Università degli Studi di Padova

Dipartimento di Ingegneria dell'Informazione

Lecture 13
Multistage games

Thomas Marchioro

November 16, 2023

Previously on game theory

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

- **Dynamic games**: introduce a temporal dimension with subsequent moves
- **Sequential games**: dynamic games where one player moves after the other
  - sequential games with no random outcomes are games of **perfect information** and can be solved via backward induction
  - "Solved" means that we can fully predict how the game will unfold if players act rationally
- **Equilibrium path**: sequence of nodes in a NE (equivalent to sequence of nodes in a NE)

Previously on game theory

- One can find NE in a dynamic game by writing down its normal form and looking for joint strategies where all players are choosing a best response
  - exactly as in static games
- However, NE are an inherently static concept and may contain irrationality in dynamic games
- For this reason, in dynamic games we are often interested in subgame-perfect NE

Previously on game theory

- **Subgame**: subset of nodes including a node x_j of the tree and all its descendants; in a proper subgame x_j must be the only node in its information set
- **Subgame-perfect Nash equilibrium (SPE)**: NE whose outcome is a NE in all (proper) subgames
- Stronger definition of equilibrium
  - In static games, players end up in a NE only if their beliefs match
  - In dynamic games of perfect information, the equilibrium path of a SPE is guaranteed to be played (e.g., in tic-tac-toe we can deterministically play an optimal strategy)

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

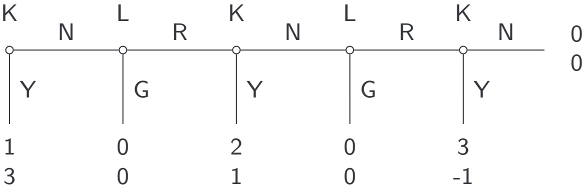
- It is the discount sales season, and Lou (L) wants to go shopping. He thinks it is best to wait until the last three days of the discount sales, because prices are cheapest. On day 1, Lou asks Karen (K) to go with him. If Karen says yes (Y), they go shopping and the game ends. If Karen says no (N), Lou can either give up (G) or request (R) again the following day. On day 2, the same happens. On day 3, if K still says N, the game ends as well and does not go into a further day. If in the end K and L do not go shopping, both of their payoffs are 0. If they do, their payoffs are computed as $u_K = d$ and $u_L = 5 - 2d$, respectively, where d is the day in which they go. All of this information about the game is common knowledge among the players.

Exercise

1. Represent the game in its extensive form.
2. How many (pure) strategies do K and L have, respectively?
3. Find the SPE of this game.
4. Find *one* NE that is not subgame-perfect.
5. Represent the game in its normal form.
6. Find *all* the NE of this game (both SPE and non-SPE).

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

1. Extensive form:



- K's payoff are on top, L's payoff are on the bottom

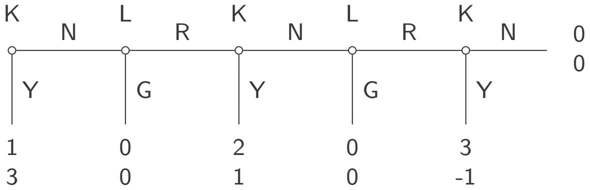
NB: 5 information sets here so, 5 subgames

Exercise: solution

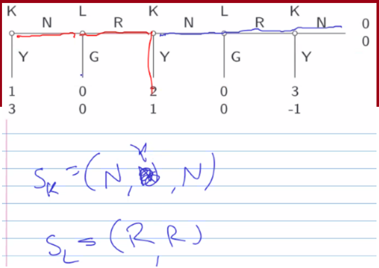
2. Number of strategies for each player:
   - K plays in 3 information sets and has 2 possible actions in each set: $2^3 = 8$ possible strategies
   - Her strategies are triplets: $s_K = (a_0, a_2, a_4)$, $a_j \in \{\text{Y, N}\}$
   - L plays in 2 information sets has 2 possible actions in each set: $2^2 = 4$ possible strategies
   - His strategies are triplets: $s_L = (a_1, a_3)$, $a_j \in \{\text{G, R}\}$

