Topic 4: Dynamic Games

Outcomes & Objectives

- Be proficient in solving Stackelberg (leader-follower) games.
    - Model real-world interactions with leader-follower dynamics in various applications.
    - Develop a solution concept called Stackelberg equilibrium using principles of backward induction to solve Stackelberg games.

- Be proficient with extensive-form games.
    - Model perfectly observable multi-stage interactions in various examples and real-world applications.
    - Develop a solution concept called subgame perfect equilibrium via extending the concept of Stackelberg equilibrium to multi-stage games.
    - Develop a solution concept (inspired from subgame perfect equilibrium) to solve Bayesian games in extensive-form.

- Be proficient in solving repeated games.
    - Investigate the effects of long-term strategic interactions, as opposed to short-term interactions.
    - Develop a solution concept which accounts for temporal dynamics (e.g. discounting behavior).

Revising Nash Equilibrium...

- Consider a two-player game where Alice and Bob choose mixed strategies $(\boldsymbol{\sigma}_a, \boldsymbol{\sigma}_b) \in \Delta(\mathcal{C}_A) \times \Delta(\mathcal{C}_B)$ at equilibrium.

- MSNE: $\boldsymbol{\sigma}_a = \underset{\boldsymbol{x} \in \Delta(\mathcal{C}_A)}{\text{arg max}} \ \boldsymbol{\sigma}_b^T U_A \boldsymbol{x}$ and $\boldsymbol{\sigma}_b = \underset{\boldsymbol{y} \in \Delta(\mathcal{C}_B)}{\text{arg max}} \ \boldsymbol{y}^T U_B \boldsymbol{\sigma}_a$.

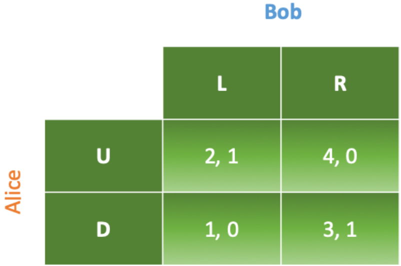
- This means that Alice and Bob choose their strategies simultaneously.

Will players choose $(\boldsymbol{\sigma}_a, \boldsymbol{\sigma}_b)$ at equilibrium,
if they choose their strategies in a leader-follower setting?

Isn't there a first mover advantage?

Revising Nash Equilibrium... (cont...)

Consider the following game:



Bob

		L	R
Alice	U	2, 1	4, 0
	D	1, 0	3, 1

- PSNE: (U, L)
- Now, say Alice leads the game via announcing a strategy.
- However, such an announcement should be made via taking Bob's response into account.

$$\begin{aligned} BR_B(U) = L \quad &\Rightarrow \quad U_A = 2 \\ BR_B(D) = R \quad &\Rightarrow \quad U_A = 3 \end{aligned}$$

The equilibrium in this leader-follower game is (D, R)!

Note that the outcome is more favorable to Alice!

Real-World Leader-Follower Interactions

- *Airport Security:* Cops are stationed strategically, and adversaries choose their attack strategy.
- *Markets:* Big firms announce their strategies, after which new startups arise in the market.
- *Recommender Systems:* Users make decisions after a recommendation is presented to them.

Heinrich Von Stackelberg (1934)

Equilibrium in Stackelberg Games

Consider a two-player game where Alice is the leader, and Bob is the follower.

- Assume the utility matrices at Alice and Bob are U_A and U_B respectively.
- Let Alice choose $\boldsymbol{x}_a \in \Delta(\mathcal{C}_A)$, and Bob choose $\boldsymbol{x}_b \in \Delta(\mathcal{C}_B)$.

Idea: Use ***backward induction***

- Maximize Alice's expected utility, while accounting for Bob's response in the next stage.

Definition

A Stackelberg equilibrium is a mixed strategy $(\boldsymbol{\sigma}_a, \boldsymbol{\sigma}_b) \in \Delta(\mathcal{C}_A) \times \Delta(\mathcal{C}_B)$ such that

$$\boldsymbol{\sigma}_a = \underset{\boldsymbol{x} \in \Delta(\mathcal{C}_A)}{\text{arg max}} \ \boldsymbol{y}^*(\boldsymbol{x})^T U_A \boldsymbol{x} \text{ and } \boldsymbol{\sigma}_b = \boldsymbol{y}^*(\boldsymbol{\sigma}_a),$$

where $\boldsymbol{y}^*(\boldsymbol{x}) = \underset{\boldsymbol{y} \in \Delta(\mathcal{C}_B)}{\text{arg max}} \ \boldsymbol{y}^T U_B \boldsymbol{x}$ is Bob's best response to Alice's strategy $\boldsymbol{x} \in \Delta(\mathcal{C}_A)$.

Theorem

Every two-player finite game admits a Stackelberg equilibrium.

