Game theory

a course for the
MSc in ICT for Internet and multimedia

Leonardo Badia

leonardo.badia @gmail.com

Time inconsistencies

Contradictory discounting

Scarce resource allocation

- A player has a fixed resource budget K=1 to allocate over N subsequent time steps
  - For simplicity, assume N=3 (can be generalized)
  - Assume a discount factor of $\boxed{\delta}$
  - Total payoff = sum of discounted partial payoffs
  - $\boxed{v(x_1, x_2, x_3) = u(x_1) + \delta u(x_2) + \delta^2 u(x_3)}$
- Problem is: $\boxed{\max v, \quad \text{s.t.: } x_1 + x_2 + x_3 = 1}$
  - note that this is a single-person optimization

Scarce resource allocation

- Take this original example: $\boxed{u(x) = \log(1+x)}$
- Take 1st-order derivative of $v(1-x_2-x_3, x_2, x_3)$ and set it to 0 to get

val. massimo

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

  - for $\boxed{\delta = 1}$, equal split
  - for $\delta = 0.8 \rightarrow x_1 = 0.6393, x_2 = 0.3115, x_3 = 0.0492$
  - Side note: $\delta > (5^{0.5} - 1)/2 = 0.618$ must hold
- Is this choice consistent? Or can the player regret it later on in the game?

Time consistency

- If the player already spent x_1, we are left with $1 - x_1$ to be split between 2 periods, $x_2 + x_3$
- $\boxed{\max w = u(x_2) + \delta\, u(x_3) = u(x_2) + \delta\, u(1-x_1-x_2)}$
  - At period 2, $u(x_2)$ is weighed 1 (time 2 = present), while $u(x_3)$ is discounted by δ.
  - We get: $\boxed{x_2 = \frac{2 - x_1 - \delta}{1+\delta} \qquad x_3 = \frac{-1 + 2\delta - \delta x_1}{1+\delta}}$
  - if $\delta = 1$ and $x_1 = 1/3$ $\rightarrow$ equal split
  - if $\delta = 0.8$ and $x_1 = 0.6393$ $\rightarrow$ $x_2 = 0.3115$, $x_3 = 0.0492$

stessi risultati

- Like before: exponential discount is **consistent**

prima δ, poi δ^2, etc.

Time consistency

- What if we have consistency issues? Assume $v(x_1, x_2, x_3) = u(x_1) + \boxed{\delta u(x_2) + \delta\, u(x_3)}$
  - Future payoffs are all discounted but with the same factor δ (no exponential stacking)
- Same procedure, we get

$$x_1 = \frac{3-2\delta}{2\delta+1} \qquad x_2 = \frac{2\delta-1}{2\delta+1} \qquad x_3 = \frac{2\delta-1}{2\delta+1}$$

  - we get $x_2 = x_3$ as they have the same discount
  - for $\delta=1$, equal split (note: now we need $\delta>0.5$)
  - for $\delta=0.8 \rightarrow x_1=0.5385, x_2=0.2308, x_3=0.2308$

Time consistency

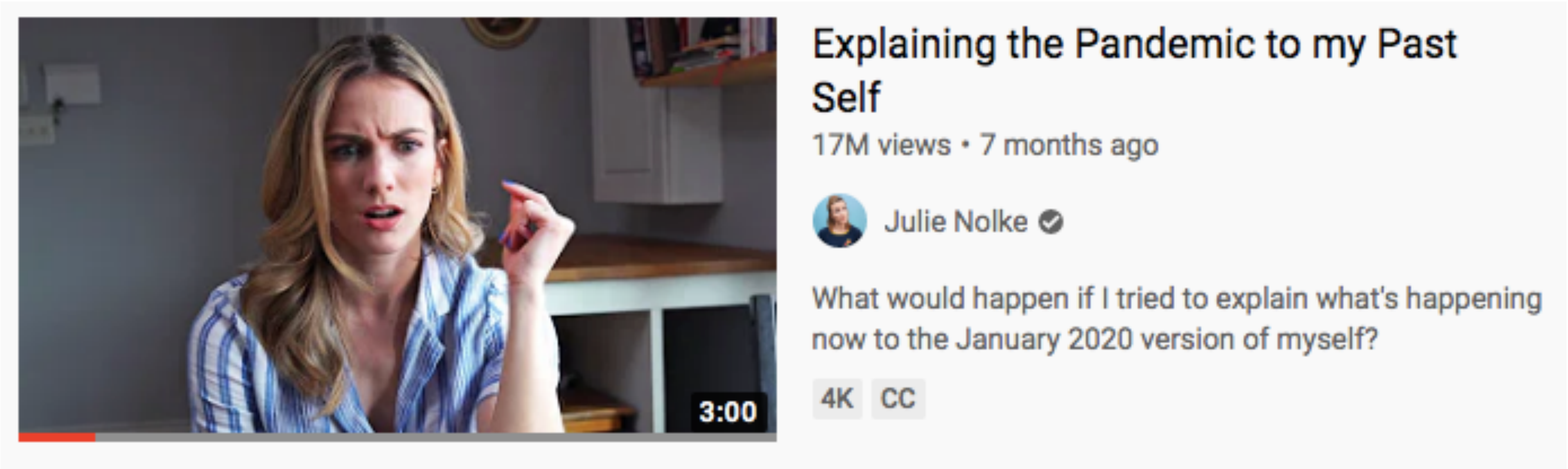
- However, at time 2 the player does not want to respect $x_2 = x_3$ (the future is discounted)
- max $w(x_2, x_3) = u(x_2) + \delta\, u(x_3)$
  - Future payoffs are all discounted but with the same factor δ (no exponential stacking)
  - we already saw that this leads to:

$$x_2 = \frac{2 - x_1 - \delta}{1+\delta} \qquad x_3 = \frac{-1 + 2\delta - \delta x_1}{1+\delta}$$

  - for $\delta = 0.8$, $x_1 = 0.5385 \rightarrow x_2 = 0.3675$, $x_3 = 0.0940$
  - <u>Inconsistent split!</u>

Time consistency

- Actually, it is even worse than that, because a fully rational player knows he/she will act strangely and wants to anticipate this
  - Struggle between: Player 1 (present-day player) and Player 2 (future self at step 2)



Time consistency

- $v(x_1, x_2, x_3) = u(x_1) + \delta u(x_2) + \delta u(x_3)$
  - for Player 1, day 2 and 3 are equally important, the choice would be $\boxed{x_2 = (1-x_1)/2,\ x_3 = (1-x_1)/2}$
  - but Player 2 instead wants

$$\boxed{x_2 = \frac{2 - x_1 - \delta}{1+\delta} \qquad x_3 = \frac{-1 + 2\delta - \delta x_1}{1+\delta}}$$

- If Player 1 anticipates this through **backward induction** and $\boxed{\max u(x_1) + \delta u(x_2) + \delta u(x_3)}$, the result will be $\boxed{x_1 = 1}$!
  - overconsumption to prevent further misuse!!

Comment

- This is actually a side-derivation to justify what we will do in the following, that is:
  - everytime we combine different discounted payoffs, we always do exponential discount
  - other choices lead to inconsistencies that are not coherent with rationality

Multistage games

Same players playing multiple games

Multistage games

- Normal form games describe well situations where players act simultaneously
- Extensive form games add a time dimension
  - But payoffs are given only at the end nodes
- Many real games have **intermediate** steps that give partial payoffs, valued on aggregate
  - 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 (that is, simultaneous move games)
  - possible extension to infinite horizon
    we will see it only in some special cases
- Total payoffs are evaluated from the sequence of outcomes of the stage games

Multistage games

- Example: a sequence of 2 stage games with same players but different action sets
  - Actions chosen in each game lead to an outcome for that game, and thus to a partial payoff $\boxed{u_i^{(j)}}$
  - Players get the same payoffs for their second decisions, whatever the outcome of the first game
  - Total payoffs are the (discounted) sums of partial payoffs for each player (discount factor δ is the same for all the users, and is common knowledge)
  - total payoff for player i: $\boxed{u_i = \sum_{j=1..T} \delta^j u_i^{(j)}}$

Example: Prisoner-Revenge

- Al and Bob play the Prisoner's Dilemma
- After that, they go out of jail and they can either join a gang (G) or remain a "loner" (L)
  - If they both stay alone, they never meet again → payoff is 0 for both
  - If they both join a gang, they fight each other → negative payoff for both
  - If only one has a gang to defend him, he gets a (small) loss, the other a (heavy) loss

Prisoner-Revenge

- Suppose the payoffs are as follows

first stage (Prisoner)

