6. Dynamic Games of Complete Information

Laurent SIMULA

ENS de Lyon

Table of Contents

Repeated games

Often interactions are repeated over time, i.e., a game is played over and over.

- A repeated game $G(T)$ results from T repetitions of a game G, called stage game.
  - It can be finite $(T < \infty)$ or infinite $(T \to \infty)$.
  - The stage game is played at each time period $t = 1, ..., T$ and at the end of each period the action choice of each player is revealed to everybody.
  - A strategy mentions players's decisions at every stage, taking into account the history of the game.
  - A history in time period t is simply a sequence of action profiles from period 1 to period $t - 1$.

Payoff functions

- Different possibilities. For example:
  - If T finite, time-average payoffs:

$$u_i = \frac{1}{T} \sum_{t=1}^{T} u_i\left(t\right).$$

  - If T infinite, discounted payoffs [with $0 < \delta < 1$]:

$$u_i = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} u_i\left(t\right).$$

Payoff functions

- In

$$u_i = (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}u_i\left(t\right),$$

$(1-\delta)$ is a normalization factor that serves to measure the stage game and the repeated game payoffs in the same units.

- Illustration:
  - Getting 2 in the stage game $= 2$.
  - Getting 2 at every period t:

$$u_i = (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\times 2 = \frac{1-\delta}{1-\delta}\times 2 = 2.$$

Remark (useful to compute the payoffs)

- Recall that
$$\sum_{k=t}^{k=T} \delta^{k-1} = \frac{\delta^{t-1} - \delta^{T}}{1-\delta}$$
because
$$\begin{array}{ll} \sum_{k=t}^{k=T} \delta^{k-1} & = \delta^{t-1} + \dots\dots\dots\dots + \delta^{T-1} \\ - \quad \delta \sum_{k=t}^{k=T} \delta^{k-1} & = \qquad + \delta^{t} + \dots\dots + \delta^{T-1} + \delta^{T} \\ \hline (1-\delta) \sum_{k=t}^{k=T} \delta^{k-1} & = \delta^{t-1} - \delta^{T} \end{array}$$
- Note also that, when $|\delta| < 1$, $\delta^{T-1} \longrightarrow 0$ when $T \longrightarrow \infty$.

Table of Contents

Playing the Grade Game several times

- Consider the following grade game or prisoner's dilemma:

1\2	C	D
C	2, 2	0, 3
D	3, 0	1, 1

  where C =cooperation and D =defection.
- Unique Nash equilibrium of the stage game = (Defect,Defect). Pareto-dominated by (Cooperation,Cooperation).
- We often observe cooperation in the real world; what must we add to our model in order that cooperation can become rational? Perhaps cooperation would be rational when players acknowledge that they are in a repeated relationship: that there is actually a *sequence* of stage games, where a player's behavior in a stage can be *conditioned* upon the treatment she/he has received from other players *in the past*.

Defection in the Finite Game

- Consider $G(T)$ with $T < \infty$.
  All players know T from the beginning. What will happen?
- In the last period, each player chooses D (because there is no next period where cooperation could be beneficially profitable).
- By backward induction, the strategy profile (D, D) is obtained as a NE at every $t = 1, ..., T$.

Is cooperation possible?

- Yes! But not always.
  - In the infinite game $[G(T)$ with $T \longrightarrow \infty]$ because there is no last period.
  - In the finite game if players don't know for sure when the game will stop.
- Of course, <u>defection no matter what</u> is a NE of the stage game and thus of the repeated game.

- <u>Basic idea to obtain endogenous cooperation as a NE:</u>
  - Cooperation must be sustained by credible rewards and threats, for every player.
  - Any player playing defection must incur a loss which is larger than what he would gain from keeping playing cooperation.

The Grim-trigger strategy

- Grim-trigger strategy:
  - Play cooperation until a player chooses defection;
  - Then always play defection.
- Assume P2 adopts GT strategy.