Stackelberg Competition in Markets

- Consider two firms with same product, with Firm 1 making the first move.
- Firm-i produces $s_i \geq 0$ quantity at a cost c_i per item.
- Unit Price: $p(s_1 + s_2) = a - b(s_1 + s_2)$
- Utility: $U_i(s_1, s_2) = p(s_1, s_2)s_i - c_i s_i$

Firm 2's Best Response: $\max_{s_2 \geq 0} \left[a - b(s_1 + s_2)\right] s_2 - c_2 s_2$
Differentiate w.r.t. s_2 and equate it to zero:

$$a - bs_1 - 2bs_2 - c_2 = 0.$$

In other words, $s_2^*(s_1) = \dfrac{1}{2b} \left[a - c_2 - bs_1\right]_+$

Firm 1's Commitment: $\max_{s_1 \geq 0} \Big[a - b(s_1 + s_2^*(s_1))\Big] s_1 - c_1 s_1$

- If $s_1 > \frac{a - c_2}{b}$, then $s_2^* = 0$.
  Differentiate w.r.t. s_2 and equate it to zero:

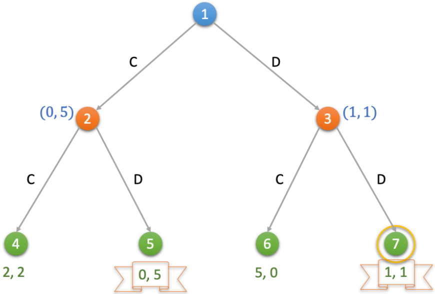
$$a - 2bs_1 - c_1 = 0. \ \Rightarrow \ s_1^* = \left[\frac{a - c_1}{2b}\right]_+.$$

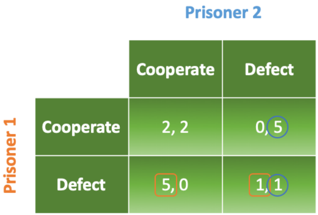
- Else, $s_2^*(s_1) = \dfrac{1}{2b} \left[a - c_2 - bs_1\right]$.
  Differentiate w.r.t. s_1 and equate it to zero:

$$a - 2bs_1 + \frac{b}{2}s_1 - \frac{1}{2}\left[a - c_2 - bs_1\right] - c_1 = 0. \ \Rightarrow \ s_1^* = \left[\frac{a - 2c_1 + c_2}{2b}\right]_+.$$

Stackelberg Prisoner's Dilemma

- Two prisoners, Alice and Bob, are interrogated sequentially.
- Alice leads and decides whether to cooperate/defect, and Bob picks a choice having seen Alice's choice, as shown below.





Prisoner 2

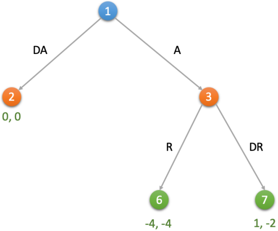
		Cooperate	Defect
Prisoner 1	Cooperate	2, 2	0, 5
	Defect	5, 0	1, 1

Alleviating First Mover's Advantage...

Can we alleviate first mover's advantage?

Follower needs to commit on their strategies, even if they do not make sense rationally!

Example: What if, in the following game, Player 2 declares to choose R if Player 1 chooses A?

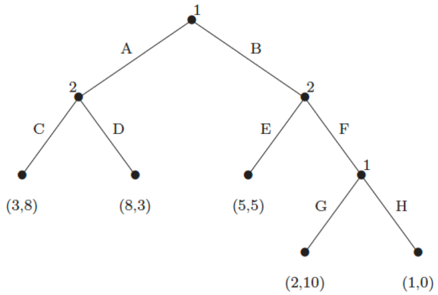


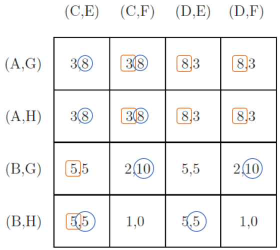
- Commitment must be observable and irreversable!
- Many real-world examples:
  - William, the Conqueror, ordered his soldiers to burn thier ships after landing to prevent men from retreating!
  - Hernn Corts sank his ships after landing in Mexico for the same reason.

> *The power to constrain an adversary depends on the power to bind oneself*
>
> – Thomas Schelling

Solving Perfect Extensive-Form Games...

