Course notes

LINMA2345 — Game Theory

Rational decisions among individuals

Professors

Raphaël Jungers

Matthew Philippe

Teaching Assistants 2021-2022

Anne Rubbens

Brieuc Pinon

UCLouvain

Louvain school of Engineering (L.S.E. - E.P.L.)

LINMA2345 Notes — Published in February 2022 — sources: `https://github.com/UCL-INMA/LINMA2345`

Contents

Intro

Since the birth of humanity, Mathematics have been the central tool for scientists and engineers. Arguably, Mathematics have not only been a language to express concepts and techniques developed in the different scientific fields, but it has been very often a key motor of the development of these fields, thriving on, and at the same time fertilizing the scientific discipline in question. This is strikingly the case in decision and game theory.

The birth of a new discipline Unlike Mechanics, Chemistry, or other traditional sciences, it was not before the middle of the 20th century that a firm Science of Decision making was laid down. It is not clear why Decision Theory appeared so late[1], but there are some clear explanations why it appeared by then: computers made their apparition at that moment, unlatching a power that allowed to tackle (some of) the extremely heavy computations needed in Decision Theory. Interestingly, computers themselves are automatic entities who have to take decisions on the fly, at extremely high pace (which instruction should it realize first?, where should it store its product?,...). With this in mind, it is not surprising that Von Neuman, one of the two founding fathers of Decision Theory (with O. Morgenstern, also at Princeton's famous Institute for Advanced Study) was also one of the main designers of the first computers. The middle of the 20th century was also a period where humanity started facing (decision) problems whose scale was reaching dimensions beyond the abilities of the human brain: companies were reaching unprecedented sizes; transport, storage, infrastructure supply were becoming critical problems in our massively populated cities/countries, and technologies were enabling various different solutions for these situations, thus implying decision problems to tackle them at best. Not to mention WW2, which was really a trigger in the development of mathematical techniques for decision making. Let us be fair and acknowledge the genius of some mathematicians of that time as another reason for the outcome of this paradigm shift. "*Beautiful minds*" like John Von Neumann, or John Nash, where instrumental in these developments.
Suddenly, people were understanding that decision could be modeled and optimized mathematically, and that it satisfied fundamental laws, just like two co-prime numbers have to follow Euler's theorem, or the Moon gravitating around the Earth has to follow Kepler's laws.

A journey from rationality to bargaining In this course, our goal will be to model, analyze, predict, and prescribe the behavior of agents that have to take decisions in order

[1] Needless to say, this statement is presumptuous, and questions on how to make decisions rationally had already been raised before, notably by Daniel Bernoulli in his 1738 treatise.

to optimize their outcome, while these decisions affect the others' welfare too.
We will first model a very general situation, and focus on the decision making problem for a single rational and intelligent agent. We will observe that, after all, we can restrict ourselves to an idealized situation where agents are just trying to optimize their expected "*payoff*": this is the fundamental Theorem of Decision Theory.
We will then study what happens when two such agents have to take concomitant decisions, with the critical fact that one's decision can have impact on the other's payoff. These situations are referred to as "*games*", and will lead us to the celebrated concept of *Nash Equilibrium*. We will then analyze the impact of temporality on decision making, and will propose refinements of our notion of Equilibria by pushing the concept of rationality as far as we can.
As a third step, we will think as engineers, and, rather than still restricting the range of possible outcomes of a game, we will ask ourselves the question of by which mechanism we can modify these sets of equilibria. This will take us to the notion of 1) *repeated games*, where repetition is the mechanism that enables this; 2) *to the notion of correlated equilibria*, where a third party, or "*mediator*", is the key mechanism; finally we will 3) analyze a particular case of games, namely auctions, where we will investigate several possible mechanisms in order to favor some desired behavior for the players. This will give us a glimpse at "*mechanism design*" problems, which are still an important field of research nowadays.
In the fourth part of the course, we will finish our journey by looking at "*bargaining problems*": we will take a step back, and look at all the possible outcomes of a given game (these possible outcomes can for instance be the result of any solution concept previously analyzed in the course); we will ask ourselves the question: *"what if, among all the equilibria, the agents have to agree on the most reasonable outcome?"*. In other words, among all the possible outcomes, how can we single out a unique one, which will be the 'most fair' (for whatever it means)?

Game Theory is an *axiomatic theory*. By this, we mean that all the results and concepts that we will see in the course rely on *axioms*, which are a priori postulates on the agents' behaviour. The concepts that we want to study (several of them we already mentioned here in this introduction) are extremely vague: fairness, rationality, intelligence, bargaining, etc... If I decide to go to a casino tonight (where my expected gains are strictly negative), rather than staying at home (where my expected gains are zero), am I intelligent? If I choose to take "*LPSP1001: Introduction to Psychology*" rather than "*LINMA 2345: Game Theory*", am I rational? Context is key in answering these questions, and it looks very ambitious to hope to model in a mathematical way the behavior of human agents, let alone predict it.
The proper way to get out of this conundrum is to adopt an axiomatic approach: we will be very careful, at the beginning of this course, to introduce a number (as small as possible) of axioms, that is, mathematical properties that we will not question, and which describe the behavior of rational and intelligent agents. Needless to say, there are situations where these axioms will be violated; but watch this: if we want to assess whether our mathematical results are of application in some given situation, the only thing we will have to check is this tiny set of elementary properties; and, as we will see, this is remarkable how such a little set of assumptions can have huge consequences. In fact, every time we will try to model a new vague concept (like, later in the course, the

notion of "*fair bargaining*"), our reflex, as mathematicians, will be to introduce new axioms in order to properly define what we mean by it, thereby laying the ground for a mathematical treatment of this vague notion.
This axiomatic approach is both the strength and the weakness of Game Theory: on the one hand, one can argue that these axioms are never completely satisfied, and this would invalidate all the results that we will work so hard to obtain; but on the other hand, we will see how powerful they are, and we will perfectly master at every point of the course the range of application of our results: they are completely valid, as soon as the axioms are satisfied. After all, axioms are nothing but a mathematical model of the true world, and as every model, it has its limits; just like a linear approximation of the effect of an inductance on an electrical circuit, or the assumption that a flow is laminar in Fluid Mechanics.

Today, Decision and Game Theory are more central than ever, with large-scale, multi-agent infrastructures where agents make decisions in an embedded and decentralized way, by then having a critical impact on the global behavior of the system. Think of the smart grid, where agents decide to produce, consume, sell or buy electricity. Think of large social networks, where people buy, or consume ads; bid for optimizing their advertising investment; "like" ads, and surf in the network, thereby generating a virtual payoff for the advertisers. Think also of high frequency trading, where several competing computers take decisions, which affect the behavior of the markets in a close loop fashion.
Often, these situations are adversary and one is interested in *"the cost of anarchy"*, or *"resilience guarantees"*. All these problems and notions can only be analyzed with a careful model of how agents behave, and a deep understanding of the potential phenomena occurring when multiple agents make decisions which not only impact their welfare, but also their fellow agents' welfare.
This is exactly what this course is about.

Thanks I want to express my deep gratitude to Romain Hollanders, Sebastian Stich, and Matthew Philippe for working hard on the first versions of these notes. Many pedagogical ideas presented in these notes come from them. In particular, I want to stress the amazing work achieved by Matthew, who was leading the project from the first drafts to the final version, not only in writing, but also for many higher level ideas. Thanks a lot!

Chapter 1

Decision theory

"To prefer evil to good is not in human nature; and when a man is compelled to choose one of two evils, no one will choose the greater when he might have the less. " — Plato

The material from this chapter is based on (Myerson, 1991, pp 1 - 23).

Game theory studies how a set of players makes decisions that may impact the welfare of one another. Throughout the course, we will assume that the individuals considered are *intelligent* and *rational*:

- A rational player makes his decisions consistently in pursuit of his own goal.
- An intelligent player is able to reflect on the situation and analyze it as well as we would, and to conclude from this analysis what are the good/bad choices available to him.

Decision theory focuses on the case where one *single* player has to take a decision in an uncertain context. This will lay down a solid foundation for the study of games with several players, where the actions of the others is part of the uncertainty any single player faces. In order to convey intuition, we start with a simple example.

Example 1.1 *In this example, we consider a human that plays a rock-paper-scissors game against a machine that has been programmed in a given way. Consider that the computer's programming is to play*

1. Rock *with a 20% probability,*
2. Paper *with a 45% probability and*
3. Scissors *with a 35% probability.*

If the human player has access to this information, what should he play against the computer?
Assuming that the player is intelligent, he can infer that playing e.g. scissors *would give him a* 45% *chance of winning the game, a* 20% *chance of losing it, and a* 35% *chance of the game resulting on a tie.*
Assuming the player is rational (and his goal is to win the game), he can infer that playing scissors *is the best move for him.*
In this case, we are therefore able to understand how the player makes his decision.

We are entitled to ask ourselves, in Example 1.1, what should the player do if he had little knowledge of the computer's programming? Or what should he do when faced with another human player? Such questions gave birth to the field of *Decision Theory*, where first and foremost, we ask ourselves *how does a rational and intelligent player take his decisions?*

1.1 Basic concepts of Decision Theory

A *player* has to make a decision (rock or paper?...) in an *uncertain* situation (what are the others going to do?...). Depending on his choice, he may or may not receive a *prize* (winning, loosing, a tie...).
The prize may not be handed to the player in a deterministic manner (we will see this in the next example). The probability of receiving a prize depends on *two* elements:

1. the decision made by the player,
2. the realization of the uncertainty.

In order to formalize this, we rely on the following notation, illustrated later in Example 1.3.

Notation 1.2: Δ

For any finite set Z, $\Delta(Z)$ is *the set of probability distributions* over the elements of Z:

$$\Delta(Z) = \left\{ q : Z \to \mathbb{R} \,\middle|\, \sum_{y \in Z} q(y) = 1, \text{ and } q(z) \geq 0, \forall z \in Z \right\}.$$

Example 1.3 *Consider the rock-paper-scissors game where, as in Example 1.1, you play against a computer that plays rock, paper or scissors according to some probability distribution.*

Your choice may well be taken as a probability distribution over playing rock, paper and scissors. Let us denote by ρ_1, π_1, and σ_1 the probabilities of playing rock, paper and scissors respectively, and write $p_1 = \{\rho_1, \pi_1, \sigma_1\}$. If the computer is programmed to play with a distribution $p_2 = \{\rho_2, \pi_2, \sigma_2\}$, then

1. *The probability of you winning the game is $w = \rho_1\sigma_2 + \pi_1\rho_2 + \sigma_1\pi_2$.*
2. *The probability of you losing the game is $\ell = \rho_1\pi_2 + \pi_1\sigma_2 + \sigma_1\rho_2$.*
3. *The probability of the game coming to a tie is $t = \rho_1\rho_2 + \pi_1\pi_2 + \sigma_1\sigma_2$.*

We can conclude that to any choice of p_1 you make, and given the configuration of the computer p_2 the outcome of the game is to be expressed as a probability distribution $z(p_1, p_2) = (w, \ell, t) \in \Delta(\{win, lose, tie\})$.

Now, suppose that you do not know at all how the computer program was implemented. Maybe there is no random mechanism for it to decide the action it will make. Maybe it is just hard-coded to play scissors; who knows?

This situation is very different from the previous one, as you have absolutely no clue from the statement of the problem on what the outcome will be. We only know that the outcome will be one element from $\{r, p, s\}$, which will only be revealed to you at the very end.

To emphasize this fact, and to represent any such situation where no probability distribution can be a priori postulated, we will call such possible outcomes possible states of the world.

More generally, one may have a combination of both types of uncertainties. Imagine for instance that there are two computer programs with two different probability distributions implemented, and someone activated one of them, but you have no clue about which program has been activated. In this most general situation, the player makes choices, and for every realization of a set of uncertain parameters, theses choices will result into some probability distribution over a set of prizes. This is formalized through the definition of a *lottery*, that describes the probability of receiving some *prize* in a given *state of the world.*

Definition 1.4: Lottery

A lottery f is a function

$$f : \Omega \to \Delta(X),$$

where

1. Ω denotes the (finite) set of all possibles *states of the world*, that are all the possible realizations of the uncertainty,
2. X denotes the (finite) set of *prizes*.

Notation 1.5: $f(\cdot|t)$, $f(x|t)$

Given a lottery $f : \Omega \to \Delta(X)$, a state $t \in \Omega$, we write $f(\cdot|t) \in \Delta(X)$ be the probability distribution on the prizes if the state of the world is t.
For a prize $x \in X$, we write $f(x|t)$ to denote the probability of receiving the prize x in the state of the world t for the lottery f.
For any $x \in X$, we write $[x]$ to denote a lottery that will always give the prize x:

$$\forall t \in \Omega,\ \forall y \in X,\ [x](y|t) = \begin{cases} 1 \text{ if } y = x, \\ 0 \text{ otherwise.} \end{cases}$$

In Example 1.3, we had $X = \{\text{win}, \text{lose}, \text{tie}\}$, and Ω is the set of all probability distributions p_2 the computer can be programmed[1] to follow. If we fix a distribution p_1 for the

[1]Note that in the definition above, we assume Ω to be finite. We can get away with our example by saying that, since we are programming a computer to play, there is a finite (but huge) number of possible ways the computer can be programmed to play (due to the finite memory of the machine). However, decision theory also extends to infinite sets of event, which is outside the scope of this chapter.

player, we can see that the outcome of the game (distribution on the prizes) depends only on p_2. Thus, by modeling p_2 as the state of the world, the outcome of the game can be expressed as a lottery as in the above definition.

In a decision making process, player choices lead to *lotteries*: once his decisions are made, the outcome is a probability distribution on a set of prizes and that distribution depends on the state of the world. Next, we discuss how we can put this notion to good use to understand the mechanism directing the choice of the player.

1.2 Preferences and the axioms of Decision Theory

We now wish to understand how a player, being able to choose between different lotteries, takes his decisions. One of the main difficulties in the decision making is the uncertainty on the *state of the world*.

In practice the player might have some prior knowledge on what could or could not happen. Taking Example 1.3, the player could have the information that the computer plays *rock* with a probability higher than the one of playing *paper*. Then, assuming the player wants to avoid loosing, he would naturally play *paper* with a higher probability than *scissors*.

Definition 1.6: $P(\Omega)$, Event

Let $P(\Omega)$ be the set containing all *non-empty* subsets of the set of states of the world Ω. An event $S \in P(\Omega)$ is a non-empty subset of Ω.

In what follows, we assume that, even though there are no fixed a priori probabilities on the states of the world, yet the player has a subjective knowledge of an event S (maybe Ω itself), and we develop the notion of preference between lotteries.

Notation 1.7: $\succsim, \sim, \succ$

Given two lotteries f and g and an event $S \in P(\Omega)$, we write $f \succsim_S g$ if the player finds f to be *at least as good as* g if the true state of the world is in S.
We write $f \sim_S g$ if $f \succsim_S g$ and $g \succsim_S f$, and we write $f \succ_S g$ if $f \succsim_S g$ but $f \nsim_S g$.

1.2.1 Axioms of Decision Theory

In order first to better understand the decision making process and second to be able to develop mathematical decision-making analysis tools, it is assumed that the preferences of a rational decision maker satisfy a list of axioms.

In the following, we let e, f, g, h be lotteries and S, T denote two events.

Axiom 1.1 (Completeness) *$f \succsim_S g$ or $g \succsim_S f$.*

Axiom 1.2 (Transitivity) *If $f \succsim_S g$ and $g \succsim_S h$, then $f \succsim_S h$.*

Axiom 1.3 (Relevance) *If $f(\cdot|t) = g(\cdot|t)\, \forall t \in S$, then $f \sim_S g$.*

Axiom 1.4 (Monotonicity) *If $f \succ_S h$ and $0 \leq \beta < \alpha \leq 1$, then $\alpha f + (1-\alpha)h \succ_S \beta f + (1-\beta)h$.*

Axiom 1.5 (Continuity) *If $f \succsim_S g$ and $g \succsim_S h$, then $\exists 0 \leq \gamma \leq 1$ such that $g \sim_S \gamma f + (1-\gamma)h$.*

Axiom 1.6 (Objective Substitution) *If $e \succsim_S f$ and $g \succsim_S h$ and $0 \leq \alpha \leq 1$ then $\alpha e + (1-\alpha)g \succsim_S \alpha f + (1-\alpha)h$.*

Axiom 1.7 (Strict Objective Substitution) *If $e \succ_S f$ and $g \succsim_S h$ and $0 < \alpha \leq 1$ then $\alpha e + (1-\alpha)g \succ_S \alpha f + (1-\alpha)h$.*

Axiom 1.8 (Subjective Substitution) *If $f \succsim_S g$ and $f \succsim_T g$ and $S \cap T = \emptyset$, then $f \succsim_{S \cup T} g$.*

Axiom 1.9 (Strict Subjective Substitution) *If $f \succ_S g$ and $f \succ_T g$ and $S \cap T = \emptyset$, then $f \succ_{S \cup T} g$.*

Note that there is some redundance in the axioms.

Axiom 1.10 (Interest) *For every state $t \in \Omega$, there exists $x, y \in X$ such that $[x] \succ_t [y]$.*

Axiom 1.11 (State Neutrality) *For any two states t and r in Ω, if $f(\cdot, t) = f(\cdot, r)$, $g(\cdot, t) = g(\cdot, r)$ and $f \succsim_r g$, then $f \succsim_t g$.*

The following two examples illustrate the relevance of such axioms for studying decision processes.

Example 1.8 *In this example, we wish to help a player take the best decision, and in order to realize the importance of the Axioms of decision theory, we are going to break one of them (for the reader to find).*
Assume a situation with $X = \{w, x, y, z\}$. The player can do the following: 1) take the prize w or 2) toss a coin pick x if heads, z if tails or 3) toss a coin and pick y if heads, z if tails.
In terms of lotteries, the player can pick

1. $[w]$,
2. $.5[x] + .5[z]$ *or*
3. $.5[y] + .5[z]$.