Exercise: solution

3 Find SPE:



- The SPE can be found via backward induction
- On the last day, K prefers Y (payoff 3) to N (payoff 0): $a_4 =$Y
- Before that, L prefers G (payoff 0) to R (payoff -1, knowing K's next choice): $a_3 =$G
- Before that, K prefers Y (payoff 1) to N (payoff 0, knowing how the rest of the game will play out): $a_2 =$Y
- ... Only SPE: $s_K = (N, Y, Y)$, $s_L = (R, G)$



4. Find *one* NE that is not a SPE:
   - Can be done without normal form
   - Any joint strategy that follows the SPE's equilibrium path is a NE (players do not regret the outcome)
   - We can just introduce irrational choices in subgames that are never played
   - E.g., $s_K = (N, Y, N)$, $s_L = (R, G)$: Karen refuses Lou's invitation on the last day even though that gives her a lower payoff; however, that scenario is never reached because Karen accepts Lou's invitation the day before

5. Normal form
6. Find *all* NE
   - Trivial and left to the reader as an exercise

Today on game theory

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

- **Multistage games**: a specific type of dynamic games, which consist in a sequence of smaller games, called "stages"
- Many games are divided in *independent* stages, each giving intermediate rewards
  - E.g., poker, tournaments, partial exams
- Payoffs in a multistage game are just the sum of all stages
- However, a **discount factor** δ can be introduced when stages are far apart in time
- **Important!** Do not confuse stages with subgames

Discount factor and time consistency

Discount factor

- In a multi-stage game with T stages, let $u_i^{(t)}$ be the utility received by player i at stage t
- Assuming **discount factor** δ, the total payoff of player i for the multi-stage game is

$$u_i = u_i^{(1)} + \delta u_i^{(2)} + \delta^2 u_i^{(3)} + \cdots + \delta^{(T-1)} u_i^{(T)}$$

- Meaning that the discount applied to the utility is an **exponential** function of the time t
- Why exponential? Why cannot we have, e.g., constant discount $u_i = u_i^{(1)} + \delta u_i^{(2)} + \cdots + \delta u_i^{(T)}$?

Discount factor

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

- Main reason: **time consistency**
- Suppose player i plans his/her strategy at stage 1 to optimize the current payoff

$$u_i = u_i^{(1)} + \delta u_i^{(2)} + \delta^2 u_i^{(3)} + \cdots + \delta^{(T-1)} u_i^{(T)}$$

- Then, at stage 2, the player re-evaluates the optimal strategy, now with total utility

$$u_i' = u_i^{(2)} + \delta u_i^{(3)} + \delta^2 u_i^{(4)} + \cdots + \delta^{(T-2)} u_i^{(T)}$$

- Does the optimal strategy change?

Example: scarce resource allocation

- Suppose we need to partition limited resources over 3 time slots
- In other words, we need to allocate x_1, x_2, x_3 with $x_1 + x_2 + x_3 = 1$
- We aim to optimize some utility function $u(x)$ (say $u(x) = \log(1 + x)$) with discount factor δ
- Clearly, if $\delta = 1$, $x_1 = x_2 = x_3 = \frac{1}{3}$

Example: scarce resource allocation

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

- If $\delta > 1$ we have that the total utility is

$$u(x_1, x_2, x_3) = \log(1 + x_1) + \delta \log(1 + x_2) + \delta^2 \log(2 - x_1 - x_2)$$

- Taking the derivatives w.r.t. x_1 and x_2 (clearly $x_3 = 1 - x_1 - x_2$), we get

$$\frac{\partial u}{\partial x_1} = \frac{1}{1 + x_1} - \frac{\delta^2}{2 - x_1 - x_2} \text{ and } \frac{\partial u}{\partial x_2} = \frac{\delta}{1 + x_2} - \frac{\delta^2}{2 - x_1 - x_2}$$

- By setting both derivatives equal to 0, we get

$$x_1^* = \frac{3 - \delta - \delta^2}{1 + \delta + \delta^2}, x_2^* = \frac{-1 + 3\delta - \delta^2}{1 + \delta + \delta^2}, x_3^* = \frac{-1 - \delta - \delta^2}{1 + \delta + \delta^2}$$

Example: scarce resource allocation

- Now, repeat the calculation at stage 2 with $x_2 + x_3 = 1 - x_1^*$ available resources
- We get again

$$x_2^* = \frac{-1 + 3\delta - \delta^2}{1 + \delta + \delta^2}, x_3^* = \frac{-1 - \delta - \delta^2}{1 + \delta + \delta^2}$$

