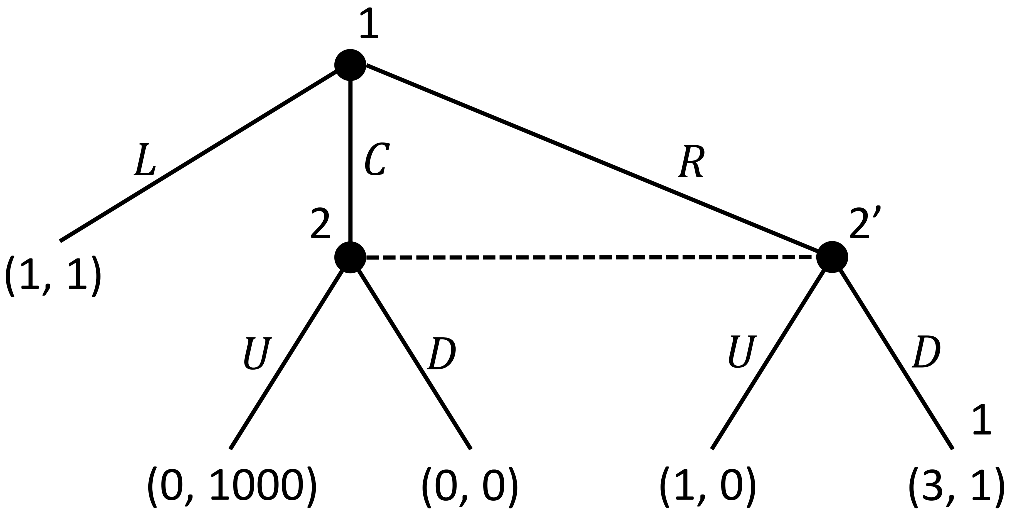
Consider the following imperfect information game in a extensive form game



- In a subgame-perfect equilibrium, we require that the strategy of each agent be a best response in every subgame
  - ✓ It cannot be applied to imperfect information game because a subgame cannot be well defined
  - ✓ What we have at each information set is a "subforest" or a collection of subgames
- Can we require that each player's strategy be a best response in each subgame in each forest?
  - ✓ Sometimes it is yes, but not in general
  - ✓ U dominates D in the left subgame, but D dominate U in the right subgame in the forest

Sequential equilibrium

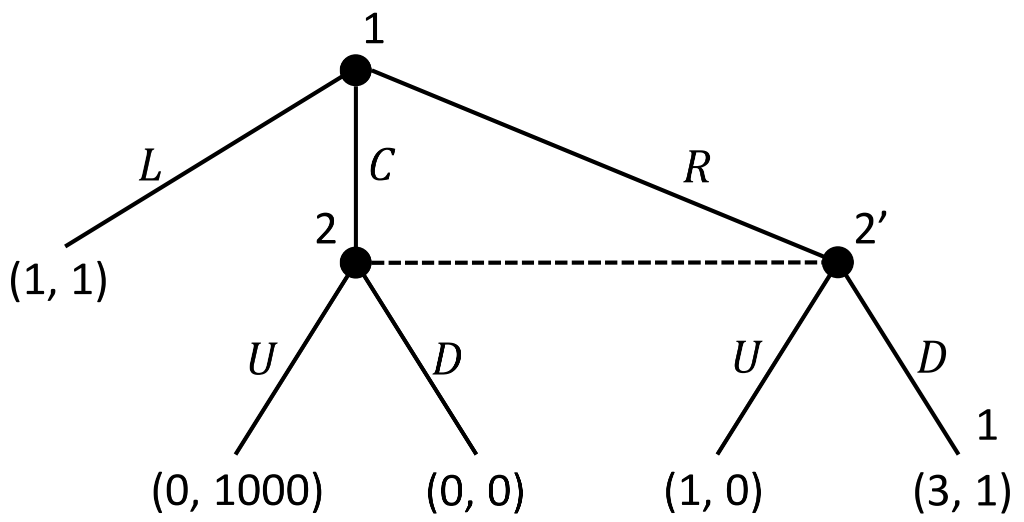
Consider the following imperfect information game in a extensive form game



- An imperfect-information extensive-form game
  ⇒ A normal-form game
- The Nash Equilibrium (both pure and mixed) concept remains the same for imperfect-information extensive-form games

	U	D
L	$\boxed{1, 1}$	1, 1
C	0, 1000	0, 0
R	1, 0	$\boxed{3, 1}$

Consider the following imperfect information game in a extensive form game



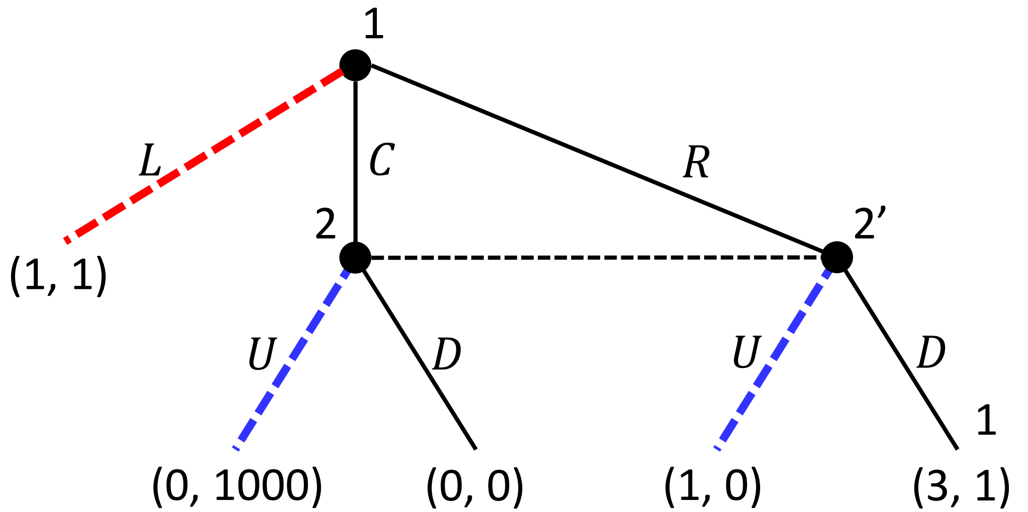
- An imperfect-information extensive-form game
  ⇒ A normal-form game
- The Nash Equilibrium (both pure and mixed) concept remains the same for imperfect-information extensive-form games

- The pure strategies of player 1 are $\{L, \mathrm{C}, \mathrm{R}\}$
- The pure strategies of paler 2 are $\{U, D\}$
- The two pure strategy Nash equilibria are”
  - (L, U) and (R, D)

	U	D
L	1, 1	1, 1
C	0, 1000	0, 0
R	1, 0	3, 1

Sequential equilibrium

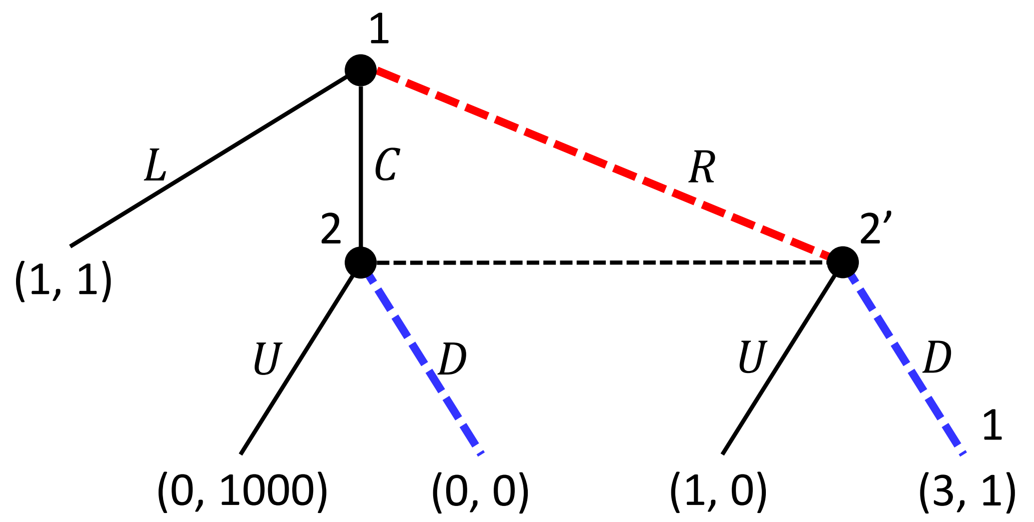
Consider the following imperfect information game in a extensive form game



- The two pure strategy Nash equilibria are: (L, U) and (R, D)
- (L, U) is not a subgame perfect equilibrium
  - ➢ Agent 2 should select D instead of U at the right node in his information set

Sequential equilibrium

Consider the following imperfect information game in a extensive form game

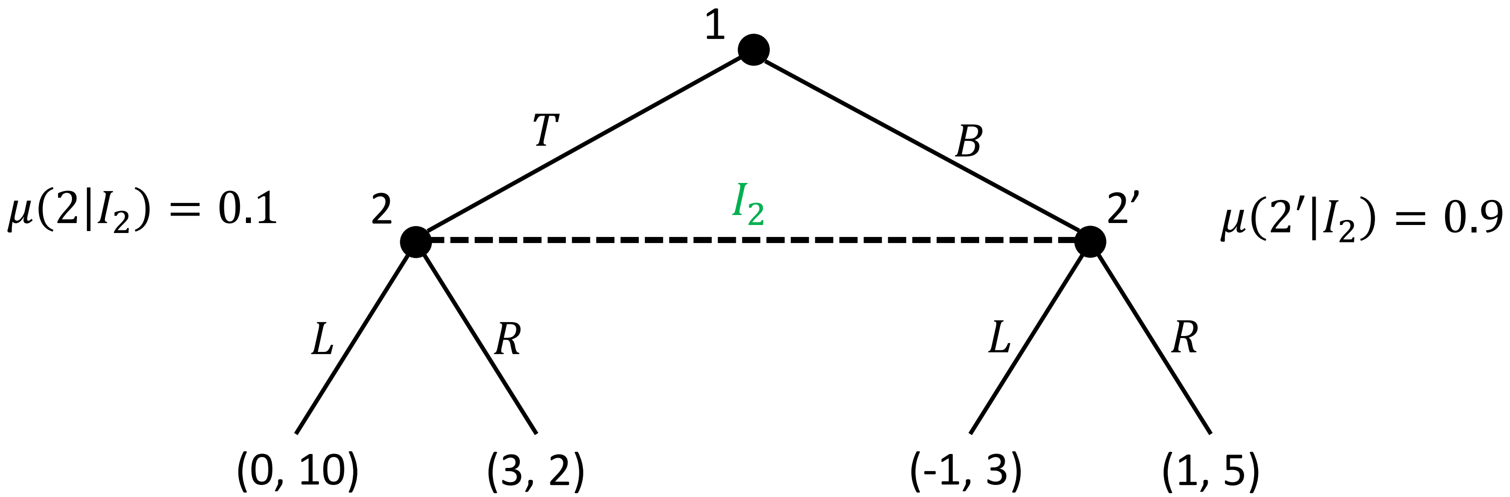


- The two pure strategy Nash equilibria are: (L, U) and (R, D)
- (L, U) is not a subgame perfect equilibrium
  - ➢ Agent 2 should select D instead of U at the right node in his information set
- (R, D) is the only **sequential equilibrium**
  - ➢ R dominates C for player 1, and player 2 knows this
  - ➢ player 2 can deduce that he is in the rightmost one based on player 1's incentives, and hence will go D : (player 2's belief)->(player 2's best action): **sequential rationality**
  - ➢ player 1 knows that player 2 can deduce this, and therefore player 1 should go R (rather than L) : player 2's belief is consistent with player 1's strategy: **consistency**

Beliefs

- A belief μ is a function that assigns to every information set a probability measure on the set of histories in the information set.

- For any information set I, the player who moves at I believes that he is at node $h \in I$ with probability $\mu(h|I)$



Definition (Sequential equilibrium)

A strategy profile s is a sequential equilibrium of an extensive-form game G if there exist probability distributions $\mu(I)$ for each information set I in G, such that the following two conditions hold:

1. Sequentially rational: The assessment (s, μ) is sequentially rational if for every player i and every information set $I_{i,j} \in I_i$ we have

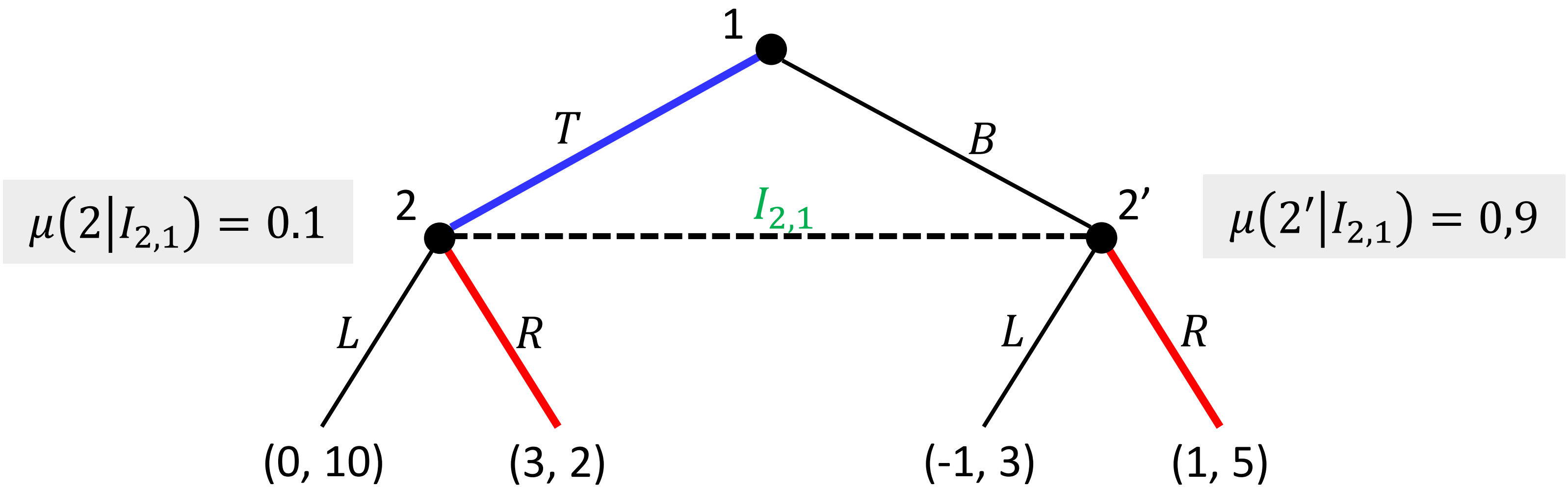
$$E\left[u_i(s_i, s_{-i}) | I_{i,j}\right] \geq E\left[u_i(s_i', s_{-i}) | I_{i,j}\right] \quad \text{for any } s_i' \neq s_i$$