The player tells you his preferences:

$$[x] \succ [y] \text{ and } .5[x] + .5[z] \prec [w] \prec .5[y] + .5[z].$$

What would be your advice to the player? Should he play choice 1, 2 or 3?
Advising him to pick $[w]$ would be nonsense, since he would clearly prefer to pick $.5[y]+.5[z]$ in this case. Should you advise him to pick $.5[y] + .5[z]$, you would again make a mistake, since he would prefer $[x]$ to $[y]$, and thus prefer $.5[x] + .5[z]$ to $.5[y] + .5[z]$...
It is thus impossible to give a recommendation to the player. Can you find which axiom is not met in this situation?

Example 1.9 *This example is taken from a discussion in (Shoham and Leyton-Brown, 2008, Chapter 3), dealing with Axiom 1.2. Given three lotteries f, g and h, what if it were the case that a player expresses the preferences $f \succ g$, $g \succ h$ but $h \succ f$?*
For the sake of the discussion, let us assume that these three lotteries are physically manifested under the form of lottery tickets, and that the player initially holds the ticket for the lottery f.
Then, clearly, the agent would be willing to exchange a (perhaps small) amount of money to get a ticket h instead of a ticket f. He would then be very happy to exchange another (small) amount of money to get g instead of f. Finally, we would be again happy to spend some money to obtain h instead of g.
Ultimately, this amounts to the agent accepting to give away some utility for free, which is not rational.

1.2.2 The utility maximization theorem

So far we have discussed the concept of lottery (which tells the player what he may get by doing something) and discussed preference relationships between lotteries.

It seems quite natural that this idea of preference should be related to the benefit that one player gets from receiving a prize. This is reflected through the usage of *utility functions*:

Definition 1.10: Utility function

A utility function is a function $u : X \times \Omega \to \mathbb{R}$, that assigns a value to each prize in each state of the world.

Naturally, a player can't make decisions using only utility functions. For example, the reader would probably be very happy winning at the national lottery. The reason he may not invest all his money into lottery tickets is because he assigns a low probability of actually wining the lottery. Now, as we have seen, there are some events that the player does not control (e.g. in this example, 'how many other people will participate to the national lottery?'). How to describe the preferences of the player if he does not have any idea of what the state of the world will be? Let us first describe formally what he would need to make a decision.

Definition 1.11: Conditional probability function

A conditional probability function on Ω is a function $p : P(\Omega) \mapsto \Delta(\Omega)$ that specifies for all $S \in P(\Omega)$ a probability distribution on Ω such that, for all $t \in \Omega$,

$$p(t|S) = 0 \text{ if } t \notin S, \text{ and } \sum_{r \in S} p(r|S) = 1.$$

If the player had this information, he would indeed be able to easily define his optimal decision, thanks to the concept of *expected utility value:*

Definition 1.12: Expected Utility Value

The expected utility value of a lottery f given an event S and a conditional probability function p is given by

$$E_p(u(f)|S) = \sum_{t\in S} p(t|S) \sum_{x\in X} u(x,t)f(x,t).$$

These definitions lead us to the most important theorem of this chapter. This theorem is fundamental in that it describes how an individual plays. *To some extent, all results from game theory can be derived as consequences of this axiomatic description of the players behavior.*

The theorem states that the behavior of a rational and intelligent decision maker is driven by two main elements. First a *subjective* conditional probability function p, encoding his beliefs on the uncertainties he is facing. Second a utility function, transcribing his preferences for the different prices once uncertainty is revealed.

Theorem 1.13: The utility maximization theorem

Consider a decision problem as described above, and the preference relation for a particular agent, denoted by $\succsim_S$. The Axioms 1.1 to 1.10 are satisfied *if and only if* there is a utility function u and a conditional probability function $p : P(\Omega) \to \Delta(\Omega)$ such that

1. The utility function satisfies $\max_{x\in X} u(x,t) = 1$ and $\min_{x\in X} u(x,t) = 0$, $\forall t \in \Omega$;

2. The conditional probability function satisfies, for all sets $R \subseteq S \subseteq T \subseteq \Omega$ with $S \neq \emptyset$,
$$p(R|T) = p(R|S)p(S|T), \tag{1.1}$$
where for two sets $A, B \subseteq \Omega$, $p(A|B) = \sum_{a\in A} p(a|B)$.

3. For any two lotteries f and g and for any event S, the preference relation $f \succsim_S g$ holds if and only if
$$E_p(u(f)|S) \geq E_p(u(g)|S).$$

The Axiom 1.11 is also satisfied on top of the others if and only if the above holds with a *state-independent* utility function, that is, there exists a function $U : X \to \mathbb{R}$ such that $u(x,t) = U(x)$ for all $t \in \Omega$.

Let us discuss the meaning of the different points of the theorem:

- Condition 1 is a simple normative assumption.

- Condition 2 is a formula which establishes how conditional probabilities assessed in one event must be related to conditional probabilities assessed in another. We discuss this in the next section.

- Condition 3 is the main one, and states that any rational player in fact behaves exactly like if he was optimizing an expected payoff.

Example 1.14 (Understanding the theorem) *Assuming that the decision maker's preferences satisfy Axioms 1.1 through 1.10, there is a mechanical way to compute a probability function p and a utility function u satisfying the theorem.*

One can build the utility function for a given price. Consider a state $t \in \Omega$. First there must exist two prizes y_t and z_t in X such that, $\forall x \in X$,

$$[y_t] \succsim_{\{t\}} [x] \succsim_{\{t\}} [z_t].$$

One can ask the decision maker to provide these two prizes. Since these prices are respectively the maximal and minimal valued, their utility is equal to $u(y,t) = 1$ and $u(z,t) = 0$.
Second, for all $x \in X$, one can ask the decision maker for which number $0 \leq \alpha \leq 1$ does the following hold,

$$[x] \sim_{\{t\}} \alpha[y_t] + (1-\alpha)[z_t],$$

and set $u(x,t) = \alpha$. This gives us a way to compute the utility function.
To build the subjective probabilities $p(t|S)$, the idea is as follows. For any state $t \in \Omega$, define the prizes y_t and z_t as before. Now build the lottery $b : \Omega \mapsto \Delta(X)$ such that

$$b(r) = [y_t] \text{ if } r = t,$$
$$b(r) = [z_t] \text{ if } r \neq t.$$

Finally, specify $p(t|S)$ as the number such that, for the decision maker,

$$b \sim_S p(t|S)[y_t] + (1 - p(t|S))[z_t].$$

The utility function u and conditional probability distribution p in Theorem 1.13 represent the *preferences* (satisfying axioms 1 through 10) expressed by a decision maker. It is however important to understand to what extent we can be sure that this accurately reproduces the mental model of the decision maker. More precisely, is this mental model the unique possible one? The following theorem tells us that the answer is affirmative, up to some linear scaling.

Theorem 1.15: Equivalent representations of preferences

Consider a *utility function* $u : \Omega \to X$ and a *conditional probability function* $p : P(\Omega) \to \Delta(\Omega)$ satisfying 1.1. A utility function v and conditional probability function q represent the *same preferences* as u and p from Theorem 1.13 if and only if there exists a constant $A > 0$ and a function $B : S \to R$ such that

$$\forall S \subseteq \Omega, \forall t \in S, \forall x \in X : q(t|S)v(x,t) = Ap(t|S)u(x,t) + B(t).$$

1.2.3 Bayesian probability systems

Theorem 1.13 (respecting the second condition of the Theorem) introduces a particular set of functions similar to conditional probability functions. It is through those that the agents assess the probability of occurrence of different events. Our focus in this section is on these functions, which we call Bayesian Probability Systems.

Definition 1.16: Bayesian probability system

A conditional probability function $p : P(\Omega) \to \Delta(\Omega)$ is a *Bayesian Conditional Probability System* on the set Ω if, $\forall S \subseteq \Omega$, $S \neq \emptyset$, $p(\cdot|S)$ is a probability distribution on Ω such that $p(S|S) = 1$ and $p(R|T) = p(R|S)p(S|T)$, $\forall R \subseteq S$, $\forall T \supseteq S$.

In other words, $p(\cdot|S)$ is a well defined probability function for all $S \in P(\Omega)$ and, on top of that, satisfies Bayes formula.

The fact that a player takes decision using a Bayesian probability system as a reference has an interesting interpretation regarding his planning of events. In particular, even if he known that an event S *should not happen* according to the fundamental theorem, he still needs to define $p(\cdot|S)$.

While this may appear as a mathematical artifact of the theory, it turns out this behavior is central in our everyday life decisions, in particular when interacting with other players (see Example 1.17).

Example 1.17 *We end this chapter with an example, that also introduces 2-player games. Consider the chess-board of Figure 1.1.*

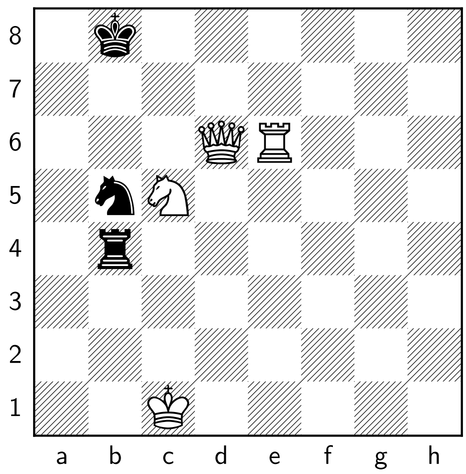


Figure 1.1: Black plays - what are his possible moves to escape the queen?

You are the white player, it is the black player's turn, and you need to guess his move. Assuming that the black player is rational, intelligent, and wants to avoid check-mate, you can assign probabilities on the moves the player will make (see Table 1.1).

Mvmt m	$P(m)$	
`king to a8`	0	checkmate in 3 steps,
`king to a7`	0	checkmate in 2 steps,
`king to c8`	0	checkmate in 2 steps,
`knight to c7`	x	avoids checkmate,
`knight to d6`	$(1-x)$	avoids checkmate.

Table 1.1: Probability of black doing something and justification.

Let Ω_b be the set of moves of the black player, and $S = \{ka8, ka7, kc8\} \subset \Omega_b$ be the set of moves to which white assigns 0 probability of occurrence. Theorem 1.13 - Condition 2 implies that $P(ka8|\Omega_b) = P(ka8|S)P(S|\Omega_b)$ should be equal to 0. Theorem 1.13 - Condition 3 implies that even in the event S, the white player should make a decision maximizing his expected utility, which implies that the player should assign a value to $p(\cdot|S)$. For example, he could consider

$$p(ka7|S) = p(ka8|S) = p(kc8|S) = 1/3,$$

even though the event S itself is considered to be of 0 probability.

In other words, we know from Theorem 1.13 that "an intelligent and rational decision maker always expects the unexpected!". This is the only possible way to mathematically express rationality: we should not make a bad move at chess against a rational decision maker because we know that he is prepared to react (even if he does not suspect us to play badly).

It remains however possible to reconcile Bayesian probability systems with classical probability distributions. In particular, for probability distributions $p \in \Delta(\Omega)$ assigning non-zero probabilities to all events in Ω, Bayes' formula defines a numerical value for $p(t|S)$ for every choice of t, S. The next theorem provides an algebraic characterization of bayesian conditional probability functions in the most general case, when some events can have zero probability.

Theorem 1.18

The conditional probability function p is a Bayesian conditional probability system on Ω if and only if there is a sequence of probability distributions $p^k \in \Delta(\Omega)$, $k = 1, 2, \ldots$, with $p^k(t) > 0\ \forall t \in \Omega$, such that

$$\begin{aligned} p(t|S) &= \lim_{k\to\infty} \frac{p^k(t)}{\sum_{r\in S} p^k(r)} \qquad \text{if } t \in S, \\ p(t|S) &= 0 \qquad \text{if } t \notin S. \end{aligned} \tag{1.2}$$

1.2.4 Limits of the Bayesian framework

In the framework developed above, the behaviour (in terms of preferences) of a player is first axiomatized. From these axioms, Theorem 1.13 states the existence of a Bayesian probability system and a utility function characterizing preferences between lotteries (thus, between decisions), through the expected utility value that can be computed from a lottery.

This allows to:

1. predict the behaviour of people: they should maximize their expected utility;

2. help people take their decisions following their preferences.

Of course, these axioms represent a formal model for us to work with and, like every model in Sciences, it has its limits. Sometimes, the situation just does not follow the axioms, and this has been well documented by economists.

Example 1.19 *Here is a famous paradox due to M. Allais.*
Let $X = \{12 \text{ millions } \$, 1 \text{ million } \$, 0\$\}$*, and let*

$$\begin{aligned} f_1 &= .10[12 \text{ millions } \$] + .9[0\$], \\ f_2 &= .11[1 \text{ million } \$] + .89[0\$], \\ f_3 &= [1 \text{ million } \$], \\ f_4 &= .10[12 \text{ millions } \$] + .89[1 \text{ million } \$\$] + 0.01[0\$]. \end{aligned} \tag{1.3}$$

What would you pick? Many people will express the preferences $f_1 \succ f_2$ *and* $f_3 \succ f_4$ *(remark that the expected payoff of* $f4$ *is greater than the expected payoff of* $f3$*)... As it turns out, even by relaxing this assumption, we violate an axiom... Indeed, let* $x = 12 \text{ millions } \$$*,* $y = 1 \text{ million } \$$*, and* $z = 0\$$ *be the prizes for our scenario. By the strict objective substitution axiom (Axiom 7), it should be the case that* $.5f_1 + .5f_3 \succ .5f_2 + .5f_4$*. However, we can see that*

$$.5f_1 + .5f_3 = .5f_2 + .5f_4 = .05[x] + .5[y] + .45[z],$$

contradicting the preference relation (since the two lotteries are qualitatively the same).

Chapter 2

Basic models

"Let's play a game!" — The puppet guy from the Saw movies.
The chapter is based on (Myerson, 1991, pages 37 to 74).

In this chapter, we begin our analysis of *games*, that are *situations where the decision of individuals (called players) impact the welfare of each other.*

Our goal here is to present the two principal tools to *model* games within well-defined mathematical frameworks. This allows us to discuss and analyze central concepts, such as the notions of *best-responses* and *domination*, and in doing so, gain a finer understanding of how we can expect rational players to behave in a game.

2.1 The extensive form and the strategic form

We first present tools that allow to provide a mathematical description from a game out of a high-level description, to a mathematical description of this game. In order to introduce the models, we begin, in Example 2.1, by describing a game to be modeled and analyzed later.

Example 2.1 (The betting game) *Benoît asks Matthew to play a game. He explains the rules as follows:*

1. *Both players put* 1 € *on a table.*
2. *Benoît picks a card randomly out of a regular deck of 52 cards. He looks at the card privately.*
3. *At this point, two options are available to him:*
   - (a) *He reveals the card to Matthew. If the card is* red, *Benoît gets all the money on the table, and if* black, *Matthew gets all the money.*
   - (b) *He chooses not to reveal the card, and raises the bet by putting an additional* 1 € *on the table. In this case, then Matthew has to decide to either*
     - i. *Pass, letting Benoît win the game, or*
     - ii. *meet the raise, adding* 1 € *on the table. In this case, the card is revealed, and Benoît wins everything if it is* red, *or Matthew wins if it is* black.

In any case, ***the game ends here****.*

At this point, we would like to encourage the reader to reflect on the game of Example 2.1. In particular, we are interested by the following questions

- What should *Benoît* do in the situation 3? When should he reveal the card, and when should he raise the bet?

- What should *Matthew* do in the situation 3(b)ii? Should he meet the raise, or fold?

Rest assured that answering these questions is not trivial at all! In fact, the answers will not come until Chapter 3.

Maybe the first step in analyzing a game is to single out the answers to the following questions:

$$\text{Relevant questions about a game:} \begin{cases} \text{1. Who are the players ?} \\ \text{2. What can each player do ?} \\ \text{3. What are the possible states of the world ?} \\ \text{4. What do they know of it?} \\ \text{5. In what \textit{sequence} are the actions taken ?} \\ \text{6. What do they get ?} \end{cases}$$

We invite you to answer these questions for the game of Example 2.1, and provide two examples next.

Example 2.2 (Answers for the game of Chess)

1. *There are at least two players,*

2. *at their turn, they can move any chess piece and take one of the other player's pieces as long as it obeys the rules,*

3. *the configuration of the board (position of each pieces on it) defines what actions can be done, and their outcome,*

4. *they both see the board,*

5. *the white player starts, and then they take turn,*

6. *a player wins if e.g. he manages to check-mate his opponent, or his opponent forfeits, …*

To the chess fans: we are leaving lots of interesting details aside.

Example 2.3 (Answers for the game of Poker)

1. *There are two or more players around a table,*

2. *at their turn, players learn about their own hands (not the others'), they may bid, raise, fold, etc...,*

3. *there are a lot of possible hands that can be dealt,*

4. *they know their own hand, how much money has been bet by all players, who has folded, etc...,*

5. *players play in a clock-wise manner, taking turn with the first player being chosen at the begining of the game,*

6. *a player that has not folded and has the better hand wins all active bets.*

To the poker fans: we are leaving lots of interesting details aside.

In order to analyze such a game, we need first to represent it in a mathematical model upon which we will be able to perform a systematic analysis.

Later, we will present the *strategic form*, which is a more *concise* (bearing less information) model for games and focuses on the players, their possible choices of actions, and the expected utility (e.g. monetary...) payoffs these actions lead to.

2.1.1 The extensive form

The first model we present below is called the *extensive form*. The reason we start with this is that it is a very natural (and almost verbatim) representation of a game. This has benefits and drawbacks, as it will contain lots of information, providing us with a complete understanding of the structure of the game. However, the amount of information contained in the model often makes the analysis more cumbersome.