- Exponential discount is **consistent over time**: the optimal strategy is the same if recomputed at different stages
- Other types of discount lead to inconsistencies

Example: scarce resource allocation

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

- E.g., with constant discount, at stage 1 we have

$$u = x_1 + \delta x_2 + \delta x_3$$

  which leads to $x_2^* = x_3^*$
- However, at stage 2 we have

$$u = x_2 + \delta x_3$$

  which leads to $x_2^* \neq x_3^*$

Multistage games

Multistage games

- Static games describe well situations where players act simultaneously
- Dynamic games add a time dimension
  - But payoffs are given only at the end nodes
- Many real games have **intermediate** steps that give partial payoffs
  - Tournaments, rounds of cards, partial exams, ...
- Can we see them as a single grand game?

Multistage games

- Define multistage games as a finite sequence of T normal form **stage games**
  - Stage games are defined independently of each other and include the same set of players
  - They are complete but imperfect information games (there are simultaneous moves)
  - Possible extensions to infinite horizon, which we will see only in some special cases
- Total payoffs are evaluated according to the sequence of outcomes in all stages

Multistage games

- **Example**: A sequence of T stage games with same players but different action sets
  - Action chosen in each game lead to an outcome for that game, and thus to a partial payoff $u_i^{(t)}$
  - The payoffs on the second stage are independent of what happened in the first stage
  - Total payoffs are the (discounted) sums of partial payoffs for each player (discount factor δ is the same for all users, and this is common knowledge)
- Total payoff for player i: $u_i = \sum_{t=1}^{T} \delta^{t-1} u_i^{(t)}$

Example: Prisoner-Revenge

- A and B play the Prisoner's dilemma
- After that, they eventually go out of jail and can either gather a gang and fight the other (G), or leave the other alone (L)
  - If they both choose to leave the other alone, they never meet again $\rightarrow$ payoff is 0 for both
  - If they both gather a gang, they get into a harsh fight that leaves both quite beaten $\rightarrow$ payoff is -3 for both
  - If one gathers a gang and the other stays alone, the loner gets beaten to a pulp ($\rightarrow$ payoff -4), but also the other gets some bruises ($\rightarrow$ payoff -1)

Prisoner-Revenge

- The normal-form representation of each stage is as follows

First stage (Prisoner's dilemma)

A \ B	m	f
M	4, 4	-1, 5
F	5, -1	1, 1

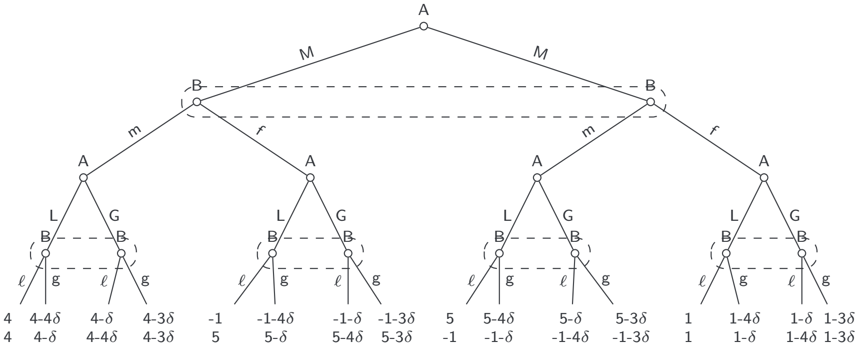
Second stage (Revenge)

A \ B	ℓ	g
L	0, 0	-4, -1
G	-1, -4	-3, -3

- Suppose the payoffs are aggregated with discount δ
- E.g., if the outcome of the first stage is (F, m) and the outcome of the second stage is (G, g), the payoffs are $u_A = 5 - 3\delta$ for player A, and $u_B = -1 - 3\delta$ for player B

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

- Extensive form



Strategies in multistage games

- A strategy for each player must specify
  - what to do in the first stage (just one action)
  - what to do in the subsequent game(s) depending on the outcome of the previous game(s)
- The Prisoner-Revenge game has already 32 possible strategies (already complex enough)
- Strategies can be thought of as "I play X at stage 1, then I play Y if the outcome of stage 1 is O"