The Grim-trigger strategy: Payoffs

Then, P1's expected payoffs are:

- If he cooperates at every period, the outcome is

$$(C, C), ..., (C, C), ...$$

  and his expected payoff is

$$u_i = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} \times 2 = 2.$$

The Grim-trigger strategy: Payoffs (continued)

Then, P1's expected payoffs are:

- If he defects at T, the outcome is

$$\underbrace{(C,C),\dots(C,C)}_{T-1 \text{ times}}\underbrace{(D,C)}_{T}(D,D)\dots.(D,D)\dots$$

and his expected payoff is

$$\begin{aligned} u_i &= (1-\delta)\left[2+\delta 2+\dots+\delta^{T-2}2+\delta^{T-1}3+\delta^{T}1+\delta^{T+1}1+\dots\right] \\ &= (1-\delta)\left[2\times\frac{1-\delta^{T-1}}{1-\delta}+3\times\delta^{T-1}+1\times\frac{\delta^{T}}{1-\delta}\right] \\ &= 2+\delta^{T-1}-2\delta^{T}. \end{aligned}$$

The Grim-trigger strategy: Best Responses (continued)

Player 1's chooses cooperation if and only if the corresponding expected payoff is larger than that obtained when defecting, i.e., iff

$$2 \geqq 2 + \delta^{T-1} - 2\delta^{T} \iff \delta \geqq 1/2.$$

By symmetry, the same result is obtained for Player 2.

Theorem

Each player choosing the Grim-trigger strategy is a NE of the infinitely repeated Prisoner's dilemma game provided $\delta \geqq 1/2$ (i.e., if both players are patient enough).

- Note that, in such a NE, every player's strategy is sustained by credible threats. Hence, nobody has an incentive to unilaterally change his behaviour.
- There are many other strategies that sustain a NE in the infinitely repeated Prisoner's Dilemma.

A second example: the tit-for-tat strategy (TFT)

= Do whatever the other player did in the previous period.
The length of the punishment depends on the behaviour of the player being punished:

- If he continues to choose Defection, then tit-for-tat continues to do so.
- If he reverts to Cooperation, then tit-for-tat reverts to C also.

Tit for tat is an English saying meaning "equivalent retaliation". It is also a highly effective strategy in game theory for the iterated prisoner's dilemma. The strategy was first introduced by Anatol Rapoport in Robert Axelrod's two tournaments, held around 1980. Notably, it was (on both occasions) both the simplest strategy and the most successful. An agent using this strategy will first cooperate, then subsequently replicate an opponent's previous action. If the opponent previously was cooperative, the agent is cooperative. If not, the agent is not. This is similar to superrationality and reciprocal altruism in biology.

Tit-for-tat (TFT)

- Under which conditions is the strategy pair in which each player uses the strategy TFT a NE?
- Assume Player 1 chooses TFT.
- Consider Player 2's behaviour, with t the first period (if any) where he chooses D. Then, Player 1 chooses D in $t+1$ and continues doing so until Player 2 reverts to C.

Tit-for-tat (continued)

- Player 2 has two options from period $t + 1$:
  - revert to C and face the same situation he faces at the start of the game;
  - continue to choose D in which case Player 1 will continue to do so too.
- Hence, if Player 2's best response to TFT chooses D in some period, then it either:
  - alternates between D and C (to get $+3$ half of the time and 0 half of the time);
  - chooses D in every period (to get $+1$ at every repetition of the game).

Tit-for-tat (continued)

- If P2 alternates between D and C:
  - P2's stream of payoffs:

$$\left(\underbrace{3}_{t=1}, 0, 3, 0, 3, 0, \ldots\right).$$

  - P2's discounted average:

$$\begin{aligned} & (1-\delta) \times 3 \times \left(1 + \delta^2 + \delta^4 + \ldots\right) \\ = \quad & 3\frac{1-\delta}{1-\delta^2} = 3\frac{1-\delta}{(1-\delta)(1+\delta)} = \frac{3}{1+\delta}. \end{aligned}$$

Tit-for-tat (continued)

- If P2 plays D at every period:
  - P2's stream of payoffs:

$$\Big(\underbrace{3}_{t=1}, 1, 1, 1, 1, 1, \ldots\Big);$$

  - P2's discounted average:

$$\begin{aligned}(1-\delta)\left[3+\delta+\delta^2+\ldots\right] &= (1-\delta)\left[3+\delta\left(1+\delta+\delta^2+\ldots\right)\right] \\ &= (1-\delta)\left(3+\frac{\delta}{1-\delta}\right) \\ &= 3(1-\delta)+\delta.\end{aligned}$$

- If P2 uses TFT, his discounted average is 2.