2. Consistency: $(s, \mu) = \lim_{m \to \infty} (s^m, \mu^m)$ for some sequence $(s^1, \mu^1), (s^2, \mu^2), \ldots,$ where s^m is fully mixed, and μ^m is consistent with s^m (derived from s^m using Bayes rule); and

- Sequential equilibrium is a pair, not just a strategy profile
- Hence, in order to identify a sequential equilibrium, one must identify
  - a strategy profile s, which describes what a player does at every information set
  - A belief assessment μ, which describes what a player believes at every information set

1. Sequentially rational: The assessment (s, μ) is sequentially rational if for every player i and every information set $I_{i,j} \in I_i$ we have

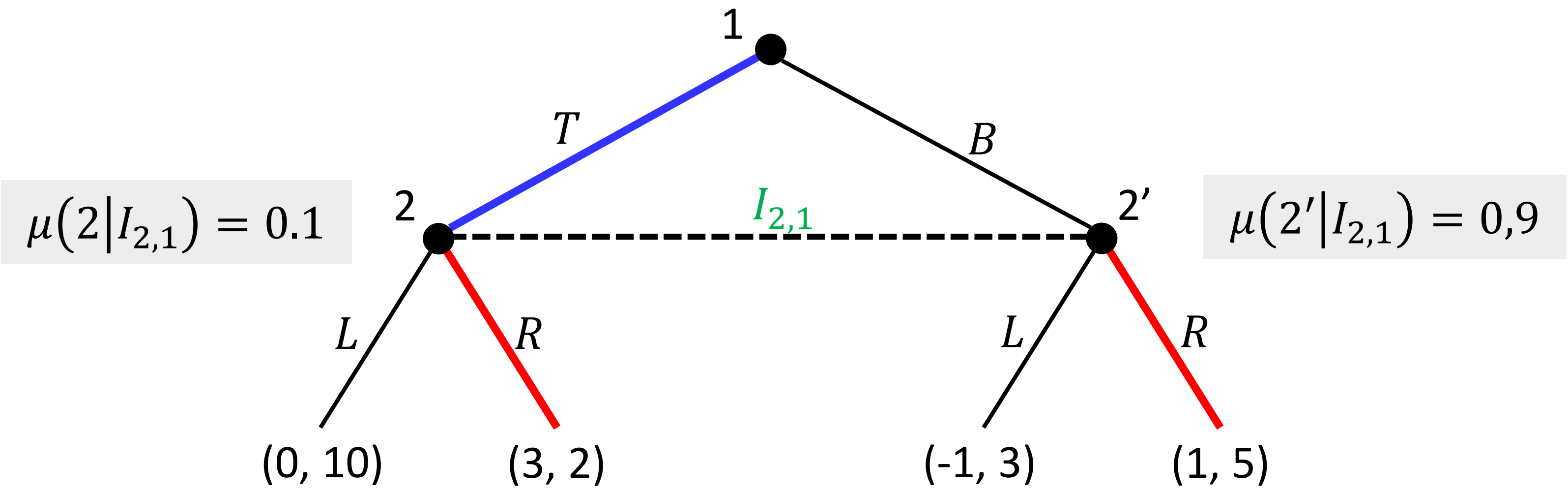
$$E\left[u_i(s_i, s_{-i})|I_{i,j}\right] \geq E\left[u_i(s_i', s_{-i})|I_{i,j}\right] \quad \text{for any } s_i' \neq s_i$$



Sequential Rationality

1. Sequentially rational: The assessment (s, μ) is sequentially rational if for every player i and every information set $I_{i,j} \in I_i$ we have

$$E\left[u_i(s_i, s_{-i})|I_{i,j}\right] \geq E\left[u_i(s_i', s_{-i})|I_{i,j}\right] \quad \text{for any } s_i' \neq s_i$$



- Player 2:

$$E\left[u_2(s_1, L)|I_{2,1}\right] = 0.1 \times 10 + 0.9 \times 3 = 3.7$$
$$E\left[u_2(s_1, R)|I_{2,1}\right] = 0.1 \times 2 + 0.9 \times 5 = 4.7$$

  - Playing R is sequential rational

- Player 1: Deterministically at node 1 (deterministic belief)

$$u_1(s_1 = T, R) > u_1(s_1 = B, R) \text{ for any } s_{-i}$$

- Is playing (T, R) Sequential Equilibrium?

Consistency (on the equilibrium path)

2. Consistency: $(s, \mu) = \lim_{m \to \infty} (s^m, \mu^m)$ for some sequence $(s^1, \mu^1), (s^2, \mu^2), \ldots,$ where s^m is fully mixed, and μ^m is consistent with s^m (derived from s^m using Bayes rule); and

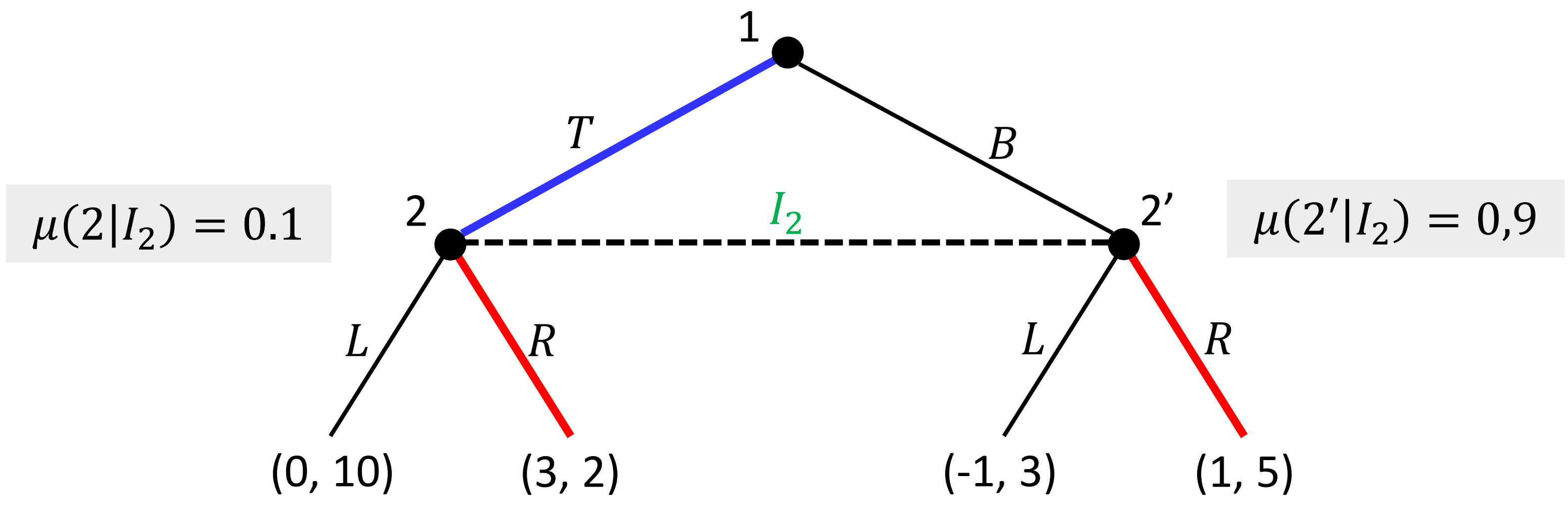
- Let discuss first a simple case (**on the equilibrium path**)
- Given any strategy profile s, belief μ, and any information set I that is reached with positive probability according to s, the beliefs $\mu(\cdot\,|I)$ at I is said to be consistent with s iff $\mu(\cdot\,|I)$ is derived using the **Bayes rule** and s. That is

$$\mu(h|I) = p(h|s) = \frac{p(s|h)p(h)}{\sum_{h'} p(s|h')p(h')}$$

  - $p(h|s)$ is the probability that we reach node h according to s

Consistency (on the equilibrium path)

2. Consistency: $(s, \mu) = \lim_{m\to\infty}(s^m, \mu^m)$ for some sequence $(s^1, \mu^1), (s^2, \mu^2), \ldots$, where s^m is fully mixed, and μ^m is consistent with s^m (derived from s^m using Bayes rule); and



- Let's go back to the example:

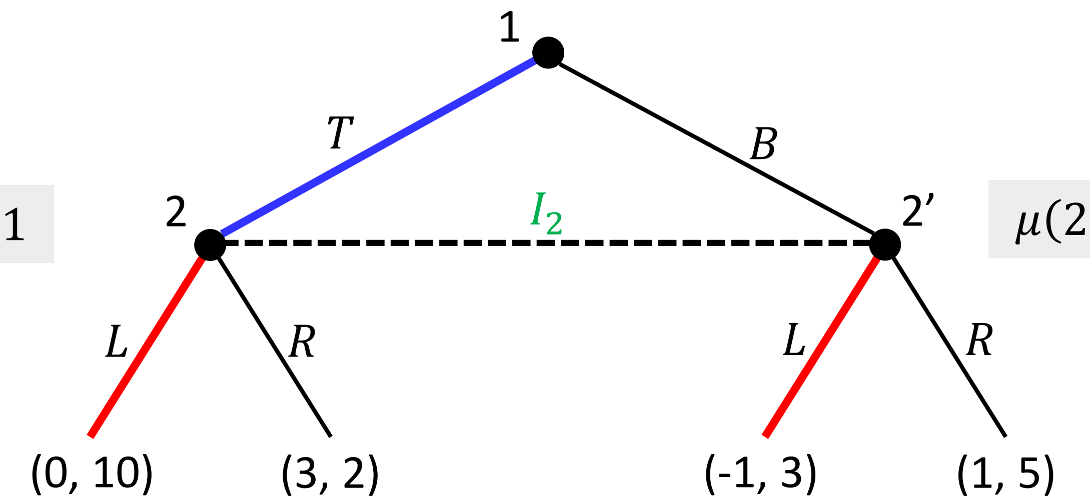
$$\mu(2|s = (T,R)) = \frac{p(s = (T,R)|2)p(2)}{p(s = (T,R)|2)p(2) + p(s = (T,R)|2')p(2')} = \frac{1}{1+0} = 1$$

Q. How was this calculated?

- This contradicts to our previous belief $\mu(2|s) = 0.1$
- Thus, s $= (T, R)$ is not SE

- A pair (s, μ) of a strategy profile s and a belief μ is said to be a sequential equilibrium if (s, μ) is sequentially rational and μ is consistent with s.

$\mu(2|I_2) = 1$



$\mu(2'|I_2) = 0$

- Based on the corrected belief

$$E[u_2(s_1, L)|I_2] = 1 \times 10 + 0 \times 3 = 10$$
$$E[u_2(s_1, R)|I_2] = 1 \times 2 + 0 \times 5 = 2$$

→Thus, player 2 will choose L
Player 1 will choose T

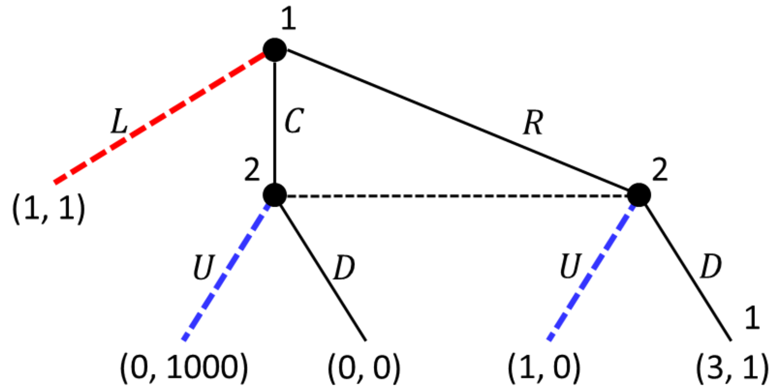
- Check the consistency of belief

$$\mu(2|s = (T, L)) = \frac{p(s = (T, L)|2)p(2)}{p(s = (T, L)|2)p(2) + p(s = (T, L)|2')p(2')} = \frac{1}{1 + 0} = 1$$

- Thus, $\left(\mathrm{s} = (T, L), \mu(\cdot\,|I)\right)$ is sequential equilibrium

