Chapter 1

Introduction and Overview

The course covers two main topics: *game theory* and *mechanism design*. We begin by introducing the fundamental concepts and topics in game theory and will then apply these to topics in mechanism design.

1.1 Game Theory

Game theory is a method for analyzing *strategic interactions*. We consider a situation to be strategic whenever the consequences of a choice by one individual depend not only on the individual's behavior alone but also on the behavior of other individuals. For example, suppose that Apple introduces a new iPhone and considers the price it will charge for it. Clearly, the demand for the new iPhone will depend on the price chosen by Apple—however, if we reasonably assume that the latest Samsung Galaxy is, at least to some extent, comparable to the new iPhone, then the two are substitutes, and the demand for the iPhone will also depend on the price chosen by Samsung for the latest Samsung Galaxy. Pricing a product in a market with few competitors (a so-called oligopoly) is a situation of strategic interaction that we can study with the tools of game theory. Notice that, in contrast, the pricing decision of a monopolist is *not* a strategic situation: A monopolist does not face competition and therefore solves a decision-theoretic rather than a game-theoretic problem.

1.1.1 Examples

There are many applications of game theory from a wide range of fields and topics.

1. How should two presidential candidates design their political agenda to maximize their chances of being elected?

2. A pedestrian wants to cross a street and a driver wants to drive through the crossing. Should the pedestrian wait for the car to pass or should the pedestrian cross the street? How should the driver of the car behave?

3. How should a football player decide where to kick the penalty? How does the goalie decide in which direction to jump?

4. Generative adversarial networks are inspired by the theory of zero-sum games: a generator generates new artificial examples (for example, fake photographs of celebrities) with the aim of fooling a discriminator. The discriminator attempts to distinguish generator-generated examples from real examples (for example, actual photographs of celebrities).

In the chapters on game theory, we will learn to describe such real-world situations as games using mathematical models. Based on this modeling, we will introduce concepts that allow us to understand individuals' behavior in games *given an economic environment*; that is, given the rules of the games.

1.1.2 Different Classes of Games

The tools of game theory have been applied to a variety of different strategic interactions. Depending on the context of the environment to be analyzed, different features of the strategic situation must be taken into account. The following is a broad overview of distinctions in the modeling framework.

Cooperative and Noncooperative Games. Game theory broadly consists of two main branches: cooperative and non-cooperative game theory. As the name suggests, the distinction is in the nature of the decision-making process of the players. In cooperative game theory, players can engage in *binding* pre-game agreements. That is, players can "sit together" before taking their actions in the actual game to be played and write a contract that they can enforce in court, specifying how they will behave in the game. As that contract is enforceable, the players will follow their agreement.

Why would players want to write such agreements? Such agreements can be beneficial because they may prevent bad outcomes that arise due to *individual incentives* to take different actions. For example, consider the decision of two countries whether to build nuclear weapons. If the presidents of the two countries can sign a binding agreement that they will never build such a nuclear weapon, there will be a nuclear-weapon-free world. If, however, such an agreement is not possible, it may be optimal, in a non-cooperative game, for each of the countries individually to build a nuclear weapon: If the other country has none, having one may help exert power over that country. If the other country has one, having one may help defend your interests viz-a-viz that country. Both countries might end up having nuclear weapons—although they would have preferred not to build one, all else equal.

The focus of our class and these notes will be on non-cooperative games. However, even in non-cooperative games, players might engage in similar cooperative agreements. To include this possibility, the negotiation process leading up to such agreements also has to be explicitly modeled.

Static and Multistage Games. One important decision when modeling a strategic situation is the timeline of the interaction and the sequence in which participants act. For example, when visiting a local bar as a tourist, it is likely that the barkeeper recognizes you as such. Therefore, the mutual understanding will be that your interaction is only of short duration (and probably even a one-time interaction only). In contrast, if you are a local visiting the same bar, the barkeeper may recognize you as a local. In this case, it is more likely that you will come back. In particular, you can condition your choice of returning to the bar based on the quality of the drinks and the service.

The preceding example illustrates the distinction between one-time interactions and more dynamic, repeated strategic situations. However, beyond this distinction, even a one-time interaction may be of a more dynamic nature. For example, as a tourist visiting a bar only once, you can decide how much to tip depending on the quality of the service in the bar. Such a setting is not static: the barkeeper first decides how friendly to be or how much effort to put into mixing a high-quality drink for you. Only after learning about the barkeeper's behavior do you decide on the tip, and, in particular, you may tip differently depending on the barkeeper's behavior.