Tit-for-tat (continued)

Lemma

For P2, TFT is a best response to P1's TFT, if

$$\left[2 \geqq \frac{3}{1+\delta} \text{ and } 2 \geqq 3\left(1-\delta\right)+\delta\right] \iff \delta \geqq \frac{1}{2}.$$

[By symmetry, the same result holds when reverting P1 and P2.]

Theorem

"Each player playing Tit-for-Tat" is a Nash equilibrium of the infinitely repeated prisoner's dilemma (with payoffs as given above) if $\delta \geqq 1/2$.

[Note that the threshold $\delta = 1/2$ depends on the payoffs given in the bimatrix of the prisoner's dilemma. But, even if they are modified, the same result holds provided players are patient enough].

Tit-for-tat (Example 1)

- Real-life example 1: BitTorrent peers use tit-for-tat strategy to optimize their download speed.
  - BitTorrent peers have a limited number of upload slots to allocate to other peers. Consequently, when a peer's upload bandwidth is saturated, it will use a tit-for-tat strategy. Cooperation is achieved when upload bandwidth is exchanged for download bandwidth. Therefore, when a peer is not uploading in return to our own peer uploading, the BitTorrent program will choke the connection with the uncooperative peer and allocate this upload slot to a hopefully more cooperating peer. Regular unchoking corresponds very strongly to always cooperating on the first move in prisoner's dilemma.
  - Periodically, a peer will allocate an upload slot to a randomly chosen uncooperative peer (unchoke). This is called optimistic unchoking. This behavior allows searching for more cooperating peers and gives a second chance to previously non-cooperating peers.

Tit-for-tat (Example 2-3)

- Real-life example 2: Prosocial behaviour of animals
  - Studies in the prosocial behaviour of animals, have led many ethologists and evolutionary psychologists to apply tit-for-tat strategies to explain why altruism evolves in many animal communities.
  - Evolutionary game theory, derived from the mathematical theories formalised by von Neumann and Morgenstern (1953), was first devised by Maynard Smith (1972) and explored further in bird behaviour by Robert Hinde. Their application of game theory to the evolution of animal strategies launched an entirely new way of analysing animal behaviour.
- Real-life example 3: Wars
  - The tit for tat strategy has been detected by analysts in the spontaneous non-violent behaviour, called "live and let live" that arose during trench warfare in the First World War.
  - Troops dug in only a few hundred feet from each other would evolve an unspoken understanding. If a sniper killed a soldier on one side, the other could expect an equal retaliation. Conversely, if no one was killed for a time, the other side would acknowledge this implied "truce" and act accordingly. This created a "separate peace" between the trenches.

Folk Theorem

- Outcomes that can be obtained as NE in the (infinitely) repeated Prisoner's dilemma.
- Multiplicity of NE known as FOLK THEOREM (name introduced by 2005 Nobel Price Laureate ROBERT AUMANN).

Table of Contents

Subgame Perfect Nash Equilibrium (SPE)

- What is the problem?
- Consider the following game of complete and perfect information:
  - P1 plays first. He can choose L or R.
  - P2 observes P1 action and then play. He can choose K or U.
  - Payoffs:

1\2	KK	KU	UU	UK
L	3,1	3,1	1,3	1,3
R	2,1	0,0	0,0	2,1

NE and Incredible Threats

- What is the extensive-form representation of the game?
- Are there NE which are not sustained by credible threats? (such threats would be irrational to carry out if a player was ever called upon to do so).