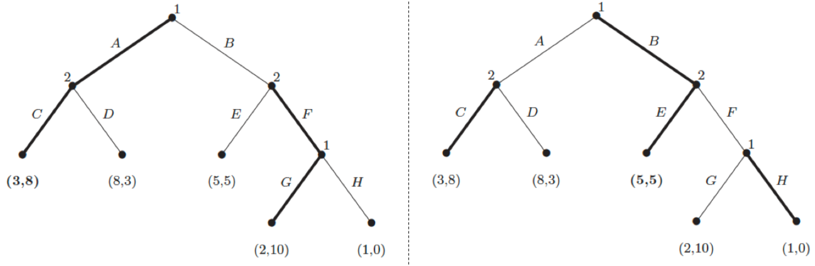
Consider the following extensive-form game:





	(C,E)	(C,F)	(D,E)	(D,F)
(A,G)	3,⑧	$\boxed{3}$,⑧	$\boxed{8}$,3	$\boxed{8}$,3
(A,H)	3,⑧	$\boxed{3}$,⑧	$\boxed{8}$,3	$\boxed{8}$,3
(B,G)	$\boxed{5}$,5	2,⑩	5,5	2,⑩
(B,H)	$\boxed{5}$,⑤	1,0	5,⑤	1,0

Not all Nash equilibria makes sense in extensive-form!



Subgame Perfect Equilibrium

Definition

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

Definition

The ***subgame-perfect equilibrium*** (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'.

Claim

Every subgame perfect equilibrium is a Nash equilibrium.

Claim

Every finite extensive-form game has at least one subgame perfect equilibrium.

This is essentially called the ***principle of optimality*** in dynamic programming.

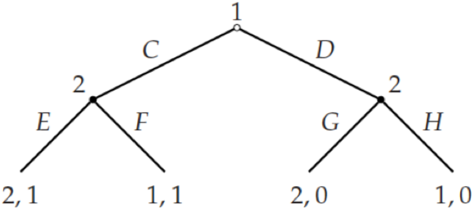
SPE and Backward Induction

The underlying philosophy of SPE is:

> Identify the equilibria in the "bottom-most" subgame trees, and assumes that those equilibria will be played as one backs up sequentially to evaluate larger trees.

This is **backward induction**[1].

Exercise:



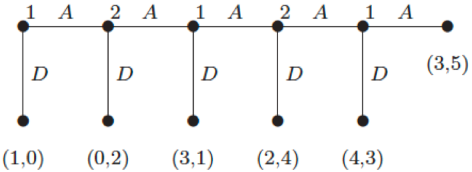
---

[1] Backward induction is also called *minimax algorithm* in two-player zero-sum games.

Backward Induction: Concerns and Challenges

SPE and Backward induction has their own share of concerns:

- Computationally infeasible in large extensive games.
    - Example: Chess ($\sim 10^{150}$ nodes.)
    - Needs gradual development of tree using a *heuristic* search algorithm!
    - Examples: ***Alpha-Beta Pruning***, ***Monte-Carlo Tree Search***
- Consider the following *centipede* game:



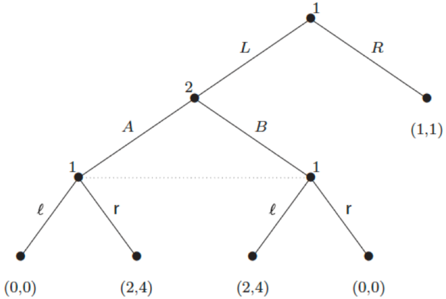
SPE ⇒ Players always choose to go down!

But, this is indeed a paradox at the second player!!!

Solving Imperfect Extensive Games...

What if, we have information sets in the game?

Consider the following example:



Note that the subgame at Player 2's node is the smallest subgame!

- Idea: Reduce this subgame into its strategic game and continue – *Inefficient*!
- Can we operate directly on the extensive-form representation?

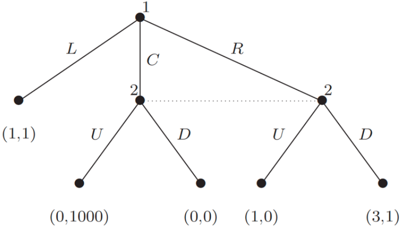
Is Subgame Perfect Equilibrium Suitable?

What do we mean by a subgame in imperfect extensive games?

What if, we define a subforest (a collection of subgames) at each information set?

Example:

- Pure strategies: $P_1 \Rightarrow \{L, C, R\}$, $P_2 \Rightarrow \{U, D\}$
- PSNE: (L, U), (R, D)
- Can either of these equilibria be considered *subgame perfect*?
  - Left subtree – U dominates D
  - Right subtree – D dominates U
- But, R dominates C at Player 1
- So, (R, D) is subgame perfect!



Lesson: *The requirement that we need best responses in all subgames is too simplistic!*

Behavioral Strategies in Extensive Games

If the set of information sets at the i^{th} player is denoted as $\mathcal{I}_i$, then

- ***Pure strategies*** *in extensive-form games are choice tuples at a given player, where each entry is picked from one of his/her information sets.*

  Notation: $\boldsymbol{c}_i = (c_{i,j_1}, \cdots, c_{i,j_L}) \in \mathcal{C}_i$, where c_{i,j_ℓ} is the ℓ^{th} strategy in the j^{th} information set in $\mathcal{I}_i$.

