Lecture 8: 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.

SAMSUNG VS

- 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

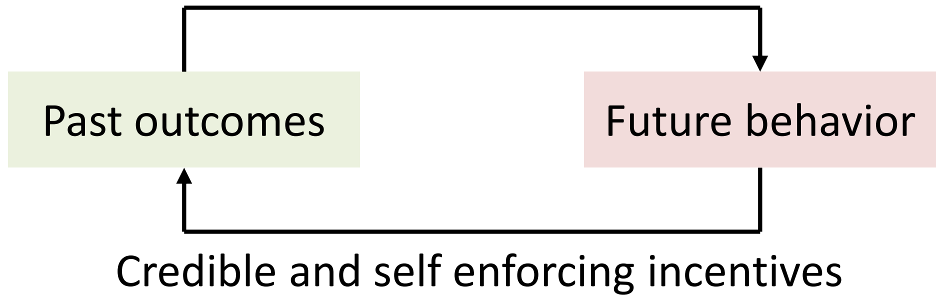
- Two questions we will answer from this lecture
- First, if the players are rational and forward looking, should they not view this sequence of games as one grand game?
- Second, if they do view these as on grand game, should we expect that their actions in the later stages will depends on the outcomes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



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

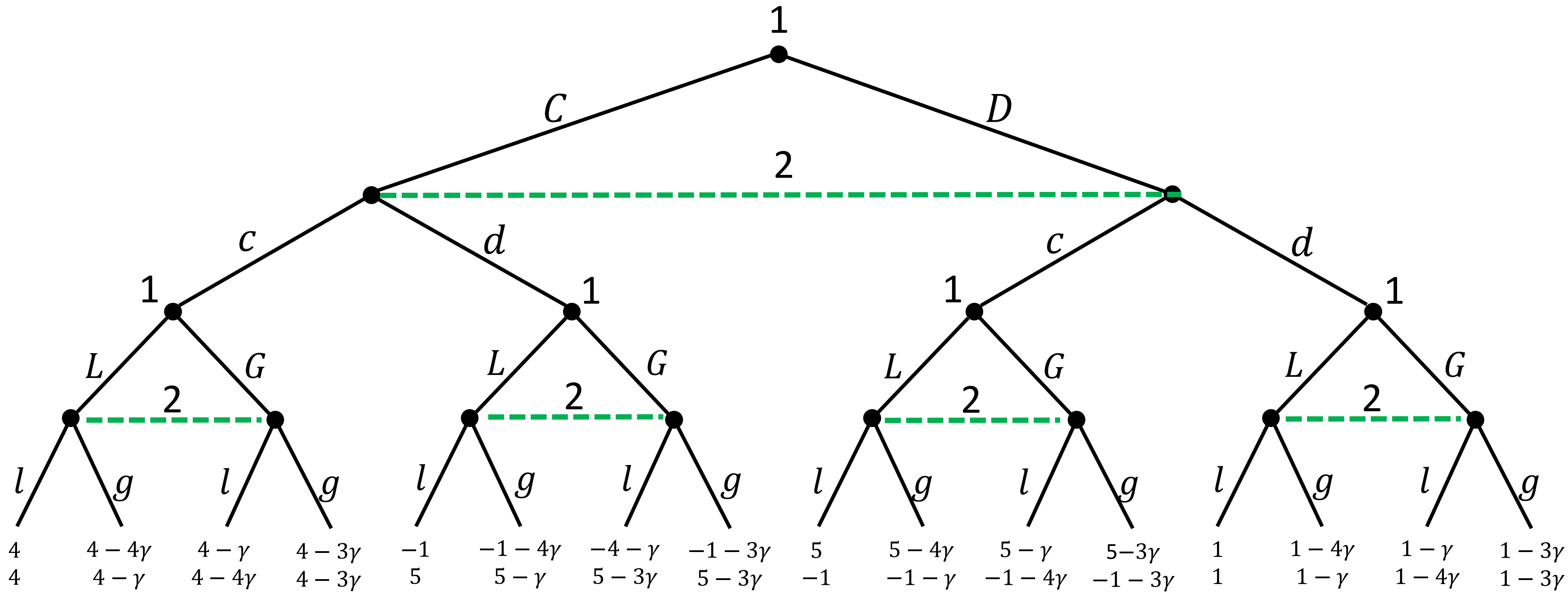
	c	d
C	$4, 4$	$-1, 5$
D	$5, -1$	$1, 1$

$t = 1$

Revenge game

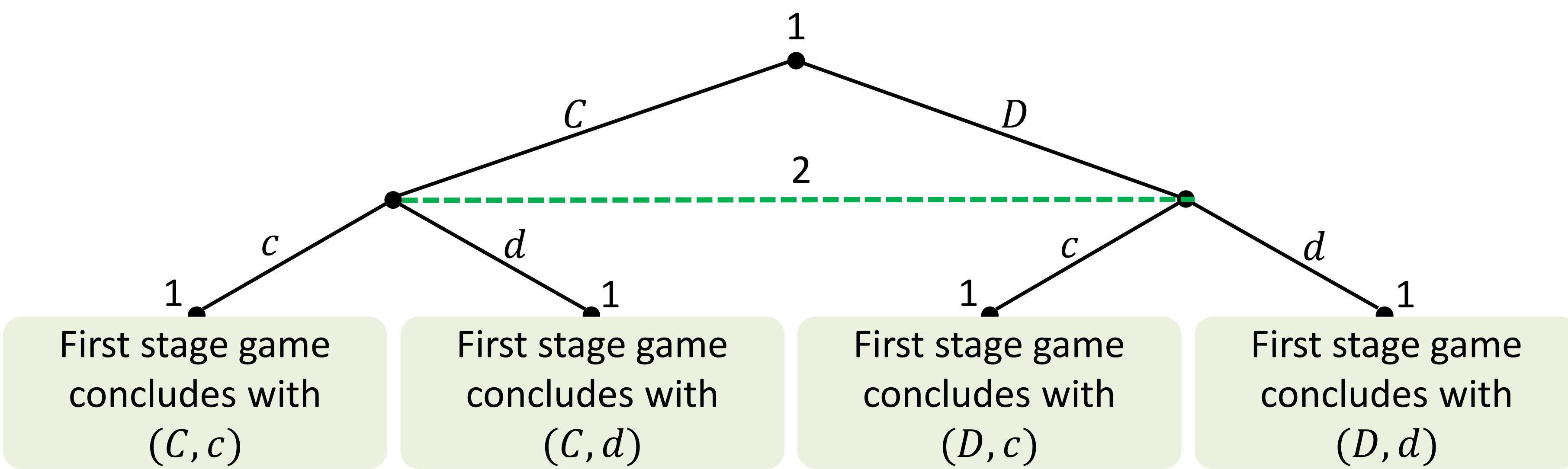
	l	g
L	$0, 0$	$-4, -1$
G	$-1, -4$	$-3, -3$