In other settings, the interaction may be simultaneous in nature. For example, consider two innovative pharmaceutical companies. Research on novel drugs is usually highly secretive, as firms want to ensure that competitors cannot free-ride on insights previous internal research has delivered. Therefore, when deciding on the focus of R&D spending in the current year, managers of different firms act without knowing what the other firms focus on. Hence, the firms act simultaneously: they cannot condition their choice on other firms' choices. Note that this is even true if the decisions do not happen literally simultaneously. The crucial point is that these decisions occur *without knowing* what other firms decided.

Is the choice of the R&D focus a static or a multistage game? That depends. If we think of pharmaceutical firms as long-lived entities competing with a similar set of firms year after year, then we should model their choices as a multistage game, where, in each stage (e.g., year), the firms *simultaneously* choose the focus of their research efforts. If we think of them as short-lived or as competing with different firms year after year, then we can model their choices as a static game (that is, a one-time interaction with simultaneous moves).

A simpler example of a static game can be the occasional Rock-Paper-Scissors game that you play with your roommates to decide who does the dishes or at the beginning of a friendly soccer match to decide on the kickoff. Usually, there is a negligible dynamic component to such games, and by definition, players act simultaneously in that game.

Perfect and Imperfect Information Games. Another crucial feature of strategic situations is the information available to players when making their decisions. In a game of perfect information, a player can observe the past moves of all other players. Chess, for example, is a perfect information game. Scotland Yard, in contrast, is a game of imperfect information, as some players cannot observe the moves of other players.

Complete and Incomplete Information Games. In a game of complete information, players have access to all the relevant information about the game. For example, in chess, a player knows which actions the other player has available and that the other player does not want to lose the game. However, in many other real-world strategic situations, a player might not know what other players aim for or other relevant pieces of information.

Poker is an example of an incomplete information game. Players do not know their opponent's hand. Hence, they are uncertain about the consequences of their actions. Similarly, a buyer of a good might not know whether the good is of high or low quality.

We will see later in the class that there is an elegant way to establish a close conceptual connection between imperfect and incomplete information games, which allows us to study them with the same tools.

1.2 Mechanism Design

Mechanism design, on the contrary, uses game theory as input and assumes that individuals behave in strategic situations according to the concepts of game theory. Based on this hypothesis, mechanism design aims to answer questions about the optimal design of a game to achieve a particular objective with limited information. For example, suppose that a seller wishes to sell an object for the highest possible revenue. Should the seller put the object with a price tag into a store and the first buyer willing to pay the price receives the object? Or should the seller rather run an auction to sell the object? In this sense, mechanism design can be viewed as reverse game theory: there is an institution that wants to achieve an objective (for example, sell an object at the highest possible revenue), and to achieve this objective, the institution designs a game played among individual agents (for example, potential buyers of an object). Then, given the behavior of the agents in the games, the institution optimizes the rules of the game so that it achieves its objective.

1.2.1 Examples

Similar to game theory, mechanism design has been applied in many different contexts and fields.

1. How should a country design its presidential election? Should citizens vote for a single candidate, or should citizens be asked for a ranking of the candidates?

2. How does the traffic law optimally punish pedestrians and car drivers after a crash? Should there be pedestrian crossings and traffic lights?

3. Should the consequences of illegally preventing a last-minute goal in soccer be more severe than a penalty and a red card (for example, by assigning a goal to the team and a red card)?
4. What is the optimal way to reduce carbon emissions? By introducing a carbon tax or by designing a market for emissions?
5. How should organ donations be allocated to patients to maximize the number of lives saved?

In this part of the course, we will introduce the fundamental concepts of mechanism design required for studying such questions about the optimal design of environments to achieve the desired objectives.

Part I

Game Theory

Chapter 2

Representation of Static Games

The simplest class of games are static games, that is, strategic situations in which all players of the game act simultaneously and only once. Consequently, in a static game, a player cannot observe the actions of other players and can, therefore, not react to others' actions or influence others with their own actions. In this chapter, we will narrow our focus even further and assume that players have access to all information that is relevant to the strategic situation—that is, we focus on static games of complete information. We will be more precise about what we mean by this as we go along.

In the first step, we introduce a formal representation of static games. In particular, we want to understand which components of the strategic situation we have to incorporate into our model to be able to discuss with the tools of game theory. In the second step, we will treat the strategic situation as a *decision problem* to understand which outcomes of the game can be expected with minimal restrictions on the players' decision-making process beyond rationality.