- ***Mixed strategies*** *are lotteries on pure strategies.*

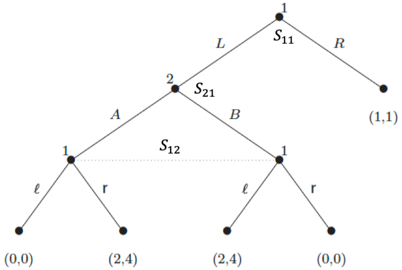
  Notation: $\boldsymbol{\sigma}_i \in \Delta(\mathcal{C}_i)$.

However, in extensive games, we can define another type of lottery, as shown below:

Definition

Given a extensive game $\Gamma = (\mathcal{N}, \mathcal{C}, G, \pi, P, \mathcal{I}, \mathcal{U})$, a ***behavioral strategy*** at the i^{th} player is a conditional lottery $\boldsymbol{\pi}_i \in \Delta(D_{i,s})$ on the choice set $D_{i,s}$ available within the state (node) s in a given information set at the i^{th} player.

Behavioral Strategies: An Example



- Information Sets: $\mathcal{I}_1 = \{S_{11}, S_{12}\}$, $\mathcal{I}_2 = \{S_{21}\}$
- Pure strategies: $\mathcal{C}_1 = \{(L, \ell), (L, r), (R, \ell), (R, r)\}$, $\mathcal{C}_2 = \{A, B\}$
- Mixed strategy: $\boldsymbol{\sigma}_1 = \{p_1, p_2, p_3, 1 - p_1 - p_2 - p_3\}$, $\boldsymbol{\sigma}_2 = \{q, 1 - q\}$
- Behavioral strategy for P_1: $\boldsymbol{\pi}_1 = \{\boldsymbol{\pi}_{11}, \boldsymbol{\pi}_{12}\}$, where
  - $\boldsymbol{\pi}_{11} = \boldsymbol{\pi}_1(S_{11}) = \{L : \alpha_{11}, R : 1 - \alpha_{11}\}$
  - $\boldsymbol{\pi}_{12} = \boldsymbol{\pi}_1(S_{12}) = \{\ell : \alpha_{12}, r : 1 - \alpha_{12}\}$
- Behavioral strategy for P_2: $\boldsymbol{\pi}_2(S_{21}) = \{A : \beta_{21}, B : 1 - \beta_{21}\}$.

Equivalence between Mixed and Behavioral Strategies

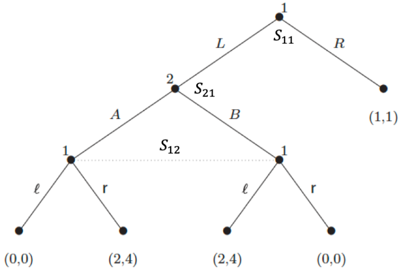
Theorem

In a game of perfect recall, for any mixed strategy, there is an outcome-equivalent behavioral strategy, and vice versa.

In the following example, we have

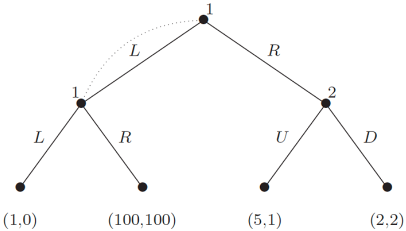
$$\boldsymbol{\sigma}_1 = \{(L,\ell) : 0.5, (R,\ell) : 0.5\} \equiv \boldsymbol{\pi}_1 = \{\boldsymbol{\pi}_{11} = \{L : 0.5, R : 0.5\},\ \boldsymbol{\pi}_{12} = \{\ell : 1, r : 0\}\}$$

since Player 2 believes that Player 1 does not play r in S_{12}, given $\boldsymbol{\sigma}_1$.



Extensive Games with Imperfect Recall

Behavioral and mixed strategies are incomparable in general.



- Pure strategies: $P_1 \Rightarrow \{L, R\}$, $P_2 \Rightarrow \{U, D\}$
- Mixed strategy for P_1: $(L : \pi, R : 1 - \pi))$ – once P_1 samples his/her mixed strategy, that strategy will be chosen in both nodes within the information state.
- Unique NE: (R, D)
- Behavioral strategy at P_1: $\{L : p, R : 1 - p\}$ (randomize afresh every time.)
    - $U_1(D) = p\,[p + 100(1 - p)] + (1 - p)2$
    - $\underset{p \in [0,1]}{\text{arg max}}\ U_1(D) = p^* = \dfrac{98}{198}.$
    - A new equilibrium in behavioral strategies: $\left\{ \left(\frac{98}{198}, \frac{100}{198} \right), (0, 1) \right\}$

Equilibrium in Perfect-Recall Games

Eliminate *nonsensical* NE using behavioral strategies!