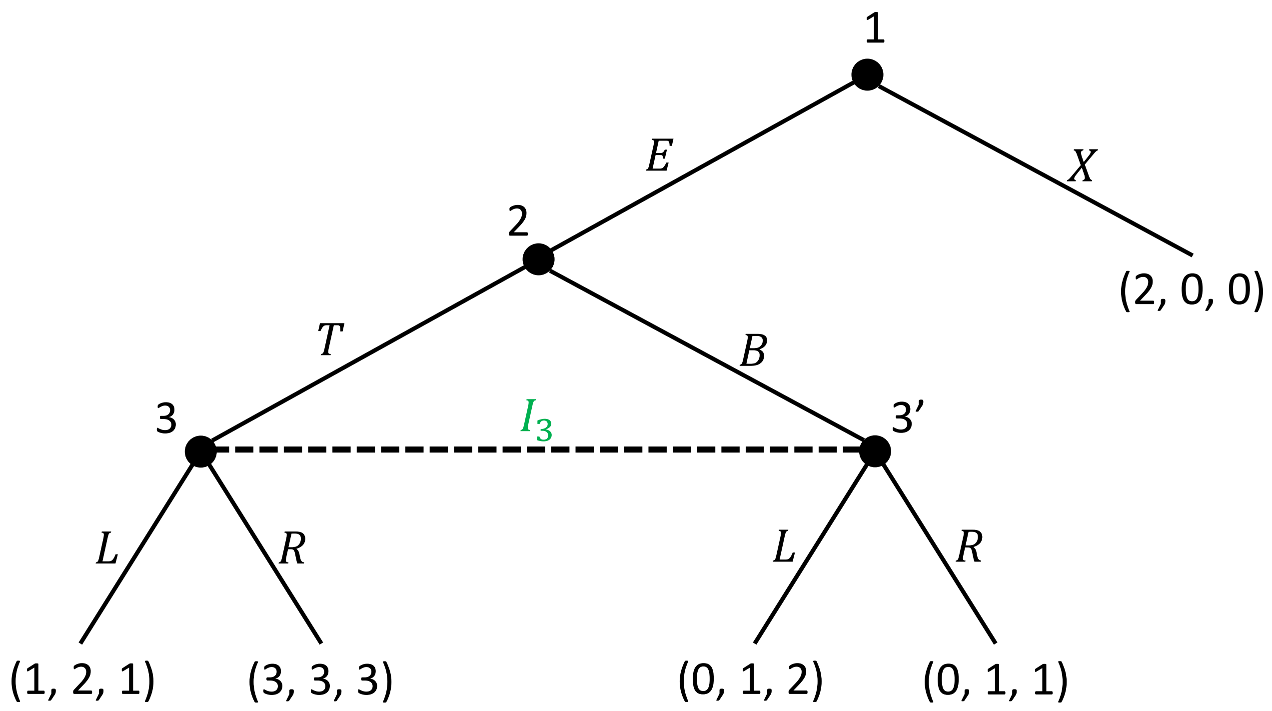
Consistency (off the equilibrium path)

- But, what about beliefs for information sets that are **off the equilibrium path**?
- We want beliefs for information sets that are off the equilibrium path to be reasonable. But what is reasonable?
- Consider the Nash Equilibrium (L, U) again:
  - Player 2's information set will not be reached at the equilibrium, because player 1 will play L with probability 1.
  - But assume that player 1 plays a completely mixed strategy, playing L, C, and R with probabilities $1-\epsilon, \frac{3\epsilon}{4}$, and $\frac{\epsilon}{4}$.
  - Then, the belief on player 2's information set is well defined. Now, if $\epsilon \to 0$, it's still well defined.



	U	D
L	1, 1	1, 1
C	0, 1000	0, 0
R	1, 0	3, 1

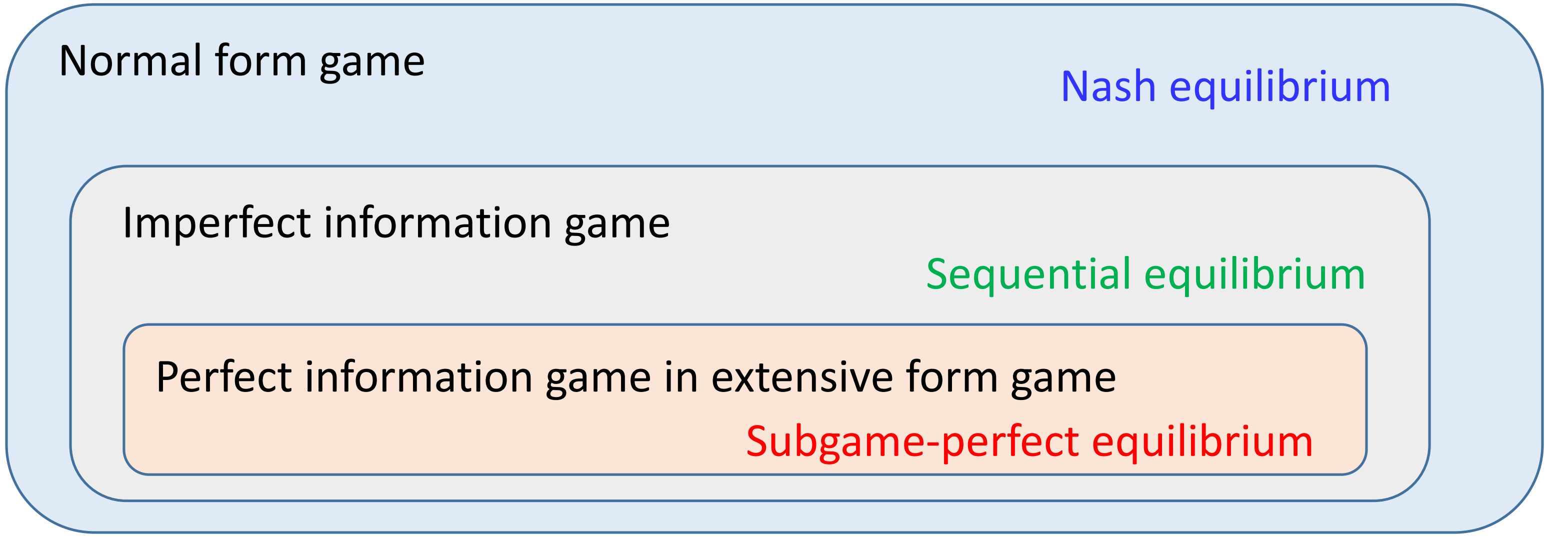
Exercise



- Find sequential equilibria

Q. Answer here is (E, T, R)?

Summary



- Every finite extensive-form game with perfect recall has a sequential equilibrium.
- A sequential equilibrium is a Nash equilibrium.
- With perfect information, a subgame perfect equilibrium is a sequential equilibrium.
- Analogous to subgame-perfect equilibria in games of perfect information, sequential equilibria are guaranteed to always exist
- In extensive-form games of perfect information, the sets of subgame perfect equilibria and sequential equilibria are always equivalent

Multistage games

Motivations

- The normal form game was a model of players choosing their actions simultaneously, without observing the moves of their opponents.

- The extensive form, a player observe the action taken by the previous player and condition his behavior on the moves of other players
  - Have payoffs delayed until the game reached a terminal node

Motivations

- The normal form game was a model of players choosing their actions simultaneously, without observing the moves of their opponents.

- The extensive form, a player observe the action taken by the previous player and condition his behavior on the moves of other players
    - Have payoffs delayed until the game reached a terminal node

- In reality dynamic play over time may be more complex, and it may not be correctly modeled by on “grand” game that unfolds over time with payoffs distributed at the end of the game.

- Players can play one game that is followed by another and receive some playoffs after each one of the games in this sequence is played

Motivations

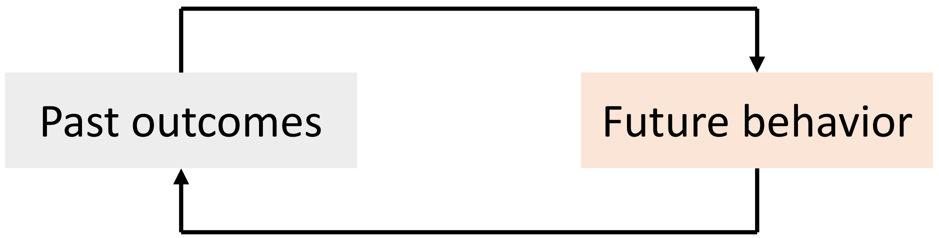
- Two questions we will answer from this lecture
- First, if the players are rational and forward looking, should they not view this sequence of games as one grand game?
- Second, if they do view these as on grand game, should we expect that their actions in the later stages will depends on the outcomes of earlier stages?

will the players be destined to play a sequence of action profiles that are Nash equilibria in each stage-game

vs.

will they be able to use future games to support behavior in the earlier stages that is not consistent with Nash equilibrium in those early stages

- If players can condition future behavior on past outcomes then this may lead to a richer set of self-enforcing outcomes



Credible and self enforcing incentives

Preliminaries

- A multistage game is defined as a finite sequence of normal-form stage-games, in which each stage-game is
  - an independent game
  - well-defined games of complete but imperfect information (a simultaneous-move game)

- These stage-games are played sequentially by the same players
  - the total payoffs from the sequence of games are evaluated using the sequence of outcomes in the games that are played

- We adopt the convention that each game is played in a distinct period, so that game 1 is played in period $t = 1$, game 2 in period $t = 2$, and so on, up until period $t = T$ (last stage)

- We will also assume that, after each stage is completed, all the players observe the outcome of that stage, and that this information structure is common knowledge

Prisoner-Revenge Game

Prisoner's dilemma

	c	d
C	$4, 4$	$-1, 5$
D	$5, -1$	$1, 1$

$t = 1$

Revenge game

	l	g
L	$0, 0$	$-4, -1$
G	$-1, -4$	$-3, -3$

$t = 2$

C: Cooperate with other prisoner
D: Defect other prisoner

- A unique Nash equilibrium (D, d)

L: loner
G: join local gang

- Two pure strategy Nash equilibria (L, l) and (G, g)
- A mixed-strategy equilibrium
- $\{(0.5L, 0.5l), (0.5G, 0.5g)\}$

- We will see that once these two games are played in sequence, equilibrium behavior can be more interesting

Payoffs

- We should evaluate the total payoffs from a sequence of outcomes in each of the sequentially played stage-games
  - ➢ We adopt the well-defined notion of present value

- Consider a multistage game in which there are T stage-games played in each of the periods $1, 2, \ldots, T$.
- Let u_i^t be player i's payoff from the anticipated outcome in the stage-game played in period t
- We denote by u_i the total payoff of player i from playing the sequence of games in the multistage game and define it as

$$u_i = u_i^1 + \gamma u_i^2 + \gamma^2 u_i^3 + \cdots + +\gamma^{T-1} u_i^T = \sum_{t=1}^{T} \gamma^{t-1} u_i^t$$

Which is the discounted sum of payoffs that the player expects to get in the sequence of games

Prisoner's dilemma

	c	d
C	$4, 4$	$-1, 5$
D	$5, -1$	$1, 1$

$t = 1$

Revenge game

	l	g
L	$0, 0$	$-4, -1$
G	$-1, -4$	$-3, -3$

$t = 2$

(C, d) at $t = 1$ and (L, g) at $t = 2$
$\rightarrow u_1 = -1 + \gamma(-4) = -4\gamma - 1$
$\rightarrow u_2 = 5 + \gamma(-1) = -\gamma + 5$

Multistage prisoner-revenge game

Prisoner's dilemma

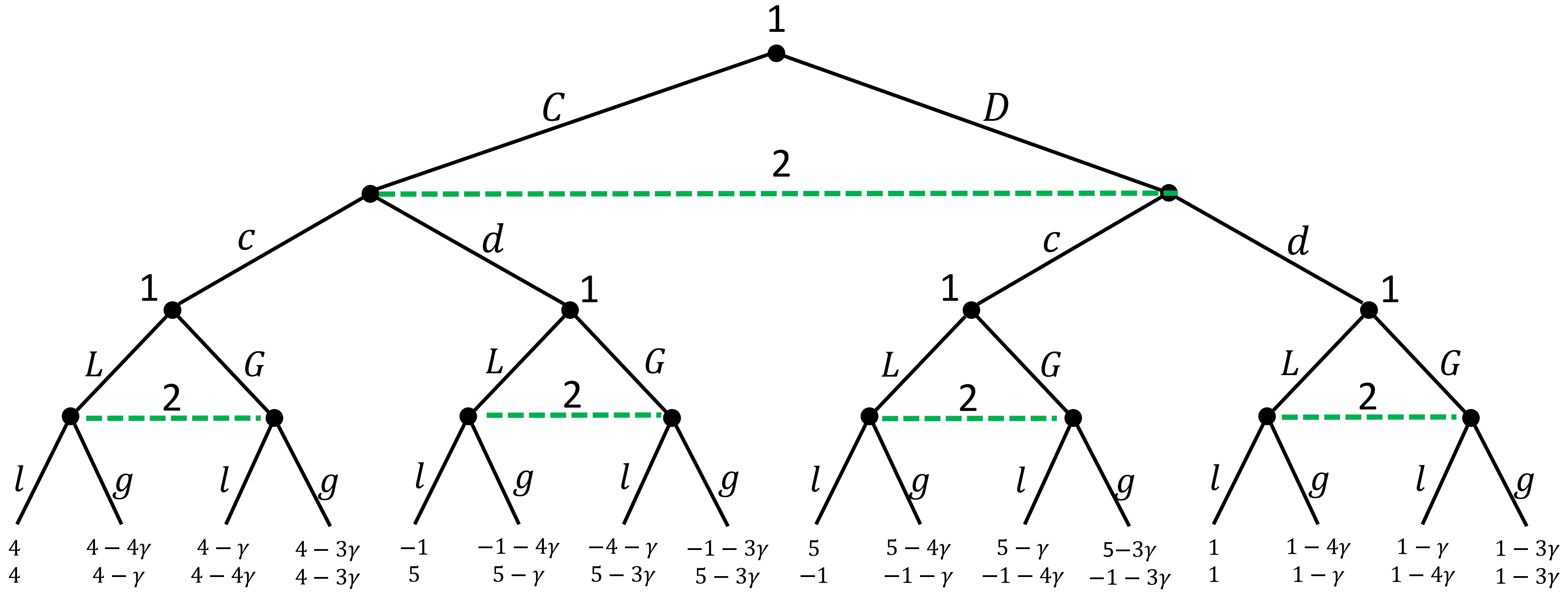
	c	d
C	$4, 4$	$-1, 5$
D	$5, -1$	$1, 1$

$t = 1$

Revenge game

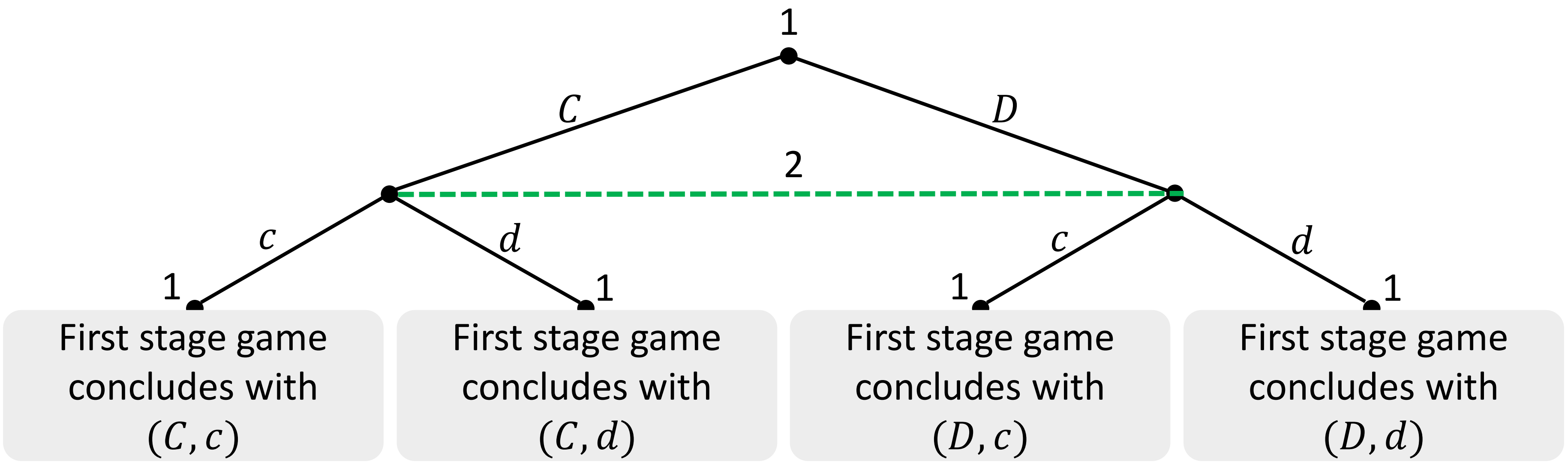
	l	g
L	$0, 0$	$-4, -1$
G	$-1, -4$	$-3, -3$