SPE in multistage games

- Remember that a SPE is a joint strategy that is a NE in every subgame
- In multistage games, each stage is independent of the others, therefore:
- **Theorem**: Suppose $s^{(t)} = (s_1^{(t)}, \dots, s_n^{(t)})$ is a NE for stage t of multistage game $\mathbb{G}$; then there is a SPE of $\mathbb{G}$ with equilibrium path $s^{(1)}, \dots, s^{(T)}$
- Specifically, a strategy where each players chooses the same NE strategy at each stage (ignoring the outcome of other stages) is always a SPE

SPE in Prisoner-Revenge

- In the Prisoner-Revenge game, (F,f) is a NE of the first stage; (G,g) and (L,ℓ) are NE of the second stage
- This means that (FLLLL,f$\ell\ell\ell\ell$) and (FGGGG,fgggg) are SPE for the multistage game, since a NE is played in both the first and second stage
- However, these strategies ignore any possible strategic link between stages
  - The stage games are played independently
  - Are there other SPE that exploit the connection between stages?

SPE in multistage games

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

- We need to start from the end of the game
  - as we did in backward induction
- **Theorem**: Any NE s^* (even if it is not a SPE) of multistage game $\mathbb{G} = (\mathbb{G}_1, \ldots, \mathbb{G}_T)$ requires that a NE is played in the last stage $\mathbb{G}_T$
- Intuition: In the last stage, players do not have future stages that influence their payoff: they just play the best responses for that stage
- **Theorem**: If stage games $\mathbb{G}_1, \ldots, \mathbb{G}_T$ all have a unique NE, then $\mathbb{G} = (\mathbb{G}_1, \ldots, \mathbb{G}_T)$ has a unique SPE

Strategic connections

- These last two theorems imply that if the last stage of a multistage game has a single NE, that is what rational players will play
  - Not much of a surprise, and nothing we can do about it
- However, what if the last stage has **multiple** NE?
  - Surprisingly, that enables non-NE strategies to be played in previous stages
  - In other words, it is possible to find SPE where players do not play the NE strategy in some of the stages

Strategic conn. in Prisoner-Revenge

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

- In Prisoner-Revenge (M,m) is not a NE for the first stage
  - in fact, strategy M is strictly dominated for both players
- However, in the second stage we have two NE: a "good" NE (L,ℓ) and a "bad" NE (G, g)
- If the discount factor δ is *large enough*, we can leverage the NE in the second stage to enforce collaboration in the first stage
  - Remember: a larger δ means that players care more about future payoffs

Strategic conn. in Prisoner-Revenge

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

- Set strategies s_A =(M,L,G,G,G) and s_B =(m,ℓ,g,g,g)
- In other words, the strategy for both players is "Play M in the 1st stage; in the 2nd stage, play L if the outcome was (M, m) in stage 1, otherwise play G"
- That means each prisoner will leave the other alone if they both keep mum in the first stage, otherwise they will join a gang and pick a fight
- We need to verify if this strategy is **sustainable**
  - Do players have an incentive to deviate from such strategy?
  - Remember: a "deviation" means that one player changes strategy while the other keeps the current one

Strategic conn. in Prisoner-Revenge

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

First stage (Prisoner's dilemma)

A \ B	m	f
M	4, 4	-1, 5
F	5, -1	1, 1

Second stage (Revenge)

A \ B	ℓ	g
L	0, 0	-4, -1
G	-1, -4	-3, -3

- In the last stage, no player has incentive to deviate unilaterally, since the strategy always involves playing a NE in it
- We need to check in stage 1 if $s_A =$(M,L,G,G,G) is a best response to $s_B =$(m,ℓ,g,g,g)
  - $u_A(s_A, s_B) = 4 + 0\delta$
  - $u_A((\text{F,L,G,G,G}), s_B) = 5 - 3\delta$ (change only the strategy at stage 1)
- The strategy is *sustainable if* $4 \geq 5 - 3\delta$, i.e. for $\delta \geq 1/3$

Strategic connections: comment

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

- Strategic connection is possible if the last stage has multiple NE that are different in terms of payoff: a "stick" and a "carrot"
- So, the SPE is created as follows:
  - play desired non-NE action in the first stage
  - reward the other player(s) with the "carrot" if they collaborate
  - otherwise, punish them with the "stick"
- The discount factor δ should be high enough for the difference between stick and carrot to have an impact