Definition

A ***extensive-form Nash equilibrium*** is a mixed strategy Nash equilibrium $\boldsymbol{\sigma}$ that is equivalent to an assessment pair $(\boldsymbol{\pi}, \boldsymbol{\mu})$, where the behavioral strategy $\boldsymbol{\pi}$ is consistent with $\boldsymbol{\sigma}$ and a set of beliefs $\boldsymbol{\mu}$ according to Bayes' rule.

Can't we operate directly on the tree representation?

Definition

A ***behavioral equilibrium*** is a pair $(\boldsymbol{\pi}, \boldsymbol{\mu})$ which satisfies:

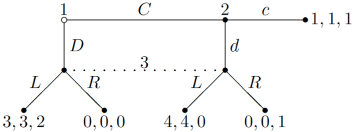
- **Sequential Rationality:** Given any alternative strategy $\boldsymbol{\pi}_i'$ at the i^{th} player and his/her belief $\boldsymbol{\mu}_{i,j_s}$ on the state j_s within an information set $\mathcal{I}_{i,j}$, we have

$$u_i(\boldsymbol{\pi} | \mathcal{I}_{i,j_s}, \boldsymbol{\mu}_{i,j_s}) \geq u_i(\boldsymbol{\pi}_i', \boldsymbol{\pi}_{-i} | \mathcal{I}_{i,j_s}, \boldsymbol{\mu}_{i,j_s}), \text{ and}$$

- **Consistency:** Assuming that all the players picked a strategy $\boldsymbol{\pi}$ until reaching a state s, there exists a belief $\boldsymbol{\mu}(s)$ that is consistent with Bayes' rule.

Example: Selten's Horse

Induced Normal-Form Game:



Nash Equilibria:

- $NE_1 : \left\{D : 1,\ c : \left[\frac{1}{3}, 1\right],\ L : 1\right\}$
- $NE_2 : \left\{C : 1,\ c : 1,\ \sigma_3(R) \in \left[\frac{3}{4}, 1\right]\right\}$

Behavioral Equilibrium:

- NE_1 is not a behavioral equilibrium (violates sequential rationality at Player 2)
- NE_2 is sequentially rational. But, how about the beliefs in $\mathcal{I}_3$?
- Let $\sigma^\epsilon = \left\{\sigma_1^\epsilon(C) = 1 - \epsilon,\ \sigma_2^\epsilon(d) = \frac{2\epsilon}{1-\epsilon},\ \sigma_3^\epsilon(R) = \sigma_3(R) - \epsilon\right\}$, for a small ϵ.
- $\mu_{3,\ell} = \dfrac{\sigma_1^\epsilon(D)}{\sigma_1^\epsilon(D) + \sigma_1^\epsilon(C) \cdot \sigma_2^\epsilon(d)} = \dfrac{1}{3}.$

Sequential Equilibrium: A Refinement

Definition

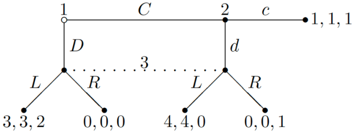
An assessment pair $(\boldsymbol{\pi}, \boldsymbol{\mu})$ is a ***sequential equilibrium*** if

1. Given any alternative strategy $\boldsymbol{\pi}'_i$ at the i^{th} player and his/her belief $\boldsymbol{\mu}_{i,j_s}$ on the state j_s within an information set $\mathcal{I}_{i,j}$, we have
$$u_i(\boldsymbol{\pi} | \mathcal{I}_{i,j_s}, \boldsymbol{\mu}_{i,j_s}) \geq u_i(\boldsymbol{\pi}'_i, \boldsymbol{\pi}_{-i} | \mathcal{I}_{i,j_s}, \boldsymbol{\mu}_{i,j_s}),$$
2. **Consistency:** Assuming that all the players picked a strategy $\boldsymbol{\pi}$ until reaching a state s, there exists a belief $\boldsymbol{\mu}(s)$ that is consistent with Bayes' rule.
3. **Convergence:** There exists a sequence $\left\{ \left(\boldsymbol{\pi}^{(n)}, \boldsymbol{\mu}^{(n)} \right) \right\}_{n=1}^{\infty}$ such that $(\boldsymbol{\pi}, \boldsymbol{\mu}) = \lim_{n \to \infty} \left(\boldsymbol{\pi}^{(n)}, \boldsymbol{\mu}^{(n)} \right)$, where $\boldsymbol{\mu}_n$ is a belief that is consistent with the behavioral strategy $\boldsymbol{\pi}_n$, for all $n = 1, 2 \cdots$.

Theorem

- Every finite game of perfect recall has a sequential equilibrium.
- Every subgame-perfect equilibrium is a sequential equilibrium, but the converse is not true in general.