The extensive form describes a *sequential (or dynamic) game*. Such a game is a sequence of *events*. Some events are due to chance, such as the random selection of the card color in Example 2.1. Other events are controlled by players, and are more commonly referred to as *moves*. They correspond to a player doing something at some point of the game. After a sequence of events, the game ends, and players receive their payoffs.

In order to capture these notions, the *extensive form game* takes to form of a *tree*. For the sake of understanding, we start by illustrating the concepts with Example 2.4.

Example 2.4 *The extensive form of the game of Example 2.1 is given in Figure 2.1 as an illustration.*

We construct it as follows. We begin by extracting information regarding the players, their actions, and their knowledge of the game.

- *First, we let $N = \{1, 2\}$ be the set of players for the game. Here, Player 1 refers to Benoît, and Player 2 refers to Matthew in Example 2.1.*

- *Then, we need to define for each player the possible* information states. *Information states correspond to the different scenarios in which players have to make decisions. We let $S_1 = \{a, b\}$ be the set of information states that Player 1 may encounter, and $S_2 = \{c\}$ the set of the second player.*

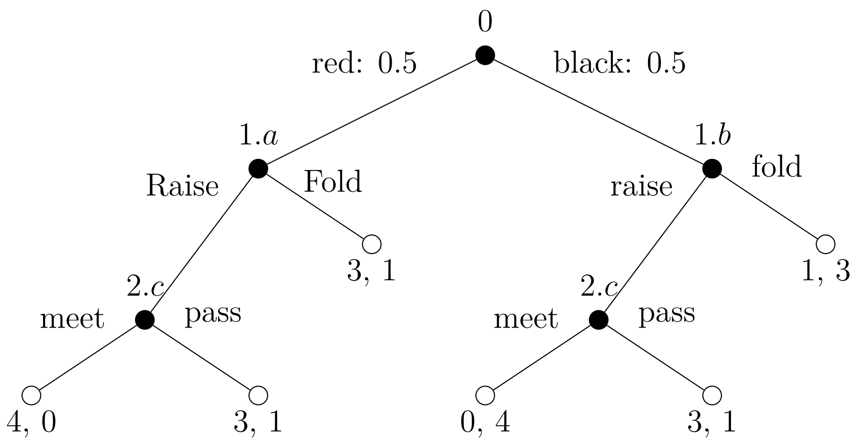


Figure 2.1: Extensive form of the game in Example 2.1

In the above, the state a corresponds to the scenario where Player 1 sees he has a red card, and b to the case where he has a black card. Note that Player 2 only has a single information state: he only knows when it is his turn to play, he has no information about the state of Player 1. ***It is critical, when modeling a game, to enumerate all relevant information states.***

- *At each information state s, we let D_s be the set of actions that can be taken at this information state.*

Next, we add structure to represent in which sequence the actions can be taken.

- *We construct a tree on a set of nodes V. These represent three concepts: random events, decisions, or termination of the game, and we assume without loss of generality that all the payoffs are collected there. We thus represent these payoffs in the order corresponding to the players.*

  1. *We let V_0 be the set of* chance nodes. *In the tree of Figure 2.1, there is one chance node at the root of the tree with a label "0". It corresponds to drawing a card, which can be either black or red with probability 1/2 each.*
  2. *For each information state s, we let Y_s be the set of* decision nodes *for the state s. Note that, in the above, Y_c has two nodes: we use this to represent the fact that when at state c, Player 2 does not have perfect information about the state of the game.*
  3. *Finally, we let X be the set of* leafs *in the tree. These correspond to the termination of the game.*

- *We then describe how we may navigate the tree.*

  1. *To each chance node $v \in V_0$, we attach a distribution $p_x \in \Delta(V)$, which corresponds to the probability of reaching some node in the graph from the node v as a result of uncertainty.*
  2. *For each state $s \in \bigcup S_i$, for each node $y \in Y_s$ and each action $d \in D_s$ that can be chosen at this state, we let $t_s(y, d) \in V$ be the destination node attained in the tree.*

Once the tree is defined (of course, the distributions p_v and transition functions t_s much be chosen such that we obtain a tree), it actually provides a sequential representation of the game.

- *The last element we need is a* utility function $w : X \to \mathbb{R}^N$*, that assigns a value to each player in each leaf of the tree. Here, we simply put the amount of money each player has in the end of the game, assuming they each begin with 2 coins in their pockets.*

Notation 2.5

If $f_1, f_2, \dots, f_K$ are scalar functions, with $f_i : S \to \mathbb{R}$, then $f = (f_i)_{i\in\{1,\dots,K\}}$ is a vector function $f : S \to \mathbb{R}^K$. For $s \in S$, $f(s) = (f_i(s))_{i\in\{1,\dots,K\}}$.

Definition 2.6: Games in extensive form

A game in extensive form is a structure $\Gamma_e = (N, S, D; V, V_0, Y, X; t, p; w)$, where

- N is a set of players[a].
- $S = \bigcup_{i\in N} S_i$ is a set of *information states*, where S_i is the set of information state for player i, and $S_i \bigcap S_j = \emptyset$ if $i \neq j$.
- $D = \bigcup_{s\in S} D_s$ is the set of *moves*, where D_s is the set of moves available for a player in information state $s \in S$.
- V is a set of *nodes*.
- $V_0 \subset V$ is the set of *chance nodes*.
- $Y = \bigcup_{s\in S} Y_s \subset V$ is the set of *action nodes*, where Y_s is the set of nodes with information state s.
- $X \subset V$ is the set of *leafs*.
- $t = (t_s)_{s\in S}$ are transitions, where for each information state s, $t_s : Y_s \times D_s \to V$ assigns to every *source node* $y \in Y_s$ and action $d \in D_s$ a destination node $x \in V$.
- $p = (p_x)_{x\in V_0}$ assigns a probability distribution $p_x \in \Delta(V)$ to each chance node $x \in V_0$.
- $w : X \to \mathbb{R}^N$ is a utility function, assigning to each leaf $x \in X$ the payoff of every players if the game ends up at that leaf.

[a]By a slight abuse of notations, we will use N to denote either the set of players, or the number of players in a game. We will clarify if needed.

Again, when presenting a game in extensive form, we prefer to rely on their graphical representation. Here are some conventions for their drawing:

Notation 2.7: Conventions for drawing an extensive form game

In order to *represent* the tree, we follow a classical set of conventions to label nodes and edges, as in Figure 2.1.

1. We put a label 0 on a chance node.
2. Under each leaf, we write a vector with one entry per player to denote the utilities of the players at this node.
3. We label each decision node with a pair of the form (player playing at the node, information state of the node).
4. We label each edge outgoing from a chance node with the probability of this transition being taken.
5. We label each edge outgoing from a node Y_s with its corresponding action in D_s.

Exercise 2.8 *Propose an extensive form model for the following games.*

- *Heads-or-tails, with 2 players each betting a dollar.*
- *Rock-Paper-Scissors., with 2 players each betting a dollar.*
- *The game of Example 2.1 where now, Benoît picks a card from a deck of 53 cards, with 26 red cards, 26 black cards, and one joker. If he gets a joker Matthew immediately wins the bet, and the game ends.*
- *The game of Example 2.1, when played twice in a row.*

Decision theoretic perspective

We make a short digression to discuss the concepts above from the perspective of Decision theory, through the concepts introduced in Chapter 1.

As a reminder, in decision theory, a player needs to make a decision which leads him to obtaining a prize. This is formalized by the concept of *lottery* (see Definition 1.4). A lottery is a function

$$f : \Omega \to \Delta(X),$$

where Ω is the set of possible states of the world, and X are the possible prizes (outcomes) the player may obtain. In this section, we take the time to equate the notions of states of the world and prizes with their parallels in the extensive form.

Let us begin with the concept of decision. In Chapter 1, we associated to each decision of the player a lottery. In our current context, the decision of a player i consists in choosing, at all $s \in S_i$, an action $d \in D_s$ to perform.
This leads us to the concept of *pure strategy*.

Definition 2.9: Pure strategy

Given a game in extensive form Γ_e, a *pure strategy* for a player $i \in N$ is an element $c \in C_i = \times_{s \in S_i} D_s$, acting as a function that associates to each information state of the player a *move* to be played at that state.

Example 2.10 *In Example 2.1, Benoît has 4 pure strategies. Indeed, he has 2 information states, with two moves at each. A possible strategy is* play Raise on a red card, play fold on a Black card*.*
Matthew has 2 pure strategies, since he has only 1 information state and 2 available moves at this state. These strategies are always raise *and* always pass*.*

Later, and in particular in Chapter 4, we consider the possibility for a player to chose a randomized strategy: he will pick at each state $s \in S_i$ a distribution $\sigma_s \in \Delta(D_s)$, and pick $d \in D_s$ with probability $\sigma_s(d)$.

Once such a decision in made, the player receives a prize depending on the realization of uncertainty at chance nodes (which represent objective probabilities) as well as the moves of other players (which are subjective in the sense that they are not defined as part of the game). Hence, the *state of the world* for a player making a decision at state Y_s, denoted Ω_s, will be an element of the set containing all the subsequent decisions played by other players at the later information states.

Interestingly, this alone teaches us something about how a rational and intelligent decision maker should reflect on a game! Indeed, the Utility Maximization Theorem states that a player should possess a *conditional probability function*, expressing thus his own beliefs on what the other players will do in the game. We will develop that idea in full extent in Chapter 4.

2.1.2 The strategic form

A *strategic form game* can be seen as a summary of an extensive form game, where we keep only the answers to the questions

Who are the players, what can they do, and what they will receive by doing so?

To answer this question, we focus on *pure strategies* and they *payoffs*.

Definition 2.11: The strategic form

A *strategic form game* is a tuple

$$\Gamma = (N, C, u),$$

where

- N is the non-empty set of players in the game,
- $C = \bigtimes_{i \in N} C_i$ is the set of all *pure strategies* for all players,

- and $u = (u_i)_{i \in N} : C \to \mathbb{R}^N$ is the payoff function. The payoff $u_i : C \to \mathbb{R}$ of player i is the expected payoff of the player.

When constructing a strategic form game from a game in extensive form, the payoffs are computed as follows:

$$u_i(c) = \sum_{\text{leaf } \ell \text{ in the tree}} p(\ell|c) w_i(\ell),$$

where $w_i(\ell)$ is the payoff of player i at a leaf denoted by ℓ in the tree, and $p(\ell|c)$ is the probability of reaching that leaf if the players picked their strategies following c.

Throughout the course, we will focus on finite games, that are games where both the sets N and C are finite.

Example 2.12 *For Example 2.1, the set of players is $N = \{1, 2\}$, the pure strategies for each players are*

$$C_1 = \begin{pmatrix} Raise/raise, \\ Raise/fold, \\ Fold/raise, \\ Fold/fold. \end{pmatrix}, C_2 = \begin{pmatrix} meet, \\ pass. \end{pmatrix},$$

and the set C is the cartesian product of C_1 and C_2. Consider now the strategy profile $c = (Raise/fold, meet)$. The payoffs are then given by

$$u(c) = \frac{1}{2} \cdot (4, 0) + 0 \cdot (3, 1) + 0 \cdot (0, 4) + 0 \cdot (3, 1) + \frac{1}{2} \cdot (1, 3) = (2.5, 1.5).$$

We now introduce the *normal representation of a game in strategic form*, a compact way to present a strategic form game. A normal representation is a table in N dimensions (one dimension per player), and represents the payoff function $u : C \to \mathbb{R}^N$. We prefer an example to a definition here:

Example 2.13 *The* normal representation *of the game of Example* 2.1 *is given at Figure 2.2. By convention, we always refer to the player on the left of the table as the* first *player, which means that the first entry in the payoff vector is always his payoff.*

Benoît vs Matthew	meet	pass
Raise/raise	2, 2	3, 1
Raise/fold	2.5, 1.5	2, 2
Fold/raise	1.5, 2.5	3, 1
Fold/fold	2, 2	2, 2

Figure 2.2: Normal representation of Example 2.1

The strategic form is built using *pure* strategies. In practice, there are games where one would rather *randomize* between strategies. For example, in rock-paper-scissor, the pure strategies are to play *rock*, *paper* or *scissors*. However, in practice, people often try to play rock, paper or scissor with probability 1/3 each.

Definition 2.14: Randomized strategy

Given a player $i \in N$ and its set of strategies C_i, a *randomized strategy* for the player is a probability distribution $\sigma_i \in \Delta(C_i)$:

$$\forall c_i \in C_i : \sigma_i(c_i) \geq 0; \qquad \sum_{c_i \in C_i} \sigma_i(c_i) = 1.$$

A randomized strategy profile $\sigma \in \Delta(C)$ is a set of one randomized strategy per player. The *expected payoff* of player $i \in N$ for the profile $\sigma \in \Delta(C)$ is given by

$$E_\sigma(u_i(c)) = \sum_{c \in C} \sigma(c) u_i(c) = \sum_{c_i \in C_i} \sigma_i(c_i) \sum_{c_{-i} \in C_{-i}} \sigma_{-i}(c_{-i}) u_i(c_i, c_{-i}),$$

where "$-i$" denotes the set of all players except i, and $\sigma_{-i}(c_{-i}) = \prod_{j \neq i} \sigma_j(c_j)$, with $c_{-i} = (c_j)_{j \neq i}$.

Given a game in extensive form, there is a unique strategic form representation for the game. The reverse is not true. In particular, if we are given a game in strategic form, we cannot retrieve the order in which the players will play the game.
This sequentiality is sometimes necessary to define the behavior of intelligent and rational agents (or, oppositely, to rule out some behaviours as being irrational). We will see that in Chapter 4.
However, it is often sufficient to study the normal representation of a game for analysis.

2.2 Best responses and domination

In the previous section, we *represented games*. We did not approach the following fundamental question: **What should (or will?) each player do in a game?**

This question is similar to the one targeted by *Decision Theory* (Chapter 1). For game theory, we ask ourselves what should be the decision taken by a single player if he wants to maximize his own gains, *and he knows that the other players too.*

We will first introduce the concept of *best response strategy*. It is a very important concept for understanding the behavior of rational players, because this simple concept, derived from our axioms of decision theory, will lead as to a (partial) solution to the above question. The rationale is the following: if a particular strategy can never be optimal, whatever the other players do, then, one can for sure rule it out of its sensible actions.

2.2.1 Best responses

Definition 2.15: Best response

Consider a player $i \in N$. If the players $-i$ play following the strategy profile

$\sigma_{-i} \in \Delta(C_{-i})$, then the set of *best response strategies* for player $i \in N$ is given by

$$\arg\max_{c_i \in C_i} \sum_{c_{-i} \in C_{-i}} \sigma_{-i}(c_{-i}) u_i(c_i, c_{-i}).$$

The fact that best responses are a central concept in game theory can be seen as a direct consequence of decision theory. If player i associates a probability $\sigma_{-i}(c_{-i})$ that other players play the strategy c_{-i}, then picking a best-response strategy corresponds to maximizing its expected utility (Theorem 1.13), which is the expected behavior of an intelligent and rational player.

Now, one can push this reasoning further: Since we can assume that every player in the game wants to pick a best response, we can iterate this argument, allowing us to refine our prediction (or recommendation) on the decision of the players. Indeed, the choice of a best response for player i is driven by what he thinks others will play: $\sigma_{-i}(c_{-i})$. However, since the other players want to pick best responses as well, player i might be able to adapt his beliefs $\sigma_{-i}(c_{-i})$.

This observation leads to a recursive way to find the *best* strategy to play. First, pick an initial strategy σ_{-i}, then repeat

- Get c_i as a best response to σ_{-i},
- Update σ_{-i} so that it corresponds to a best response from the other players to c_i.

Sadly, things are not that easy. It is straightforward to build examples where this approach does not converge...

Example 2.16 *Take again the game of example 2.1. Benoît wants to find the best strategy. We advise you to follow the reasoning by referring to the normal representation of the game at Figure 2.2.*

- *Assume first Matthew is going to "pass" everytime. Then Benoît should play "Raise/raise" or "Fold/raise", which guarantees a payoff of* 3$.
- *If Benoîts plays as defined above, Matthew should play "meet", which gives him an expected payoff of* 2.25$.
- *In that case Benoît should play "Raise/fold"!...*
- *...to which the best response from Matthew is "pass", which brings us back at the beginning of the analysis.*

We have to be more clever...

2.2.2 Domination

A rational player always plays what he believes is a best response to the other moves. Knowing this, we can sometimes rule out the possibility that one player plays a particular strategy c_i, in particular when we can show that c_i can never be a best response.

Definition 2.17: Strong domination

For a player $i \in N$, a strategy d is *strongly* dominated if it is *never* a best response:

$$\forall \sigma_{-i} \in \Delta(C_{-i}),\ d \notin \arg\max_{c_i \in C_i} \sum_{c_{-i} \in C_{-i}} \sigma_{-i}(c_{-i}) u_i(c_i, c_{-i}).$$

Example 2.18 *Corentin and François play the game in normal representation at Figure 2.3. Have a closer look to strategies B and D, and observe that for any pure strategy of Corentin,* François always has a greater payoff by playing D instead of B. *François being an intelligent and rational player, he will never play B, which cannot be a best response and is strongly dominated. We could even erase B from the game!*

Corentin vs François	A	B	C	D
a	8, 5	4, 6	1, 6	7, 8
b	5, 6	7, 7	4, 4	2, 9
c	6, 7	5, 5	2, 6	3, 6
d	5, 4	5, 2	6, 5	7, 5

Figure 2.3: A two player game with 4 pure strategies each. Corentin plays a,b,c,d; François plays A,B,C,D.

There is an equivalent definition of strong domination, which is often easier to check. With the following, we also define the idea of *weak domination*: in some cases, a player might play a weakly dominated strategy, but there is always another strategy at least as good than the first one, better in some cases.

