Introduction to Game Theory

Part 1: Static Games

Gabriel Ziegler

3rd Summer School & Workshop on Experimetrics & Behavioral Economics

Soleto, July 2025

Agenda

1. Introduction to Static Games
   - Game Form (the rules)
   - Exogenous personal features (tastes)
2. Beliefs: 1st-order and 2nd-order
3. Best replies
4. Predictions: Iterated Elimination of Never-Best Replies

Key References for Game Theory

- BATTIGALLI, P. (2025): *Mathematical Language and Game Theory.* Typescript, Bocconi University. [Downloadable from Pierpaolo Battigalli's webpage]
- BATTIGALLI, P., E. CATONINI, AND N. DE VITO (2025): *Game Theory: Analysis of Strategic Thinking.* Typescript, Bocconi University. [Downloadable from Pierpaolo Battigalli's webpage]

Key References for Psychological Game Theory (PGT)

- Battigalli, P., and M. Dufwenberg (2022): "Belief-Dependent Motivations and Psychological Game Theory," *Journal of Economic Literature*, 60, 833–882.
- Battigalli, P., and M. Dufwenberg (2009): "Dynamic Psychological Games," *Journal of Economic Theory*, 144, 1–35.
- Battigalli, P., R. Corrao, and M. Dufwenberg (2019): "Incorporating Belief-Dependent Motivation in Games," *Journal of Economic Behavior & Organization*, 167, 185–218.

Game theory

Game theory is the formal analysis of *interactive decision making*

- interactive situations among *n* individuals called **players**
- some (or all) take **actions**, which affect the **outcome/consequences** for everybody

  We will focus on monetary outcomes only!

We differentiate two classes of games:

- **Static Games:** Active players move *simultaneously once and for all*
- **Dynamic Games:** moves are *sequential* (some of them may be simultaneous)

  As the game unfolds, new information leads to belief-updating & revision.

Game Form vs. Game

It is crucial to distinguish:

- the description of the "*rules of the game*" (in experiments, controlled by the experimenter), called **game form**, from
- the description of the *exogenous personal features* of the participating individuals, such as their "tastes" (e.g., preferences over lotteries of outcomes).

Game = (Game form) + (description of players' exogenous features)

Examples of static game forms (not games!)

Simple static game forms are used to model/simulate social dilemmas and are often implemented in the lab (all numbers are monetary outcomes of the players)

Prisoner's dilemma:

Player 2

Player 1		c_2	d_2
	c_1	€3, €3	€0, €4
	d_1	€4, €0	€1, €1

- selfish behavior (defection) yields inefficiency (unlike perfectly competitive markets!)

Coordination:

Player 2

Player 1		l_2	r_2
	l_1	€1, €1	€0, €0
	r_1	€0, €0	€1, €1

- find a convention to coordinate [e.g., drive on the right (EU), or on the left (UK)].

Examples of static game forms (not games!) cont.

Stag Hunt:

		Player 2	
		s_2	h_2
Player 1	s_1	€5, €5	€0, €3
	h_1	€3, €0	€3, €3

- Originated from philosopher Jean-Jacques Rousseau in his "Discourse on Inequality"
- Coordination on a risky action (hunt for a stag together) achieves first best
- or, alternative, go for the hare alone, which is a safe option

From a game form to a game: enter personal features

Standard economics describes tastes by means of utilities of outcomes.

That is, letting Y denote the set of outcomes (often $Y \subseteq \mathbb{R}^n$), a utility function for player i is

$$v_i : Y \to \mathbb{R}.$$

Examples for $y = (y_1, \ldots, y_n)$ denoting the *profile* of monetary outcomes:

- $v_i(y) = v_i(y_1, \ldots, y_n) = y_i$ (selfish and risk-neutral)
- $v_i(y) = V_i(y_i)$, $V_i : \mathbb{R} \to \mathbb{R}$ is strictly increasing and concave (selfish and risk-averse)
- $v_i(y) = y_i + \gamma_i \sum_{j \neq i} y_j$, with $\gamma_i \in (0, 1)$ (risk-neutral and partially altruistic)

From a game form to a game: beyond standard economics

Non-standard economics allows utility to depend on more than the *material* outcomes

Examples:

- Chosen actions matter per se (as in, e.g., “worm glow of giving”)

  modelled via $u_i(a)$ where $a = (a_1, \ldots, a_n)$ is an *action profile*

  Can use standard GT as we will see!

- In **psychological game theory (PGT)** also beliefs matter, including those of others.

  Will have something like

  $$u_i(a_1, \text{belief}_1, \ldots, a_n, \text{belief}_n)$$

  Even in static games: need to differentiate **initial** vs **ex-post** beliefs. Here, we will only consider initial beliefs, but ex-post ones are important for some psychological reasons.