$t = 2$



Strategies and conditional play

- Create a strategic link between the stage-games
  - Players can use strategies of the form "If such-and-such happens in games $1, 2, \dots, t-1$ then I will choose action s_t in game t."
- It is important to determine correctly the information sets of each player
  - Because the outcomes of every stage are revealed before the next stage is played, the number of information sets at any stage t must be equal to the number of possible outcomes from the previously played stag-games $1,2,\dots,t-1$



- Each player knows exactly how the fist-stage game concluded
- Define a strategy player i in the multistage Prisoner-Revenge game as a quintuple:

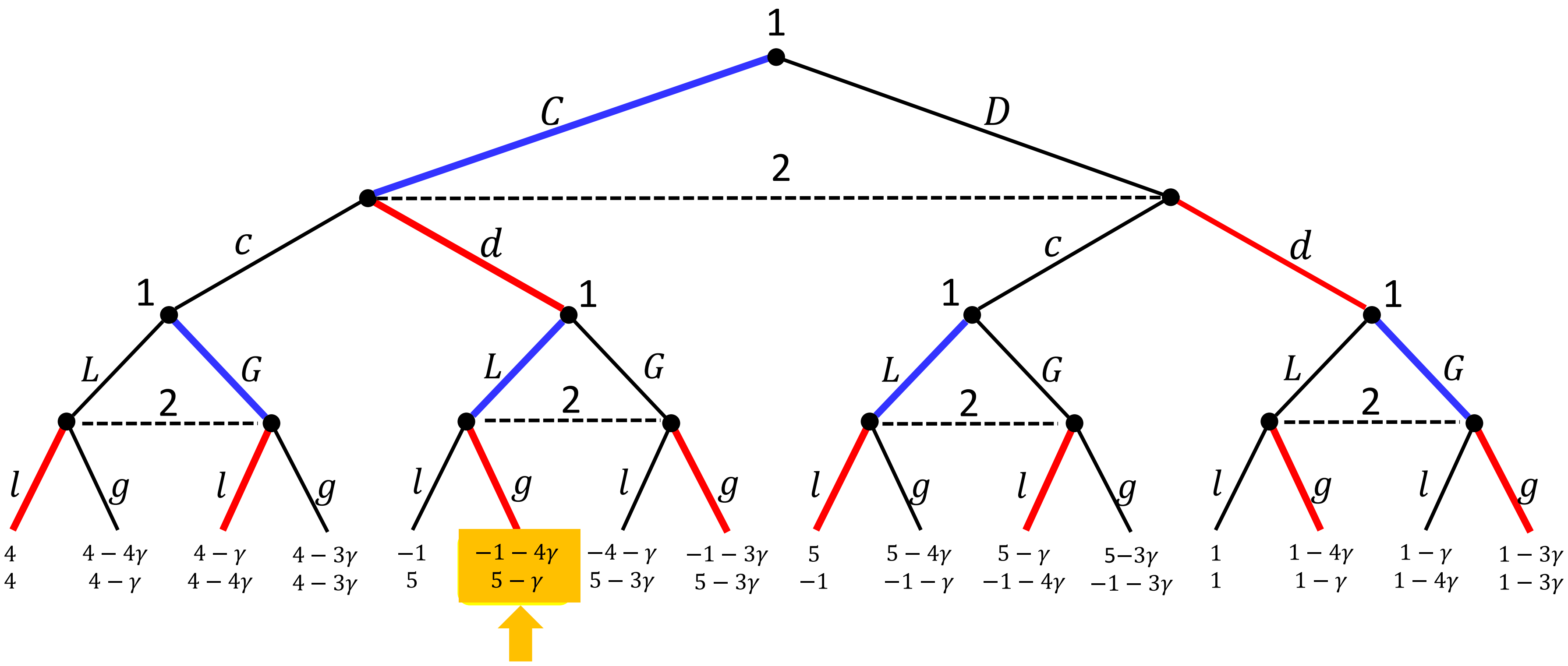
$$s_i = \left(s_i^1, s_i^2(Cc), s_i^2(Cd), s_i^2(Dc), s_i^2(Dd)\right)$$

s_i^1 is the action taken at the first stage, $s_i^2(Cc)$ the action taken at the second stage with Cc result at $t=1$

Strategies and conditional play

$s_1^1 \in \{C, D\}, s_1^2(xy) \in \{L, G\}$
$s_1 = (s_1^1, s_1^2(Cc), s_1^2(Cd), s_1^2(Dc), s_1^2(Dd)) = (C, GLLG)$

$s_2^1 \in \{c, d\}, s_2^2(xy) \in \{l, g\}$
$s_2 = (s_2^1, s_2^2(Cc), s_2^2(Cd), s_2^2(Dc), s_2^2(Dd)) = (d, lglg)$



Strategies and conditional play

- In a multistage game that consists of T stage-games, a pure strategy of player i will be a list of conditional pure strategies of the following form:

$$S_i = \{s_i^1, s_i^2(h_1), \dots, s_i^t(h_{t-1}), \dots, s_i^T(h_{T-1})\}$$

  - h_{t-1} is a particular outcome that occurred up to period t from all possible histories (or outcomes) H_{t-1}
  - $s_i^t(h_{t-1})$ is an action for player i from the t-th stage

- As an example, consider a game with n firms that are choosing prices in a sequence of markets:
  - If each firm selects its price p_i^t in period t, a pure strategy for firm i is a list of prices $\left(p_i^1, p_i^2(h_1), \dots, p_i^t(h_{t-1}), \dots, p_i^T(h_{T-1})\right)$
  - Each history h_{t-1} is the sequence of previously chosen prices:
  - ✓ $h_1 = (p_1^1, p_2^1, \dots, p_n^1)$
  - ✓ $h_2 = \left((p_1^1, p_2^1, \dots, p_n^1), (p_1^2, p_2^2, \dots, p_n^2)\right) = (h_1, (p_1^2, p_2^2, \dots, p_n^2))$
  - ✓ $h_{t-1} = \left((p_1^1, p_2^1, \dots, p_n^1), (p_1^2, p_2^2, \dots, p_n^2), \dots, (p_1^{t-1}, p_2^{t-1}, \dots, p_n^{t-1})\right) = \left(h_{t-2}, (p_1^{t-1}, p_2^{t-1}, \dots, p_n^{t-1})\right)$

- Strategies are defined as a complete list of (mixed or pure) actions for each player at each of his information set that is associated with the history of play from the previous stage-games

Subgame-perfect equilibria for multistage games

- Because multistage games are dynamic in nature, and because past play is revealed over time, we use the notion of a subgame-perfect equilibrium
  - ➢ Rational players should paly sequentially rational strategies
- **Proposition:** Consider a multistage game with T stages, and let s^{t*} be a Nash equilibrium strategy profile for the tth stage-game. There exists a subgame-perfect equilibrium in the multistage game in which the equilibrium path coincides with the path generated by $s^{1*}, s^{2*}, \dots, s^{T*}$.

Prisoner's dilemma

	c	d
C	$4, 4$	$-1, 5$
D	$5, -1$	$\boxed{1, 1}$

$t = 1$

Revenge game

	l	g
L	$\boxed{0, 0}$	$-4, -1$
G	$-1, -4$	$\boxed{-3, -3}$

$t = 2$

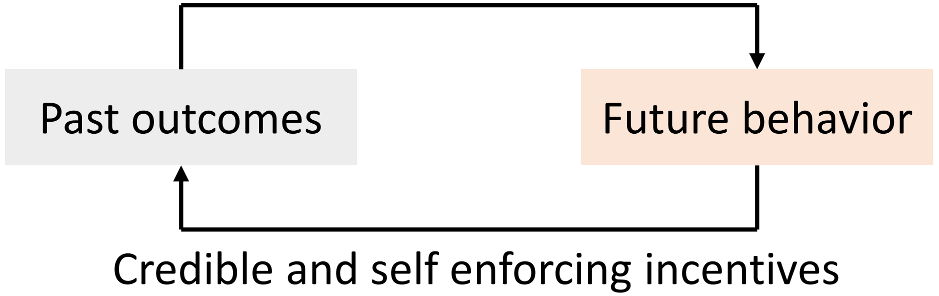
$$s_1 = (s_1^1, s_1^2(Cc), s_1^2(Cd), s_1^2(Dc), s_1^2(Dd)) = (D, L, L, L, L)$$
$$s_2 = (s_2^1, s_2^2(Cc), s_2^2(Cd), s_2^2(Dc), s_2^2(Dd)) = (d, l, l, l, l)$$

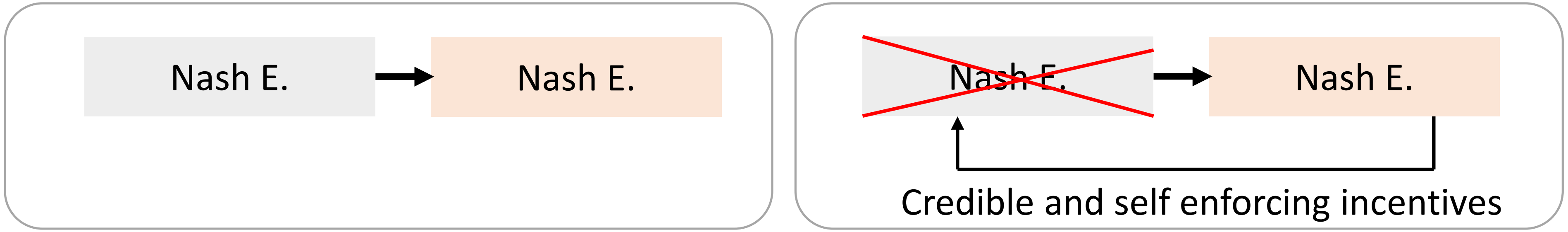
$$s_1 = (s_1^1, s_1^2(Cc), s_1^2(Cd), s_1^2(Dc), s_1^2(Dd)) = (D, G, G, G, G)$$
$$s_2 = (s_2^1, s_2^2(Cc), s_2^2(Cd), s_2^2(Dc), s_2^2(Dd)) = (d, g, g, g, g)$$

Subgame perfect equilibrium

Subgame-perfect equilibria for multistage games

- How to consider a strategic linkage between the games?
- If players were to condition their future play on past paly in a sequentially rational manner, would be able to support behavior in early stages that is not a Nash equilibrium in the early stage-games?





- Yes, the trick for strategically linking the stage-games will be to condition the behavior in later stage-games on the actions taken in earlier stage-games
  - ➢ This will be possible only when some of the stage-games in later periods have multiple Nash equilibria

Subgame-perfect equilibria for multistage games

Prisoner's dilemma

	c	d
C	$\boxed{4, 4}$	$-1, 5$
D	$5, -1$	$\boxed{1, 1}$

Revenge game

	l	g
L	$\boxed{0, 0}$	$-4, -1$
G	$-1, -4$	$\boxed{-3, -3}$

- It would be nice for the players to paly (C, C) in the first period, even though it is not Nash equilibrium at the first stage game.

- To support cooperative behavior (C, c) as part of an equilibrium path of paly in a subgame-perfect equilibrium, the players must find a way to give themselves incentives to stick to (C, c)
  - If they cooperate, give an incentive to them by playing (L, l)
  - If they defect, punish them by playing (G, g)

  Both are N.E

Player 1
Stage 1: Play C in
Stage 2: Play L if (C, c) was played in stage 1, and play G if anything but (C, c) was played

Player 2
Stage 1: Play c in
Stage 2: Play l if (C, c) was played in stage 1, and play g if anything but (C, c) was played

Subgame-perfect equilibria for multistage games

Prisoner's dilemma

	c	d
C	$\boxed{4, 4}$	$-1, 5$
D	$5, -1$	$\boxed{1, 1}$

Revenge game

	l	g
L	$\boxed{0, 0}$	$-4, -1$
G	$-1, -4$	$\boxed{-3, -3}$

- It would be nice for the players to paly (C, C) in the first period, even though it is not Nash equilibrium at the first stage game.
- To support cooperative behavior (C, c) as part of an equilibrium path of paly in a subgame-perfect equilibrium, the players must find a way to give themselves incentives to stick to (C, c)
  - If they cooperate, give an incentive to them by playing (L, l)
  - If they defect, punish them by playing (G, g)

Player 1

$$s_1 = (s_1^1, s_1^2(Cc), s_1^2(Cd), s_1^2(Dc), s_1^2(Dd)) = (C, \underbrace{L}_{\text{incentive}}, \underbrace{G, G, G}_{\text{punishment}})$$

Player 2

$$s_2 = (s_{12}^1, s_2^2(Cc), s_2^2(Cd), s_2^2(Dc), s_2^2(Dd)) = (c, \underbrace{l}_{\text{incentive}}, \underbrace{g, g, g}_{\text{punishment}})$$

Subgame-perfect equilibria for multistage games

- We need to check if this pair of strategies is a subgame-perfect equilibrium!