$t = 2$



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 $1, 2, \dots, t-1$



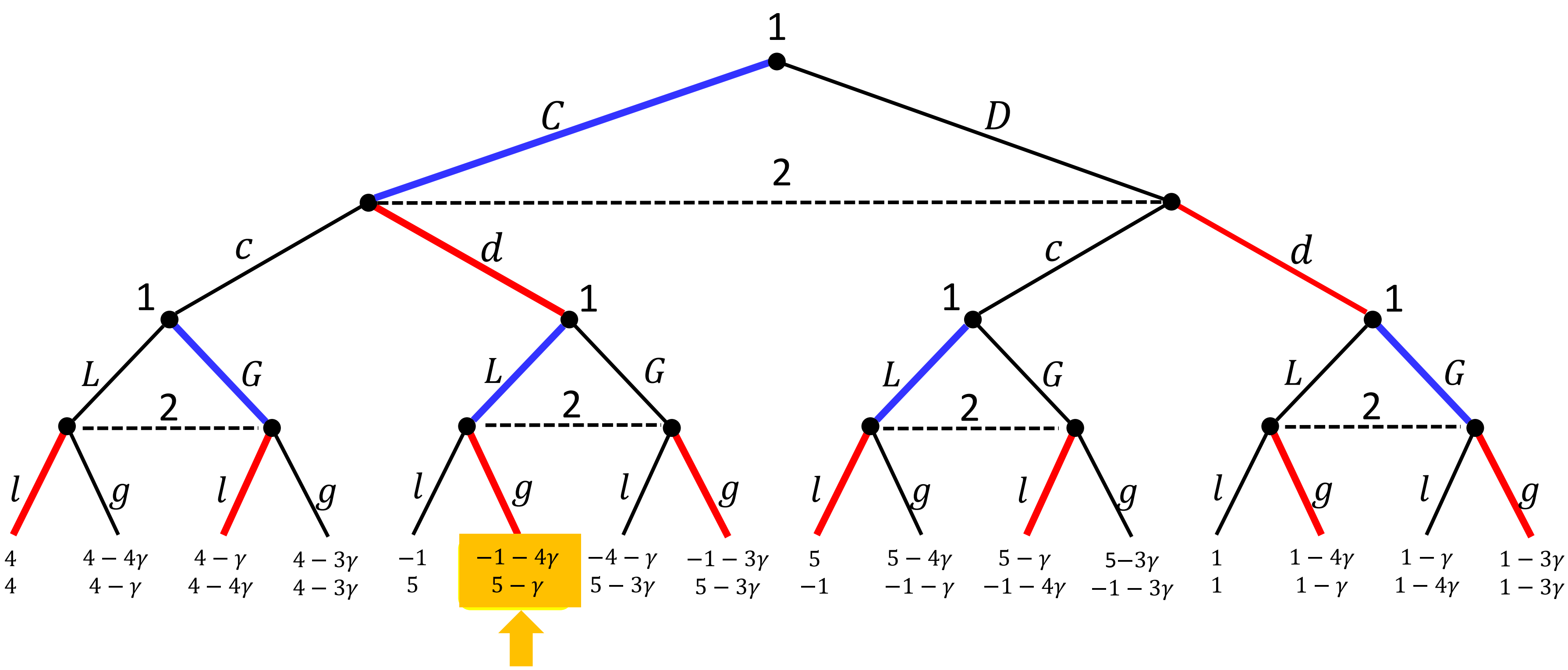
- Each player knows exactly how the fist-stage game concluded
- Define a strategy player i in the multistage Prisoner-Revenge game as a quintuple:

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

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 $t = 1$

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

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



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 $s^{1*}, s^{2*}, \ldots, s^{T*}$.

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	$-1, -4$	$\boxed{-3, -3}$

$t = 2$

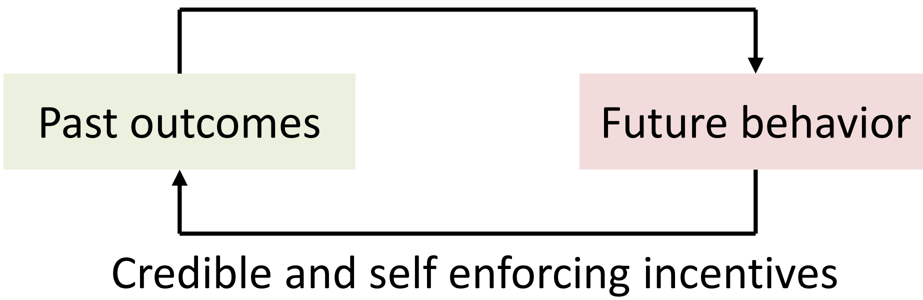
$$s_1 = (s_1^1, s_1^2(Cc), s_1^2(Cd), s_1^2(Dc), s_1^2(Dd)) = (D, L, L, L, L)$$
$$s_2 = (s_2^1, s_2^2(Cc), s_2^2(Cd), s_2^2(Dc), s_2^2(Dd)) = (d, l, l, l, l)$$

