Repeated Games: An Elementary Analysis

Pierpaolo Battigalli
Bocconi University
Game Theory: Analysis of Strategic Thinking

November 13, 2023

Abstract

We present and illustrate some elementary results about the uniqueness, or multiplicity of subgame perfect equilibria in a special class of multistage games with observable actions: repeated games with perfect monitoring.

[These slides summarize Chapter 13.1-2 of GT-AST. For the OD Principle in infinite games see Ch. 10.5 of GT-AST.]

Repetition of the Prisoners' Dilemma

- The Prisoners' Dilemma (PD) is the simplest stylized example of social dilemma whereby—unlike perfectly competitive markets without externalities—the pursuit of individual interests leads to a loss for the group (but maybe not for society at large: the group could be a set of firms that try to collude, or even a criminal organization):

G :

1\\$2$	C	D
C	**4**, *4*	**0**, *5*
D	**5**, *0*	**2**, *2*

- Is defection an inevitable result? It depends:
  - Is the PD played only for a finite (commonly known) number of times, or—at least potentially—infinitely often?
  - The role of time is essential: Here, stages and periods coincide; within periods, instantaneous payoffs are realized. How much do players care about future payoffs?

Finitely Repeated PD

- *If the PD game G is played finitely many times* (with a commonly known end), with one-period—possibly discounted—payoffs cumulated in time, *then BI implies permanent defection*:
- In the last period, players must choose the one-period dominant action D.
- Suppose players' expect that (D, D) will be played in the last k periods. Then, at each h with $L(\Gamma(h)) = k + 1$, they expect that *future payoffs are independent of their current actions*. Hence, they choose the one-period dominant action.
- [**Note:** Strong rationalizability yields the same result.]

Infinitely Repeated PD

- *If the PD game G is played infinitely many times* (*with discounting*) *and players are sufficiently patient, then there is a multiplicity of SPEs:*
- Obviously, "always defect" is a SPE: if future payoffs are expected to be independent of current actions, players choose the one-period best reply, D.
- Consider the symmetric "**Nash reversion**" strategy pair s^* whereby
  - players start with C, and keep playing C as along as (C, C) was played in the past;
  - if there is (at least one) deviation D, then they switch forever to the one-period dominant action D.
- If δ (discount factor) is high enough, this is a SPE. Key insight: if (C, C) in the past, playing D triggers (D, D) forever.
  - Relevant comparison in expected present value: $4/(1-\delta)$ if C *vs* $5 + \delta(2/(1-\delta))$ if D.
  - Such s^* is an SPE iff $\frac{4}{1-\delta} \geq 5 + \delta\frac{2}{1-\delta}$ iff $\delta \geq \frac{1}{3}$.

PD Augmented with Punishments

- Now add to the PD a "punishment" action:

G' :

1\2	*C*	*D*	*P*
C	**4**, *4*	**0**, *5*	**-1**, *0*
D	**5**, *0*	**2**, *2*	**-1**, *0*
P	**0**, *-1*	**0**, *-1*	**0**, *0*

- Even if G' is finitely repeated, initial cooperation is SPE-possible.
  - Key observation: G' has two (Pareto ranked) equilibria.
  - Start playing (C, C), then play a one-period eq. in the last k periods, play (P, P) forever after a deviation. A "punishing switch" from playing (D, D) to playing (P, P) in the last k periods triggered by a deviation from (C, C) is consistent with SPE in the last k periods.
  - If such switch is expected at histories h with $L(\Gamma(h)) = k + 1$, the relevant present-value comparison is $\overbrace{4 + \delta \frac{2(1-\delta^{k+1})}{1-\delta}}^{(\text{if } C)} \gtrless \overbrace{5 + 0}^{(\text{if } D)}$.