- At the second stage game, we are playing either one of two Nash equilibria
  - Already part of subgame-perfect equilibrium

- Thus, we need to show that players would not want to deviate from plying (C, c) at the first stage-game
  - That is, we need to show that plying (C, c) is a best response to what each player believes about other players given the continuation play in the second-period game
    - ✓ $u_1(C, s_2) = 4 + \gamma \times 0$
    - ✓ $u_1(D, s_2) = 5 + \gamma \times (-3)$
    - ✓ C is a best response if an only if $4 \geq 5 - 3\gamma$ or $\gamma \geq 1/3$

- Thus, if the discount factor is not too small, then we can support the behavior of (C, c) in the first-stage game even though it is not a Nash equilibrium in the stand alone Prisoner's Dilemma game.

- The fact that the Revenge Game has two different Nash equilibria, one of which is significantly better than the other, allows the players to offer a self-enforcing incentive scheme that supports cooperative behavior in the first-stage game

Subgame-perfect equilibria for multistage games

- **Two requirements** that are crucial to supporting behavior in the first stage (or early periods in general) that is not a Nash equilibrium are:

  1. There must be at least two distinct equilibria in the second stage: a “stick” and a “carrot”

  2. The discount factor has to be large enough for the difference in payoffs between the “stick” and the “carrot” to have enough impact in the first sage of the game

Verify the optimality

- If there are three stage game, it would be more complicated to check that a profile of strategies constitutes a subgame-perfect equilibrium

- If we have two actions available at each of three information sets
- If LLL is optimum given other player's fixed strategy s_{-i} , we need to check

$$u_i(LLL, s_{-i}) \geq u_i(LLR, s_{-i})$$
$$u_i(LLL, s_{-i}) \geq u_i(LRL, s_{-i})$$
$$u_i(LLL, s_{-i}) \geq u_i(LRR, s_{-i})$$
$$u_i(LLL, s_{-i}) \geq u_i(RLL, s_{-i})$$
$$u_i(LLL, s_{-i}) \geq u_i(RLR, s_{-i})$$
$$u_i(LLL, s_{-i}) \geq u_i(RRL, s_{-i})$$
$$u_i(LLL, s_{-i}) \geq u_i(RRR, s_{-i})$$

- That is, maybe they could gain by combining several deviations in separate stages of the game.

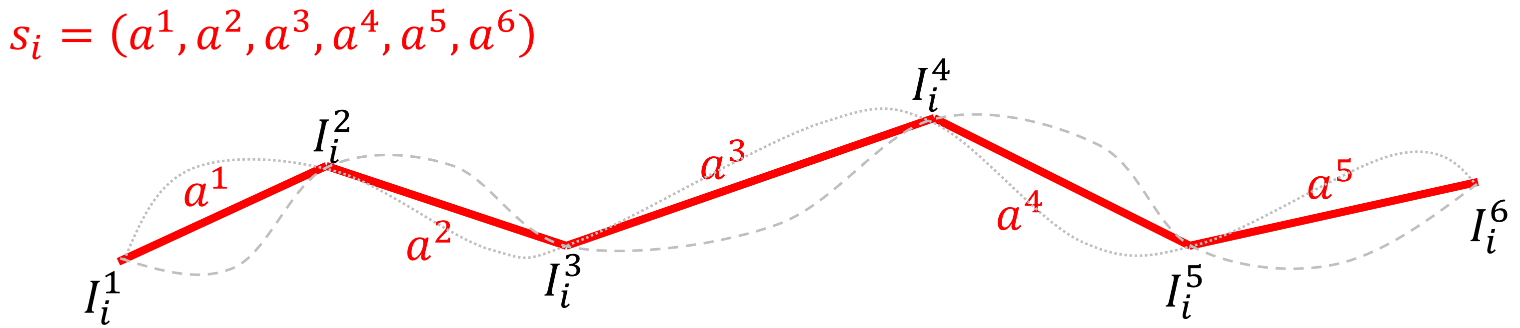
- If there are 10 nodes? $2^{10} = 1024$ combinations!

- Do we need to check every possible combination to check the optimality of a strategy?

Verify the optimality

- Treat every information set I_i as a node in the single-player decision tree induced by other players' strategy s_{-i} (considered to be fixed)

- Then, we can define $u_i(s_i, I_i)$ to be the expected payoff of player i from the information set I_i onward by playing s_i



- We say that a strategy s_i is optimal if there is no strategy s_i' and information set I_i, such that

$$u_i(s_i', I_i) > u_i(s_i, I_i)$$

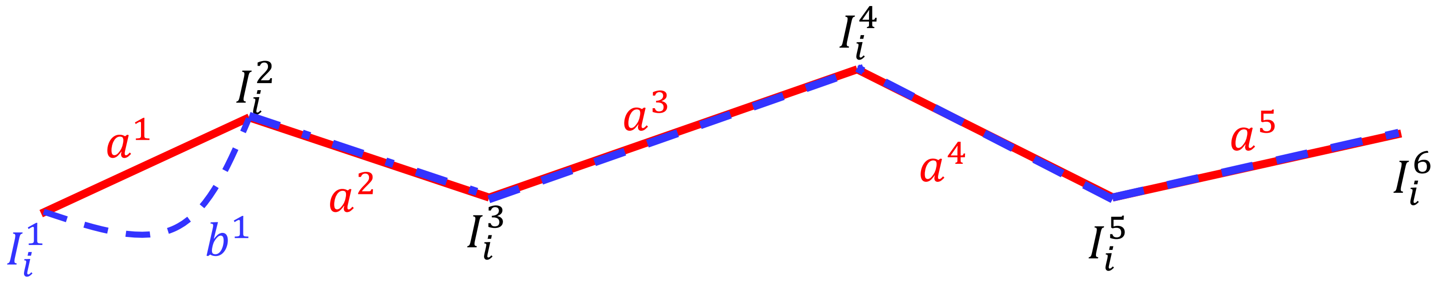
- Checking every possible s_i' would be daunting tasks!

Definition (One-stage unimprovable)

A strategy s_i is one-stage unimprovable if there is no information set I_i, action $b \in \chi(I_i)$, and the corresponding strategy s_i^{b,I_i}, such that $u_i(s_i^{b,I_i}, I_i) > u_i(s_i, I_i)$

s_i^{b,I_i} is the strategy that is identical to s_i everywhere except at I_i.

$s_i = (a^1, a^2, a^3, a^4, a^5, a^6)$ $\quad$ $s_i^{b^1,I_i^1} = (b^1, a^2, a^3, a^4, a^5, a^6)$



- Deviation at I_i^1 : $a^1 \to b^1$

$$u_i(s_i, I_i^1) \geq u_i\left(s_i^{b^1,I_i^1}, I_i^1\right)$$

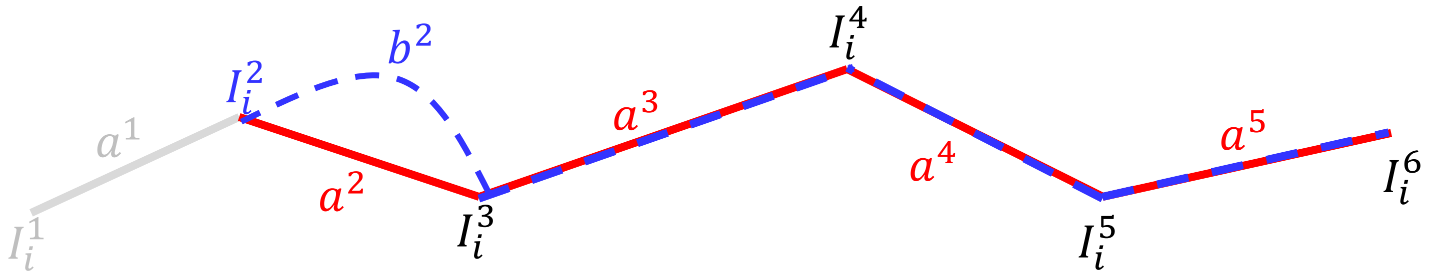
The one-stage deviation principles

Definition (One-stage unimprovable)

A strategy s_i is one-stage unimprovable if there is no information set I_i, action $b \in \chi(I_i)$, and the corresponding strategy s_i^{b,I_i}, such that $u_i\left(s_i^{b,I_i}, I_i\right) > u_i(s_i, I_i)$

s_i^{b,I_i} is the strategy that is identical to s_i everywhere except at I_i.

$s_i = (a^1, a^2, a^3, a^4, a^5, a^6)$ $\quad s_i^{b^2,I_i^2} = (b^1, b^2, a^3, a^4, a^5, a^6)$



- Deviation at I_i^2 : $a^2 \to b^2$

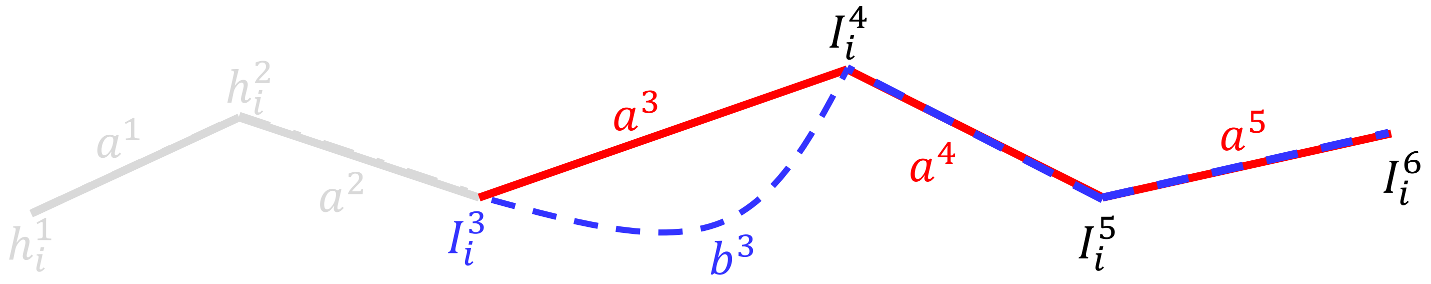
$$u_i(s_i, I_i^2) \geq u_i\left(s_i^{b^2,I_i^2}, I_i^2\right)$$

Definition (One-stage unimprovable)

A strategy s_i is one-stage unimprovable if there is no information set I_i, action $b \in \chi(I_i)$, and the corresponding strategy s_i^{b,I_i}, such that $u_i(s_i^{b,I_i}, I_i) > u_i(s_i, I_i)$

s_i^{b,I_i} is the strategy that is identical to s_i everywhere except at I_i.

$s_i = (a^1, a^2, a^3, a^4, a^5, a^6)$ $\quad s_i^{b^3,I_i^3} = (a^1, a^2, b^3, a^4, a^5, a^6)$



- Deviation at I_i^3 : $a^3 \to b^3$

$$u_i(s_i, I_i^3) \geq u_i\left(s_i^{b^3,I_i^3}, I_i^3\right)$$

Theorem (One-stage deviation principle)

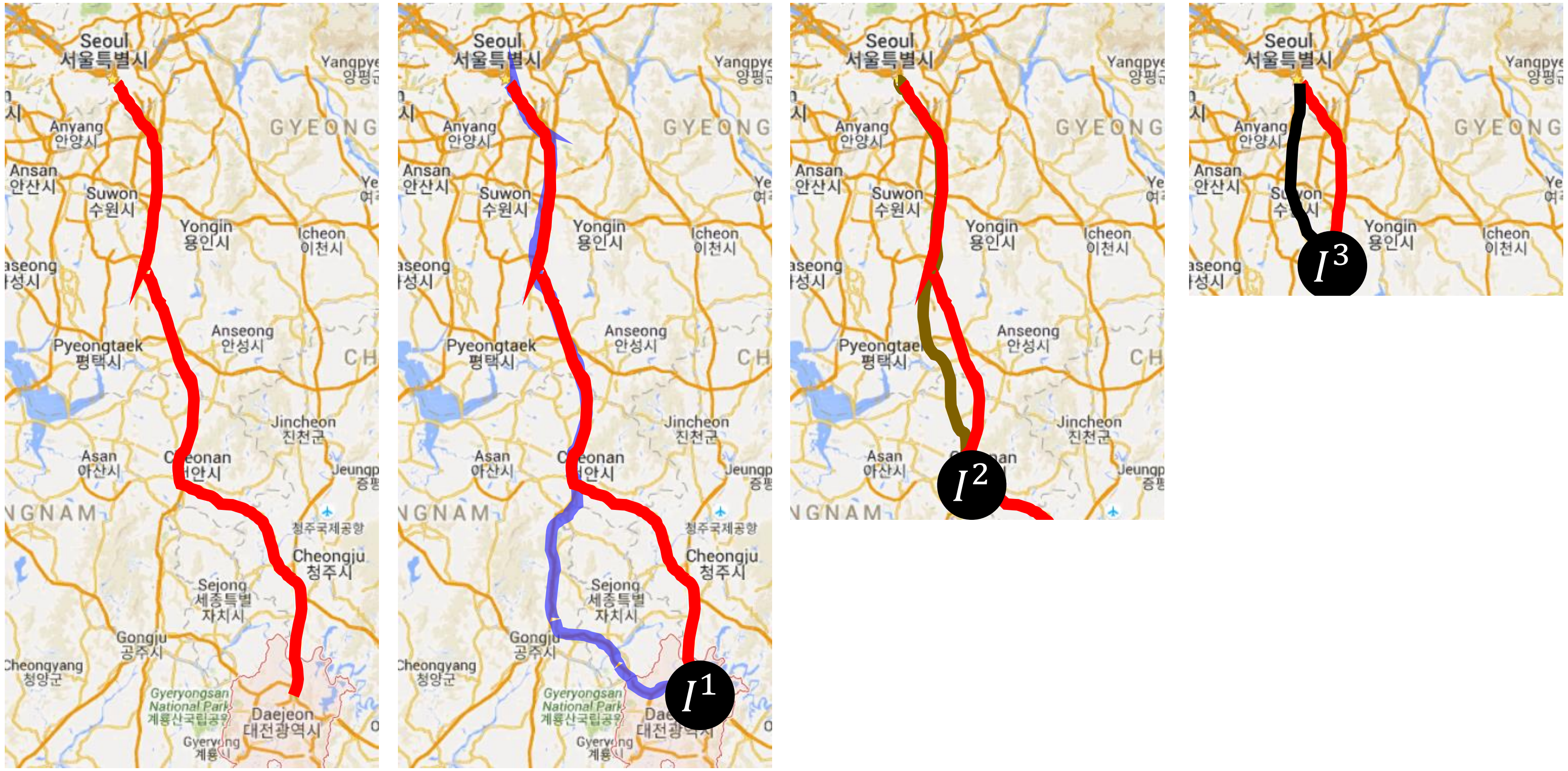
For infinite horizon multi-stage games with observed actions, s^* is a subgame perfect equilibrium if and only if for all i, t and I_i^t, we have

$$u_i\left(s_i^*, s_{-i}^* \middle| I_i^t\right) \geq u_i\left(s_i, s_{-i}^* \middle| I_i^t\right)$$