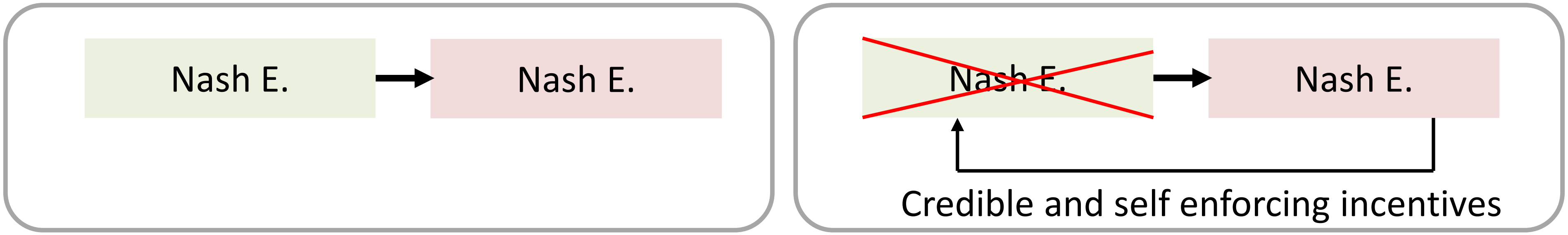
$$s_1 = (s_1^1, s_1^2(Cc), s_1^2(Cd), s_1^2(Dc), s_1^2(Dd)) = (D, G, G, G, G)$$
$$s_2 = (s_2^1, s_2^2(Cc), s_2^2(Cd), s_2^2(Dc), s_2^2(Dd)) = (d, g, g, g, g)$$

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	$4, 4$	$-1, 5$
D	$5, -1$	$1, 1$

Revenge game

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

- 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	$4,4$	$-1,5$
D	$5,-1$	$1,1$

Revenge game

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

- 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

$$s_1 = (s_1^1, s_1^2(Cc), s_1^2(Cd), s_1^2(Dc), s_1^2(Dd)) = (C, L, G, G, G)$$

incentive punishment

Player 2

$$s_2 = (s_{12}^1, s_2^2(Cc), s_2^2(Cd), s_2^2(Dc), s_2^2(Dd)) = (c, l, g, g, g)$$

incentive 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!

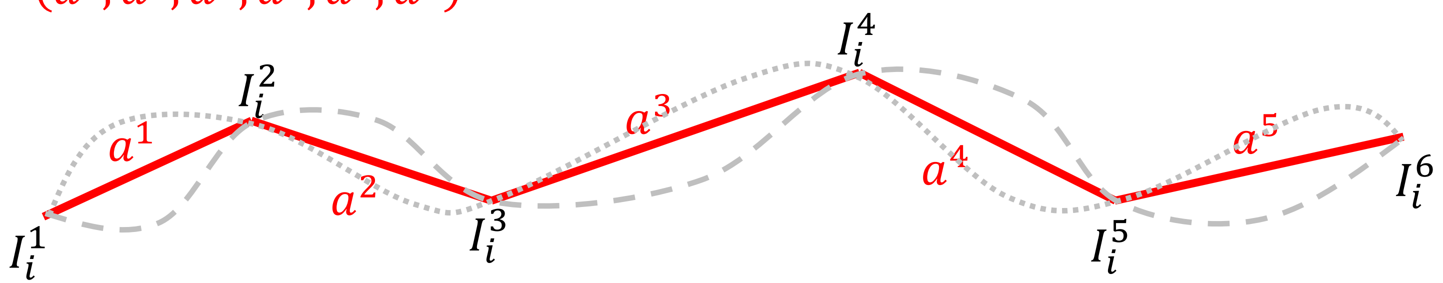
- Do we need to check every possible combination to check the optimality of a strategy?

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

$$s_i = (a^1, a^2, a^3, a^4, a^5, a^6)$$



- 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!