Repeated Games with Discounting

- Fix a static game $G = \langle I, (A_i, v_i)_{i\in I} \rangle$ with v_i *bounded* for each $i \in I$. The T-**repeated game with** (perfect monitoring, and) **discount factor** $\delta \in (0,1)$ (with $T \in \mathbb{N} \cup \{\infty\}$) is the multistage game with observable actions $\Gamma^{\delta,T}(G) = \langle I, (A_i, \mathcal{A}_i(\cdot), u_i)_{i\in I} \rangle$ with
  - $\mathcal{A}_i(h) = A_i$ for every $i \in I$ and $h \in A^{<\mathbb{N}_0}$ with $\ell(h) < T$;
  - $\mathcal{A}_i(h) = \emptyset$ for every $i \in I$ and $h \in A^T$, if $T < \infty$ (hence, $Z = A^T$);
  - $u_i\left((a^t)_{t=1}^T\right) = \sum_{t=1}^T \delta^{t-1} v_i(a^t)$ for every $i \in I$ and $(a^t)_{t=1}^T \in Z = A^T$.
- **Observations:** To avoid trivialities, let $T \geq 2$; then:
  - Time is key: stages are periods, one-period payoffs realize at the end of each period and are aggregated *via* discounting (each u_i is well defined even if $T = \infty$, because v_i is bounded).
  - $\Gamma^{\delta,T}(G)$ is: meaningless if $\delta = 0$; meaningful if $\delta = 1$ and $T < \infty$.
  - If G is compact-continuous, so is $\Gamma^{\delta,T}(G)$.
  - *The OD principle holds (see below).*

- Consider *any multistage game* Γ. Suppose that $A \subseteq \mathbb{R}^n$ is *bounded*. Fix $\delta \in (0,1)$. For each $T \in \mathbb{N} \cup \{\infty\}$, endow A^T with the following "**discounting metric**":

$$d_T\left(\left(a^t\right)_{t=1}^T, \left(\bar{a}^t\right)_{t=1}^T\right) = \sum_{t=1}^{T} \delta^{t-1} d\left(a^t, \bar{a}^t\right)$$

(d is the metric in $\mathbb{R}^n$; by boundedness and $0 < \delta < 1$, d_T is a metric even if $T = \infty$). Thus, $\left(A^T, d_T\right)$ is a *metric space*. Let $Z_T := Z \cap A^T$ be the set of terminal histories of length T.

Definition

Game Γ is **compact-continuous** if Z_T is *compact* in metric space $\left(A^T, d_T\right)$ for each $T \in \mathbb{N} \cup \{\infty\}$ and u_i is *continuous* on Z_T for each $T \in \mathbb{N} \cup \{\infty\}$ and $i \in I$.

- [**Recall:** A subset K of a metric space is compact if, for every cover of K with open sets, there is a finite sub-cover of K.]

Intermezzo I: Infinite Games 2, One-Step Optimality

- We take folding-back (FB) optimality as our basic notion of rational planning. But, by definition, the *FB algorithm cannot be applied to infinite-horizon games.*
- If the game has *finite horizon*, but it is infinite (because some feasible action set $\mathcal{A}_i(h)$ is infinite), then maximizations may be impossible (we will study a prominent example concerning bargaining).
- But the definitions (with sup) still apply (as written, if each $\beta^i(\cdot|h)$ has finite/countable support) and versions of the *FB, Optimality, and OD principles hold.*
- With this, we take the *one-step optimality* (OSO) as our *general* characterization of *rational planning*. [**Note**: OSO is also relevant for sophisticated agents with dynamically inconsistent preferences, e.g., because of non-exponential discounting.]

Intermezzo I: Infinite Games 3, OD Principle

- The following result extends the OD principle to compact-continuous games.

Theorem

(**Generalized OD principle**) *In every compact-continuous game the OD principle holds: for every i, s_i, and β^i, strategy s_i is sequentially optimal given conjecture β^i IFF s_i is one-step optimal given β^i.*

- **Intuition** (by *contraposition*): If s_i is not sequentially optimal given β^i in the compact-continuous game Γ, then we can find a finite-horizon approximation of Γ, viz. $\bar{\Gamma}$, such that the restriction of s_i to $\bar{\Gamma}$ is not sequentially optimal in $\bar{\Gamma}$ given (the restriction of) β^i; hence (by the OD principle for finite-horizon games), it fails one-step optimality in $\bar{\Gamma}$. Given that $\bar{\Gamma}$ is a sufficiently good approximation of Γ, s_i must fail one-step optimality in Γ. ♡
    - Futher generalization: it is enough that Γ satisifes the weaker property of **"continuity at infinity"** (see book).