1\2	KK	KU	UU	UK
L	3,1	3,1	1,3	1,3
R	2,1	0,0	0,0	2,1

Subgame Perfect Nash Equilibrium (SPE): Principles

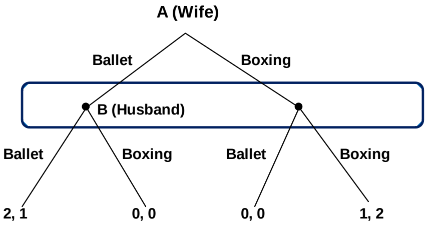
- Defined for extensive-form games.
- Discards NE associated with players' uncredible threats.
- Rests on an extension of backward induction.
- May be viewed as a combination of backward induction and NE.

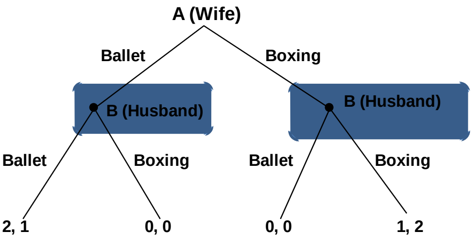
Subgames: definition

A subgame of an extensive-form game has the following properties:

- It begins at a node of the tree corresponding to an information set reduced to a singleton.
- It encompasses all the parts of the tree following the starting node.
- It never divides an information set.

Subgames: Example





SPE: Definition

Definition

A SPE is a NE of every subgame of the game.

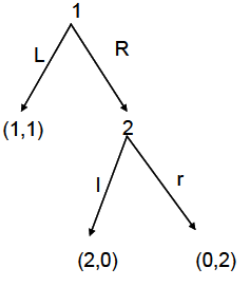
- A SPE is a NE of every subgame of the game, EVEN IF THE SUBGAME IS NOT REACHED!
- Every SPE is a NE.
- Every finite extensive-form game of perfect information has at least one pure strategy SPE, and if no ties occur in players' payoff, then there is a unique SPE.

SPE

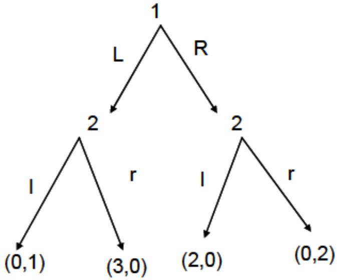
- 1994 Nobel Price Laureate Reinhard Selten proved that any game which can be broken into "sub-games" containing a subset of all the available choices in the main game will have a subgame perfect Nash Equilibrium strategy.

SPE and Backward induction

- SPE can be found through bachward induction [in game of complete information].
- Examples:



Game 1



Game 2

SPE and repeated games

- The concept of SPE can be applied to repeated games.
- When we were examining whether the strategy profiles (TFT, TFT) and (GT, GT) were NE of the repeated prisoner's dilemma, we examined whether they were optimal for each player, given the other's strategy, and whether punishment was credible.
- Hence, we have shown the following:

Theorem

The strategy pairs (TFT, TFT) and (GT, GT) are SPE of the infinitely repeated Prisoner's dilemma (with payoffs as given above) if $\delta \geqq 1/2$.

Example

Army 1, of country 1, must decide whether to attack army 2, of country 2, which is occupying an island between the two countries. In the event of an attack, army 2 may fight, or retreat over a bridge to its mainland. Each army prefers to occupy the island than not to occupy it; a fight is the worst outcome for both armies.

1. Model this situation as an extensive game with perfect information.
2. Show that army 2 can increase its subgame perfect equilibrium payoff (and thus reduce army 1's payoff) by burning the bridge to its mainland (an act which is assumed to entail no cost), eliminating its option to retreat if attacked.

Example: Continued

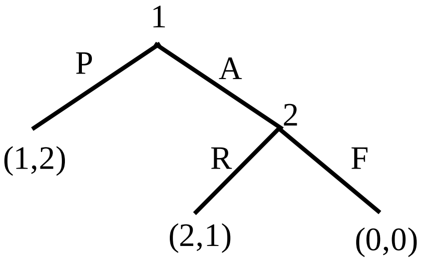


Figure: Burning a Bridge: Extensive Form Game. $P = Peace$; $A = Attack$; $R = Retreat$; $F = Fight$.