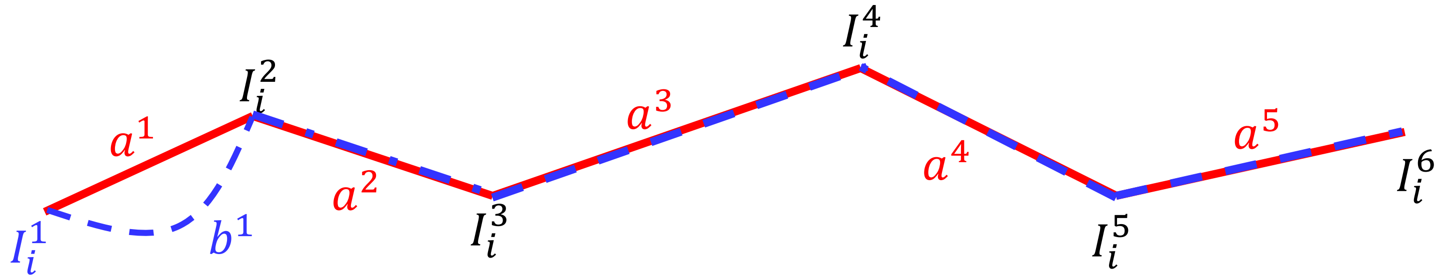
The one-stage deviation principles

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.

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



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

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

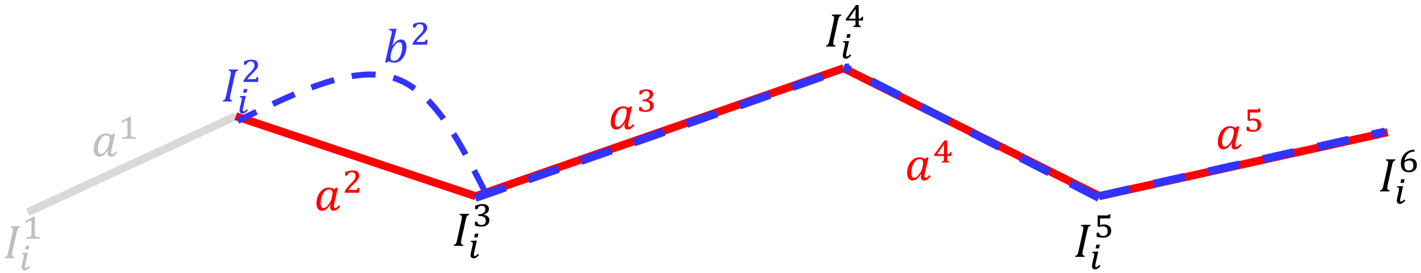
The one-stage deviation principles

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 : $a^2 \to b^2$

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

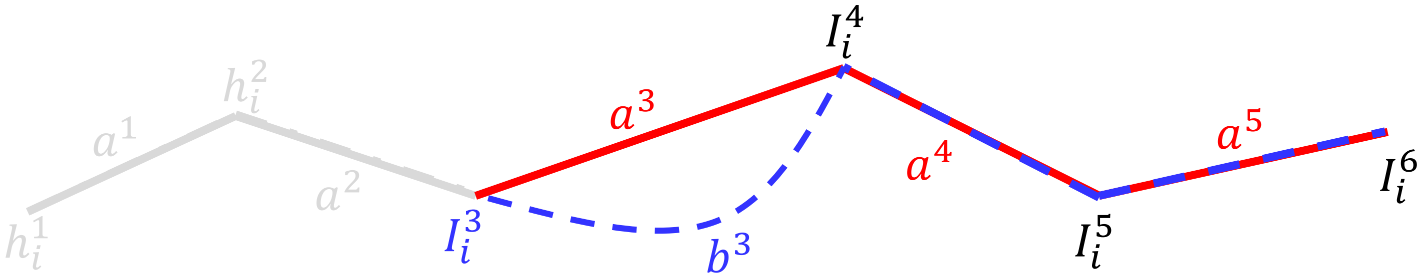
The one-stage deviation principles

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.