- $s_i\left(I_i^t\right) \neq s_i^*(I_i^t)$ (action taken at information set I_i^t)
- $s_i\left(I_i^{t+k}\right) = s_i^*\left(I_i^{t+k}\right)$ for all $k > 0$

- Informally, s is a subgame perfect equilibrium (SPE) if and only if no player i can gain by deviating from s in a single stage and conforming to s thereafter.
- The proof of one-stage deviation principle for finite horizon games relies on the idea that if a strategy satisfies the one stage deviation principle then that strategy cannot be improved upon by a finite number of deviations.
- One-stage deviation principle is essentially the principle of optimality of dynamic programming.

The one-stage deviation principles



- When there are three information sets, we can check the optimality of the trajectory by comparing three routes, each of which has one stage deviation.

Repeated games

Lessons from multi-stages game

- First, when players play a sequence of games over time, it will be to their benefit to use conditional strategies in later stage-games to support desirable behavior in early stage-games
- Second, the future that the players face must be important enough to support these dynamic incentives as self-enforcing
    - Using so-called reward-and-punishment strategies to sustain static non-best-response behavior is possible only if the palyers do not discout the futre too heavily

Repeated games

- A repeated game is simply a multistage game in which the same stage-game is being player at every stage
- A repeated game is a special case of multistage games
- A repeated game can be used to model interactions occurring more than once:
  - Firms in a marketplace
  - Political alliances
  - Friends (favor exchange...)
  - Workers (team production...)
- How to model such repeated conflicts and find an equilibrium strategy?

Examples

- **OPEC: Oil Prices**
  - 20$/bbl or less from 1930-1973 (2008 dollars)
  - 50$/bbl by 1976
  - 90$/bbl by 1982
  - 40$/bbl or less from 1986 to 2002
  - 100$/bbl by late 2008 ..

- Cooperative Behavior: Cartel is much like a repeated Prisoner's Dilemma
  - Need to easily observe each other's plays and react (quickly) to punish undesired behavior
  - Need patient players who value the long run (wars don't help!)
  - Need a stable set of players and some stationarity helps
    - constantly changing sources of production can hurt, but growing demand can help...

https://blogs.cornell.edu/info2040/2015/09/10/game-theory-and-oil-prices/

Setups

- Questions we'll need to answer before analyzing games:
  - what will agents be able to observe about others' play?
  - how much will agents be able to remember about what has happened?
  - what is an agent's utility for the whole game?

- Some of these questions will have different answers for
  - finitely-repeated games
  - infinitely-repeated games.

Finitely Repeated Games

Definition (Finitely repeated game)

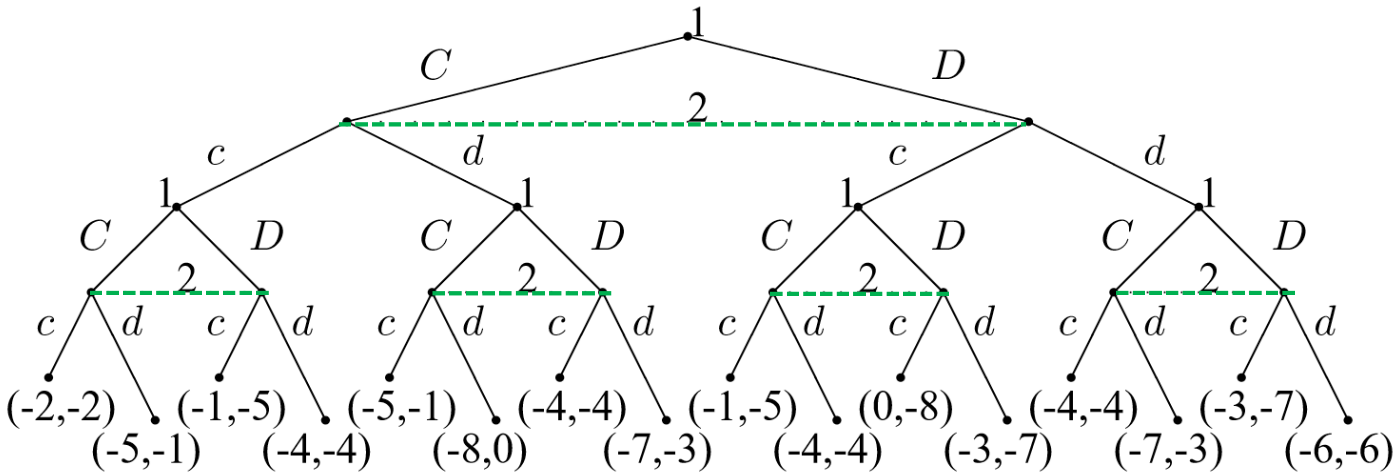
Given a stage game G, $G(T, \gamma)$ denotes the finitely repeated game in which the stage-game G is played T consecutive times, and γ is the common discount factor.

Finitely Repeated Games

- Everything is straightforward if we repeat a game a finite number of times
- We can write the whole thing as an extensive-form game with imperfect information
  - at each round players don't know what the others have done; afterwards they do
  - overall payoff function is additive: sum of payoffs in stage games

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$

	C	D
C	$-1, -1$	$-4, 0$
D	$0, -4$	$-3, -3$



Finitely Repeated Games: Example

	m	f	r
M	4, 4	-1, 5	0, 0
F	5, -1	**1, 1**	0, 0
R	0, 0	0, 0	**3, 3**

	m	f	r
M	4, 4	-1, 5	0, 0
F	5, -1	**1, 1**	0, 0
R	0, 0	0, 0	**3, 3**

- There are two pure-strategy Nash equilibria, (R, r) and (F, f)
  - (R, r) can serves as "carrot"
  - (F, f) can serves as "stick"

  can be used to discipline first-period behavior

  - Implies that for a high enough discount factor, we may be able to find SPE

Finitely Repeated Games: Example

	m	f	r
M	4, 4	-1, 5	0, 0
F	5, -1	**1, 1**	0, 0
R	0, 0	0, 0	**3, 3**

→

	m	f	r
M	4, 4	-1, 5	0, 0
F	5, -1	**1, 1**	0, 0
R	0, 0	0, 0	**3, 3**

- Convicne yourself that for a discount factor $\gamma \geq 1/2$, the following strategies constitutes SPE
  - Player 1:
    - stage 1: Play M
    - stage 2: play R if (M, m) was played in stage 1, and play F if anything but (M, m) was played in stage 1
  - Player 2:
    - stage 1: Play m
    - stage 2: play r if (M, m) was played in stage 1, and play f if anything but (M, m) was played in stage 1

	m	f	r
M	4, 4	-1, 5	0, 0
F	5, -1	**1, 1**	0, 0
R	0, 0	0, 0	**3, 3**

	m	f	r
M	4, 4	-1, 5	0, 0
F	5, -1	**1, 1**	0, 0
R	0, 0	0, 0	**3, 3**

Answer:

- This repeated game has nine outcomes at the first game
- The strategy for player 1 is

$$s_1^* = \left(s_1^1, s_1^2(h_1)\right)$$

where $s_1^1 = M$ and $s_1^2(h_1) = \begin{cases} R & \text{if } h_1 = (M, m) \\ F & \text{if } h_1 \neq (M, m) \end{cases}$

- To show this strategy is SPE, we need to show
  - In the second stage the players are clearly playing a Nash equilibrium regardless of the history of play
  - Players would not want to deviate from M in the first stage of the game:

$$u_1(M, s_2) = 4 + \gamma \underset{\text{carrot}}{3} \geq u_1(F, s_2) = 5 + \gamma \underset{\text{punishment}}{1} \text{ when } \gamma \geq \frac{1}{2}$$

Finitely Repeated Games: Example

- The difference between this example and the Prisoner-Revenge Game of the previous chapter is
  - The same game is repeated twice
- It is the multiplicity of equilibria in the stage-game that is giving the players the leverage to use conditional second-stage strategies of the reward-and-punishment kind.

Proposition

If the stage-game of a finitely repeated game has a unique Nash equilibrium, then the finitely repeated game has a **unique** subgame-perfect equilibrium.

- The proof is the same for multi-stage games
- Illustrative example:
  - Repetition of Prisoner's Dilemma game 500 times **(this is a finite number !!)**
  - What is PSE?

Finitely Repeated Games: Example

- The difference between this example and the Prisoner-Revenge Game of the previous chapter is
  - The same game is repeated twice
- It is the multiplicity of equilibria in the stage-game that is giving the players the leverage to use conditional second-stage strategies of the reward-and-punishment kind.

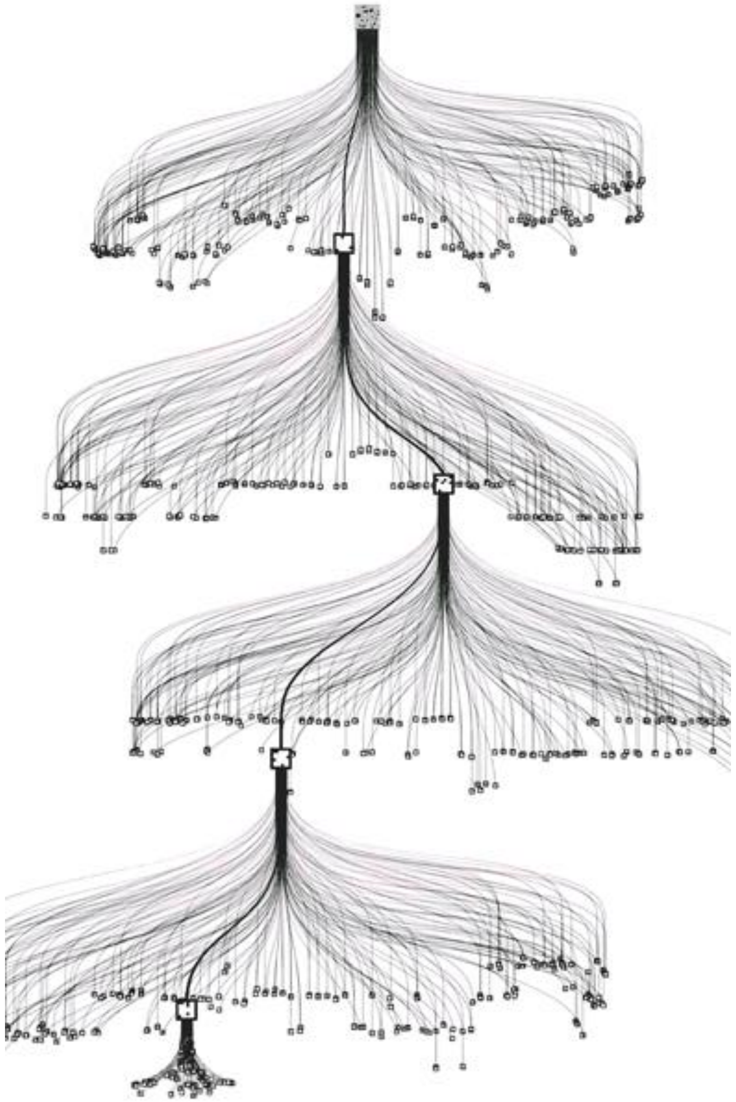
Proposition

If the stage-game of a finitely repeated game has a unique Nash equilibrium, then the finitely repeated game has a **unique** subgame-perfect equilibrium.

- The proof is the same for multi-stage games
- Illustrative example:
  - Repetition of Prisoner's Dilemma game 500 times **(this is a finite number !!)**
  - What is PSE?

- To give an incentives to cooperate, the players must be able to construct **reward-and-punishment continuation** equilibrium strategies
- These continuation strategies themselves must be equilibrium strategies and hence must rely on **multiple equilibria** in the continuation of the repeated game

Infinitely repeated Games: Motivations



Infinitely repeated games

Proposition

If the stage-game of a finitely repeated game has a unique Nash equilibrium, then the finitely repeated game has a **unique** subgame-perfect equilibrium.

- What would happen if we assume the game does not have a final period? That is, what would happen if players find that there is always a “long” future ahead of them
- This slight but critical modification allows player to chose an strategy that is not a static Nash equilibrium of the stage-game even when the stage-game has a unique Nash equilibrium
  - ➢ The players will have the freedom to support a wide range of behaviors that are not consistent with a static Nash equilibrium in the stage game.