Intermezzo II: Strategies and Automata

- Back to *repeated games:* the set of **non-terminal histories** is
$$H = A^{<\mathbb{N}_0} \text{ if } T = \infty \text{ and } H = A^{<T} := \bigcup_{t=0}^{T-1} A^t \text{ if } T < \infty.$$
- The set of strategies of i is $S_i = (A_i)^H$. (How many strategies does i have if $\Gamma^{\delta,T}(G)$ is finite?)
- Convenient representation of strategy profiles (and, similarly, strategies), especially if $T = \infty$, with **automata**, i.e., structures $(\Psi, \psi_0, \gamma, \varphi)$ where
  - Ψ is a set of **states** (interpret as players' "moods");
  - $\psi_0 \in \Psi$ is the **initial state**;
  - $\gamma : \Psi \to A$ is the **behavioral rule**;
  - $\varphi : \Psi \times A \to \Psi$ is the **transition function**.
- **Example**: The "Nash reversion" strategy pair s^* in the infinitely-repeated PD is represented with $\Psi = \{\mathbf{c}, \mathbf{d}\}$, $\psi_0 = \mathbf{c}$, $\gamma(\mathbf{c}) = (C, C)$, $\gamma(\mathbf{d}) = (D, D)$, $\varphi(\mathbf{c}, (C, C)) = \mathbf{c}$, $\varphi(\mathbf{c}, a) = \mathbf{d}$ if $a \neq (C, C)$, and $\varphi(\mathbf{d}, a) = \mathbf{d}$.

Sequences of One-Period Equilibria and SPE

Theorem

*(**One-period NEs**) Let $(a^t)_{t=1}^T \in NE(G)^T$ and let $\bar{s}$ be defined by $\bar{s}(h) = a^t$ for all t and h with $1 \leq t < T$ and $h \in A^{t-1}$. Then $\bar{s}$ is an SPE of $\Gamma^{\delta,T}(G)$.*

- **Proof:** Apply the OD principle. Note: the *behavior described by* $\bar{s}$ may depend on calendar time, but it *is independent of past actions* $\Rightarrow$ *future behavior is expected to be independent of current choice.*
    - No incentive to deviate if, for all $1 \leq t < T$, $h \in A^{t-1}$, and $i \in I$,

$$\forall a_i \in A_i,\ v_i(a^t) + \sum_{k=t+1}^{T} \delta^{k-t} v_i(a^k) \geq v_i(a_i, a_{-i}^t) + \sum_{k=t+1}^{T} \delta^{k-t} v_i(a^k)$$

      (where $\sum_{k=t+1}^{T} \delta^{k-t} v_i(a^k) = 0$ if $t = T < \infty$).
    - The 2nd terms of both sides cancel out: $\forall i \in I, \forall a_i \in A_i$, $v_i(a^t) \geq v_i(a_i, a_{-i}^t)$ satisfied because $a^t \in NE(G)$.
    - By the OD principle, $\bar{s}$ is a SPE. ■

Unique SPE in Finitely Repeated Games

Theorem

(**Unique SPE**) *Suppose that G has a unique Nash equilibrium a° and $T<\infty$; then the unique SPE of $\Gamma^{\delta,T}(G)$ is the profile s° with $s^\circ(h)=a^\circ$ for every $h\in H$.*

- **Proof:** Recall: by def., sequential optimality⇒One-Step Optimality.
  - By **Theorem One-Period NEs**, s° is a SPE. Prove by induction on $L(\Gamma(h))$ that $(s\ \mathrm{SPE})\Rightarrow s=s^\circ$. Let s *be* a *SPE*; then each s_i *satisfies One-Step Optimality* given s_{-i}.
  - *Basis step:* $L(\Gamma(h))=1$ $(h\in A^{T-1})$. By the OSO of each s_i given s_{-i}, $s(h)\in NE(G)=\{a^\circ\}$, that is, $s(h)=a^\circ=s^\circ(h)$.
  - *Inductive step:* Suppose that $s(h')=a^\circ$ for each h' with $L(\Gamma(h'))\le k$ (IH). Let $L(\Gamma(h))=k+1$. By IH and OSO of each s_i given s_{-i}, $\forall i\in I,\forall a_i\in A_i$, $v_i(s(h))+\sum_{\ell=1}^k\delta^\ell v_i(a^\circ)\ge v_i(a_i,s_{-i}(h))+\sum_{\ell=1}^k\delta^\ell v_i(a^\circ)$.
  - The 2nd terms of both sides cancel out: $\forall i\in I,\forall a_i\in A_i$, $v_i(s(h))\ge v_i(a_i,s_{-i}(h))$. Thus, $s(h)\in NE(G)=\{a^\circ\}$, that is, $s(h)=a^\circ=s^\circ(h)$. ■