Example: Selten's Horse



Nash Equilibria:

- $NE_1 : \left\{ D : 1, \ c : \left[\frac{1}{3}, 1 \right], \ L : 1 \right\}$
- $NE_2 : \left\{ C : 1, \ c : 1, \ \sigma_3(R) \in \left[\frac{3}{4}, 1 \right] \right\}$

Behavioral Equilibrium:

- NE_1 is not a behavioral equilibrium (violates sequential rationality at Player 2)
- NE_2 is sequentially rational with $\mu_{3,\ell} = \dfrac{1}{3}$.

Sequential Equilibrium:

- NE_1 is not a sequential equilibrium (violates sequential rationality at Player 2)
- For every equilibrium of type NE_2, there exists a sequential equilibrium with the following sequence:
  - $\sigma^\epsilon = \left\{ \sigma_1^\epsilon(C) = 1 - \epsilon, \ \sigma_2^\epsilon(d) = 2\epsilon, \ \sigma_3^\epsilon(R) = \sigma_3(R) - \epsilon \right\}$.
  - $\mu_{3,\ell} = \dfrac{\sigma_1^\epsilon(D)}{\sigma_1^\epsilon(D) + \sigma_1^\epsilon(C)\sigma_2^\epsilon(d)} = \dfrac{1}{3 - 2\epsilon} \xrightarrow{\epsilon \to 0} \dfrac{1}{3}$

Perfect Bayesian Extensive-Form Games

- Let Θ_i denote the set of types of the i^{th} player with a prior belief p_i.
- Let $\boldsymbol{p} = \{p_i\}_{i \in \mathcal{N}}$ be the profile of prior beliefs.
- Perfect Bayesian equilibrium ⇒ A generalization of *behavioral equilibrium*.

Definition

A pair $(\boldsymbol{\pi}, \boldsymbol{\mu})$ is a ***perfect Bayesian equilibrium*** if

1. The mixed strategy profile $\boldsymbol{\pi}$ is ***sequentially rational***, given $\boldsymbol{\mu}$.
2. There exists a belief system $\boldsymbol{\mu}$ that is ***consistent*** with the mixed strategy profile $\boldsymbol{\pi}$ and the prior belief about agents' type $\boldsymbol{p}$.

Theorem

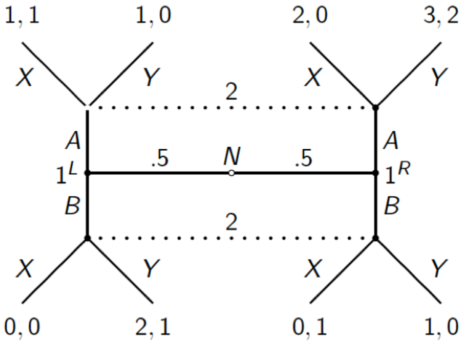
Every finite Bayesian extensive game has a perfect Bayesian equilibrium.

Theorem

Every perfect Bayesian equilibrium is a Nash equilibrium.

Signaling Games

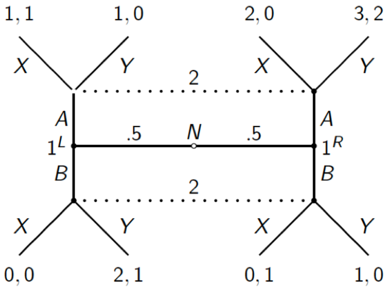
Consider the following sender-receiver (signaling) game, where the sender is characterized by one of the two types.



Sender's strategies:

- **Pooling Strategies:** AA, BB – *Sender does not reveal its type*
- **Separating Strategies:** AB, BA – *Sender reveals its type*

Signaling Games (cont...)



Two perfect Bayesian equilibria: Proof will be provided in a separate handout.

- ***Pooling Equilibrium:*** (AA, YX) when $\mu(L|A) = 0.5$ and $\mu(L|B) \leq 0.5$
    - How did we compute $\mu(L|B)$ if Player 1 plays AA?
    - Note: X is the best response to B only when $\mu(L|B) \leq 0.5$
- ***Separating Equilibrium:*** (BA, YY) when $\mu(L|A) = 0$ and $\mu(L|B) = 1$.

Pooling in e-Bay markets: *Buyers do not trust sellers who always signal high quality products, regardless of their true type.*

Repeated Games

Repeated interactions stimulate agents to track players' reputation over time and design strategies either to retaliate, or to act prosocially.

- Why participate in free crowdsourcing platforms (e.g. Wikipedia, Google's Crowdsource) even though workers do not get paid?
- Why look after neighbor's house when they are away?

Two types:

- Finite-Horizon Repeated Games – similar to extensive-form games
- Infinite-Horizon Repeated Games – no outcome nodes in the game!

Our Focus: ***Infinitely repeated games***

How to define choices and utilities in an infinitely repeated game?

