Introduction to Game Theory

Part 1: The basics

Christian Hilbe, christian.hilbe@it-u.at
Version May 6, 2025

0 Preamble: About this class

Remark 0.1 (Background and motivation of this course)
This course is part of the PIC *Analyzing Social Behavior*. It should expose students to different techniques to analyze strategic decision making. I did not *have to* give this course; I *wanted to*. With it, I want to achieve at least three goals:

1. Most immediately, it should provide a gentle introduction to game theory. Students will learn how to interpret real decision problems as a *game*. Then they will learn how to identify solutions to these decision problems (e.g., to compute the game's *equilibria*).
2. More generally, the course should serve as an example of mathematical modeling. Students should have in mind that the framework of game theory (the formalism, the solution concepts) are human constructs. They should understand why smart people at some point found it useful to come up with this formalism; they should be able to critically question the underlying assumptions and methods. Ideally, when taking this course, students should put themselves into the shoes of a smart social scientist, a hundred years ago (when game theory was not yet an established formalism). Ideally, they try to come up with some of the formalism and some of the solution concepts themselves. They will also learn how to test some of these solution concepts empirically.
3. I hope the course will also just serve as brainfood. Students should enjoy to hear more about some scientific discipline that they might find stimulating, but that they would not hear about otherwise.

Remark 0.2 (Prerequisites)
A fully rigorous game theory course requires quite a bit of math (e.g., matrices, derivatives, fixed point theorems). However, this year's group is rather heterogeneous, and people have diverse interests. To accommodate these two constraints, this course opts for a compromise. During classes, we will entirely focus on intuitions. We will introduce the key concepts without much mathematical formalism. However, I will suggest some reading material that interested students could use for further reading.

Remark 0.3 (Some final administrative remarks)

- <u>Overall structure.</u> The course will consists of four parts. Thursday morning will cover the basics, such as the elements of a game, and how to analyze normal-form games). Thursday afternoon, we will explore learning in games. On Friday morning, we will discuss more complicated game models (involving sequential decision-making and incomplete information). Friday afternoon, we will discuss how to test game theory experimentally in the lab, and we will have some final group conversation about game theory.

- How to "pass" this course. Grading will be based on participation (comments/questions during the course, solving smaller exercises, taking part in the final group conversation).

1 A primer in game theory

1.1 An overview

Remark 1.1 (What is game theory?)
Game theory is a scientific field that explores strategic decision-making in interactions with multiple decision-makers who jointly determine each others' outcomes. In particular, not all decision-making is covered by game theory; as an example, the following questions are not covered: *How to play roulette? Should I buy a given flat at a given price? Should I do a PhD?* Instead, a strategic component is important. So, the following variations of the previous questions are in fact covered by game theory: *How to play poker? How to bargain the price of a flat you want to buy? How to apply to a PhD program?* In the following, we will ask two questions: (1) How can we formalize such strategic interactions in a mathematical model? (2) Given we have a model, what does it mean to *solve* a game?

Example 1.2 (Some examples to form intuitions)
Are the following questions covered by game theory (and if so, why?): (1) *Should I actively participate in this course?* (2) *Should I be modest or bold?* (3) *Should I apply to a given job advertisement?*

Example 1.3 (Some further examples to form intuitions)

1. Guess 2/3 of the average. All participants secretly choose a number between 0 and 100. The participant that is closest to 2/3 of the average wins the competition.

2. Split or Steal. Split or Steal used to be the final stage of the UK game show *Golden Balls*. It involves two participants, each having two balls, marked 'Split' and 'Steal'. Participants independently choose one of the balls. If they both split, they share the prize money. If one splits and the other steals, the latter gets all the prize money. If they both steal, no one gets anything. Prior to making a decision, participants are allowed to talk to each other. Some examples:
   `https://www.youtube.com/watch?v=TKaYRH6E36U`
   `https://www.youtube.com/watch?v=S0qjK3TWZE8&t=159s`

Remark 1.4 (Elements of a game)
When formalizing strategic interactions as a game, it is important to capture the following five elements:

- Players. Who are the relevant individuals that make (strategic) decisions?
- Actions. Between which actions can the players decide?
- Order of moves. Who gets to make a decision at which point?
- Information. What is it that players know when making their decisions?

- Payoffs. Depending on the players' decisions, what is the eventual outcome for each player? [e.g., measured in money, or happiness, or *utility*]

Example 1.5 (Identifying the game elements in various strategic decisions)
What are the players, actions, order of moves, information, and payoffs in (*i*) Guess 2/3 of the average, (*ii*) Split or steal, (*iii*) Poker, (*iv*) PhD application game?