Multiplicity of SPEs: Examples

- If either assumption of **Theorem Unique-SPE** fails, then there may be a multiplicity of SPEs, where some SPEs s^* prescribe $s^*(h) \notin NE(G)$ at least in early periods, provided players are patient.
- Consider $G = PD$ and $\Gamma^{\delta,\infty}(G)$. Then, if $\delta \geq 1/3$, the strategy pair s^* described by $\Psi = \{\mathbf{c}, \mathbf{d}\}$, $\psi_0 = \mathbf{c}$, $\gamma(\mathbf{c}) = (C, C)$, $\gamma(\mathbf{d}) = (D, D)$, $\varphi(\mathbf{c}, (C, C)) = \mathbf{c}$, $\varphi(\mathbf{c}, a) = \mathbf{d}$ if $a \neq (C, C)$, $\varphi(\mathbf{d}, a) = \mathbf{d}$ for each $a \in A$ is a SPE.
- Consider $G' = PD + \text{punishment}$ and $\Gamma^{\delta,2}(G')$. Let $s^*(\varnothing) = (C, C)$, $s^*((C, C)) = (D, D)$, $s^*(a) = (P, P)$ if $a \neq (C, C)$. Then s^* is an SPE if $4 + 2\delta \geq 5$, that is, $\delta \geq 1/2$.

Multiplicity in Infinitely Repeated Games: Nash Reversion

Theorem

*(**Nash-Reversion**) Let G be such that for some $a^\circ \in NE(G)$ and $a^* \in A$, a^* strictly Pareto-dominates a°, that is,*

$$\forall i \in I,\ v_i(a^*) > v_i(a^\circ).$$

Consider $\Gamma^{\delta,\infty}(G)$ and the profile s^* described by $\Psi = \{\mathbf{c}, \mathbf{d}\}$, $\psi_0 = \mathbf{c}$, $\gamma(\mathbf{c}) = a^*$, $\gamma(\mathbf{d}) = a^\circ$, $\varphi(\mathbf{c}, a^*) = \mathbf{c}$, $\varphi(\mathbf{c}, a) = \mathbf{d}$ if $a \neq a^*$, $\varphi(\mathbf{d}, a) = \mathbf{d}$ for each $a \in A$. Then, s^* is an SPE if and only if,

$$\forall i \in I,\ v_i(a^*) \geq (1-\delta) \sup_{a_i \in A_i} v_i\left(a_i, a^*_{-i}\right) + \delta v_i(a^\circ), \tag{IC}$$

that is, iff

$$\forall i \in I,\ \delta \geq \bar{\delta}_i(a^\circ, a^*) := \frac{\sup_{a_i \in A_i} v_i\left(a_i, a^*_{-i}\right) - v_i(a^*)}{\sup_{a_i \in A_i} v_i\left(a_i, a^*_{-i}\right) - v_i(a^\circ)}.$$

Comments on Nash-Reversion Theorem

- It is an abstract, general version of the result about the ∞-repeated PD.
- The hypotheses imply: threshold $\bar{\delta}_i(a^\circ, a^*) \in [0,1)$.
- (The automaton describing) s^* starts with a^* (cooperative state $\mathbf{c}$, $\gamma(\mathbf{c}) = a^*$) and switches forever to a° as soon as a deviation from a^* occurs (non-cooperative state $\mathbf{d}$, $\varphi(\mathbf{c}, a^*) = a^*$, $\varphi(\mathbf{c}, a) = \mathbf{d}$ if $a \neq a^*$, $\varphi(\mathbf{d}, a) = \mathbf{d}$ for all a). Thus deviations from a^* *"trigger"* permanent defection.
- Most economists interpret SPE as a *self-enforcement requirement* for non-binding, self-enforcing agreements (to play a strategy profile).
- With this, the result is widely used to analyze cooperation, e.g., among sovereign states, collusion among firms, and organized crime.
- G is *not* assumed compact-continuous: it may be a Bertrand oligopoly (discontinuous). That is why we write sup instead of max.