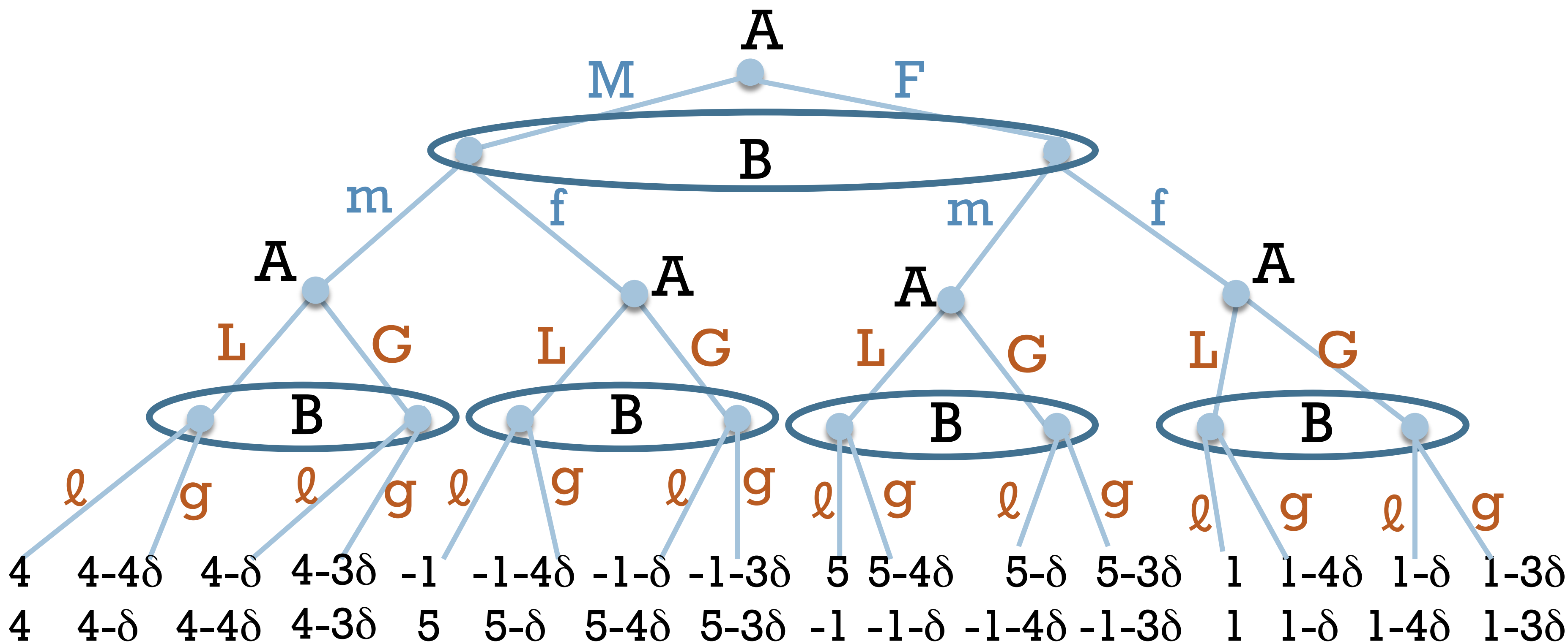
Al \ Bob	m	f
M	4, 4	-1, 5
F	5, -1	1, 1

second stage (Revenge)

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

- And they are aggregated with discount δ

Prisoner-Revenge



Strategies of 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 start by playing X, then I play Y if this happens"

Subgame perfect equilibria

- Remember that a SPE is a joint strategy such that a NE is played in every subgame
- The stage games are independent, thus:
- **Theorem 1**. If $\boxed{s_j^*}$ is a NE strategy profile for the jth stage game, then there exists a SPE whose equilibrium path is $\boxed{s_1^*, s_2^*, \ldots, s_T^*}$

uno

→ #stage

  - **Proof**. Consider a strategy where each player is allowed to only play what s_j^* states at stage j. This implies a NE is achieved in every subgame

Prisoner-Revenge

- Remember that (F,f) is a NE of the first stage
- This means that A playing (F, $\overbrace{\text{L, L, L, L}}$) and B playing (f, ℓ, ℓ, ℓ, ℓ) must be a SPE because (L, ℓ) is a NE of the second stage game

  azione | outcome di 1° stage

- Similarly, (F, G, G, G, G) - (f, g, g, g, g) is another SPE, as (G,g) is a NE of stage 2
- Note that we removed any strategic link
  - The games are played independently
  - Is there an alternative with strategic connection?

Subgame perfect equilibria

- We need to start from the end of the game
  - Same as we did for backward induction!
- **Theorem 2.** Any NE $\boxed{s^*}$ (even if it is no SPE) of a multistage game $\boxed{(G_1, G_2, \dots, G_T)}$ ~~must dictate~~ implies a NE is played in stage game G_T
  - **Proof.** Stage T is the last one, and this is common knowledge. No future to influence the actions of the players: they play only best responses
- **Theorem 3.** If $\boxed{G_1, G_2, \dots, G_T}$ all have a unique NE, then $\boxed{(G_1, G_2, \dots, G_T)}$ has a unique SPE

Strategic connection

- Theorems 2 and 3 imply that if the last stages have only one NE, this will be played
  - Not much of a surprise, and nothing we can do
- What if the Tth stage has multiple NE?
- Surprisingly, this enables non-NE to be played (in other stages of course)
  - This means that SPE can be built, where some of the intermediate stages have non-NE strategies that are played!