Theorem 2.19: Strong domination

A pure strategy d_i is strongly dominated if there exists a randomized strategy $\sigma_i \in \Delta(C_i)$ such that, $\forall c_{-i} \in C_{-i}$:

$$u_i(d_i, c_{-i}) < \sum_{e_i \in C_i} \sigma_i(e_i) u_i(e_i, c_{-i}). \tag{2.1}$$

The above theorem is a simple (and beautiful) application of duality of Linear Programming, and its proof is left for the students.

Definition 2.20: Weak domination

A pure strategy d_i is *weakly dominated* if there exists a randomized strategy $\sigma_i \in \Delta(C_i)$ such that, $\forall c_{-i} \in C_{-i}$:

$$u_i(d_i, c_{-i}) \leq \sum_{e_i \in C_i} \sigma_i(e_i) u_i(e_i, c_{-i}), \tag{2.2}$$

with the inequality being strict for at least one[a] strategy in c_{-i}.

[a]naturally, if the inequality is never strict, then the strategies are equivalent.

Even though the concept of weak domination also enjoys a dual definition as in Theorem 2.19, we have to be more careful when discarding a strategy on the grounds that it is weakly dominated, as we will see in the next section.

2.3 Equivalent games

In this section, we characterize different notions of *equivalence* between games. Equivalence here relates to the similar notion present in Decision Theory (see Chapter 1, Theorem 1.15). Intuitively, we will seek for a notion of equivalence that encompasses the fact that players behave the same in both games. So, two games might have very different available payoffs but yet be equivalent, if these payoffs can never occur in practice (i.e. in situations where players are rational.)

Definition 2.21: Full equivalence

The two games $\Gamma = (N, C, u)$ and $\hat{\Gamma} = (N, C, \hat{u})$ are *fully equivalent* if

$$\forall i \in N, \exists A_i > 0, B_i : \hat{u}_i(c) = A_i u_i(c) + B_i, \forall c \in C.$$

Example 2.22 *Adeline and Emilie want to model the rock-paper-scissor game with a normal representation.*
Adeline's version is as follows:

	R	*P*	*S*
r	*0, 0*	*-1, 1*	*1, -1*
p	*1, -1*	*0, 0*	*-1, 1*
s	*-1, 1*	*1, -1*	*0, 0*

,

Emilie's version is as follows:

	R	*P*	*S*
r	*1, 1*	*-1, 3*	*3, -1*
p	*3, -1*	*1, 1*	*-1, 3*
s	*-1, 3*	*3, -1*	*1, 1*

.

Let us now reflect on two questions:

- *Which one would you rather play?*
- *Would you behave differently in playing one game or the other?*

The first one is a tricky question and id depends on the utility scale in each game.
The second question however, admits a clear answer: a rational and intelligent decision maker will make the same decisions in both games, that is, the games are fully equivalent. Indeed, he payoff in the second game are obtained by multiplying the ones in the first by 2, and then adding 1.

Full equivalence is a property that is difficult to satisfy, and is mostly used in proofs. The next definition allows for an even more succint representation of the game, with a focus on important strategies.

Definition 2.23: Best-response equivalence

Two games $\Gamma = (N, C, u)$ and $\hat{\Gamma} = (N, C, \hat{u})$ are *best-response* equivalent if and only if they have the same best response sets. That is, $\forall i \in N$, $\forall \sigma_{-i} \in \Delta C_{-i}$:

$$c_i' \in \arg\max_{c_i \in C_i} \sum_{c_{-i} \in C_{-i}} \sigma_{-i}(c_{-i}) u_i(c_i, c_{-i}) \Leftrightarrow c_i' \in \arg\max_{c_i \in C_i} \sum_{c_{-i} \in C_{-i}} \sigma_{-i}(c_{-i}) \hat{u}_i(c_i, c_{-i}).$$

Full equivalence implies best-response equivalence. Remember that decision makers will always pick what they believe is a best response. Once they have fixed an assumption on the actions of the other players, they play a best responses to maximize expected payoff. When two games are best-response equivalent, the decision making is therefore preserved.

Example 2.24 *Vladimir and Nikos play the two following games:*

$$\Gamma = \left\{ \begin{array}{c|cc} & x_2 & y_2 \\ \hline x_1 & \mathit{9,\ 9} & \mathit{0,\ 8} \\ y_1 & \mathit{8,\ 0} & \mathit{7,\ 7} \end{array} \right\}, \qquad \hat{\Gamma} = \left\{ \begin{array}{c|cc} & x_2 & y_2 \\ \hline x_1 & \mathit{1,\ 1} & \mathit{0,\ 0} \\ y_1 & \mathit{0,\ 0} & \mathit{7,\ 7} \end{array} \right\},$$

and we wonder whether they will behave differently in one game or another. We will show that Γ and $\hat{\Gamma}$ are best response equivalent, and therefore the players should behave the same in both game. To do so, we exploit the symmetry of the game.
We put ourselves in the shoes of Player 1, and assume that Player 2 plays x_2 with probability α and y_2 with probability $1 - \alpha$.
In the first game, the expected payoffs are

$$u_1(x_1, \alpha x_2 + (1-\alpha) y_2) = 9\alpha, \qquad u_1(y_1, \alpha x_2 + (1-\alpha) y_2) = \alpha + 7.$$

Therefore, we play x_1 if and only if $9\alpha \geq \alpha + 7$, or $\alpha \geq 7/8$ (plain lines in Figure 2.4).
In the second game, we have

$$\hat{u}_1(x_1, \alpha x_2 + (1-\alpha) y_2) = \alpha, \qquad \hat{u}_1(y_1, \alpha x_2 + (1-\alpha) y_2) = 7 - 7\alpha.$$

Again, we will play x_1 if and only if $\alpha \geq 7/8$ (dotted lines in Figure 2.4), and we can conclude best response equivalence.

2.3.1 Reducing a game

As always when studying a problem, the complexity of the analysis grows as the size of the problem grows. In this subsection, we give techniques that allow to reduce the size of a game by removing strategies.

Elimination of redundant strategies.

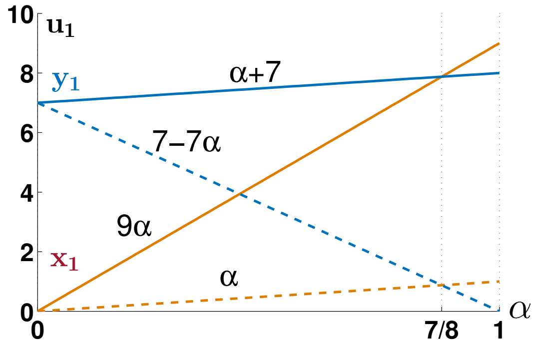


Figure 2.4: Best responses illustrated.

Definition 2.25: Redundant strategies

We say that the pure strategy $c_i \in C_i$ is *randomly redundant* or *payoff equivalent* if there is a random strategy $\sigma_i \in \Delta(C_i \setminus c_i)$ such that $\forall c_{-i} \in C_{-i}$ and for all players $j \in N$

$$u_j(c_i, c_{-i}) = \sum_{e_i \in C_i} \sigma_i(e_i) u_j(e_i, c_{-i}).$$

For a player i who has a redundant strategy, this strategy can be removed and replaced by a randomized one with the same effect.
Take another player now, and say he believes the player i will play c_i with probability α. Then, the best response set to that situation is the same to the one where player i plays the equivalent strategy with probability α, and c_i with probability 0.
If a randomly redundant strategy is found, then we can remove it from the game without changing the set of best responses (except, maybe, by removing from it redundant strategies). The normal representation obtained after removing all redundant strategies from a game is called the *fully reduced normal representation*.

Elimination of dominated strategies.

In order to further reduce a game, we may try to construct a *best response equivalent game* which would be *as small as possible*. A natural approach is to remove all the strategies that can never be best-responses, which are, by definition, *strongly dominated strategies*.
The following recursive procedure allows to eliminate all *strongly dominated* strategies from a game.

Procedure 2.26: Elimination of dominated strategies

Consider a strategic form game Γ. The following procedure converges:

1. Let $\Gamma^0 = \Gamma$ be a game in strategic form, an let $k = 0$.

2. Let Γ^{k+1} be the game formed by removing all (or at least some) strongly

dominated strategies from Γ^k.

3. If $\Gamma^{k+1} = \Gamma^k$, stop, and let $\Gamma^\infty = \Gamma^k$. Else, do $k := k+1$, and go back to step 2.

It is easy to show that any agent will always play a strategy of the game Γ^∞. Indeed, for any $k \geq 0$, a player should always play a strategy from the game $k+1$, to avoid strong domination.
The above procedure is certain to converge since we consider finite games. There can only be a finite number of strategies to remove.
A less obvious fact is that the game Γ^∞ is unique. In fact, we can consider another procedure where, at step 2, only one strongly dominated strategy is picked and removed. We would still converge to the same Γ^∞.
If we try to eliminate *weakly* dominated strategies, then the solution is no longer unique, as illustrated in the next example.

Example 2.27 *Consider the game of Example 2.18. The game Γ^0 is the one represented in Figure 2.3. We already saw that strategy B was dominated by strategy D, so Γ^1 does not contain B. Removing this strategy, we obtain:*

$$\Gamma^1 = \left\{ \begin{array}{c|ccc} \textit{Corentin vs François} & A & C & D \\ \hline a & 8,5 & 1,6 & 7,8 \\ b & 5,6 & 4,4 & 2,9 \\ c & 6,7 & 2,6 & 3,6 \\ d & 5,4 & 6,5 & 7,5 \end{array} \right\}.$$

Inspecting Γ^1, we see that if Corentin was to randomize between a and d, he might get higher payoff than with strategy c. Let us verify this: if he plays a with probability α and d with probability $(1-\alpha)$ then c is strongly dominated if

$$8\alpha + 5(1-\alpha) > 6;\ \alpha + 6(1-\alpha) > 2;$$

which holds true for $1/3 < \alpha < 4/5$.
Moreover, strategy b is also strongly dominated by $\beta a + (1-\beta)d$ for all $0 < \beta < 2/5$.
We therefore obtain

$$\Gamma^2 = \left\{ \begin{array}{c|ccc} \textit{Corentin vs François} & A & C & D \\ \hline a & 8,5 & 1,6 & 7,8 \\ d & 5,4 & 6,5 & 7,5 \end{array} \right\}.$$

We can now see that A is strongly dominated by D, and obtain the game

$$\Gamma^\infty = \Gamma^3 = \left\{ \begin{array}{c|cc} \textit{Corentin vs François} & C & D \\ \hline a & 1,6 & 7,8 \\ d & 6,5 & 7,5 \end{array} \right\}.$$

Remark that 1) by inspecting Γ^1 only, we cannot see the domination of A by D, which justify the recursive elimination procedure (Procedure 2.26).

Observe now that there is some weak *domination in Γ^∞. However, if we wanted to eliminate those strategies, we would not obtain an unique solution. For example, removing a in favor of d, we have*

$$\Gamma^\infty \backslash \{a\} = \left\{ \begin{array}{c|cc} \textit{Corentin vs François} & C & D \\ \hline d & \textit{6, 5} & \textit{7, 5} \end{array} \right\},$$

but we could also remove C in favor of D:

$$\Gamma^\infty \backslash \{C\} = \left\{ \begin{array}{c|c} \textit{Corentin vs François} & D \\ \hline a & \textit{7, 8} \\ d & \textit{7, 5} \end{array} \right\}.$$

In any case, once we have removed one, we cannot remove the other.

2.4 Common knowledge and private information

In many games, players have access to private information. Examples include the color of the card Benoît sees in example 2.1, a poker hand, the maximum amount of money one is willing to put on an auction... When we think of it, in many games involving some kind of "bluff", we often try to convince other players that our private information is the most advantageous to us. In example 2.1, Matthew would rather play "meet" if he thinks Benoît has a black card, and "pass" if he has a red card. Therefore, Benoît would prefer Matthew to believe he has a red card when he in fact has a black card, and vice-versa...

This being said, we understand how this concept of private information comes to play a very important role in the analysis of games. Let us first define formally what private information, and its opposite *common knowledge* means.

Definition 2.28: Common knowledge

A *private information* is any knowledge on the state of the world that is not a *common knowledge.* An information X is *common knowledge* if the statement

$$(\textit{everybody knows that})^K \ \textit{everybody knows } X$$

holds for all $K = 1, 2, \ldots$.

The somewhat recursive definition of common knowledge seems to be complicated. It may appear that "I know X" and "You know X" should be enough so that X is common knowledge. The following well known example shows otherwise.

Example 2.29 *We present here the famous* paradox of the generals*. Two generals, Raphaël and Julien plan on capturing an enemy fort in a valley. Their respective encampments are located at opposite sides of the valley. The only viable tactic is a coordinated attack: they must attack on both sides of the fort at the same time. Else, the fort defences will organize and easily defeat them.*
The generals need to agree on the hour of the attack. To do so, they send messengers through the valley.

- *Assume first that there is a 100 % chance that a messenger sent by one side arrives at the other side. Then, if Raphaël sends a messenger, he knows Julien will receive the message. If Raphaël says* I'm attacking before lunch*, then Julien will necessarly follow since not doing so implies a defeat.*

- *Assume now that both Julien and Raphaël* believe[1] *that there is a non zero probability that the messenger is intercepted during the carrying of the message.*

  1. *When Raphaël sends a message, he cannot be certain Julien receives it.*
  2. *Thus, Julien should confirm the reception of the message to Raphaël, by sending a messenger back.*
  3. *If Raphaël receives this message, then he knows that Julien knows the hour of the attack. But Julien will not move unless he is certain that Raphaël knows that Julien knows the hour of the attack. Thus, Raphaël sends another messenger to Julien.*
  4. *Now, Julien received this messenger, and knows that Raphaël knowns that Julien knows the hour of the attack. But...*

  At the end, what the story says is that if both generals want to be certain to win the battle, they will have to send an infinite number of messengers. This is because they want the hour of the attack to be a common knowledge.

2.4.1 Modelling games with incomplete information

A game with incomplete information is a game in which some player possesses private information on the game before it even begins. To represent these informations, we assign to the players a *type* that identifies the information. For example, players in an auction know before the auction the amount of money they are willing to spend. Then, we could say that a player is of type "I'll spend at most 1000 $", etc...

In what follows, we assume that there is a finite number of possible types for each player. The set of types for player i is written T_i.

A natural way to model a game is to assume that the types will be assigned before the beginning of a game with all the appropriate randomness. This leads us to an extensive form representation of the game, using a chance node at the root of the tree.

Example 2.30 *Consider example 2.1 with a twist. Matthew has a dangerous gambling addiction, and sees Benoît with a card in his pocket. Without hesitation, he proposes to Benoît a bet on the color of the card, where Benoît, as in example 2.1, can raise the bet if he wishes so.*
A way to model this game is to use the same extensive form to that of Figure 2.1. The difference is that the first chance node, instead of representing the drawing of the card, represents the realization of the type *of Benoît, that is the color of the card in his pocket.*

Modeling a game with private information using an extensive-form-like structure has some drawbacks. A first practical drawback concerns the size of the tree. For an N player game, the set of all type-profiles $T = (T_i)_{i \in N}$ can be quite big - and we need as many branches per type profile in T to represent the game.
A second drawback appears when considering the temporal aspects of the game. An extensive form presents a clear, sequential, step-by-step game, where the root to the tree is the beginning of the game. When representing the possible types in extensive form, this root corresponds to a moment *before* the beginning of the game...

[1] It does not need to be a fact!

In order to model games with incomplete information, a generalisation of the *strategic form*, called a *bayesian game* is often preferred to the extensive form.

Definition 2.31: Bayesian game

A *bayesian game* is a tuple

$$\Gamma^b = (N, C, T, p, u),$$

where

- N is the non-empty set of players,
- $C = (C_i)_{i \in N}$ is the set of all pure actions for all players,
- $T = (T_i)_{i \in N}$ is the set of all player types;
- $p = (p_i)_{i \in N}$ are *beliefs functions*,

$$p_i : T_i \to \Delta(T_{-i}),$$

- and $u = (u_i)_{i \in N}$ are the payoff functions,

$$u_i : C \times T \to \mathbb{R}.$$

The game is finite if N, C and T are finite.

Example 2.32 *Back to example 2.30.*

Benoît has two types: $T_1 = \{red, black\}$. Matthew has one type $T_2 = \{gambler\}$. Therefore, the set T contains only two entries. The payoff function can be represented by these two tables:

red	*meet*	*pass*
Raise	*4/0*	*3/1*
Fold	*3/1*	*3/1*

,

black	*meet*	*pass*
Raise	*0/4*	*3/1*
Fold	*1/3*	*1/3*

.

The belief function of Benoît is $p_1(gambler|red) = p_1(gambler|black) = 1$. The belief function of Matthew is $p_2(red|gambler) = \alpha$, $p_2(black|gambler) = 1 - \alpha$.

Representing the belief functions of each player can be cumbersome. This can be made easier is the beliefs are *consistent*:

Definition 2.33: Consistent beliefs

We say that the belief functions $p = (p_i)_{i \in N}$ are *consistent* if there exists an a-priori probability distribution $P \in \Delta(T)$ such that for all $t_{-i} \in T_{-i}$, $t_i \in T_i$,

$$p_i(t_{-i}|t_i) = \frac{P(t_i, t_{-i})}{\sum_{s \in T_{-i}} P(t_i, s)}$$

It turns out, we can show that any bayesian game is *equivalent* to a bayesian game with consistent beliefs.

In fact, bayesian games are not essentially different from general standard games. To see this, one can see that, given any bayesian game, one can build an equivalent standard game, named the 'type-agent representation' of the bayesian game.

