Modeling Human Behavior

An Introduction to game theory – Part 2: Elimination of dominated strategies

Christian Hilbe, christian.hilbe@it-u.at
Version May 26, 2026

Remark 1.8 (A recapitulation of last week's class)
Last week, we introduced the basic setup of game theory. In short, we talked about what games are and how we can translate real-world problems into formal games. In particular, we discussed the following questions:

1. What is game theory about? (In particular, how is a game different from some ordinary investment decision, or some other simple decision problem?)
2. Suppose you want to model a strategic decision problem as a game. What are the five elements/aspects that you need to take into account?

With the tools you have learned, you should be able to address the following questions: What are the incentives people face when doing teamwork? How do two (or multiple) firms compete in a market (e.g. Boeing vs Airbus)? What are some strategic considerations people may have when using agentic AI?

Remark 1.9 (Plan for today's class)
After discussing how to formalize such questions, in the next two classes we talk about what it means to solve a game. Roughly speaking, we want to predict how rational individuals in such games might act. As a reminder: Ideally, in the following you do not merely consume the introduced concepts. Rather you should aim to anticipate and/or to question them (this is a course about mathematical modeling as much as it is a course on game theory).

1.2 Normal-form games

Remark 1.10 (Different classes of games)
There are various ways how to classify games. One common distinction is whether all players decide once and simultaneously, or whether they make (possibly multiple) decisions over time (think of chess, for example). The respective game classes are called *static games* and *dynamic games*, respectively. Another distinction is whether players know all relevant aspects of the game. Here, the distinction is between *games with complete information* and *games with incomplete information* (think of chess vs poker).

Remark 1.11 (Normal-form games)
For most of this course, we will focus on the simplest possible case: static games with complete information among two players. That is, players make only one decision, and they need to decide simultaneously.

When making their decision, they know all relevant aspects of the game (e.g., the possible actions and the feasible payoffs). The only thing they do not know is the decision of the other players. These games are called *normal form games*. They can be represented by a table (*payoff matrix*).

Example 1.12 (**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. Ignoring the pre-game communication state, 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.

In-Class Exercise 1.13 (**Deriving payoff matrices, part I: Volunteer's dilemma**)
There is a task that needs to be done (e.g., providing first aid after an accident, or washing the dirty dishes in the kitchen sink). There are two individuals who need to decide whether or not to volunteer for doing the task. Both individuals prefer that the task is done, but they both prefer the other one to do it. Assume if at least one person volunteers, both get a benefit b. However, anyone who volunteers pays some small cost $c < b$.

Homework Exercise 1.14 (**Deriving payoff matrices, part II: 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. All by yourself, you cannot get a stag and you would get a payoff of 0 if you tried. Hunting a hare can be done individually, without any risks, leading to a safe (but smaller) payoff of 4, independent of the other tribe member.

1.3 Elimination of dominated actions

Remark 1.15 (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 1.16 (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 1.17 (On dominated actions and rationality)
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 1.18 (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 1.19 (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.

In-Class Exercise 1.20 (Traveler's dilemma)
Consider the following scenario by Basu (1994): 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 5 EUR to the traveler who made the smaller claim.

In this game, who are the players, what are their possible actions, and what are their possible payoffs? Can we solve this game by elimination of dominated actions? Can we solve it by iterated elimination of dominated actions?

Remark 1.21 (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.
3. Empirical observations on 'Split or Steal'. ven den Assem et al. (2011) analyzed participant's behavior in the game show. Among other findings, they report that on average participants split in roughly 53% of the instances. There is some limited support that cooperation rates decrease with increasing stakes. Also, men are somewhat more likely to steal, but old men are not.

References

Basu, K. (1994). The traveler's dilemma: Paradoxes of rationality in game theory. *American Economic Review*, 84:391–395.

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.

ven den Assem, M. J., van Dolder, D., and Thaler, R. H. (2011). Split or Steal? Cooperative behavior when the stakes are large. *Management Science*, 58:2–20.