Remark 1.6 (Actions vs strategies)
Later, it will sometimes be important to clearly distinguish actions and strategies. An *action* is an option that the player can choose. A *strategy* is a rule that tells the player which action to choose, depending on the information the player has. For example, in Poker, *to fold* is an action. But *fold if the cards are bad and raise if the cards are good* is a strategy, and so is *Always fold*.

Remark 1.7 (An overview on different kinds of game theories)
There are by now several different approaches to game theory:

1. Classical/Standard game theory. This kind of game theory asks what rational players would do. Being rational here means that players completely understand all aspects of the game, and that they wish to maximize their payoffs.
2. Evolutionary game theory. Here, players do not need to be rational, but they try to improve their payoffs over time.
3. Epistemic game theory. This kind of game theory tries to model how individuals form expectations/beliefs if players are no longer assumed to be rational, or if players no longer believe in their co-players' rationality.
4. Algorithmic game theory. Here, researchers are concerned with the computational complexity of solving games, and in identifying optimal algorithms for doing so.
5. Experimental game theory. Here the question is: How do actual people behave in strategic interactions?

We will mostly focus on classical game theory, but we'll also have a brief look at evolutionary and experimental game theory (that's also what we do in our research group).

1.2 Normal-form games

Remark 1.8 (Normal-form game)
To get intuitions, we will look at the simplest possible case. We consider a scenario with two players only. Those players have complete knowledge of the possible actions and each other's possible payoffs. They each make one decision, and they need to decide simultaneously. These so-called normal form games can be represented by a table (*payoff matrix*).

Example 1.9 (Split or Steal)
Let's revisit the game 'Split or Steal'. There are two players, who can choose among two actions. Depending on the players' actions, they either share the prize P, one of the players gets all, or they both get

nothing. The situation can be represented by the table

	Split	Steal
Split	$P/2, P/2$	$0, P$
Steal	$P, 0$	$0, 0$

Here, one of the players (the blue player) chooses a row of the matrix and the other (the red player) chooses a column. The entries of the table represent the payoff of the respective player. Players wish to maximize their realized payoff.

Group Exercise 1.10 (**Deriving payoff matrices**)
Derive the payoff matrix for the following scenarios

1. Volunteer's dilemma. There is a task that needs to be done (e.g., calling the police when witnessing a suspicious event), and two individuals that need to decide whether or not to volunteer for doing the task. Both individuals prefer that the task is done, but they would prefer the other one to do it. Assume if someone volunteers, both get a benefit b, but the one who volunteers pays some small cost c.
2. Stag-hunt game. There are two members of a tribe who need to decide whether to hunt stag or hare. To successfully hunt a stag, it takes two committed hunters. In that case, each hunter gets a payoff of 10. Hunting a hare can be done individually, without any risks, leading to a safe (but smaller) payoff of 4.

Remark 1.11 (**Solving games**)
So far, we looked at how we can represent strategic interactions in more formal way. Next, we ask: What does it mean to solve such games (roughly: how should we predict what rational players would do). In this course, we will discuss two solution concepts: (1) (Iterated) elimination of dominated actions, and (2) Nash equilibria.

1.3 Elimination of dominated actions

Remark 1.12 (An intuitive solution for (a modified) Split and Steal)
Consider a Split and Steal game with a prize of 10,000 EUR. Moreover, suppose players feel really unhappy when they split while the opponent steals (represented by an emotional cost of 100 EUR). Then the resulting game can be represented by

	Split	Steal
Split	$5,000, 5,000$	$-100, 10,000$
Steal	$10,000, 0$	$0, 0$

In this game, it is noteworthy that while splitting might seem to be the social choice, we probably would predict that rational players would steal. The reason is: even if I'm unsure about what the other player is doing, by stealing I am always better off.

Definition 1.13 (Dominated actions)

An action a is weakly dominated if there is some other action a' which always yields at least the payoff of a, irrespective of the co-player's action. Action a is strictly dominated if a' always yields a strictly larger payoff.

Remark 1.14 (On dominated actions)
The notion of dominated actions captures some kind of minimum requirement that we would expect rational players to do: They should never use actions that are never optimal, for no possible belief on what the co-player might be doing. Hence, we should expect rational players to *eliminate* dominated actions. If, after elimination of all dominated actions, there is only one action per player left, we may call the remaining strategy pair a solution of the game. In this sense, the above 'Split and Steal' can be solved, and the solution is for both players to steal (even though that outcome leaves everyone worse off on a collective level). However, not all games can be solved in that sense.