Definition 2.34: Type-agent representation

Given a Bayesian game

$$\Gamma^b = (N, C, T, p, u),$$

the type-agent representation the following strategic form game:

- The set of agents is $T_1 + T_2 + \ldots + T_N$. We write $t_k \in T_i$ to denote the fact that *agent* k belongs to player i.
- If $t_k \in T_i$, then the pure strategies for t_k are C_i.
- The payoff function of agent $t_k \in T_i$ is given by

$$v_{t_k}(c_{t_1}, \ldots, c_{t_T}) = \sum_{t_{-i}} p_i(t_{-i}|t_k) u_i(c, (t_k, t_{-i})).$$

Example 2.35 *The following table shows the payoffs for the type-agent representation of the game of example 2.30. We assume that the probabilities that Benoît has a red or black card are the same. These payoffs are to be read*

$$(\textit{Matthew's payoff, red Benoît's payoff, black Benoît's payoff}).$$

	red:Raise		*red:Fold*	
	black:raise	*black:fold*	*black:raise*	*black:fold*
meet	*2, 4, 0*	*1.5, 4, 1*	*2.5, 3, 0*	*2, 3, 1*
pass	*1, 3, 3*	*2, 3, 1*	*1, 3, 3*	*2, 3, 1*

The entries of the table are computed as follows:

1. *The first entry of the triple in each cell corresponds to Matthew's payoff. As the probability that Benoît has a red or black card is* $1/2$*, the payoff is the average between the payoff in the first table of example 2.30 an the payoff in the second table. For instance, in the entry (meet, Raise, raise), we compute* $4/2 + 0/2 = 2$.
2. *The second entry of the triple in each cell corresponds to Benoît's payoff with a red card, it is simply the payoff in the first table of example 2.30.*
3. *Similarly, as the third entry of the triple in each cell corresponds to Benoît's payoff with a black card, it is the payoff in the second table of example 2.30.*

Chapter 3

The Nash Equilibrium

"People usually call them Nash equilibria, but I just call them equilibria." — John Nash (1928 – 2015).

Chapter based on pages 91-108 and 122-127 of the book "Game theory - Analysis of conflict" by R. Myerson.

The utility maximization theorem of the first chapter *quantifies* the behaviour of a rational and intelligent agent, having to make decisions in an uncertain environment. We can always expect an agent to perceive the uncertainties through a *conditional probability function*, and have a *utility function*, which quantifies the value of a *prize* obtained as a result of his decision for every realization of the uncertainty.

When considering a game with two or more players, the *strategic form* introduced in Chapter 2 allows to abstract the objective probabilities, that is, the one that can be objectively described as part of the rules of the game (the outcome of a dice roll, the flip of a coin, ...). As a consequence, the outcome of a game considered to be given by a simultaneous choice of the players. However, the outcome of such a game is far from clear, even though we were able to get rid of these objective probabilities. Indeed, for a particular player to select his action (what we called a "lottery" in the framework of decision theory), one must specify the subjective probabilities for every "state of the world" (the events that are possible, but for which no a priori probabilities are given). The Utility Maximization Theorem does not enable us to compute the best decision to make without these important elements.

We now push further our analysis by incorporating the simultaneity in the actions of the different players. Given a game in strategic form, we wish to provide any kind of characterization on what will be the outcome of this game. This characterization will take the form of (necessary and/or sufficient) conditions on which strategy the players will select. We will call such a characterization a *solution concept*, that is, a model, an idea of how the players are going to play.

In chapter 2, we already exhibited an important rule that must be followed by an intelligent and rational player, which is that *he should always play a best response to what he believes the others should play.*

The concept of *Nash Equilibrium* is a natural application of this principle, and it is probably one of the most important concepts of game-theory. A Nash equilibrium is a situation where the actions of all the players are all consistent with the fact that they

are rational and intelligent. The payoff each agent obtains is the highest he can obtain given what he thinks the other players will do. It must then be considered as a necessary condition on the outcome of the game, if we accept the axioms of decision theory.
The goal of this chapter is to further define, formalize and analyse the concept of Nash Equilibrium, as well as to give tools for computing Nash equilibria of games in strategic form.

3.1 Nash equilibria for strategic form games

We are considering strategic form games $\Gamma = (N, (C_i)_{i \in N}, (u_i)_i) \in N$, as defined in Chapter 2. The players in the game each decide on a *strategy* they will play, which assigns a different probability of playing each of their available actions (in the set C_i for player i). Recall that we use $\Delta(C_i)$ to denote the set of all probability distributions on C_i; that is, the set of all randomized strategies player i could chose.

A *strategy profile* σ is a choice of one randomized strategy per player: $\sigma = (\sigma_1, \dots, \sigma_N)$, with $\sigma_i \in \Delta(C_i)$. A strategy profile is said to be a Nash equilibrium if, roughly speaking, for each player $i \in N$, when considering that all the other players play according to σ_{-i}, playing σ_i maximizes the payoff of player i. The following is a more formal definition of a Nash equilibrium[1]. It also introduces the notion of *support* of the equilibrium.

Definition 3.1: Nash Equilibrium

Consider a game in strategic form $\Gamma = (N, (C_i)_{i \in N}, (u_i)_{i \in N})$. A Nash Equilibrium is a (randomized) strategy profile $\sigma_1, \dots, \sigma_N$, with $\sigma_i \in \Delta(C_i)$ and with support $D_i \subseteq C_i$, satisfying to the following conditions:

1. For all $i \in N$, σ_i is a randomized strategy with *support* D_i:

$$\sum_{c_i \in D_i} \sigma_i(c_i) = 1, \tag{3.1}$$

$$\forall d_i \in D_i : \quad \sigma_i(d_i) \geq 0, \tag{3.2}$$

$$\forall e_i \in C_i \backslash D_i : \quad \sigma_i(e_i) = 0. \tag{3.3}$$

2. For all $i \in N$, all the strategies $d_i \in D_i$ are *best response* to σ_{-i} and as a consequence, their payoffs are equal:

$$\forall d_i \in D_i : \quad \sum_{c_{-i} \in D_{-i}} \left(\prod_{j \in N-i} \sigma_j(c_j) \right) u_i(c_{-i}, d_i) = w_i, \tag{3.4}$$

$$\forall e_i \in C_i \backslash D_i : \quad \sum_{c_{-i} \in D_{-i}} \left(\prod_{j \in N-i} \sigma_j(c_j) \right) u_i(c_{-i}, e_i) \leq w_i, \tag{3.5}$$

[1]Can you see how this definition is a mere consequence of the fundamental theorem of Decision Theory?

The first set of conditions (3.1, 3.2, 3.3) simply state that a player will play a randomized strategy obtained by picking a strategy at random within a subset of his available strategies (the support).
The fourth condition (3.4) is more interesting. The left hand side of the inequality is nothing but the *expected payoff* of player $i \in N$ when playing $d_i \in D_i$ if all other players follow the strategy dictated by σ (here, they play σ_{-i}) [2]. The condition states that this payoff (w_i) remains the same if the player was to change his strategy for another in his support D_i.
The fifth condition (3.5) translates the fact that for all player $i \in N$, the randomized strategy σ_i is indeed a best response to σ_{-i}; it makes sure that no player has an advantage in playing a strategy *outside* of his support.

3.1.1 Existence of Nash Equilibria

For *any* game in strategic form, there always exists at least one Nash equilibrium. This is a non-trivial result, proven by Nash himself in 1951.

Theorem 3.2

Given any finite game Γ in strategic form, there exists at least one Nash equilibrium (cfr Definition 3.1) $\sigma_1, \ldots, \sigma_N$, with $\sigma_i \in \Delta(C_i)$.

Proof 3.3 (Theorem 3.2) *Consider a game* $\Gamma = (N, (C_i)_{i\in N}, (u_i)_{i\in N})$.
The set of all randomized strategy profiles, i.e the cartesian product $\times_{i\in N}\Delta C_i$*, is a non-empty, convex, closed and bounded subset of* $\mathbb{R}^{\sum_{i\in N}|C_i|}$*, where* $|C_i|$ *is the number of different pure strategies available to player i.*
For each player $j \in N$*, we define the following* best response *map, which maps strategies of j's adversaries,* $\sigma_{-j} \in \times_{i\in N-j}\Delta C_i$*, to a* set of strategies $R_j(\sigma_{-j}) \subset \Delta(C_j)$*:*

$$R_j(\sigma_{-j}) = \arg\max_{\tau_j \in \Delta C_j} u_j(\sigma_{-j}, \tau_j) \subset \Delta(C_j).$$

By definition, $R_j(\sigma_{-j})$ *is the set of all best responses from j to* σ_{-j}*. It thus contains some pure strategies, and all convex combinations of these pure strategies*[3]*. Therefore, the set* $R_j(\sigma_{-j})$ *is a convex subset of* ΔC_j*.*

From these different maps we construct the map $R : \Delta(C) \to 2^{\Delta(C)}$*, where* $2^{\Delta(C)}$ *is the set of all sets included in* $\Delta(C)$ [4]*, defined as follows:*

$$R(\sigma) = \times_{i\in N} R_i(\sigma_{-i}).$$

A Nash equilibrium $\sigma \in \times_{i\in N}\Delta C_i$*, of* Γ *can be expressed through the map R as follows:*

$$\sigma \textit{ is a Nash equilibrium} \leftrightarrow \sigma \in R(\sigma),$$

[2]Indeed, the probability for the other players to play a combination of pure strategies $c_{-i} = (c_j)_{j\in N-i} \in D_{-i}$ when following σ_{-i} is given by $\prod_{j\in N-i}\sigma_j(c_j)$.

[3]Indeed, if τ_j^1 and τ_j^2 are such that $u_j(\sigma_{-j}, \tau_j^1) = u_j(\sigma_{-j}, \tau_j^2)$, then for all $0 \le \alpha \le 1$, $u_j(\sigma_{-j}, \tau_j^1) = u_j(\sigma_{-j}, \alpha\tau_j^1 + (1-\alpha)\tau_j^2) \triangleq \alpha u_j(\sigma_{-j}, \tau_j^1) + (1-\alpha)u_j(\sigma_{-j}, \tau_j^2)$.

[4]Equivalently, $S \in 2^{\Delta(C)} \Leftrightarrow S \subset \Delta(C)$.

in other words, if σ is a fixed-point *of the map R. Thus, by showing that R has a fixed point, we also prove the existence of a Nash equilibrium. To do so, we rely on the* Kakutani fixed-point theorem*:*

Theorem 3.4: Kakutani

Let S be any nonempty, convex, bounded and closed subset of a finite-dimensional vector space. Let $F : S \to 2^S$ be any upper-hemicontinuous *map such that, for every $x \in S$, $F(x)$ is a non-empty convex subset of S. Then there exists $\bar{x} \in S$ such that $\bar{x} \in F(\bar{x})$.*

Definition 3.5

The map $F : S \to 2^S$ is upper-hemicontinuous *if for all sequence $(a_n)_{n=1,2,\ldots} \in S$ and for all sequence $(b_n \in F(a_n))_{n=1,2,\ldots}$, if both sequences converge, i.e. $\lim_{n\to\infty} a_n = a$, $\lim_{n\to\infty} b_n = b$, then $b \in F(a)$.*

Therefore, if we can prove that the best-response map R is upper-hemicontinuous, Theorem 3.4 allows us to conclude the existence of a fixed point of the map R.

Take a sequence of randomized strategy profiles, $(\sigma^k)_{k=1,2,\ldots}$, and a sequence $\tau^k = R(\sigma^k)$. Assume that both sequences converge respectively to $\bar{\sigma}$ and $\bar{\tau}$. We wish to show that this implies $\bar{\tau} \in R(\bar{\sigma})$.
To do so, note that for any player j and any strategy $\rho \in \Delta C_j$, and all $k = 1, 2, \ldots$

$$u_j(\sigma^k_{-j}, \tau^k_j) \geq u_j(\sigma^k_{-j}, \rho).$$

By continuity of the utility function u_j on $\times \Delta C_i$, this implies that for all player j and all $\rho \in \Delta C_j$,

$$u_j(\bar{\sigma}_{-j}, \bar{\tau}_j) \geq u_j(\bar{\sigma}_{-j}, \rho),$$

which indeed confirms that $\bar{\tau}_j$ is a best response to $\bar{\sigma}_{-j}$, and thus $\bar{\tau} \in R(\bar{\sigma})$. This allows to conclude the proof by using Kakutani's fixed-point theorem: there exists a strategy profile σ such that $\sigma \in R(\sigma)$, and consequently, there exists a Nash equilibrium in Γ.

3.1.2 Computing Nash Equilibria

We now turn ourselves to the following question: given a game in strategic form, $\Gamma = (N, (C_i)_{i\in N}, (u_i)_{i\in N})$, provide the set of all the Nash Equilibria of the game. We know that at least one equilibrium exists from Theorem 3.2.
The algorithm presented below follows directly from the definition of the Nash equilibrium. It checks, for every combination of support $(D_i)_{i\in N}$, whether there exists a randomized strategy profile σ satisfying the equations (3.1, 3.2, 3.3, 3.4, 3.5).

Procedure 3.6

- For all player $i \in N$, for all pure strategy $c_k \in C_i$, define the variable

$$0 \leq \sigma_i(c_k) \leq 1 \text{ (the probability that } i \text{ plays } c_k).$$

- For all choices of *support* $D = (D_i)_{i \in N}$, $D_i \subseteq C_i$,

  1. Support: For each player $i \in N$ and strategy $c \in C_i$, set

  $$\sigma_i(c) = 0 \text{ if } c \in C_i \backslash D_i, \text{and } \sigma_i(c) \geq 0 \text{ if } c \in D_i,$$

  with $\sum_{c \in D_i} \sigma_i(c) = 1$.

  2. Payoffs: For all player $i \in N$, Eq. (3.4) should be satisfied:

  $$\forall d_i \in D_i : \qquad \sum_{c_{-i} \in D_{-i}} \left(\prod_{j \in N-i} \sigma_j(c_j) \right) u_i(c_{-i}, d_i) = w_i.$$

  3. Best Response: For all player $i \in N$, Eq. (3.5) should be satisfied:

  $$\forall e_i \in C_i \backslash D_i : \qquad \sum_{c_{-i} \in D_{-i}} \left(\prod_{j \in N-i} \sigma_j(c_j) \right) u_i(c_{-i}, e_i) \leq w_i.$$

A strategy profile σ is a Nash-Equilibrium on a support D if and only if it satisfies the steps 1, 2 and 3 above.

Applying the procedure can be long (and boring), which may lead to making mistakes such as forgetting to consider a support, etc... In practice, it is often interesting to use the structure (symmetries, dominations, etc...) of the game considered to shortcut the procedure. You should always start by reducing the game by removing strongly dominated strategies. In some scenarios, you may also remove weakly dominated strategies (after showing, for example, that they can never be part of an equilibrium). Symmetries are useful in that the conclusions for a choice of support might carry on to other choices.
Overall, a good understanding of the notion of best-response may considerably speed-up the search for Nash equilibria.
In the following examples, we compute Nash equilibria, and sometimes use shortcuts to fasten the process (discarding supports in the process). Applying Procedure 2 directly would also lead to the same conclusions, feel free to do so yourselves and compare the results.

Example 3.7 (Prisoner's Dilemma) *Prisoner's Dilemma are a paradigm of games where when the players play rationally, they will most likely end up in a bad (as in sub-optimal) situation.*
*The usual story behind it is the following: two guys have been made prisoners for a crime that was apparently committed by at least one of them. The police is interrogating each one of them separately. The prisoners have the choice either to accuse the other one of being guilty (*betray*), or to remain silent (*cooperate*). If one betrayed the other but the second remained silent, then the betrayer will be set free, and the other is sent to jail alone. If both betray, they will both go to jail. If both remain silent, then they will conserve a good reputation, which is good for both of them.*
An example in strategic form is given on table 3.1.

Prisoner's Dilemma	b	c
B	1, 1	4, 0
C	0, 4	3, 3

Figure 3.1: Prisoner's Dilemma. "b" and "B" stand for betray, "c" and "C" stand for cooperate.

There is a direct way to compute the Nash equilibria of the game. Indeed, notice that B strongly dominates C, and b strongly dominates c. As a conclusion, in any case, it is never in the interest of the players to play C or c. Thus, the only possible equilibrium is for both players to betray.

For the sake of the exercise[5]*, let us compute the Nash Equilibria of the game in a* direct way*, applying Procedure 3.6.*
Since both players have 2 possible actions, there are 9 distinct support combinations to check: four using only pure strategies, four when one player plays a pure strategy, and the other a randomized one, and a last one when both randomize.
We now apply the procedure, by exploring each choice of support.

- *Pure strategy supports.*

  When considering pure strategies, step 2 of the procedure is immediate to check (D_i contains only one element).

  - ***{C}*** $\times$ ***{c}***.
    In step 2, the payoffs for each player, are $w_1 = w_2 = 3$.
    It is easy to see that if the first player played B instead of C, we have

    $$3 = w_1 < u_1(B, c) = 4.$$

    This means that eq. (3.5) does not hold, so C vs c is not a equilibrium.

  - ***{B}*** $\times$ ***{c}***.
    In step 2, the payoffs for each player, are $w_1 = 4$, $w_2 = 0$. We notice again that

    $$0 = w_2 < u_2(B, b) = 1.$$

    This means that eq. (3.5) does not hold, so B vs b is not a equilibrium.

  - ***{C}*** $\times$ ***{b}***.
    In that case, applying the procedure will not give us more information than what we already have. Indeed, the case is symmetric *with the case studied just above, B vs c. The current set of strategies do not constitute an equilibrium since Player 1 should play B instead of C.*

  - ***{B}*** $\times$ ***{b}***.
    In this last case, we do have an equilibrium. Indeed, $w_1 = w_2 = 1$. We check that eq. (3.5) does hold:

    $$w_1 \geq u_1(C, b) = 0, \text{ and } w_2 \geq u_2(B, c) = 0.$$

[5] This takes considerably longer than the argument based on domination.