Example 1. Consider the following simple example of a strategic interaction (the *grade game*) between two classmates: Ann and Bob. Ann and Bob individually contemplate whether to collaborate in the preparation for a game theory exam. Whenever both cooperate, they share their class notes and knowledge with each other and will write a good exam; say, their final grade will be a 29. However, if they both decide not to cooperate, they will be less well prepared and each receives a grade of 28. In the case that one of them shares his/her notes and knowledge, but the other one does not, then the one who did not cooperate will have a lot of knowledge and will do relatively better than the other one leading to a grade of 30L. The player who cooperated does not benefit from the other one's knowledge and only gets a grade of 27.

How can we describe this situation as a game? First, a formal description of a game has to include the set of players that are involved in the game. In this case, the set of players is $I = \{Ann,\ Bob\}$. Second, the description of a game has to state the set of available actions that each player can take. In this case, the set of available actions to both players is $A_{Ann} = A_{Bob} = \{Cooperate\ (C), Don't\ Cooperate\ (D)\}$. Third, we need to describe the possible outcomes of the game. In this case, the possible outcomes are a grade for each of the classmates, $y \in Y = \{(29, 29), (28, 28), (30L, 27), (27, 30L)\}$.[1] Fourth, a game description requires a mapping from action profiles—that is, the vector of actions taken by each player—into outcomes. In the grade game, the outcome function is

$$g(a_{Ann}, a_{Bob}) = \begin{cases} (29, 29) & , \text{ if } (a_{Ann}, a_{Bob}) = (C, C) \\ (28, 28) & , \text{ if } (a_{Ann}, a_{Bob}) = (D, D) \\ (27, 30L) & , \text{ if } (a_{Ann}, a_{Bob}) = (C, D) \\ (30L, 27) & , \text{ if } (a_{Ann}, a_{Bob}) = (D, C). \end{cases}$$

[1] We let the first entry in each outcome denote Ann's and the second entry Bob's grade.

Fifth, we need to assign preferences over outcomes for each player. We assume for now that players care only about their own grades and that they prefer better grades. For example, we could assume

$$u_{Ann}(y) = u_{Bob}(y) = \begin{cases} 1 & , \text{ if } y = (29, 29) \\ 0 & , \text{ if } y = (28, 28) \\ 3 & , \text{ if } y = (30L, 27) \\ -1 & , \text{ if } y = (27, 30L). \end{cases}$$

We often represent games with few players and few actions in a matrix as in Table 3.3.[2] Rows correspond to actions by the "row player"; that is, Ann in this case. Columns correspond to actions by the "column player"; that is, Bob in this case. Each entry of the matrix then corresponds to a realized action profile. The first number indicates the payoff of the row player given the realized action profile, the second number indicates the payoff of the column player given the realized action profile.

		Bob	
		Cooperate	Don't Cooperate
Ann	Cooperate	$1, 1$	$-1, 3$
	Don't Cooperate	$3, -1$	$0, 0$

Table 2.1: *The grade game represented in a game matrix.*

□

Example 1 suggests that a complete description of a static game with complete information requires: A description of

1. the participants in the game;
2. what each participant can do;
3. what are the possible outcomes of the strategic situation;
4. how the actions affect the potential outcomes;
5. how the participants evaluate the different outcomes.

The next section formalizes this description.

2.1 Formal Representation of Static Games

We now define the first formal representation of static games with complete information.[3] We will simplify this definition in the next step.

Definition 2.1. *A static game is a list $G = \langle I, (A_i)_{i \in I}, Y, g, (v_i)_{i \in I} \rangle$, where*

- *I is the set of players,*
- *A_i is the set of possible actions for player i,*
- *Y is the set of possible outcomes,*

[2] This game is more commonly known as *Prisoner's Dilemma*. We will introduce it as such later on and return to it frequently during the class.

[3] Note that this representation is often denoted as the *normal form* or *strategic form* representation of a game. We will come back to this notion later.

- *$g : \times_{i \in I} A_i \to Y$ is the outcome function,* [4]
- *$v_i : Y \to \mathbb{R}$ is the von Neumann-Morgenstern utility function of player i.*

We denote a player's action by $a_i \in A_i$ and define an action profile $a = (a_i)_{i \in I} \in A = \times_{i \in I} A_i$ as the vector of all players' actions. In addition, we use the notation $-i$ whenever we refer to the set of all players *except* player i. For example, $a_{-i} \in A_{-i}$ is a vector of actions chosen by the set players $j \in I \setminus \{i\}$.

Note that the outcome function g can be stochastic. That is, an action profile $a = (a_i)_{i \in I}$ can lead to different outcomes with corresponding probabilities.