Negative aspects:

- Consider an infinitely repeated game in extensive form:
  - an infinite tree!
- Thus, payoffs cannot be attached to terminal nodes, nor can they be defined as the sum of the payoffs in the stage games (which in general will be infinite).

Definition (Future discounted reward)

Given an infinite sequence of payoffs $u_i^1, u_i^2, \ldots$ for player i and discount factor γ with $0 \leq \gamma \leq 1$, i's future discounted reward is

$$u_i = \sum_{t=1}^{\infty} \gamma^{t-1} u_i^t.$$

- The interpretations for future discounted reward
  - the agent cares more about his well-being in the near term than in the long term
  - the agent cares about the future just as much as the present, but with probability $1 - \gamma$ the game will end in any given round.

Definition (Average reward 1)

Given an infinite sequence of payoffs $u_i^1, u_i^2, \ldots$ for player i, the average reward of i is

$$\bar{u}_i = (1 - \gamma) \sum_{t=1}^{\infty} \gamma^{t-1} u_i^t.$$

- The average payoff from a sequence is a normalization of the net present value
  - We are scaling down the net present value by a factor of $(1 - \gamma)$
  - This is convenient because of the following mathematical property

$$\bar{u}_i = (1 - \gamma)\{v + \gamma v + \gamma^2 v + \cdots\} = (1 - \gamma)\frac{v}{1 - \gamma} = v$$

➢ i.e., the average payoff of an infinite fixed sequence of some value v is itself equal to v

Definition (Average reward 2)

Given an infinite sequence of payoffs $u_i^1, u_i^2, \ldots$ for player i, the average reward of i is

$$\lim_{k \to \infty} \frac{1}{k} \sum_{t=1}^{k} u_i^t$$

Strategy space

- What is a pure strategy in an infinitely-repeated game?
  - Player’s strategy is a complete contingent play that specifies what the player will do in each information set

- For the extensive form representation of an infinitely repeated game
  - Game expand in “length” and “depth”
  - Infinite number of information set
  - Which requires an infinite number of actions!

- Representing a pure strategy using conventional way is impossible

Strategy space

- Every information set of each player is identified by a unique path of play or history that was played in the previous sequence
- For example, we play Prisoner's Dilemma four times, there will be 64 unique information sets, each of which corresponds to a unique path of play, or history, in the first three stages
  - There is one-to-one relationship between information sets and histories of play
- Let's define history more formally

Definition (history)

Consider an infinitely repeated game. Let H_t denote the set of all possible histories of length t, $h_t \in H_t$, and let $H = \cup_{t=1}^{\infty} H_t$ be the set of all possible histories. A pure strategy for player i is a mapping $s_i: H \to S_i$ that map histories into actions of the stage-game.

- Some famous strategies (for repeated PD):
  - Tit-for-tat: Start out cooperating. If the opponent defected, defect in the next round. Then go back to cooperation.
  - Grim trigger: Start out cooperating. If the opponent ever defects, defect forever.

Subgame-perfect equilibria for infinitely repeated games

Definition (Subgame-perfect equilibria for infinitely repeated games)

A profile of pure strategies $\left(s_1^*(\cdot), s_2^*(\cdot), \dots, s_n^*(\cdot)\right)$, $s_i: H \to S_i$ for all $i \in N$, is a subgame-perfect equilibrium if the restriction of $\left(s_1^*(\cdot), s_2^*(\cdot), \dots, s_n^*(\cdot)\right)$ is a Nash equilibrium in every subgame. That is, for any history of the game h_t, the continuation play $\left(s_1^*(\cdot), s_1^*(\cdot), \dots, s_n^*(\cdot)\right)$ is a Nash equilibrium.

- How could we check that a profile of strategies is a Nash equilibrium for any history h_t?
- Let's take a simple and familiar case

Proposition

Let $G(\gamma)$ be an infinitely repeated game, and let $(a_1^*, a_2^*, \dots, a_n^*)$ be a (static) Nash equilibrium strategy profile of the stage-game G. Define the repeated-game strategy for each player i to be the history-independent Nash strategy, $s_i^*(h) = a_i^*$ for all $h \in H$. Then, $\left(s_1^*(\cdot), s_1^*(\cdot), \dots, s_n^*(\cdot)\right)$ is a subgame-perfect equilibrium in the repeated game for any $\gamma < 1$

- That is, keep playing a static Nash equilibrium of a stage game is SPE for the whole game (infinitely repeated game), regardless of histories uncounted during the game.

Subgame-perfect equilibria for infinitely repeated games

Definition (Subgame-perfect equilibria for infinitely repeated games)

A profile of pure strategies $\left(s_1^*(\cdot), s_2^*(\cdot), \ldots, s_n^*(\cdot)\right)$, $s_i: H \rightarrow S_i$ for all $i \in N$, is a subgame-perfect equilibrium if the restriction of $\left(s_1^*(\cdot), s_2^*(\cdot), \ldots, s_n^*(\cdot)\right)$ is a Nash equilibrium in every subgame. That is, for any history of the game h_t, the continuation play $\left(s_1^*(\cdot), s_1^*(\cdot), \ldots, s_n^*(\cdot)\right)$ is a Nash equilibrium.

- How could we check that a profile of strategies is a Nash equilibrium for any history h_t?
- Let's take a simple and familiar case

Proposition

Let $G(\gamma)$ be an infinitely repeated game, and let $(a_1^*, a_2^*, \ldots, a_n^*)$ be a (static) Nash equilibrium strategy profile of the stage-game G. Define the repeated-game strategy for each player i to be the history-independent Nash strategy, $s_i^*(h) = a_i^*$ for all $h \in H$. Then, $\left(s_1^*(\cdot), s_1^*(\cdot), \ldots, s_n^*(\cdot)\right)$ is a subgame-perfect equilibrium in the repeated game for any $\gamma < 1$

	c	d
C	$4, 4$	$-1, 5$
D	$5, -1$	$1, 1$

	c	d
C	$4, 4$	$-1, 5$
D	$5, -1$	$1, 1$

	c	d
C	$4, 4$	$-1, 5$
D	$5, -1$	$1, 1$

	c	d
C	$4, 4$	$-1, 5$
D	$5, -1$	$1, 1$

…

Playing the static Nash equilibrium

→ subgame-perfect equilibrium

Subgame-perfect equilibria for infinitely repeated games

- The more interesting question that remains is whether or not we can support other types of behavior as part of a subgame-perfect equilibrium

	c	d
C	4, 4	−1, 5
D	5, −1	1, 1

	c	d
C	4, 4	−1, 5
D	5, −1	1, 1

	c	d
C	4, 4	−1, 5
D	5, −1	1, 1

	c	d
C	4, 4	−1, 5
D	5, −1	1, 1

…

→ subgame-perfect equilibrium

Playing the static Nash equilibrium strategy

	c	d
C	4, 4	−1, 5
D	5, −1	1, 1

	c	d
C	4, 4	−1, 5
D	5, −1	1, 1

	c	d
C	4, 4	−1, 5
D	5, −1	1, 1

	c	d
C	4, 4	−1, 5
D	5, −1	1, 1

…

→ subgame-perfect equilibrium?

Playing the static Nash equilibrium strategy

+

Playing non Nash equilibrium strategies

- The more interesting question that remains is whether or not we can support other types of behavior as part of a subgame-perfect equilibrium

	c	d
C	4, 4	−1, 5
D	5, −1	1, 1

	c	d
C	4, 4	−1, 5
D	5, −1	1, 1

	c	d
C	4, 4	−1, 5
D	5, −1	1, 1

	c	d
C	4, 4	−1, 5
D	5, −1	1, 1

… → subgame-perfect equilibrium

Playing the static Nash equilibrium strategy

	c	d
C	4, 4	−1, 5
D	5, −1	1, 1

	c	d
C	4, 4	−1, 5
D	5, −1	1, 1

	c	d
C	4, 4	−1, 5
D	5, −1	1, 1

	c	d
C	4, 4	−1, 5
D	5, −1	1, 1

… → subgame-perfect equilibrium?

Playing the static Nash equilibrium strategy

+

Playing non Nash equilibrium strategies

Answer is yes.

- By applying reward-and punishment strategies

SPE for infinitely repeated games: Examples

		Player 2	
		c	d
Player 1	C	$4, 4$	$-1, 5$
	D	$5, -1$	$1, 1$

- Consider the path of play in which the two players choose (C, c) in every period
  - The average payoffs are $(\bar{u}_1, \bar{u}_2) = (4, 4)$

- Is this path of play can be supported as subgame-perfect equilibrium?
  - No, because deviation from C to D will give a higher payoff 5 which is larger than 4
  - Deviation is profitable

- To make playing (C, c) continuously to be equilibrium, we need to find some way to "punish" deviation

SPE for infinitely repeated games: Examples

		Player 2	
		c	d
Player 1	C	$4, 4$	$-1, 5$
	D	$5, -1$	$1, 1$

- Consider the following strategy
  - **Player 1:**
    - stage 1: $s_1^1 = C$
    - for any stage $t > 1$: $s_1^t(h_{t-1}) = \begin{cases} C \text{ iff } h_{t-1} = \{(C,c),(C,c),\dots,(C,c)\} \\ D \text{ iff } h_{t-1} \neq \{(C,c),(C,c),\dots,(C,c)\} \end{cases}$
  - **Player 2:**
    - stage 1: $s_2^1 = c$
    - for any stage $t > 1$: $s_2^t(h_{t-1}) = \begin{cases} c \text{ iff } h_{t-1} = \{(C,c),(C,c),\dots,(C,c)\} \\ d \text{ iff } h_{t-1} \neq \{(C,c),(C,c),\dots,(C,c)\} \end{cases}$
- For any deviation from cooperation in the past, the players will revert to playing defect, and by the definition of the strategies they will stick to defect thereafter (forever)
- This strategy is referred to as grim-trigger strategies
- Is this strategy SPE?

SPE for infinitely repeated games: Examples

- To verify that the grim-trigger strategy pair is a subgame-perfect equilibrium we need to check that their is no profitable deviation in any subgame.

Proposition

In an infinitely repeated game $G(\gamma)$, a profile of strategies $s^* = (s_1^*, \dots, s_n^*)$ is subgame-perfect equilibrium if and only if there is no player i and no single history h_{t-1} for which player i would gain from deviation from $s_i(h_{t-1})$

- The definition may not seem helpful because there are still an infinite number of histories
- There is hope! Whenever we are trying to support one kind of behavior forever with the threat of resorting to another kind of behavior,
  - we have two "states" in which the players can be
    - On the equilibrium path
    - Off the equilibrium path
  - We need only to check that they would not want to deviate from each of these states

SPE for infinitely repeated games: Examples

- To verify that the grim-trigger strategy pair is a subgame-perfect equilibrium we need to check that their is no profitable deviation in any subgame.

- **On the equilibrium path**
  - Category of histories that are consecutive sequences of (C, c)
  - That is, $h_{t-1} = \{(C, c), (C, c), \dots, (C, c)\}$
  - If a player chooses to play C, his average payoff is

$$\bar{u}_i = 4 + \gamma 4 + \gamma^2 4 + \cdots = \underbrace{4}_{\text{Today's payoff}} + \underbrace{\frac{4\gamma}{1-\gamma}}_{\text{future's payoff}}$$

  - If a player chooses to play d, he gets 5 instead of 4 in the immediate stage of deviation, followed by his continuation payoff, which is infinite sequence of 1s

$$\bar{u}_i' = 5 + \gamma 1 + \gamma^2 1 + \cdots = 5 + \frac{1\gamma}{1-\gamma}$$

- We can conclude that a player will not deviate from the equilibrium path if

$$\bar{u}_i = 4 + \frac{4\gamma}{1-\gamma} > 5 + \frac{1\gamma}{1-\gamma} = \bar{u}_i'$$

$$\Rightarrow \gamma \geq \frac{1}{4}$$

SPE for infinitely repeated games: Examples

- To verify that the grim-trigger strategy pair is a subgame-perfect equilibrium we need to check that their is no profitable deviation in any subgame.

- **Off the equilibrium path**
  - Category of histories that are **not** consecutive sequences of (C, c)
  - That is, $h_{t-1} \neq \{(C, c), (C, c), \dots, (C, c)\}$

- In any subgame that is off the equilibrium path, the proposed strategies recommend that the players play (D, d)

- Then, we need to show that no player would want to choose $C\,(\text{or } c)$ instead of $D\,(\text{or } d)$ because
  - ✓ Given his belief that his opponent will play defect, such deviation from defect to cooperation will cause him a loss of -2 at the current and subsequent stages.

Player 2

Player 1	c	d
C	$4, 4$	$-1, 5$
D	$5, -1$	$1, 1$

Some notes

- Basic logic:
  - Play something with relatively high payoffs, and if anyone deviates
  - Punish by resorting to something that
    - Has lower payoffs (at least for that player)
    - and is credible: It is an equilibrium in the subgame

- We see the value of patience. If the players are sufficiently patient, so that the future carries a fair amount of weight in their preference, then there is a reward-and-punishment strategy that will allow them to cooperate forever.

- Recall the following property in multi-stage game (finite repetition):

  In multi-stage game (finite repletion) These continuation strategies themselves must be equilibrium strategies and hence must rely on **multiple equilibria** in the continuation of the repeated game

  - Where are the multiple equilibria coming from in the current example?
    - This is where the infinite repetition creates “magic” through bootstrapping
    - From the unique equilibrium of the stage-game, we get multiple equilibria of the repeated game

SPE for infinitely repeated games: Tacit Collusion example

- **Tacit collusion** occurs where firms undergo actions that are likely to minimize a response from another firm, e.g. avoiding the opportunity to price cut an opposition.
- Put another way, two firms agree to play a certain strategy without explicitly saying so.