Proof of the Nash-Reversion Theorem

- We apply the OD principle (which holds even if G is not compact-continuous, because $\Gamma^{\delta,\infty}(G)$ satisfies continuity at infinity). We only need to check that there are no incentives for one-shot deviations.
- There are two types of finite histories: those without defections, $h = \varnothing$ and $h = (a^*, ..., a^*)$, and the others.
- If a defection occurred in h, then $s^*(h') = a^\circ$ for each (finite) $h' \succeq h$. Thus, no incentive to deviate (see **Theorem One-period NEs**).
- If no defection occured in h, then there is no incentive to deviate iff

$$\forall i \in I, \forall a_i \in A_i, \ \frac{1}{1-\delta} v_i(a^*) \geq v_i\left(a_i, a^*_{-i}\right) + \frac{\delta}{1-\delta} v_i(a^\circ),$$

which is equivalent to condition (IC). ■

References

Battigalli, P., E. Catonini, and N. De Vito (2023): *Game Theory: Analysis of Strategic Thinking*. Typescript, Bocconi University.

Battigalli, P. (2023): *Mathematical Language and Game Theory*. Typescript, Bocconi University.

REPEATED GAMES AND AUTOMATA

Note Title 4/23/2020

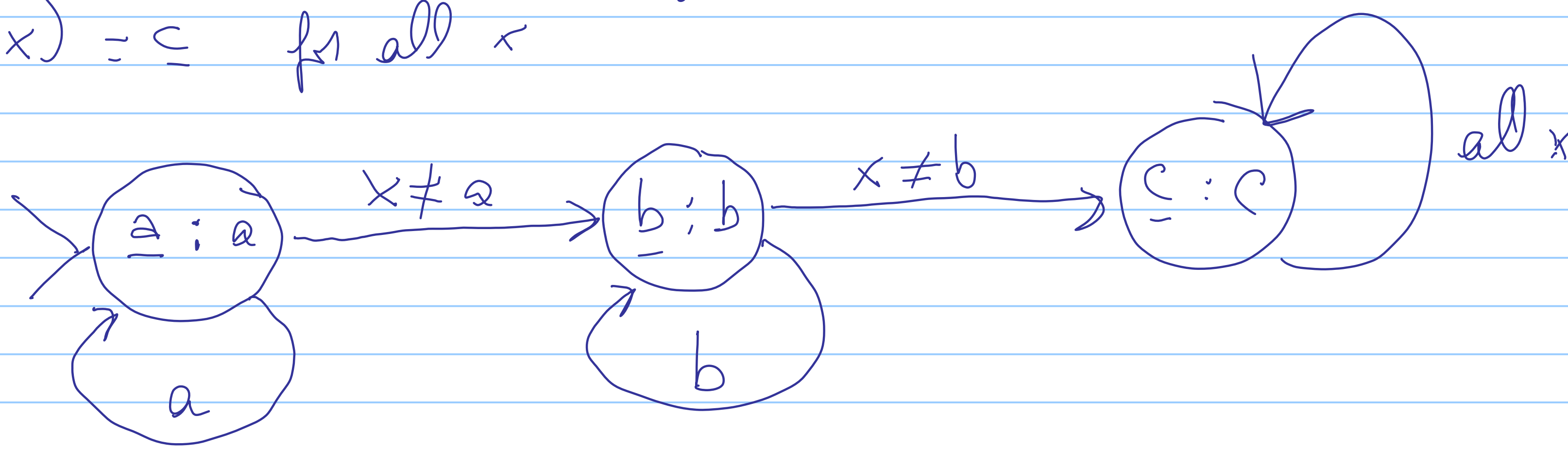
Example of automaton with 3 states $\Psi = \{\underline{a}, \underline{b}, \underline{c}\}$

$\psi_0 = \underline{a}$ $\quad \gamma(\underline{a}) = a$, $\gamma(\underline{b}) = b$, $\gamma(\underline{c}) = c$

$\varphi(\underline{a}, a) = \underline{a}$, $\varphi(\underline{a}, x) = \underline{b}$ if $x \neq a$

$\varphi(\underline{b}, b) = \underline{b}$, $\varphi(\underline{b}, x) = \underline{c}$ if $x \neq b$

$\varphi(\underline{c}, x) = \underline{c}$ for all x



TRIGGER STRATEGIES

$a^\circ \in NE(G)$ $\quad \forall i \in I,\ v_i(a^*) > v_i(a^\circ)$

$\Psi = \{\underline{c}, \underline{d}\}$ $\quad \psi_0 = \underline{c}$, $\quad \gamma(\underline{c}) = a^*$, $\gamma(\underline{d}) = a^\circ$

$\varphi(\underline{c}, a^*) = a^*$, $\quad \varphi(\underline{c}, a) = \underline{d}$ if $a \neq a^*$

$\varphi(\underline{d}, a) = \underline{d}$ for all a