- *Randomized vs pure strategy supports. This corresponds to the case when e.g. the first player randomizes between B and C, while the second sticks either to b or to c.*
  - ***{B, C}*** $\times$ ***{b}***.
    Thus, for Player 1, the system of eq. 3.4 reads $u_1(B,b) = u_1(C,b) = w_1$, which cannot hold since $u_1(B,b) > u_1(C,b)$. Thus, the support does not lead to a Nash equilibrium.
  - ***{B, C}*** $\times$ ***{c}***.
    Similarly to the above, the support will not lead to a Nash equilibrium. The system of eq. 3.4 requires $u_1(B,c) = u_1(C,c) = w_1$, which fails to be true.
  - ***{B}*** $\times$ ***{b, c}*** *and* ***{C}*** $\times$ ***{b,c}***.
    These are symmetric to the two cases above. No equilibria.
- *Fully randomized strategy:* ***{B, C}*** $\times$ ***{b, c} vs {B, C}*** $\times$ ***{b, c}***.
  For Player 1, eq. 3.4 reads
  $$\sigma_2(b)u_1(B,b) + \sigma_2(c)u_1(B,c) = \sigma_2(b)u_1(C,b) + \sigma_2(c)u_1(C,c) = \\ \sigma_2(b) + 4\sigma_2(c) = \sigma_2(b)0 + 3\sigma_2(c),$$
  to which we add the equation
  $$\sigma_2(b) + \sigma_2(c) = 1.$$
  Overall, we obtain a linear system with two unknowns, $\sigma_2(b)$ and $\sigma_2(c)$.
  However, we notice in our case that the system has no solutions, since $\sigma_2(b)+\sigma_2(c)$ should both equal 1 and 0 at the same time.
  We conclude that the support does not yield a Nash equilibrium.

We have inspected 9 supports, and found only one Nash equilibrium, corresponding to the actions where both players betray each other. Symmetry was heavily exploited.

Example 3.8 (The card game) *We consider the game of Example 2.16. Its strategic form representation, after a quick relabelling of the actions for simplified notations, is given at Table 3.2.*

	m	p
Rr	2, 2	3, 1
Rf	2.5, 1.5	2, 2
Fr	1.5, 2.5	3, 1
Ff	2, 2	2, 2

Figure 3.2: Normal representation of the game in Example 2.1

There are dominated strategies in the game, and we should exploit this in order to ease the search for equilibria.
More precisely, [Ff] is strongly dominated (see Definition 2.17) by a randomized strategy in Δ([Rf],[Rr]) (e.g. 0.5[Rf] + 0.5[Rr]). Therefore, it is never going to be a best response,

and can be removed from the game. Thus, we may reduce the game by removing the strongly dominated strategy[6].
In other words, we can safely analyze the following reduced game presented in Figure 3.3 without loosing any Nash Equilibria.

	m	p
Rr	2, 2	3, 1
Rf	2.5, 1.5	2, 2
Fr	1.5, 2.5	3, 1

Figure 3.3: Removing strongly dominated strategies from the game in Example 2.1

It appears that the reduced game still contains a weakly dominated *strategy, [Fr]. In general, we can not guarantee that weakly dominated strategies cannot be part of an equilibrium.*[7] *However, in the present case, we will see that [Fr] is indeed not part of an equilibrium.*
In order to show this, observe that if Player 2 has a non-zero probability of playing [m], then player 1 *has no interest in playing [Fr] (as he would rather play [Rr]). Thus, let us first consider the* supports *where Player 2 only plays [p].*
The best response from Player 1 to [p] is to play a randomized strategy in Δ*([Rr], [Fr]). However, if Player 1 plays such a strategy, the best response of Player 2 is to play [m], which is outside the support currently under consideration. Thus, there are no equilibria where Player 2 plays [m] with zero probability.*
As a conclusion, we can further reduce the game by eliminating [Fr], leading us to consider the game of Figure 3.4.

	m	p
Rr	2, 2	3, 1
Rf	2.5, 1.5	2, 2

Figure 3.4: Removing all dominated strategies from the game in Example 2.1

We can now proceed to look for the equilibria of the reduced game by checking all supports.

Case 1: *Equilibria with pure strategies (4 supports to check).*

There are no equilibrium in pure strategies for this game. Indeed, the best response to [m] is [Rf], the best response to [Rf] is [p], the best response to [p] is [Rr].

Applying the procedure, step 3 would reject the existence of equilibria.

Case 2: *One of the players randomizes, the other doesn't (4 supports to check).*

Here again, there will be no equilibrium, and the reason is exactly the same as before. For example, if Player 2 played [m], Player 1 would have no interest in randomizing, since the best response to [m] is [Rf].

Applying the procedure, step 2 would reject the existence of equilibria.

[6]The formal proof of this result is left as an exercise.
[7]We will investigate the elimination of weakly dominated strategies as part of an exercise.

Case 3: *Both players randomize.*

This case is more interesting and we may apply both step 2 and step 3 of the procedure. We did not find tricks to avoid doing so.
Remark that we are already certain that this last support carries an equilibrium. This is due to Theorem 3.2: there must be an equilibrium in this game, and we are now inspecting the last support remaining.

First, we need to solve the following linear systems (step 2, for Player 1 and 2 respectively)

$$\begin{aligned} 2\sigma_2(m) + 3\sigma_2(p) &= 2.5\sigma_2(m) + 2\sigma_2(p)(= w_1); \\ \sigma_2(m) + \sigma_2(p) &= 1; \end{aligned} \tag{3.6}$$

$$\begin{aligned} 2\sigma_1(Rr) + 1.5\sigma_1(Rf) &= \sigma_1(Rr) + 2\sigma_1(Rf)(= w_2); \\ \sigma_1(Rr) + \sigma_1(Rf) &= 1; \end{aligned} \tag{3.7}$$

The unique solution to this system is given by $\sigma_2(m) = 2/3$; $\sigma_2(p) = 1/3$; $\sigma_1(Rr) = 1/3$; $\sigma_1(Rf) = 2/3$.

We conclude that the one and only Nash equilibrium of the game is the randomized strategy profile

$$(1/3[Rr] + 2/3[Rf],\ 2/3[m] + 1/3[p]),$$

with the payoffs for each players being $(w_1, w_2) = (7/3, 5/3)$.

3.2 Two-player Zero-Sum games

As illustrated in the previous section, computing Nash equilibria is not straightforward. However, some games have a structure which can be exploited to ease the computation.

Two-player zero-sum games are such games. These games describe situations in which two individuals are in pure opposition to each other, where one's gain is always the other one's loss. More formally, the utility functions u_1 and u_2 of Player 1 and 2 (resp.) are such that $u_1 = -u_2$.
As a consequence, the goal of Player 2, in maximizing his own gain, is equivalent to the minimization of the gains of Player 1.

The Nash equilibria for these games are characterized by the following result:

Theorem 3.9

Given a two-person zero-sum game $\Gamma(\{1,2\}, (C_1, C_2), (u_1, -u_1))$, the strategy profile $\sigma = (\sigma_1, \sigma_2)$ is a Nash equilibrium if and only if

$$\sigma_1 \in \arg\max_{\tau_1 \in \Delta(C_1)} \min_{\tau_2 \in \Delta(C_2)} u_1(\tau_1, \tau_2)$$

and

$$\sigma_2 \in \arg\min_{\tau_2 \in \Delta(C_2)} \max_{\tau_1 \in \Delta(C_1)} u_1(\tau_1, \tau_2).$$

Furthermore, if (σ_1, σ_2) is an equilibrium, then

$$u_1(\sigma_1, \sigma_2) = \max_{\tau_1 \in \Delta(C_1)} \min_{\tau_2 \in \Delta(C_2)} u_1(\tau_1, \tau_2) = \min_{\tau_2 \in \Delta(C_2)} \max_{\tau_1 \in \Delta(C_1)} u_1(\tau_1, \tau_2).$$

There are a few interesting observations to extract from the theorem.

1. There might be multiple equilibria $\sigma^1, \sigma^2, \ldots$. However, their payoffs are all equal, i.e. $\forall i \in N$, $u_i(\sigma^k) = u_i(\sigma^1)$, for $k = 2, \ldots$.

2. When considering games in strategic form, we assume that players play simultaneously. In the present case, this can be relaxed a little. At an equilibrium, Player 1 will act as if Player 2 sees his (mixed) strategy and responds in the worst way for Player 1. The result states that the equilibrium strategy of Player 1 also maximizes his payoff under this pessimistic assumption.

3. Theorem 3.9 allows to solve two-player zero-sum games (i.e. to compute their Nash Equilibria) efficiently, thanks to, e.g., Linear Programming.

Last, let us point out that the result naturally carries on to all games that are *equivalent* (see Section 2.3) to a two-player zero-sum game, as illustrated in the following example.

Example 3.10 (The card game is a two-player zero-sum game.) *Take the game* $\Gamma = (N, (C_i)_{i \in N}, (u_i)_{i \in N})$ *of Example 3.8, Figure 3.4, and consider the game*

$$\Gamma' = (N, (C_i)_{i \in N}, (u_i' \triangleq u_i - 2)_{i \in N}).$$

	m	p
Rr	0/0	1/-1
Rf	0.5/-0.5	0/0

Figure 3.5: Equivalent zero sum game.

The game we obtain is by construction equivalent (Definition 2.21) to the original game.

Now, let x be the probability for Player 1 to play [Rr], and let y be the probability for Player 2 to play [m]. The payoff of Player 1 is given by

$$u_1'(x, y) = x(1 - y) + 0.5(1 - x)y.$$

Following the theorem, we wish to compute the function

$$w_1(x) = \min_y u_1'(x, y),$$

and then compute the value of x at the equilibrium by taking $x = \arg\max w_1(x)$. The derivative of u_1' according to y is $0.5 - 1.5x$. Therefore,

1. *If $0 \leq x < 1/3$, the derivative is positive, so Player 2 should play $y = 0$ (pure strategy [p]) to minimize the payoff of player $u_i(\sigma^k)$. In this case, we get $w_1(x) = x$ for $x \in [0, 1/3[$.*

2. *If $1/3 < x \leq 1$, the derivative is negative, so Player 2 should play $y = 1$ (pure strategy [m]). In this case, we get $w_1(x) = 0.5(1 - x)$ for $x \in]1/3, 1]$.*

3. *If $x = 1/3$, then the derivative is 0. From the expression of u_1', Player 2 should play $y = 1$. Thus, $w_1(x) = 1/3$ when $x = 1/3$.*

Overall, the maximum of $w_1(x)$ is attained at $x = 1/3$ (see Figure 3.6), which turns out to be exactly the value of $\sigma_1([Rr])$ (probability of playing [Rr]) computed for the Nash equilibrium in Example 3.8.

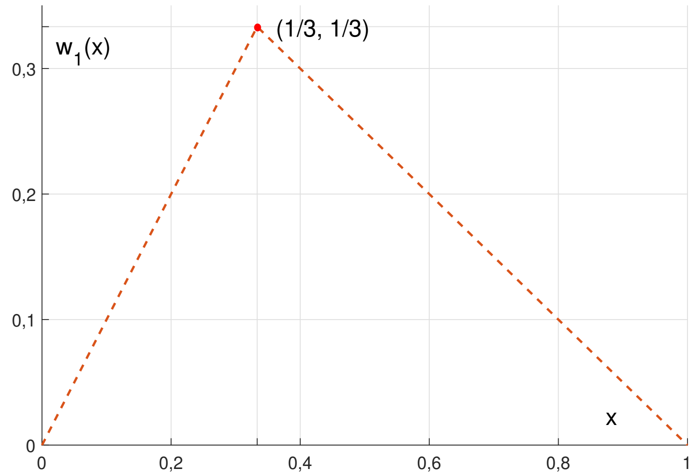


Figure 3.6: The payoff of Player 1 as a function of x.

Contrary to general games in standard form, equilibria of 2P0S games can be computed efficiently. Indeed, remark that Player 1's problem, namely

$$\max_{\sigma_1} \min_{\sigma_2} \sigma_1^T A \sigma_2,$$

can be shown to be simply an LP! (In the above, σ_1, σ_2 are the vectors of the randomized stragegies, and A is the payoff matrix of player 1.) The proof consists in replacing the inner minimization (which is an LP given σ_1 fixed) by a maximization, this can be done using the dual program by example.

Chapter 4

Sequential Equilibria

"Life is a journey, not a destination." — Ralph Waldo Emerson.
Chapter based on pages 154 to 177 of the book "Game theory - Analysis of conflict" by R. Myerson.

The Nash equilibrium is a powerful tool for the analysis of games. When *intelligent and rational* players have to make decisions, they will make sure to maximize their own payoffs (rationality) knowing that the other players will do the same (intelligence and rationality of the others). Thus, every player will play at a Nash Equilibrium.
Nash equilibria are computed from the strategic form of a game. The main idea behind the use of the strategic form is that it provides a concise summary of the game, highlighting which strategies can be played, and what are the payoffs resulting of a choice of strategies. Since the aim of rational players is to maximize their payoffs, we are tempted to believe that the study of the strategic form is sufficient for understanding the outcome of games. In this chapter, we show that this is not the case. A game is a sequential process. A strategic form ignores sequentiality, and hides away some features of the game. The main consequence of this is that, even if when studying the game *a priori* a strategy seems to be a best response to the opponents' strategies, *its implementation may not be rational.* This is due to the fact that rational players wish to maximize their payoff *at every move.* This is essentially due to the fact that our rational players actually take several decisions. And, if we certainly expect that their different decisions have to be coherent with one another, we cannot deduce from this assumption that the decisions about every actions they will take are taken at once, at the beginning of the game. Even more, we can certainly not deduce that their *beliefs* about the state of the world can be decided at once, at the beginning of the game, for every future moment where they will have to make a decision, without taking into account the sequentiality of these decisions.
In this chapter, we begin with examples showing that some Nash equilibria may translate into decisions (in the initial sequential game) which are irrational, and thus should naturally be excluded (Section 4.1). Then, we will define several notions relative to extensive form games allowing us to generalize the notion of rationality to sequential scenarios (Section 4.2). In particular, we are going to be using the concept of *behavioural strategies.* Finally, we will study *sequential rationality* and, the key solution concept of the chapter, the *sequential equilibrium.*

4.1 Rational and irrational equilibrium

Using simple examples, we are going to bring intuition for answering the question *what are the rational moves in a game?*

Our first example represents a "good" scenario, where the Nash equilibrium is rational.

Example 4.1 (The card game: bluff or no bluff?) *Let us go back to the card game of Example 2.1. We computed the Nash equilibrium of the game in Example 3.8, which is*

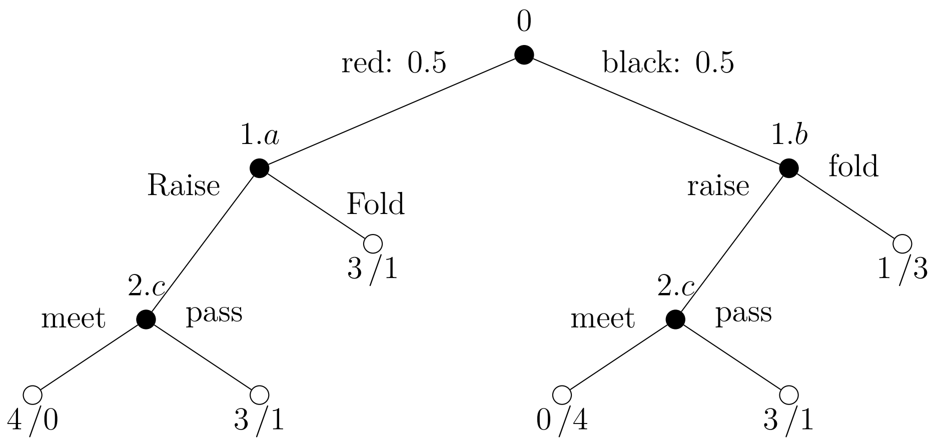


Figure 4.1: Extensive form of the game in Example 2.1.

given by the mixed strategy profile

$$(1/3[Rr] + 2/3[Rf], 2/3[m] + 1/3[p]).$$

The equilibrium corresponds to the following decisions:

1. *At information state* 1.*a (Player 1 has a red card), the move "Raise" is chosen.*
2. *At information state* 1.*b (Player 1 has a black card), the move "raise" is chosen with probability* $\frac{1}{3}$*, and "fold" with probability* $\frac{2}{3}$*.*
3. *At information state* 2.*c, the move "meet" is chosen with probability* $\frac{2}{3}$ *and the move "pass" is chosen with probability* $\frac{1}{3}$*.*

Now, let us put ourselves in the shoes of Player 2. Imagine for the moment that we had no idea *whatsoever of the strategy of player* 1*. An idealistic strategy for us would be to "pass" on a red card, and to "meet" on a black card. This strategy is however unavailable to us - since we have no information on the card's color. This is translated in the game by the fact that information state* 2.*c is shared between two decision nodes.*
However, it is quite easy to see (left as exercise) that if we believed that the card was red with probability α*, and black with probability* $1 - \alpha$*, then our rational move would be to*

- *play "meet" if* $\alpha < \frac{3}{4}$*,*
- *play "pass" if* $\alpha > \frac{3}{4}$*,*

- *and play anyone of "meet" or "pass" if* $\alpha = \frac{3}{4}$.

In our case, the probability $\alpha = \frac{3}{4}$ *is actually enforced by the strategy of player* 1 *at the equilibrium. Indeed, the probability of reaching information state* 2.*c is given by*

$$p(2.c) = \frac{1}{2}p(R|1.a) + \frac{1}{2}p(r|1.b) = \frac{1}{2} + \frac{1}{2}\frac{1}{3} = \frac{2}{3}.$$

Thus, the probability of being at the upper branch of the game and at node 2.*c is*

$$p(2.c|red) = \frac{\frac{1}{2}p(R|1.a)}{p(2.c)} = \frac{3}{4},$$

where the above is nothing else that an application of Bayes *formula.*

In this example, we saw that the move of Player 2 would need to depend on the probability that he believes to be either in the upper branch (red card) of the game or the lower branch (black card). In any scenario, a rational move can be computed. The Nash Equilibrium of the game corresponds to a case where Player 2 believes there are 3 chances over 4 that Player 1 has a red card should he raise.