Choices in Infinitely Repeated Games

Definition

Assuming that the players only observe strategy profiles at the end of each repetition stage, any ***choice in an infinitely repeated game*** is of the form

$$\boldsymbol{c} = \{\boldsymbol{c}_1, \boldsymbol{c}_2, \cdots, \boldsymbol{c}_k, \cdots\} \in \mathcal{C}_\infty,$$

where $\boldsymbol{c}_i \in \mathcal{C}_1 \times \cdots \times \mathcal{C}_N$ is the strategy profile chosen in the i^{th} iteration.

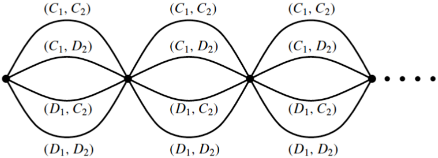
Consider the following infinitely repeated prisoner's dilemma:

Prisoner 2

Prisoner 1

	C_2	D_2
C_1	2, 2	0, 3
D_1	3, 0	1, 1

⇒



- ***Defection Strategy:*** $\boldsymbol{c}_i = (D_1, D_2)$, for all $i = 1, 2, \cdots$
- ***Grim (Trigger) Strategy:*** At the j^{th} player, we have

$$c_{i,j} = \begin{cases} D_j, & \text{if } c_{t,-j} = D_{-j} \\ C_j, & \text{otherwise} \end{cases}, \text{ for all } i = t+1, t+2, t+3, \text{ for any } t = 1, 2, \cdots.$$

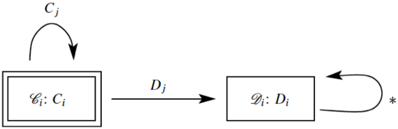
Representing Choices as Finite Machines

- Uncountably infinite strategy spaces ⇒ High strategic complexity
- **Finite automata** ⇒ tractable way to study infinitely repeated choices.
- ***Moore machine:*** Current strategy at a given player is a function of his current state, which in turn is computed using a transition function of the player's previous state and the strategy profile in the previous iteration.

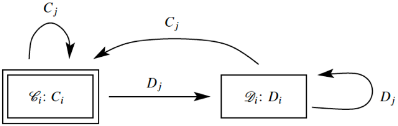
$$s_{i,t} = h\left(s_{i,t-1}, \boldsymbol{c}_{t-1}\right)$$

Examples:

- ***Grim (Trigger) Strategy:*** Both players start playing C



- ***Tit-for-Tat (TfT):*** Both players start playing C



Average Utilities

How should we define choice utilities in an infinitely repeated game?

Definition

Given an infinite sequence of one-stage utilities $u_{i,1}, u_{i,2}, \cdots$ at the i^{th} player, the ***average utility*** of the i^{th} player is defined as

$$\bar{u}_i = \lim_{k \to \infty} \frac{1}{k} \sum_{j=1}^{k} u_{i,j}.$$

Claim

If the choice is represented as a Moore machine with the longest cycle T, then the ***average utility*** of the i^{th} player can be computed as

$$\bar{u}_i = \frac{1}{T} \sum_{j=1}^{T} u_{i,j}.$$

Discounted Utilities

What if, the players build frustration with time?

Definition

Given an infinite sequence of one-stage utilities $u_{i,1}, u_{i,2}, \cdots$ at the i^{th} player, and a discounting factor $\beta \in [0,1]$, the ***discounted utility*** of the i^{th} player is defined as

$$\bar{u}_i = \lim_{k \to \infty} \frac{1}{k} \sum_{j=1}^{k} \beta^{j-1} u_{i,j}.$$

Claim

If the choice is represented as a Moore machine with the longest cycle T, then the ***discounted utility*** of the i^{th} player can be computed as

$$\bar{u}_i = \frac{1}{(1-\beta)T} \sum_{j=1}^{T} \beta^{j-1} u_{i,j}.$$

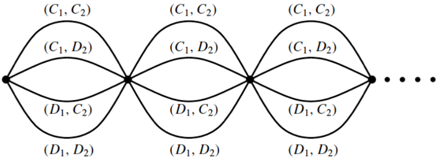
Equilibrium: Repeated Prisoner's Dilemma

Prisoner 2

Prisoner 1

	C_2	D_2
C_1	2, 2	0, 3
D_1	3, 0	1, 1

⇒



Claim

$(Grim, Grim)$ is a Nash equilibria in repeated Prisoner's Dilemma with β-discounted utilities, when $\beta \geq \frac{1}{2}$.

- Assume P_{-i} plays C_{-i} for the first T times.
- Let P_i choose C_i for the first $T-1$ times and then choose D_i at time T.
- Then, P_i's best response utility is

$$\begin{aligned} u_i &= \sum_{t=1}^{T-1} \beta^t 2 + 3\beta^T + \sum_{t=T+1}^{\infty} \beta^t 1 \\ &= 2\frac{1-\beta^T}{1-\beta} + 3\beta^T + \beta^{T+1}\frac{1}{1-\beta} \end{aligned}$$

Equilibrium: Repeated Prisoner's Dilemma

But, Grim includes the possibility where C_i can be played against C_{-i} forever!