Remark 1.15 (A Split and Steal game with remorse)
Consider a version of the above Split and Steal where the row-player would have a bad conscience when stealing against a splitting opponent. In that outcome, the row-player would experience some emotional cost of 8,000 EUR. Then the resulting payoff matrix is

	Split	Steal
Split	$5,000, 5,000$	$-100, 10,000$
Steal	$2,000, -100$	$0, 0$

Now, the row-player is actually open to splitting the amount (provided the co-player goes along). Nevertheless, if both players can be assumed to be rational, there are still reasons to believe they would both steal eventually. The reason is: For the column-player, splitting is still a dominated action; hence that player will certainly prefer to steal. However, if the row-player is aware of its opponent's rationality, the row-player will anticipate the opponent's stealing. Given that belief, it also becomes better for the row-player to steal.

Remark 1.16 (Iterated elimination of dominated actions)
The above procedure to arrive at a prediction is called iterated elimination of dominated actions. Here, we first eliminate actions that are dominated for either of the two players. In a next step, we consider the reduced game that only contains actions that haven't been eliminated yet. Then we also eliminate all actions that are dominated in this reduced game. We iterate this process until no action can be eliminated anymore. If in the end, both players only have a single action, then the game is solvable by iterated elimination of dominated actions (such games are also called 'dominance-solvable'). The above Split and Steal with Remorse is dominance-solvable. However, we note that the required reasoning on part of the players is now more complex. To arrive at this solution, not only does the row-player need to be rational (try to maximize her own payoff). She also needs to believe that her opponent is rational, and she needs to use that belief to correctly anticipate the opponent's behavior.

Example 1.17 (Traveler's dilemma)

Consider the following scenario: an airline lost two identical pieces of luggage by two different travelers. It makes the following offer: It will ask both travelers independently the value of the luggage (within a range between 180 EUR and 300 EUR). Then it will pay the lower of the two claims. But in case the two claims are different, it will also pay a small reward of $R = 5$ EUR to the traveler who made the smaller claim. Can we solve this game by elimination of dominated actions? Can we solve it by iterated elimination of dominated actions?

Remark 1.18 (Connection to the other game theories)

1. Epistemic game theory. Epistemic game theory models how rational individuals need to be, and how strongly they need to believe in their co-player's rationality, for certain solution concepts to make sense. For the traveler's dilemma, we can come up with a unique solution if we accept that each player needs to be rational, needs to be aware of its co-player's rationality, needs to be aware that the co-player is aware of one's own rationality, etc. That seems quite restrictive!
2. Experimental game theory. Goeree and Holt (2001) have tested the traveler's dilemma experimentally. The two travelers need to claim a value between 180 and 300. There are two treatments, either with a small reward ($R = 5$) or a large reward ($R = 180$). While the magnitude of the reward does not affect the theoretical prediction by dominance solvability, it has huge effects on the outcome. When the reward is large, most people ($\sim 80\%$) indeed make the smallest possible claim of 180. When the reward is small, most people (again $\sim 80\%$) make the largest possible claim of 300.

1.4 Nash equilibrium

Remark 1.19 (Optimal behavior in the stag-hunt game)
Unfortunately, it is fairly easy to come up with games that are not dominance-solvable. One example is the stag-hunt game with payoff matrix

	Stag	Hare
Stag	10, 10	0, 4
Hare	4, 0	4, 4

Opting for a stag-hunt is not dominated – after all, it is the best thing I can do if I expect my co-player to hunt stag as well. Similarly, hunting hare is not dominated either. Still, some outcomes seem less likely than others. For example, it would be odd to consider (Stag, Hare) a 'solution' (i.e., for the row-player to go for a stag-hunt, while the column-player hunts a hare). Here, at least one of the players would immediately want to deviate from that outcome. In that regard, an outcome like (Hare, Hare) or (Stag, Stag) seems much more reasonable.

Definition 1.20 (Nash equilibrium)
A pair of actions (a_R, a_C) for the two players (one action for the row-player and one for the column-player) is a Nash equilibrium if neither player can improve its payoff by unilaterally deviating.

Example 1.21 (Two examples)

1. What are the Nash equilibria of the stag-hunt game?
2. Penalty kicks. Consider a game between a striker and a goalkeeper. The striker needs to choose whether to shoot left or right, the goalkeeper chooses in which direction to jump. To simplify matters, suppose the striker scores if and only if the goalkeeper jumps into the wrong direction. Moreover, if we assume the striker is the row-player, we can represent this game as follows,

	Left	Right
Left	$0, 1$	$1, 0$
Right	$1, 0$	$0, 1$

Can we find a pair of actions that is a Nash equilibrium in that game?