Strategic connections: comment

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

- The value of δ relates to the credibility of a threat
- If $\delta \to 0$, players do not care about future; therefore, threatening punishment with the "stick" is not credible
- The punishment is effective only if short-term gains are not worth compared to long-term losses
  - Note that the latter is weighted on δ

- The carrot-and-stick strategy can be used to create a SPE with any joint outcome in the first stage (as long as it is sustainable)
- For example, we can create a SPE that supports (F, m) as first moves (the rest of the strategy is identical: friendly NE if all players comply, gang NE otherwise)
- This joint strategies is denoted as (s_1, s_2) with $s_1 =$(F,L,G,G,G) and $s_2 =$(m,ℓ,g,g,g)
- Always sustainable for player A (best response played in both stages)
- Sustainable for player B if $u_B(s_1, (\text{m},\ell,\text{g},\text{g},\text{g})) \geq u_B(s_1, (\text{f},\ell,\text{g},\text{g},\text{g}))$

Strategic conn. in Prisoner-Revenge

First stage (Prisoner's dilemma)

A \ B	m	f
M	4, 4	-1, 5
F	5, -1	1, 1

Second stage (Revenge)

A \ B	ℓ	g
L	0, 0	-4, -1
G	-1, -4	-3, -3

- To find δ such that the strategy is sustainable (and thus a SPE), we solve $u_B(s_1, (\text{m},\ell,\text{g},\text{g},\text{g})) \geq u_B(s_1, (\text{f},\ell,\text{g},\text{g},\text{g}))$ for δ
- We obtain $-1 + 0\delta \geq 1 - 3\delta$, meaning that (s_1, s_2) is a SPE if $\delta \geq 2/3$

One-stage deviation principle

- Prisoner-Revenge has 2 stages
  - Deviations are possible only at stage 1
  - Stage 2 is the last: players must have a NE there
- How does this generalize when we have more stages? Maybe in a 5-stage game players are interested in deviating by changing their strategy in two stages at once
- We need to follow the **one-stage deviation principle**
- But first some definitions:
- **Optimal strategy**: A strategy s_i is optimal for player i if for each information set h_i there is no way to improve it
  - i.e., no s_i' such that $u_i(s_i'|\{h_i\}) > u_i(s_i|\{h_i\})$
- **One-stage unimprovable strategy**: A strategy s_i is one-stage unimprovable if there is no s_i' that differs in one single stage such that $u_i(s_i'|\{h_i\}) > u_i(s_i|\{h_i\})$ for some h_i

One-stage deviation principle

- Clearly, optimum $\Rightarrow$ one-stage unimprovable
- Interestingly, the opposite is also true
- **Theorem**: A one-stage unimprovable strategy must be optimal
- *Proof*: Proceed by contradiction, assuming that there is a strategy s_i that is *one-stage unimprovable* but not optimal.
  - Then, there must be a strategy s_i' that deviates from s_i by one step or more.
  - Consider the deviation on the last stage, and consider the subgame having the corresponding stage as root
  - In that game, strategy s_i' must also be better than s_i.
  - However, in that subgame there is only a one-stage deviation.
  - Therefore, s_i is *one-stage improvable* (contradiction). $\square$

Self-assessment

- Consider multistage game $\mathbb{G} = (\mathbb{G}_1, \mathbb{G}_2)$, with $\mathbb{G}_1$ being the first stage, and $\mathbb{G}_2$ being the second (and last) stage:
  - Is it possible to find a SPE for $\mathbb{G}$ where a non-NE is played in $\mathbb{G}_2$?
  - Is it possible to find a NE for $\mathbb{G}$ where a non-NE is played in $\mathbb{G}_2$?
  - Is it possible to find a SPE for $\mathbb{G}$ where a non-NE is played in $\mathbb{G}_1$?
  - Is it possible to find a SPE for $\mathbb{G}$ where a strictly dominated strategy is played in $\mathbb{G}_1$?
  - What is the minimum number of NE in stage game $\mathbb{G}_2$ to enable a carrot-and-stick SPE in $\mathbb{G}$? What characteristics should these NE have?

Sorry, gotta bounce!
Send me questions via e-mail