Formally: Game Forms (the rules)

Definition

A **game form** is a mathematical structure $(I, (A_i)_{i \in I}, Y, g)$, where

1. I is the set of players
2. A_i is the set of actions for player i (Notation: $A = \prod_{j \in I} A_j = A_1 \times \cdots \times A_n$)

   $a_i \in A_i$ is player i's action, $a = (a_1, \ldots, a_n) = (a_j)_{j \in I} \in A$ is an *action profile*
3. Y is the set of outcomes/consequences
4. $g : A \to Y$ is the outcome function

For example, in PD we have $g(c_1, c_2) = (€3, €3)$

Formally: Games (rules + players' exogenous personal features)

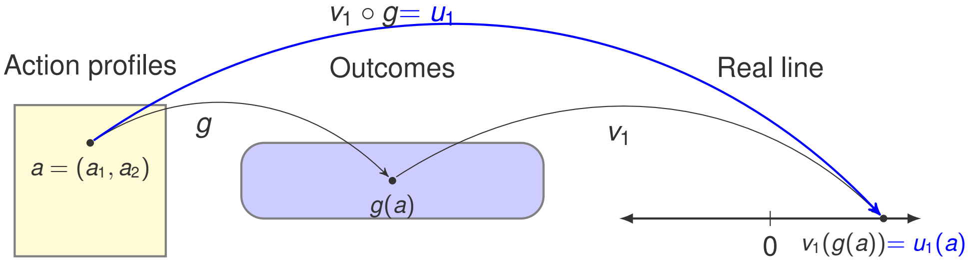
Definition

A **game** is a mathematical structure $(I, (A_i)_{i \in I}, Y, g, (v_i)_{i \in I})$, where

1. $(I, (A_i)_{i \in I}, Y, g)$ is a game form and
2. $v_i : Y \to \mathbb{R}$ is player i's utility function

PGT: $v_i : Y \times \prod_{j \in I} B_j \to \mathbb{R}$, where B_j is the set of player j's "beliefs"

Reduced game



Definition

A (reduced) **game** is a $G = (I, (A_i)_{i \in I}, (u_i)_{i \in I})$, where I and A_i as before and

$u_i : A \to \mathbb{R}$ is player i's (reduced form) utility function aka **payoff function**

PGT: $u_i : A \times \prod_{j \in I} B_j \to \mathbb{R}$

Beliefs as subjective probability measures

In order to decide what to do, players form **subjective beliefs** (subjective probability measures) about the relevant unknowns, such as the actions of others.

For this a quick recap on probabilities and the simplex: Finite set $X = \{x_1, \ldots, x_m\}$, want the set of all probability measures(here: can look at probability mass functions) on X

- $\mu : X \to [0, 1]$ is a probability mass function if $\sum_{x \in X} \mu(x) = \sum_{i=1}^{m} \mu(x_i) = 1$

  Note: μ is the same as specifying numbers $\big(\mu(x_1), \ldots, \mu(x_m)\big) \in [0, 1]^m$
- The $(m-1)$-**simplex** is $\Delta(X) := \left\{\mu : X \to [0, 1] \,\middle|\, \sum_{x \in X} \mu(x) = 1\right\}$
  - Allows for $\mu(x_k) = 0$ for some x_k. On the other hand, $\mu(x_k) = 1 \mathrel{\hat{=}} \delta_{x_k}$ means **certainty** of x_k
  - Can be generalized to infinite X

First-order beliefs

(Probabilistic) beliefs about "primitive uncertainty," such as: "How will the game be played?"

At the planning stage, i is uncertain and we model this via probabilistic beliefs

- **1st-order belief about *others*:** $\alpha_{i,-i} \in \Delta(A_{-i})$

  GT notation: $-i$ means "players other than i" and $A_{-i} := \prod_{j \neq i} A_j$

- **1st-order belief about *oneself* (plan of i):** $\alpha_{i,i} \in \Delta(A_i)$

  Typically, $\alpha_{i,i}(a_i) = 1$ for some $a_i \in A_i$: i is certain of her (planned) action

- **1st-order belief:** $\alpha_i \in \Delta(A)$ s.t. $\alpha_i(a_i, a_{-i}) = \alpha_{i,i}(a_i) \cdot \alpha_{i,-i}(a_{-i})$ (i.e., independence)

- Δ_i^1 denotes the set of all of i's 1st-order beliefs

Recall: We only consider **initial beliefs** here, not ex-post!

Best replies in standard GT

Given payoff function $u_i : A \to \mathbb{R}$ and 1$^{\text{st}}$-order belief about others $\alpha_{i,-i}$,

calculate EU for action $a_i \in A_i$:

$$u_i(a_i, \alpha_{i,-i}) = \sum_{a_{-i} \in A_{-i}} \alpha_{i,-i}(a_{-i}) u_i(a_i, a_{-i})$$

The **best-reply correspondence** associates each 1$^{\text{st}}$-order belief about others, $\alpha_{i,-i}$, with the *set* $BR_i(\alpha_{i,-i}) \subseteq A_i$ of actions that maximize EU given $\alpha_{i,-i}$:

$$BR_i(\alpha_{i,-i}) := \arg\max_{a_i \in A_i} u_i(a_i, \alpha_{i,-i})$$

Exercise: Find the best-reply correspondence for the three games above

with $v_i(y) = y_i$ and $v_i(y) = y_i + \frac{1}{2} y_{-i}$

Towards best replies in PGT

PGT allows the utility depends on outcomes/actions of everybody to depend and on beliefs of everybody. We just consider dependence on first-order beliefs: $u_i : \prod_{j \in I} \left(A_j \times \Delta_j^1\right) \to \mathbb{R}$.