Example 4.2 (A delicate situation) *Let us now consider the extensive form game presented in Figure 4.2.*

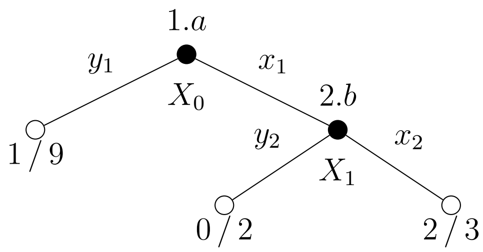


Figure 4.2: This game has two Nash Equilibria: ($[x1], [x2]$) and ($[y1], [y2]$). The second equilibrium assumes an irrational move from Player 2...

The strategic form is given by

	x2	*y2*
x1	*2, 3*	*0, 2*
y1	*1, 9*	*1, 9*

Of the two Nash Equilibria, ($[y_1], [y_2]$) *is disputable. Indeed, let us put ourselves in the shoes of Player 1, and consider the game again through the eye of decision theory. The uncertainty here is one move picked by Player* 2*; so let us say Player 2, at information state* 2.*b, plays* x_2 *with probability* α, *and* y_2 *with probability* $1 - \alpha$. *We need to compare our two moves:*

- y_1 *wins us a payoff of* 1,
- x_1 *wins us a payoff of* 2α.

Thus, we prefer x_1 over y_1 if $\alpha > \frac{1}{2}$. Now, let us try and get more information on the possible values of α. If Player 2 is at information state 2.b then his choice is trivial: playing x_2 wins him a payoff of 3, dominating the decision y_2! Consequently, Player 1, knowing that Player 2 is rational, should infer $\alpha = 1$, and thus prefer x_1 over y_1.
We conclude that as long as both players are rational and intelligent, the equilibrium (y_1, y_2) cannot occur.

Remark that this conclusion would be hidden from us if we only studied the strategic form. Here, the extensive form contains information allowing us to discard an equilibrium in the name of rationality. It is hopeless to attempt to obtain the same conclusion from using the strategic form alone. Indeed, the game of Figure 4.3 has the same strategic form, and here both equilibria seem acceptable. Moreover, by playing [y2], Player 2 now guarantees himself a payoff of 9 instead of 3, which is thus the rational thing to do here.

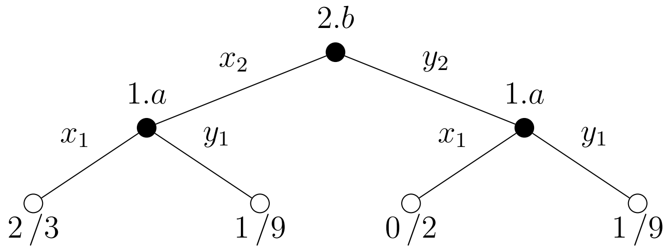


Figure 4.3: A game with the same strategic form as the one of Figure 4.2, for which $([y1], [y2])$ is a Nash Equilibrium which is rational to implement.

Most of the problems will come from imperfect information in sequential games. If we have perfect information, then by the *backward analysis* we can easily describe the sequential equilibria of a game, as stated by the following theorem.

Theorem 4.3: Zermelo's theorem

If Γ^e is an extensive form game with perfect information, then there exists at least one sequential equilibrium of Γ^e in pure strategies (so every move probability is either 0 or 1). Furthermore, for all generic games with perfect information, there is exactly one sequential equilibrium.

Throughout this chapter, we will study games having the *perfect recall* property. At any step of the game, all the players of the game are always able to recall the whole history of their past moves.

4.2 Behavioural strategies

The example of the previous section shows that Nash Equilibria may or may not be rational when analyzed through the extensive form. Indeed, a rational agent is going to play a *rational move at every information state.*

We will now introduce a formalism for studying this behaviour. Our formalism needs to enable us to access several elements. First and foremost, we will need players, information states and decisions. Additionally, for a player at a given information state, there is uncertainty about his exact position in the extensive form tree. Thus, we also need to access the different nodes of the game. To do so, we recall the following relevant objects for a game in extensive form Γ_e from Chapter 2.

Definition 4.4: Game in extensive form

For a game in extensive form Γ_e, we define the following:

- N is the set of players.
- $S = \bigcup_{i \in N} S_i$ is the set of all the information states, where S_i are the information states of $i \in N$, and $S_i \cap S_j = \emptyset$ if $i \neq j$.
- $D = \bigcup_{s \in S} D_s$ is the set of all possible moves, where D_s is the set of moves available at information state $s \in S$.

Regarding the structure of the tree, we have

- Y_s is the set of nodes having information state $s \in S$.
- $X = \bigcup_{s \in S} Y_s$ is the set of nodes in the tree.
- $\Omega \subset X$ is the set of leafs of the tree, and $X^0 \in X$ is its root.

Again, there is a clear distinction between information states (S) and nodes in the tree (X).

Example 4.5 *Consider the game in extensive form of Figure 4.4.*

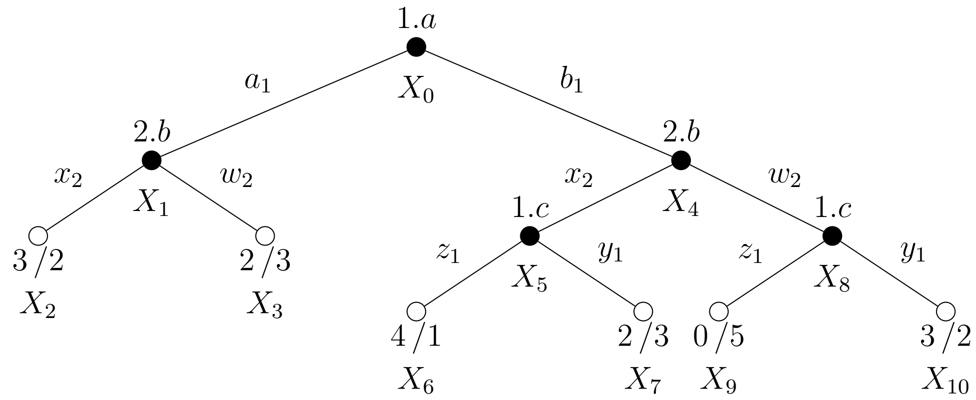


Figure 4.4: Information state are above the nodes, and the nodes' names are below.

We have

1. $N = \{1, 2\}$,

2. $S_1 = \{1.a, 1.c\}$, $S_2 = \{2.b\}$, *and* $S = S_1 \bigcup S_2$.

3. $D_{1.a} = \{a_1, b_1\}$, $D_{2.b} = \{x_2, w_2\}$, $D_{1.c} = \{z_1, y_1\}$.

Regarding the nodes, we have:

4. $X = \{X_0, \ldots, X_{10}\}$,

5. $Y_{1.a} = \{X_0\}$, $Y_{2.b} = \{X_1, X_4\}$, $Y_{1.c} = \{X_5, X_8\}$,

6. $\Omega = \{X_2, X_3, X_6, X_7, X_9, X_{10}\}$ *and* $X^0 = X_0$.

Next, we are going to define notions that relate to the players strategies. Recall that in Chapter 2, a *pure strategy* was defined as a choice of one move per information state. More formally, the set of pure strategies for player $i \in N$ is

$$C_i = \times_{s \in S_i} D_s.$$

For each $c_i \in C_i$ and each information state $s \in S_i$, we can then define $c_i(s)$ has the move played at information state s for the pure strategy c_i. Then, by considering possible randomizations between strategies, we define the set of *mixed strategies* for player $i \in N$ as

$$\Delta C_i = \Delta\left(\times_{s \in S_i} D_s\right), \tag{4.1}$$

and finally, the set of all mixed strategy profiles as $\times_{i \in N} \Delta C_i = \times_{i \in N} \left(\Delta\left(\times_{s \in S_i} D_s\right)\right)$.

Definition 4.6

A *behavioural strategy* for player $i \in N$ specifies, at each information state $s \in S_i$, a probability distribution on the moves D_s of the player. That is, the set of all behavioural strategies for $i \in N$ is

$$\times_{s \in S_i} \Delta D_s. \tag{4.2}$$

For $s \in S_i$, $d \in D_s$ and $\tau_i \in \times_{s \in S_i} \Delta D_s$, $\tau_{i.s}(d)$ is the probability of occurrence of the move d at information state s when implementing τ_i.
We define the set of all *behavioural strategy profiles* as

$$\times_{i \in N} \times_{s \in S_i} \Delta D_s.$$

Mathematically speaking, mixed strategies (4.1) and behavioural strategies (4.2) have a different nature - one is a set of distributions on a cartesian product, the other is a cartesian product on distributions. However, there are some clear links between the two notions.

Example 4.7 *Consider the game of Example 4.5. We take for the example the following mixed strategy for Player 1:*

$$(0.5[a1, y1] + 0.5[b1, z1]).$$

When implementing its strategy, Player 1 is first going to randomize between $a1$ *and* $b1$*. Then, if he did do* $b1$*, he will have to play again at information state* $1.c$*, where he should play* $z1$*. Thus, the mixed strategy actually corresponds to the following behavioural strategy:*

$$(0.5[a1] + 0.5[b1], [z1]),$$

where the first part corresponds to the move at informations state $1.a$*, and the second to the move at information state* $1.b$*.*
From a mixed strategy profile, it is quite easy to compute a corresponding behavioural strategy. However, a same behavioural strategy may correspond to several strategy profiles.

Definition 4.8

Consider an extensive form game Γ^e. For any behavioural strategy profile τ and two nodes $x, y \in X$, we define

1. $P(y|x, \tau)$ as the probability to reach node y from node x by implementing τ.
2. $P(y|\tau) = P(y|x^0, \tau)$ as the probability of reaching node y from the root of the tree by implementing τ.

Example 4.9 *Consider again the game of Example 4.5. Take a general behavioural strategy profile of the form*

$$\tau_1 = (\alpha[a_1] + (1-\alpha)[b_1],\ \beta[y_1] + (1-\beta)[z_1]),\ \tau_2 = (\gamma[x_2] + (1-\gamma)[w_2]).$$

The implementation of this strategy actually infers that each path in the extensive form tree is going to be followed with some probabilities. This is represented in Figure 4.5. Here we have for example:

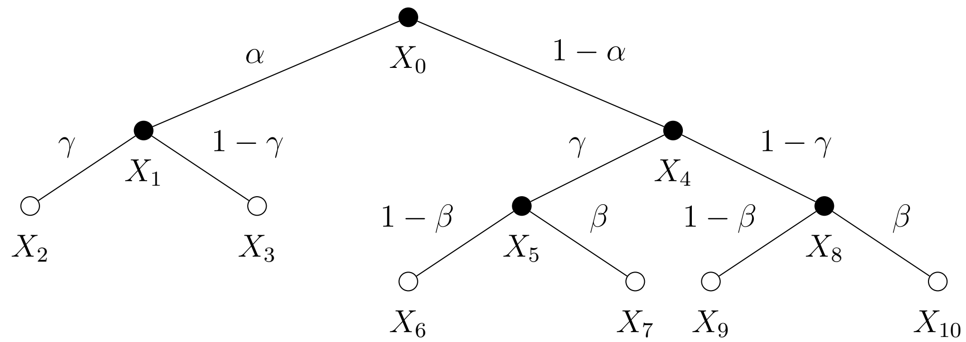


Figure 4.5: Implementing a behavioural strategy.

1. *$P(X_{10}|\tau) = (X_{10}|X_0, \tau) = (1-\alpha)(1-\gamma)(\beta)$, which is simply the probability of occurrence of the path from X_0 to X_{10}.*

2. *$P(X_7|X_4, \tau) = \gamma\beta$. Note that this can be understood as the probability of reaching X_7 conditioned upon the fact that we visited X_4.*

4.2.1 Behavioural and Mixed strategies representation.

In this section, we formalize the link between the behavioural and mixed representation of a same strategy. Intuitively, these two ways of describing the probabilistic behaviour of an agent must be coherent with the laws of probability, and in particular, their relation must satisfy Bayes' formula.

Definition 4.10

Consider an extensive form game Γ^e. For any *mixed* strategy profile σ and two nodes $x, y \in X$, we define

1. $\tilde{P}(y|x, \sigma)$ as the probability to reach node y from node x by implementing σ.
2. $\tilde{P}(y|\sigma) = \tilde{P}(y|x^0, \sigma)$ as the probability of reaching node y from the root of the tree by implementing σ.

Definition 4.11: Compatible pure strategies

For $s \in S_i$, the set of pure strategies compatible with s is defined as

$$C_i^*(s) = \{c_i \in C_i : \exists c_{-i} \in C_{-i} : \sum_{x \in Y_s} \tilde{P}(x|(c_{-i}, c_i)) > 0\}.$$

For $s \in S_i$ and $d \in D_s$, the set of pure strategies compatible with s and using d is

$$C_i^{**}(d, s) = \{c_i \in C_i^*(s) : c_i(s) = d\}.$$

Definition 4.12: Behavioural representation of mixed strategy

For a player $i \in N$, let $\sigma_i \in \Delta C_i$ be a mixed strategy. We say that a behavioural strategy τ_i is a *behavioural representation* of σ_i if,

$$\forall s \in S_i, d \in D_s,\ \tau_{i.s}(d) \cdot \left(\sum_{e_i \in C_i^*(s)} \sigma_i(e_i) \right) = \sum_{c_i \in C_i^{**}(d,s)} \sigma_i(c_i).$$

Conversely, we say that σ is a mixed representation of τ if

$$\forall c_i \in C_i,\, \sigma_i(c_i) = \prod_{s \in S_i} \tau_s(c_i(s)).$$

Definition 4.13: Mixed representation of behavioural strategy

For a player $i \in N$, let $\tau_i \in \times_{s \in S_i} \Delta D_s$ be a behavioural strategy. We say that a mixed strategy σ_i is a *mixed representation* of τ if,

$$\forall c_i \in C_i,\, \sigma_i(c_i) = \prod_{s \in S_i} \tau_{i.s}(c_i(s)).$$

Proposition 4.14

Consider a game in extensive form Γ^e and its strategic form Γ. At a leaf node $x \in \Omega$, let $w_i(x)$ be the payoff for player i in Γ^e. Let τ be a behavioural strategy and σ be a mixed representation of τ. Then

$$\forall i \in N, \sum_{x \in \Omega} P(x|\tau) w_i(x) = u_i(\sigma),$$

that is, the payoff of τ in the extensive form is the same than the one of σ in the normal form.

Theorem 4.15: Kuhn

Any two mixed strategies in $\Delta(C_i)$ that are behaviourally equivalent are also payoff equivalent.

The theorem above has some interesting consequences. If rational and intelligent players only seek to maximize their payoffs, then this justifies the focus on strategic form games as in Chapter 3. However, the Nash Equilibrium alone does not allow us to discard all irrational strategies. We require something more, a notion of equilibrium where players are rational at every information state of the game.

4.3 Sequential Equilibrium

Our goal is now to construct a solution concept allowing to understand the behaviour of rational and intelligent agents engaging in a sequential game. More formally, we would like to understand which are the behavioural strategy profiles that correspond to rational and intelligent decisions from all players in a game.

4.3.1 Multi-agent representation

A first approach would be to adapt the concept of Nash equilibrium so that it applies to behavioural strategies instead of mixed strategies. This can be done almost trivially simply by describing an extensive form game Γ^e through its *multi-agent* representation.

Definition 4.16: Multi-agent representation

Given a game in extensive form Γ^e with players N, information states $S = \bigcup_{i\in N} S_i$ and moves $D = \times_{s\in S} D_s$ the multi-agent representation is a normal-form game $\Gamma(S, (D_s)_{s\in S}, (v_s)_{s\in S})$ where

- The set of players of the multi-agent representation is the set of information states S,
- The set of pure strategies for type $s \in S$ is the set of moves D_s,
- For any player $i \in N$, for any strategy profile $d = (d_s)_{s\in S}$, for any $s \in S_i$, the payoff for type s is given by

$$v_s(d) = u_i(c(d)),$$

where $c(d)$ is a mixed representation of d and $u_i(c(d))$ is the payoff of player i for this mixed-strategy.

Example 4.17 *The multi-agent representation of the card game of Example 4.1 is given by*

	1.a: Raise		*1.a: Fold*	
	1.b: raise	*1.b: fold*	*1.b: raise*	*1.b: fold*
2.c: meet	*2, 2, 2*	*2.5, 2.5, 1.5*	*1.5, 1.5, 2.5*	*2, 2, 2*
2.c: pass	*3, 3, 1*	*2, 2, 2*	*3, 3, 1*	*1, 1, 3*

.

Theorem 4.18

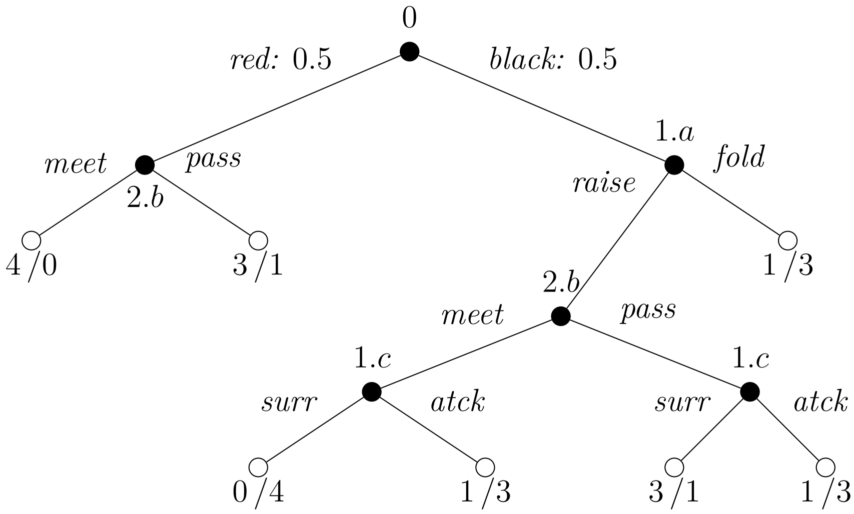
Given an extensive form game Γ^e and a Nash Equilibrium σ of its strategic representation, any behavioural representation of σ is a Nash Equilibrium of its multi-agent representation.