Nash equilibrium examples: Cournot Duopoly

- Two identical firms, players 1 and 2, produce some good
- Firm i produces quantity q_i
- Cost for production is $c_i(q_i) = 10q_i$
- Price is given by $d = 100 - q = (100 - q_i - q_j)$
- The profit of company i given its opponent chooses quantity q_j is

$$u_i(q_i, q_j) = (100 - q_i - q_j)q_i - 10q_i = -q_i^2 + 90q_i - q_jq_i$$

- Nash-Cournot equilibrium
  - $q_1^*, q_2^* = 30;$
  - $p = 100 - (30 + 30) = 60$
  - $u_1^N = u_1(q_1^*, q_2^*) = u_2^N = 900$

- Monopoly
  - $q_1^* + q_2^* = 45;$
  - $p = 100 - (45) = 55$
  - $u_1(q_1^*, q_2^*) + u_2(q_1^*, q_2^*) = 2025;$

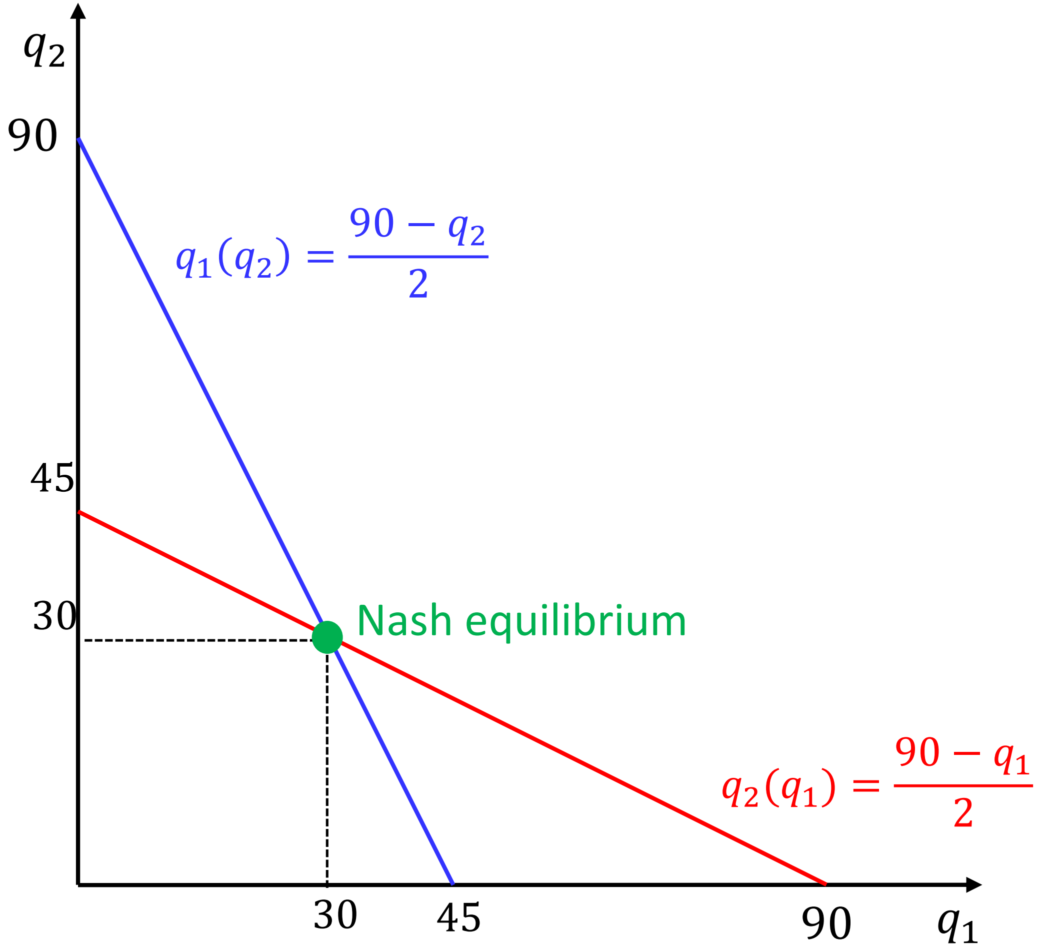
- If the two firms agree on producing $q_1^* + q_2^* = 45$, they can make more money in total
- Depending on how they set their production, q_1^* and q_2^*, various way of splitting profit become possible
  - $u_1(q_1^*, q_2^*) = u_2(q_1^*, q_2^*) = 1012.5$ when $q_1^* = q_2^* = 22.5$

Nash equilibrium examples: Cournot Duopoly

- In case there are two firms, we have two best-response equations:

$$q_1 = \frac{a - bq_2 - c_1}{2b} \quad \text{and} \quad q_2 = \frac{a - bq_1 - c_2}{2b}$$

$a = 100, b = 1, c_1 = c_2 = 10$



SPE for infinitely repeated games: Tacit Collusion example

- We now proceeds to see how reward-and-punishment strategies will allow our firms to coordinate on monopoly profit in self–enforcing subgame perfect equilibrium without the need to use binding contract

- First, we have to decide how the firms will cooperate to split the monopoly profit
  - $q_1^c + q_2^c = 45$, which results in $d = 55$
  - Let's assume $q_1^c = 22, q_2^c = 23$
  - $u_1^c = u_1(q_1^c, q_2^c) = (100 - 45)q_1^c - 10q_1^c = 990$
  - $u_2^c = u_2(q_1^c, q_2^c) = (100 - 45)q_2^c - 10q_2^c = 1035$

- Second, Specifies firms' strategies, following the logic of the infinitely repeated PD game
  - **Firm 1:**
    - stage 1: $q_1^1 = q_1^c$
    - for any stage $t > 1$: $q_1^t(h_{t-1}) = \begin{cases} q_1^c \text{ iff } h_{t-1} = \{(q_1^c, q_2^c), (q_1^c, q_2^c), \dots, (q_1^c, q_2^c)\} \\ q_1^N \text{ iff } h_{t-1} \neq \{(q_1^c, q_2^c), (q_1^c, q_2^c), \dots, (q_1^c, q_2^c)\} \end{cases}$
  - **Firm 2:**
    - stage 1: $q_2^1 = q_2^c$
    - for any stage $t > 1$: $q_2^t(h_{t-1}) = \begin{cases} q_2^c \text{ iff } h_{t-1} = \{(q_1^c, q_2^c), (q_1^c, q_2^c), \dots, (q_1^c, q_2^c)\} \\ q_2^N \text{ iff } h_{t-1} \neq \{(q_1^c, q_2^c), (q_1^c, q_2^c), \dots, (q_1^c, q_2^c)\} \end{cases}$

  Where (q_1^N, q_2^N) is Nash equilibrium

- Third, we need to check that no firm wants to deviate from the proposed strategies

$$(q_1^c = 22, q_2^c = 23\ )$$

- **On the equilibrium path**
  - Category of histories that are consecutive sequences of (q_1^c, q_2^c)
  - That is, $h_{t-1} = \{(q_1^c, q_2^c), (q_1^c, q_2^c), \dots, (q_1^c, q_2^c)\}$
  - If a player i chooses to play q_i^c, his average payoff is

$$u_i^c + \gamma u_i^c + \gamma^2 u_i^c + \cdots = u_i^c + \frac{\gamma u_i^c}{1-\gamma}$$

  - If a player chooses to play $q_i^d = BR\left(q_j^c\right)$, he gets the following average payoff

$$\underbrace{u_i^d}_{\text{Immediate increase}} + \underbrace{\gamma u_i^N + \gamma^2 u_i^N + \cdots}_{\text{Continuous reduced payoff}} = u_i^d + \frac{\gamma u_i^N}{1-\gamma}$$

- We can conclude that a player will not deviate from the equilibrium path if

$$u_i^c + \frac{\gamma u_i^c}{1-\gamma} > u_i^d + \frac{\gamma u_i^N}{1-\gamma} \Rightarrow \gamma \geq \frac{u_i^d - u_i^c}{u_i^d - u_i^N}$$

Q. How do they do this? REVIEW THIS

- For example, if player 1 deviate from $q_1^c = 22.5$ to $q_1^d = BR(q_2^c = 23) = \frac{90-23}{2} = 33.5.$
- $u_1^d = u_1\left(u_1^d = 33.5, q_2^c\right) = (100 - 33.5 - 23)33.5 - 10 \times 33.5 = 1125.25$
- $\gamma_1 \geq \frac{u_i^d - u_i^c}{u_i^d - u_i^N} = \frac{1122.5-990}{1122.5-900} = 0.595$

SPE for infinitely repeated games: Examples

- To verify that the grim-trigger strategy pair is a subgame-perfect equilibrium we need to check that their is no profitable deviation in any subgame.

- **Off the equilibrium path**
  - Category of histories that are **not** consecutive sequences of(q_1^c, q_2^c)
  - That is, $h_{t-1} \neq \{(q_1^c, q_2^c), (q_1^c, q_2^c), \dots, (q_1^c, q_2^c)\}$

- In any subgame that is off the equilibrium path, the proposed strategies recommend that the players play $(q_1^N, q_1^N) = (30{,}30)$

- Then, we need to show that no player would want to choose $q_i \neq q_i^N$ instead of q_i^N
- because
  - ✓ Given his belief that his opponent will play q_j^N, such deviation will only decrease his payoff due to the definition of Nash equilibrium

Folk Theorem

- Consider any n-player game $G = (N, A, u)$ and any payoff vector $r = (r_1, r_2, \dots, r_n)$.
- Let $v_i = \min_{s_{-i} \in S_{-i}} \max_{s_i \in S_i} u_i(s_{-i}, s_i)$
  - i's minmax value is the amount of utility i can get when $-i$ play a minmax strategy against him

Definition (Enforceable payoff)

A payoff profile r is **enforceable** if $r_i \geq v_i$.

Definition (Feasible payoff)

A payoff profile r is **feasible** if there exist rational, non-negative values α_a such that for all i, we can express r_i as $\sum_{a \in A} \alpha_a u_i(a)$ with $\sum_{a \in A} \alpha_a = 1$.

- Feasible: a convex, rational combination of the outcomes in G.

Motivation for Folk Theorem

- With an infinite number of equilibria, what can we say about Nash equilibria?
    - Nash's theorem only applies to finite games
    - we won't be able to construct an induced normal form and then appeal to Nash's theorem to say that an equilibrium exists
- Furthermore, with an infinite number of strategies, there could be an infinite number of pure-strategy equilibria!
- It turns out we can characterize a set of payoffs that are achievable under equilibrium, without having to enumerate the equilibria.

Theorem (Folk Theorem)

Consider any n-player game G and any payoff vector $(r_1, r_2, \ldots, r_n)$.

1. If r is the payoff in any Nash equilibrium of the infinitely repeated G with average rewards, then for each player i, r_i is enforceable.

2. If r is both feasible and enforceable, then r is the payoff in some Nash equilibrium of the infinitely repeated G with average rewards

Payoff in Nash ⇒ enforceable

- Suppose r is not enforceable, i.e. $r_i < v_i$ for some i.
- Then consider an alternative strategy for i: playing $BR(s_{-i}(h))$, where $s_{-i}(h)$ is the equilibrium strategy of other players given the current history h and $BR(s_{-i}(h))$ is a function that returns a best response for i to a given strategy profile s_{-i} in the (unrepeated) stage game G
- By definition of a minmax strategy, player i will receive a payoff of at least v_i in every stage game if he adopts $BR(s_{-i}(h))$
- So i's average reward is also at least v_i.
- Thus, if $r_i < v_i$ then s cannot be a Nash equilibrium.

(strategy s_i can be deviated to $BR(s_{-i}(h))$ in order to get a higher pay off v_i than r_i)

Feasible and Enforceable ⇒ Nash

- Since r is a feasible payoff profile, we can write it as $r_i = \sum_{a \in A} \left(\frac{\beta_a}{\gamma}\right) u_i(a)$ where β_a and γ are non-negative integers with $\gamma = \sum_{a \in A} \beta_a$
  - Recall that if feasible $r_i = \sum_{a \in A} \alpha_a u_i(a)$ with $\sum_{a \in A} \alpha_a = 1$
- We're going to construct a strategy profile that will cycle through all outcomes $a \in A$ of G with cycles of length γ, each cycle repeating action a exactly β_a times.
- Let (a^t) be such a sequence of outcomes. Let's define a strategy s_i of player i to be a trigger version of playing(a^t):
  - if nobody deviates, then s_i plays a_i^t in period t.
  - However, if there was a period t' in which some player $j \neq i$ deviated, then s_i will play minmax strategy to minimize u_j such that
    - $v_j = \min_{s_{-j} \in S_{-j}} \max_{s_j \in S_j} u_j(s_{-i}, s_i)$
    - Player i will player $\left(s_{-j}\right)_i$

- **First**, observe that if everybody plays according to s_i, then, by construction, player i receives average payoff of r_i (look at averages over periods of length γ).

$$u_i(a^1), u_i(a^2), u_i(a^2), u_i(a^3), u_i(a^3), u_i(a^3) \rightarrow r_i = \frac{1}{6} u_i(a^1) + \frac{2}{6} u_i(a^2) + \frac{3}{6} u_i(a^3),$$

Feasible and Enforceable ⇒ Nash

- Since r is a feasible payoff profile, we can write it as $r_i = \sum_{a \in A} \left(\frac{\beta_a}{\gamma}\right) u_i(a)$ where β_a and γ are non-negative integers with $\gamma = \sum_{a \in A} \beta_a$
  - Recall that if feasible $r_i = \sum_{a \in A} \alpha_a u_i(a)$ with $\sum_{a \in A} \alpha_a = 1$
- We're going to construct a strategy profile that will cycle through all outcomes $a \in A$ of G with cycles of length γ, each cycle repeating action a exactly β_a times.
- Let (a^t) be such a sequence of outcomes. Let's define a strategy s_i of player i to be a trigger version of playing(a^t):
  - if nobody deviates, then s_i plays a_i^t in period t.
  - However, if there was a period t' in which some player $j \neq i$ deviated, then s_i will play minmax strategy to minimize u_j such that
    - $v_j = \min_{s_{-j} \in S_{-j}} \max_{s_j \in S_j} u_j(s_{-i}, s_i)$
    - Player i will player $\left(s_{-j}\right)_i$

- **Second**, this strategy profile is a Nash equilibrium.
  - Suppose everybody plays according to s_i, and player j deviates at some point.
  - Then, forever after, player j will receive his minmax payoff $v_j \leq r_j$, rendering the deviation unprofitable.