1. **Own-plan-*in*dependence:** u_i does not depend on i's plan $\alpha_{i,i}$

   Example: other things equal, i dislikes to make others earn less material payoff than they expect (not to live up to others' expectations)

2. **Own-plan-dependence:** u_i does depend on i's plan $\alpha_{i,i}$

   Example: other things equal, i dislikes to be disappointed (to get less material payoff than she expected)

   - NOTE: How much i expects to get depends also on her plan, because i's payoff depends also on what she does

Need 2nd-order beliefs

A 2nd-order beliefs is *i*'s subjective probability measures about primitive uncertainty **and** others' 1st-order beliefs (*i* knows her own belief by introspection):

$$\beta_i \in \Delta\left(A_i \times \prod_{j\neq i} (A_j \times \Delta_j^1)\right)$$

- Recover 1st-order beliefs from 2nd-order one: for each action profile $a = (a_1, \ldots, a_n)$:

$$\alpha_i(a) = \beta_i\left(\{a_i\} \times \prod_{j\neq i} (\{a_j\} \times \Delta_j^1)\right)$$

- The really important part of β_i is its marginal $\beta_{i,-i} \in \Delta\left(\prod_{j\neq i} (A_j \times \Delta_j^1)\right)$
- **Actions and beliefs of others are not independent!**

Best replies in PGT

- The really important part of β_i is its marginal $\beta_{i,-i} \in \Delta\Big(\prod_{j \neq i} (A_j \times \Delta_j^1)\Big)$
- **Actions and beliefs of others are not independent!**

Define (a special case of) the **BR correspondence** (given *own-plan independence*) in a PGT:

- Here: $u_i : A_i \times \prod_{j \neq i} (A_j \times \Delta_j^1) \to \mathbb{R}$
- Given $\beta_{i,-i}$ and action a_i, can calculate expected utility as usual and get $u_i(a_i, \beta_{i,-i})$
- Then, the BR correspondence is given by

$$BR_i(\beta_{i,-i}) := \arg\max_{a_i \in A_i} u_i(a_i, \beta_{i,-i})$$

Best replies in PGT: Exercise

Let $[x]^+ = \max\{0, x\}$, and let δ_a denote the **deterministic belief** that $a = (a_i)_{i \in I}$ is played with certainty (i.e. with probability 1). Suppose that, in the *Prisoner's dilemma* with

$$u_1(a_1, a_2, \alpha_2) = g_1(a_1, a_2) - \frac{1}{2}\left[\mathbb{E}_{\alpha_2}(g_2) - g_2(a_1, a_2)\right]^+,$$

where $g_i(a)$ indicates the material payoff of player i for a given action profile a.

Compute the best reply of Player 1 to any 2$^{\text{nd}}$-order belief $\beta_{1,2}$ such that

$$\beta_{1,2}\left(c_2, \delta_{(c_1,c_2)}\right) + \beta_{1,2}\left(d_2, \delta_{(c_1,d_2)}\right) = 1.$$

Predictions

GT derives predictions from **solution concepts** that associate each game with a corresponding set of possible action (and belief) profiles $(a_i)_{i \in I}$ (or $(a_i, \alpha_i)_{i \in I}$ or $(a_i, \beta_i)_{i \in I}$).

Examples:

- **Nash equilibrium (pure)**: In traditional GT, any profile $(a_i^*)_{i \in I}$ satisfying

$$(\forall i \in I) \quad a_i^* \in BR_i(a_{-i}^*) = \arg\max_{a_i \in A_i} u_i(a_i, a_{-i}^*)$$

- **Iterated elimination of never best replies (rationalizability)**:
  - Step 1. Eliminate all actions that are not best replies to any belief.
  - Step $n > 1$. Eliminate all the (remaining) actions that are not best replies to beliefs consistent with steps $1, \ldots, n-1$ (hence, which assign probability 0 to eliminated actions).

Predictions

We are interested in experiments, where there is little reason to use NE to make predictions. Sometimes predictions are derived from 2–3 rounds of iterated elimination of never best replies. Often we “elicit” (i.e., measure) players’ beliefs and make predictions based on BR correspondences.

Iterated elimination of never best replies: Traditional GT

Consider the following game form and suppose that for each $i = 1, 2$, $v_i(y) = y_i$:

	~~L~~	M	~~R~~
T	~~€1, €0~~	€1, €2	~~€0, €1~~
~~B~~	~~€0, €3~~	~~€0, €1~~	~~€2, €0~~

1. If P2 is rational, *R* is deleted, because *R* is never a best reply

(given any belief of P2, *M* is better than *R*)

2. If P1 maintains believes that P2 is rational, then *B* is never a best reply and is deleted.

3. If P2 reasons as above about P1, *L* is never a best reply, and is deleted.