Remark 1.22 (**Mixed strategies.**)

- The game of penalty kicks suggests that sometimes, players may wish to be unpredictable (there are many other such examples, such as in Rock-Scissors-Paper, or in warfare). We can incorporate this by saying that players do not choose a single action with certainty; rather they play their actions according to some probability distribution (x_L, x_R). In that case, we say that the player adopts a 'mixed strategy'. [In the special case that the mixed strategy says one action should be played with certainty, e.g., $x_L = 1$ and $x_R = 0$, we also speak of a 'pure strategy'].
- One can define a payoff for mixed strategies by considering their expected value. For example, in the penalty kicks game, suppose the striker chooses 'Left' with 60% probability; moreover, suppose the goalkeeper jumps left with 90% probability. Then the striker's payoff is $0.6 \cdot 0.1 + 0.4 \cdot 0.9 = 0.42$.
- One can extend the previous definition of Nash equilibria to mixed strategies: a pair of strategies is an equilibrium if neither player has an incentive to deviate. For penalty kicks one can show: if both players set $x_L = x_R = 1/2$, that is a Nash equilibrium.

Theorem 1.23 (**Nash (1950)**)
Any normal-form game has an equilibrium point (possibly in mixed strategies).

Remark 1.24 (**Justifying Nash equilibria**)

- There is a relationship between the Nash equilibrium concept and dominance solvability. One can show: if a game is dominance-solvable, then the respective solution needs to be a Nash equilibrium.
- Is it really unthinkable that rational players in the stag-hunt game will ever end up playing (Stag, Hare)? Not necessarily – coordination failures are common-place. However, one might justify a Nash equilibrium outcome in the following way: If players had a chance to engage in pre-play communication to agree on a certain outcome, any sensible agreement they come up with has to be a Nash equilibrium [they probably would not agree on (Stag, Hare)!].
- We might also hope that if individual actions are the result of some learning process, individuals would eventually end up learning some Nash equilibrium outcome. More about that later!

Remark 1.25 (**On the status of the Nash equilibrium concept**)

- It is fair to say that the Nash equilibrium is the standard concept to 'solve' normal-form games.
- In some situations, human behavior is also fairly well predicted by the Nash equilibrium. For example, empirical papers suggest quite a good agreement for penalty kicks in soccer (Chiappori et al., 2002) or for tennis serves (Walker and Wooders, 2001). [Bonus question: Why do researchers use sports to compare the Nash equilibrium concept to data?]
- However, it is also fairly easy to construct artificial games in which human subjects considerably deviate from the Nash equilibrium predictions. For example, consider the following two examples by Goeree and Holt (2001):

	Left	Right
Top	80, 40	40, 80
Bottom	40, 80	80, 40

	Left	Right
Top	320, 40	40, 80
Bottom	40, 80	80, 40

  In both games, there is a unique Nash equilibrium. According to that equilibrium, the row-player should choose both options 50-50 in each of the two games. This prediction is experimentally confirmed for the left hand game. However, it is refuted in the right hand game; there, 96% of participants choose Top.

Remark 1.26 (A summary)
In this first part of this course, you should have learned

- what game theory is about (about strategic decision-making when individuals need to take others' possible decisions into account)
- how to formalize a game in a mathematical model (by identifying the players, their actions, the order of moves, the information they have, and the payoffs)
- how to 'solve' that model (we considered two solution concepts: iterated elimination of dominated strategies, and the Nash equilibrium).

Group Exercise 1.27 (Possible exercises)

- Discuss from a game theory viewpoint: Financial bubbles; polarization; self-commitment; groupwork; strategic interactions in the literature or in movies
- Volunteer's timing dilemma: Suppose the volunteer's dilemma is played over two periods (an early period and a late period). Individuals can decide to volunteer in either period, or not at all. Now, some individuals might want to strategically procrastinate their own volunteering to the late period – if they see their co-player volunteers to do the task in the early period, they can save the volunteering costs altogether. Can you come up with a payoff matrix for this scenario?
- Write an algorithm that finds the Nash equilibria of 2×2 normal form games

References

Chiappori, P.-A., Levitt, S., and Groseclose, T. (2002). Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. *American Economic Review*, 92(4):1138–1151.

Goeree, J. K. and Holt, C. A. (2001). Ten little treasures of game theory and ten intuitive contradictions. *American Economic Review*, 91(5):1402–1422.

Nash, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences USA*, 36:48–49.

Walker, M. and Wooders, J. (2001). Minimax play at Wimbledon. *American Economic Review*, 91:1521–1538.