Strategic connection

- See for example the Prisoner-Revenge game
- In the second stage:
  - two NEs: (L, ℓ) "friendly" and (G,g) "gang"
  - (M,m) is not a NE in the first stage
  - If a static Prisoner game is played, joint strategy (M,m) cannot be supported (it is dominated)
  - However, we can enforce it to be played if the discount factor is high enough

Strategic connection

- Set strategy $s_1 = (M, L, G, G, G)$ for player A and similarly, $s_2 = (m, \ell, g, g, g)$ for player B
- In other words, both players are adopting a strategy described as "In stage 1, I mum. Then if the first outcome is (M,m) I play loner, otherwise I play gang"
- Such a joint strategy (s_1, s_2) is a SPE if the discount factor δ is "high enough" (see later)

Strategic connection

- **Proof**. Clearly no player wants to deviate in the second stage. They also play a NE in each subgame (proper). Thus, if (s_1, s_2) gives a NE in the whole game we prove that it is an SPE
- We need to check whether in stage 1, s_1 is a best response to s_2
  - All that s_1 does in stage 1 is to play M
  - $u_1(M, s_2) = 4 + 0\,\delta$, $u_1(F, s_2) = 5 - 3\,\delta$
  - M is a best response if $4 > 5 - 3\,\delta \rightarrow$ $\boxed{\delta \geq 1/3}$

Comment

- Strategic connection is possible if the last stage has multiple NEs that are considerably different: a "stick" and a "carrot"
- So, the SPE is created as follows:
  - Play desired non-NE action in the first stage
  - Reward opponents with carrot if they collaborate
  - Otherwise... threaten opponents with stick!
- δ must be high enough for the different payoffs of "carrot" and "stick" to have impact

Comment

- The value δ relates to credibility of threats
  - For example, if $\delta = 0$, the players do not care about the future; thus, threatening punishment with stick → non credible
- Effective punishment if short-term gains are not worth compared to long-term losses
  - Note that the latter are weighted on δ
- The example shown is complex enough to apply the theorems

Strategic connection

- The carrot-and-stick procedure can work to create a SPE where the first move is whatever
  - For example, we can create a SPE that supports the initial play of (F, m)
  - (the rest of the strategy is identical: friendly NE if all players comply, gang NE otherwise)
  - However, Bob may complain (if he does not, Al also keeps quiet!). Bob likes this SPE if

$u_2(s_1, m) = -1 + 0\,\delta \;,\; u_2(s_1, f) = 1 - 3\,\delta \;\rightarrow\; \delta \geq 2/3$

(higher discount factor is needed)

One-stage deviation principle

- Does Prisoner-Revenge capture everything?
  - Deviations were possible only at stage 1
  - Stage 2 is the last: players must have a NE there
- One may wonder what happens if more stages are present
  - Maybe if the game is five-stage, they may want to deviate from their gameplay at stage 1 and 3, but not individually
- Check the **one-stage deviation principle**

One-stage deviation principle

- Principle used in constrained optimization → however, backward induction is the same!
- A strategy s_i is **optimal** if there is no way to improve it for every information set h_i
  - I.e., no s_i' and h_i for which $u_i(s_i', h_i) > u_i(s_i, h_i)$
- A strategy s_i is **one-stage unimprovable** if there is no way to improve it by changing an action done in a given information set h_i

One-stage deviation principle

- Denying $u_i(s_i', h_i) > u_i(s_i, h_i)$ implies:
  - if s_i' is generic: the strategy is optimal
  - if s_i' is very similar to s_i, just changes an action: the strategy is one-stage unimprovable
- Clearly optimum ⇒ one-stage unimprovable
  - Interestingly, also the converse statement is true
- **Theorem 4**. A one-stage unimprovable strategy must be optimal

One-stage deviation principle

- For simplicity: proof by contradiction
  - assume $\boxed{s_i}$ is 1-step unimprovable but not optimal: then it exists $\boxed{s_i'}$ that deviates in 2 steps or more
  - if s_i' deviates from s_i under information set h_i, it must have a finite number of "deviations" that differentiate it: take the **last** of them
  - take the subgame starting at that point (if not a singleton, take the first parent node that is)
  - in this subgame, there is a single deviation improving the payoff of player i → contradiction