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



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

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

The one-stage deviation principles

Theorem (One-stage deviation principle)

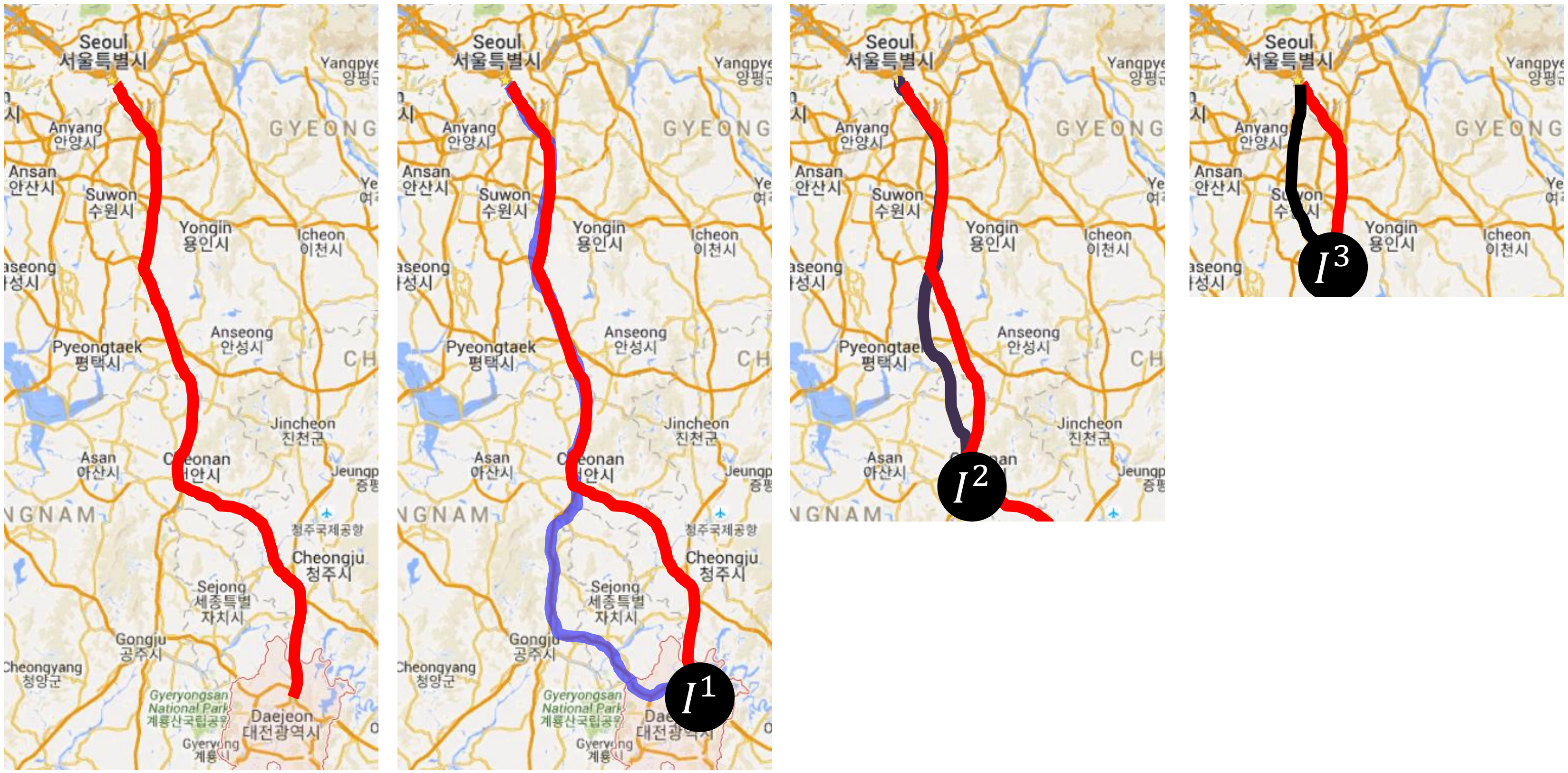
For infinite horizon multi-stage games with observed actions, s^* is a subgame perfect equilibrium if and only if for all i, t and I_i^t, we have

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

- $s_i\left(I_i^t\right) \neq s_i^*(I_i^t)$ (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.