Is C_i a best response to C_{-i} as well?

- Note that if P_i continued to play C_i for all $t \geq T$, P_i's best response utility is

$$\begin{aligned} u_i &= \sum_{t=1}^{T-1} \beta^t 2 + 2\beta^T + \sum_{t=T+1}^{\infty} \beta^t 2 \\ &= 2\frac{1-\beta^T}{1-\beta} + 2\beta^T + \beta^{T+1}\frac{2}{1-\beta} \end{aligned}$$

- C_i is the best response to C_{-i} only when

$$2\beta^T + \beta^{T+1}\frac{2}{1-\beta} \geq 3\beta^T + \beta^{T+1}\frac{1}{1-\beta}, \text{ for any } T.$$

- Upon simplification, Grim is the best response to Grim only when $\beta \geq \frac{1}{2}$.

In other words, both players should be patient enough for $(Grim, Grim)$ to be a Nash equilibrium!

Enforceability and Feasibility

Is there an easier way to validate, if a machine tuple is NE?

For that, we need to define two properties of utility profiles:

Definition

Given the minimax value of the i^{th} player as $v_i = \min_{\boldsymbol{c}^{(-i)}} \max_{\boldsymbol{c}^{(i)}} u_i(\boldsymbol{c}^{(i)}, \boldsymbol{c}^{(-i)})$, the utility profile $u = \{u_1, \cdots, u_n\}$ is ***enforceable***, if $u_i \geq v_i$ holds true for all $i \in \mathcal{N}$.

Definition

A utility profile $u = \{u_1, \cdots, u_n\}$ is ***feasible*** if there exists a lottery $\alpha \in \Delta(\mathcal{C}_\infty)$ such that, for all i, we have

$$u_i = \sum_{\boldsymbol{c} \in \mathcal{C}_\infty} \alpha_{\boldsymbol{c}} u_i(\boldsymbol{c}).$$

Folk's Theorem

Folk's theorem is actually a class of theorems, which characterizes equilibria in different types of infinitely repeated games...

Theorem

Consider any n-player normal-form game Γ, which has an average utility profile $\boldsymbol{u} = \{u_1, u_2, \cdots, u_n\}$, when repeated over an infinite time-horizon.

- If $\boldsymbol{u}$ is the utility profile for any Nash equilibrium $\boldsymbol{c}^*$ of the infinitely repeated Γ, then $\boldsymbol{u}$ is enforceable.
- If u is both feasible and enforceable, then u is the utility profile for some Nash equilibrium $\boldsymbol{c}^*$ of the infinitely repeated Γ.

Theorem

Consider any n-player normal-form game Γ, which has an discounted utility profile $\boldsymbol{u} = \{u_1, u_2, \cdots, u_n\}$ with some $\beta \in [0, 1]$, when repeated over an infinite time-horizon.

- If $\boldsymbol{u}$ is the utility profile for any Nash equilibrium $\boldsymbol{c}$ of the infinitely repeated Γ, then $\boldsymbol{u}$ is enforceable.
- If u is both feasible and enforceable, then u is the utility profile for some Nash equilibrium of the infinitely repeated Γ.

Bounded Rationality in Repeated Games

- Best response analysis ⇒ Uncountably infinite comparisons. . .
- Can we define preference orders on Moore machines?

Definition

Given two machine tuples $(M_1, \cdots, M_N)$ and $(M'_1, \cdots, M'_N)$, we define a ***preference order*** at the i^{th} player as $(M_1, \cdots, M_N) \succeq_i (M'_1, \cdots, M'_N)$, if

$$(u_i(M_1, \cdots, M_N), -|M_i|) \ \succeq_L \ \Big(u_i(M'_1, \cdots, M'_N), -|M'_i|\Big)$$

where $\succeq_L$ defines a lexicographical order in $\mathbb{R}^2$.

Definition

The tuple $(M_i, \boldsymbol{M}_{-i})$ is said to be a ***Nash-Rubinstein equilibrium*** in a repeated game, if

$$(M_i, \boldsymbol{M}_{-i}) \succeq_i (M'_i, \boldsymbol{M}_{-i}),$$

for any M'_i, for all $i = 1, \cdots, N$.

Summary

- *Stackelberg Games:* How to define equilibria in leader-follower games?
- *Perfect Extensive Games:* How to solve perfect extensive games via the notion of subgame perfect equilibrium?
- *Imperfect Extensive Games:* Subgame perfect equilibrium is no longer sufficient!

  Then, how?
    - Behavioral Equilibrium
    - Sequential Equilibrium
- *Perfect Bayesian Games:* What if, there are chance nodes (due to unknown agent types) in the game?
- *Repeated Games:* How to define choices and utilities in an infinitely repeated game, and solve it?