We can simplify Definition 2.1. Observe that we can represent the players' preferences directly over action profiles instead of over outcomes that are determined by the action profiles. For stochastic outcome functions, preferences are defined over lotteries. For a definition of lotteries and probability measures over finite domains, see Appendix A.

Second, define the payoff function $u_i : A \to \mathbb{R}$ as a function representing the players' preferences as a function from action profiles directly into the reals. To see that such a payoff function can capture both the outcome function $g : A \to Y$ and the player's von Neumann-Morgenstern utility function v_i, denote by μ a given distribution over action profiles. As $v_i : Y \to \mathbb{R}$ represents player i's preferences $\succeq_i$ over lotteries over outcomes, $\lambda_1, \lambda_2 \in \Delta(Y)$, the following obtains

$$\lambda_1 \succeq_i \lambda_2 \Leftrightarrow \sum_{y \in Y} \lambda_1(y) v_i(y) \geq \sum_{y \in Y} \lambda_2(y) v_i(y). \tag{2.1}$$

Denote by $\hat{g} : \Delta(A) \to \Delta(Y)$ the pushforward function that describes the lottery over the set of outcomes induced by a distribution over the set of action profiles, μ, as

$$\hat{g}(\mu)(y) = \mu\left(g^{-1}(y)\right) = \sum_{a \in g^{-1}(y)} \mu(a). \tag{2.2}$$

In words, the likelihood of event y happening is determined by adding up the probabilities of actions within the action profile that result in event y where these probabilities are induced by μ.

Defining the payoff function $u_i : A \to \mathbb{R}$ as the composition $u_i = v_i \circ g$, we obtain for any two distributions over action profiles $\mu_1, \mu_2 \in \Delta(A)$

$$\begin{aligned}
\mu_1 \succeq_i \mu_2 \Leftrightarrow \sum_{y \in Y} \hat{g}(\mu_1)(y) v_i(y) &\geq \sum_{y \in Y} \hat{g}(\mu_2)(y) v_i(y) \\
\sum_{y \in Y} \sum_{a \in g^{-1}(y)} \mu_1(a) v_i(g(a)) &\geq \sum_{y \in Y} \sum_{a \in g^{-1}(y)} \mu_2(a) v_i(g(a)) \\
\sum_{a \in A} \mu_1(a) u_i(a) &\geq \sum_{a \in A} \mu_2(a) u_i(a).
\end{aligned}$$

Hence, the payoff functions $(u_i)_{i \in I}$ achieve the goal we wanted to achieve: it represents player i's preferences over action profiles. Using these payoff functions, we obtain the classic more concise representation of static games in the following definition.

Definition 2.2. *A static game is a list $G = \langle I, (A_i, u_i)_{i \in I} \rangle$ where*

- *I is the set of players,*
- *A_i is a nonempty set of possible actions for player i,*
- *$u_i : A \to \mathbb{R}$ is player i's payoff function.*

[4] $\times_{i \in I} A_i$ denotes the Cartesian product of the players' set of actions. That is, for $I = \{1, \ldots, n\}$, $\times_{i \in I} A_i = A_1 \times A_2 \times \cdots \times A_n$.

In this class, most of our examples fall into one of two classes of games, finite and compact-continuous games. We define these classes in the following two definitions.

Definition 2.3. *A static game G is finite if, for all players $i \in I$, the action set A_i is finite.*

Example 1 is a simple definition of a finite game. Each player can choose from two actions (cooperating and not cooperating).

Definition 2.4. *A static game G is compact-continuous if, for all players $i \in I$, A_i is a compact subset of a Euclidean space $\mathbb{R}^{k_i}$ where $k_i \in \mathbb{N}$ and $u_i : A \to \mathbb{R}$ is continuous.*

The following is an example of a compact-continuous game that is a classical economic example, which we will return to frequently. Notably, this game is one of the earliest games that have been studied with game-theoretic concepts by Augustin Cournot in Cournot (1838) long before game theory had been established as a research field.

Example 2. Consider the following game of two firms, $I = \{1, 2\}$, competing in a market for a homogenous good (that is, consumers do not care which firm they purchase the good from). Each firm i chooses a quantity q_i to produce; that is, $q_i \in A_i = [0, \infty)$. The cost of producing q_i units is $c_i(q_i) = c\, q_i$; that is, there is a constant marginal cost of production, $c \geq 0$. The inverse demand function determines the price at which the goods sell and is given by $p(q_1, q_2) = \max\{0, 1 - (q_1 + q_2)\}$. Thus, the payoff functions are $u_i(q_1, q_2) = q_i(1 - (q_i + q_{-i}) - c)$. □