The above theorem somehow represents a negative result. It means that our attempt to represent sequentiality by 'splitting' an agent into a set of virtual agents, has failed: all the Nash equilibria of the strategic representation turn out to be Nash equilibria of the multi-agent representation. Thus, we were not able to rule out the "bad" equilibria, which do not take into account sequentiality.
The converse, as shown in the next example, is not true.

Example 4.19 *Consider the following variation of the card game of Example 4.1.*

- *Both players bet 1 EUR.*
- *Player 1 picks a red or a black card. If it is a red card, he raises the bet and puts an additional 1 EUR on the table. If it is a black card, either he raises by 1 EUR, or folds. If he folds, Player 2 wins the 1 EUR of Player 1, and the game ends.*
- *If Player 1 raises, then Player 2 privately decides whether he meets the bet or passes by putting 1 EUR in a closed envelope, or not.*
- *Then, Player 1 shows his card. If the card is red, then he wins the envelope. If it is black, then he has one additional move to make. Either he surrenders, at which point he takes back 1 EUR and leaves the rest to Player 2. Or he attacks, and if Player 2 met the raise, then Player 2 wins everything and if Player 2 passed on the bet, Player 1 wins everything.*

After a few simplifications, we can model this game with the following extensive form.



The multi-agent representation is given by

	1.a: raise		1.a: fold	
	1.c: surr	*1.c: atck*	*1.c: surr*	*1.c: atck*
2.b: meet	*2, 2, 2*	*2.5, 2.5, 1.5*	*2.5, 2.5, 1.5*	*2.5, 2.5, 1.5*
2.b: pass	*3, 3, 1*	*2, 2, 2*	*2, 2, 2*	*2, 2, 2*

.

where the payoffs are written as $(v_{1.a}/v_{1.c}/v_{2.b})$. *Observe that the behavioural strategy profile (fold, pass, atck) is a Nash Equilibrium for the multi-agent representation. The equilibrium brings a payoff of 2 to Player 1. However, if Player 1 chooses the behavioural strategy (raise, surr) instead of (fold, atck), then he would obtain a payoff of 3!. As a conclusion, the mixed representation of (fold, pass, atck) is not a Nash Equilibrium for the strategic form of the game, even though it is an equilibrium of the multi-agent representation.*

4.3.2 Sequential rationality

We now set out to develop the notion of sequential rationality, for characterizing the behavioural strategies that may be played by rational players. We want to capture the fact that a player will make a sequence of moves, and that each move in the sequence is a rational move. Roughly speaking, this is going to be translated into the need for each move in the sequence to be a move that maximizes the expected utility of the player.

Definition 4.20: Expected Utility at node x

Consider a game in extensive form Γ^e. Given a player $i \in N$, a node $x \in X$ and a behavioural strategy profile τ, we let

$$U_i(\tau, x) = \sum_{y \in \Omega} P(y|\tau, x) w_i(y),$$

where $w_i(y)$ is the payoff of i at the terminal node $y \in \Omega$. $U_i(\tau, x)$ denotes the expected utility of i at x following τ.

Now of course, at any given node $x \in X$ if a player can make a decision, this player wants to maximize his utility. However, players only have access to their information states, they don't necessarily know at which node they are. As it is the case in Example 4.1, the behaviour of a player will then depend of a probability distribution he assigns to nodes sharing an information state. We call these *belief vectors*.

Definition 4.21: Belief vectors

Consider a game in extensive form Γ^e with players N, types $S = (S_i)_{i \in N}$ and nodes X. Let $Y_s \subset X$ be the set of nodes of the tree that have information state s. A belief vector π is any element of the set

$$\times_{s \in S} \Delta Y_s.$$

For $\pi \in \times_{s \in S} \Delta Y_s$, $i \in N$ $s \in S_i$, and $x \in Y_s$, $\pi_s(x)$ is the probability assigned by player i to be at node x given the fact that i knows his information state s.

Because players know only their information state when playing a game, instead of maximizing their expected utility at every node (which they cannot do), the next best thing to do is to use their belief vector and maximize their *sequential values.*

Definition 4.22: Sequential value

Consider a game in extensive form Γ^e, a belief vector π for this game and a behavioural strategy profile τ. For a player $i \in N$ at an information state $s \in S_i$, the *sequential* *value* of the move $d \in D_s$ is given by

$$U_i(d|s, (\tau, \pi)) = \sum_{x \in Y_s} \pi_s(x) U_i((\tau_{-i}, d), x).$$

We can finally define the concept of sequential rationality:

Definition 4.23: Sequential Rationality

Consider a game in extensive form Γ^e, a belief vector π for this game and a behavioural strategy profile $\tau = (\tau_i)_{i \in N}$. The strategy τ_i is sequentially rational for player $i \in N$ at information state $s \in S_i$ according to π if

$$\tau_i(s) \in \arg\max_{d \in D_s} U_i(d|s, (\tau, \pi)).$$

We say that τ is sequentially rational if it is rational at every information state according to π.

Clearly, sequential rationality needs to be part of any reasonable definition of an equilibrium for extensive form games.

Example 4.24 *Consider the game of Example 4.2. Let a be the node at information state $1.a$, and b be the node at information state $2.b$. Clearly, we have $\pi_{1.a}(a) = 1$ and $\pi_{1.b}(b) = 1$, since there are one node per information state. These beliefs vectors are actually independent of the strategy played. Consider the profile $\tau = (y1, y2)$. We already focused on this profile in Example 4.2, as it is a Nash Equilibrium, but it is not rational. We can indeed verify that*

$$\arg\max_{d \in \{x_2, y_2\}} U_2(d|2.b, (\tau, \pi)) = \{x_2\}.$$

The example above shows an interesting feature of belief vectors: even if we consider the strategy profile (y_1, y_2), for which information state $2.b$ is not encountered, $\pi_{1.b}$ is still well defined. In the next section, we investigate what constitutes valid *belief vectors*.

4.3.3 Weak consistency: Rationality at states that are believed to occur

How does a rational player compute his believe vector? It is clear that belief vectors and behavioural strategy profiles need to be related. A fairly natural approach would be to use Bayes formula and to compute belief vectors from strategy profiles. More precisely, one's first idea is to compute the belief vector through the formula:

$$\pi_s(y) = \frac{P(y|\tau, x^0)}{\sum_{x \in Y_s} P(x|\tau, x^0)}.$$

However, the above can fail, simply because the denominator could be equal to zero. The above fails, because some elements of the belief vector may not be well defined. This is the case in Example 4.24, where $\pi_{2.b}(b) = 1$ but $\sum_{y \in Y_{2.b}} P(y|\tau) = 0$.

Bayes formula cannot be used when $\sum_{x \in Y_s} P(x|\tau, x^0) = 0$, corresponding to cases where a given information state cannot be reached (see Example 4.24).

Definition 4.25: Weak Consistency

A belief vector π is *weakly consistent* with a behavioural strategy profile τ if it satisfies

$$\forall s \in S, \forall y \in Y_s, \left(\sum_{x \in Y_s} P(x|\tau, x^0) \right) \pi_s(y) = P(y|\tau, x^0).$$

Theorem 4.26

Consider a game in extensive form Γ^e. Let τ be an equilibrium in behavioural strategy [a] and π be a belief vector *weakly consistent* with τ. For all player $i \in N$, let

$$S_i^0 = \{s \in S_i : \sum_{x \in Y_s} P(x|\tau) > 0\}$$

be the set of all information states occurring with positive probability. The equilibrium strategy profile τ is *sequentially rational* at all information states in $\bigcup_{i \in N} S_i^0$ according to the belief vector π.

[a] The behavioural strategy profile τ is an equilibrium in behavioural strategies iff it is an equilibrium of the multi-agent representation such that the mixed representation of τ is also an equilibrium of the normal representation.

However, the scope of the result remains limited. When applied to Example 4.24 for the strategy $([y_1], [y_2])$, one can verify through the theorem that the equilibrium is rational for Player 1. But it does not tell us anything for Player 2.
In order to tackle information states that occur with probability 0, we need a stronger notion of consistency for belief vectors.

4.3.4 Sequential equilibrium: Rationality at all information states

As explained above, all of the difficulties in computing beliefs lies in dealing with information states that arise with zero probability.
But how can we characterize the sequential rationality at nodes of probability zero? Bayes' formula is not useful, because the zero denominator prevents us to compute an actual value. However, this does not imply that any arbitrary value is admissible. The following definition characterizes the admissible belief vectors, not only at positive probability nodes, but also at zero-probability nodes.

Definition 4.27

A belief vector π is *strongly consistent* with a behavioural strategy profile τ in the game Γ^e if there exists a sequence of profiles $(\tau^k)_{k=1}^{\infty}$ such that

- $\tau^k \in \times_{s \in S} \Delta^0 D_s, \forall k = 1, 2, \ldots$

(non-zero probabilities on every move),

- $\tau(d) = \lim_{k\to\infty} \tau^k(d), \forall i \in N, \forall s \in S_i, \forall d \in D_s$
(sequence converges to the strategy),

- $\pi_s(x) = \lim_{k\to\infty} \frac{P(x|\tau^k,x^0)}{\sum_{y\in Y_s} P(y|\tau^k,x^0)}, \forall s \in S, \forall x \in Y_s$,
(belief vector is the limit of the sequence of beliefs given by Bayes rule).

The sequences appearing in the definition can be understood as the result of an iterative process , where players go through a succession of strategies that do not discard any moves, for eventually converging to a given behavioural strategy profile. At every step of the process, one can compute his belief vector simply by applying Bayes formula. A belief vector strongly consistent with a strategy profile is then obtained as the limit of the intermediate profiles.

The following is easily obtained from the definition of strong consistency.

Proposition 4.28

If a belief vector π is strongly consistent with a strategy τ, it is also weakly consistent.

We are finally able to define the concept of *sequential equilibria*:

Definition 4.29

A *sequential equilibrium* of an extensive form game Γ^e is a pair (τ, π) where $\tau \in \times_{s\in S}\Delta D_s$ is sequentially rational according to π, and $\pi \in \times_{s\in S}\Delta Y_s$ is a belief vector strongly consistent with τ.

We can show that a sequential equilibrium is always an equilibrium in behavioural strategy.

Theorem 4.30

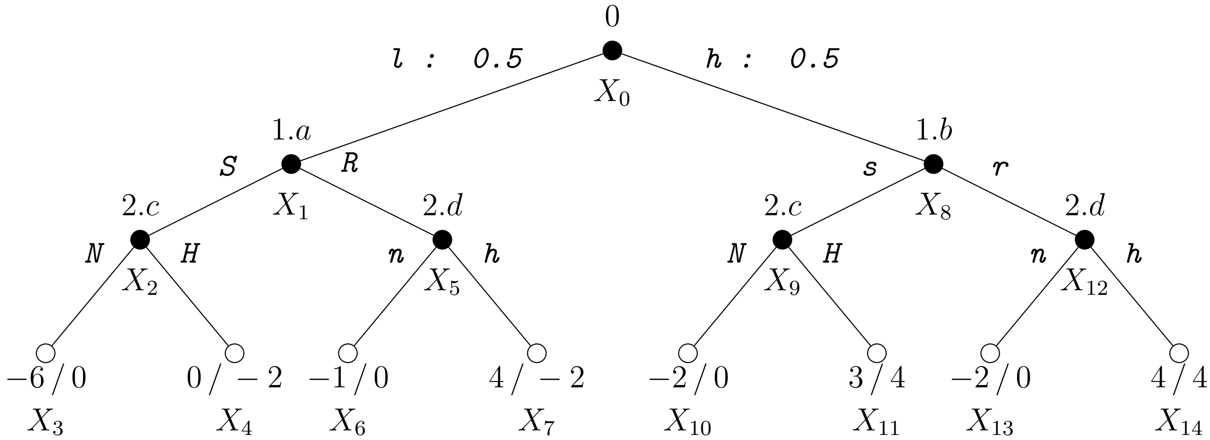
If (τ, π) is a sequential equilibria, then τ is an equilibrium in behavioural strategy.

Theorem 4.31

For any finite extensive form game, the set of sequential equilibria is non-empty.

Example 4.32 (The signalling game) *We consider a two-player game. Player 1 is a student seeking to be hired for a job. He has some private information - he is either highly motivated by the job and will be very productive (type h for* high profile candidate*) or maybe he just wants to get the money and do the least amount of job possible (type l*

for low profile candidate*).*
Based on the result of an exam, Player 2 needs to decide whether to hire Player 1 (H) or not (N).
Regarding this test, Player 1 can decide to study for it (S) or decide to relax instead (R). The game is represented by the following extensive form.



The strategic form is given by:

	Nn	*Nh*	*Hn*	*Hh*
Ss	*-4, 0*	*-4, 0*	*1.5, 1*	*1.5, 1*
Sr	*-4, 0*	*-1, 2*	*-1, -1*	*2, 1*
Rs	*-1.5, 0*	*1, -1*	*1, 2*	*3.5, 1*
Rr	*-1.5, 0*	*4, 1*	*1.5, 0*	*4, 1*

A preliminary analysis of the game[1] *reveals that Nn is dominated: the employer is trying to hire someone. So is Sr.*

After reduction, we have the following game.

	Nh	*Hn*	*Hh*
Ss	*-4, 0*	*1.5, 1*	*1.5, 1*
Rs	*1, -1*	*1, 2*	*3.5, 1*
Rr	*4, 1*	*1.5, 0*	*4, 1*

The strategies Nh and Ss are only weakly dominated here, and thus they are conserved. From this game, we will highlight candidate sequential equilibria, and check whether or not they are indeed equilibria.
The first step is to figure our what are the Nash equilibria of the game. There are only three of them, and they are pure strategy profiles: $([Ss],[Hn])$, $([Rr],[Hh])$ and $([Rr],[Nh])$. Then, we translate these equilibria in behavioural strategies, and look for strongly consistent belief vectors. Finally, we verify whether or not these behavioural representations are sequentially rational for these belief vectors. When this is the case, we obtain sequential equilibria.

1. *Pure equilibrium $([Ss],[Hn])$.*

[1] Which may have been written by a relaxed student...

The behavioural representation of this equilibrium is $\tau_1 = ([S],[s])$, $\tau_2 = ([H],[n])$. *The belief vectors for information states* 1.*a*, 1.*b*, *and* 2.*c* *are uniquely defined:*

$$\pi_{1.a}(X_1) = \pi_{1.b}(X_8) = 1,\ \pi_{2.c}(X_2) = \pi_{2.c}(X_9) = \frac{1}{2}.$$

For this behavioural strategy, the information state 2.*d* *cannot be reached:*

$$P(X_5|\tau) + P(X_{12}|\tau) = 0.$$

By setting for example $\pi_{2.d}(X_5) = \pi_{2.d}(X_{12}) = \frac{1}{2}$, *the obtained belief vector is* weakly consistent. *Thus, we know from Theorem 4.26 that the strategy is sequentially rational at information states* $1.a, 1.b$ *and* 2.*c*.
Yet the action at information state 2.*d* *is very important. Indeed, if it is not rational for Player 2 to play the move* n, *then we fall back to a situation alike the one of Example 4.2, where the Nash Equilibrium is not rational.*
Thus, assume player 1 *plays* $\tau_1^{\alpha,\beta} = (1-\alpha[S]+\alpha[R], 1-\beta[s]+\beta[r])$, *and let* $\tau = (\tau_1^{\alpha,\beta}, \tau_2)^2$. *For all* $\alpha, \beta \in [0,1]$, *we can compute*

$$\pi_{2.d}(X_5) = \frac{P(X_5|\tau)}{P(X_5|\tau) + P(X_{12}|\tau)} = \frac{\alpha}{\alpha+\beta}.$$

For Hn *to be sequentially rational at information state* 2.*d*, *there must be* α, β *small such that the* sequential value *of the move* n *should be greater or equal to that of* h. *Here, we have*

$$U_2(n|2.d, (\tau,\pi)) = \frac{0\alpha + 0\beta}{\alpha+\beta} = 0,$$

whereas

$$U_2(h|2.d, (\tau,\pi)) = \frac{-2\alpha + 4\beta}{\alpha+\beta}.$$

We wish to investigate the case where it would be rational for player 2 *to play* n *over* h *at information state* 2.*d*. *Thus, we require* $-2\alpha + 4\beta \le 0$, *which is the case when* $\beta \le 0.5\alpha$. *For this to be true, we may set* $\beta = \alpha/2$, *in which case we obtain*

$$\pi_{2.d}(X_5) = \frac{2}{3},\ \pi_{2.d}(X_{12}) = \frac{1}{3}.$$

So far, we have seen that the behavioural representation of our considered equilibrium was sequentially rational *when following the belief vector defined above. In order to verify the definition of* sequential equilibrium, *we just need to show that the belief vector is strongly consistent.*

To do so, we set

$$\tau_1^k = (1-\alpha_k[S] + \alpha_k[R], 1-\beta_k[s] + \beta_k[r]), \tau_2^k = (\epsilon_k[N] + 1 - \epsilon_k[H], \epsilon_k[h] + 1 - \epsilon_k[n]),$$

and we wish to find expressions for $\alpha_k, \beta_k, \epsilon_k$ *such that, first,* $(\tau_1^k, \tau_2^k) \to (\tau_1, \tau_2)$ *as* $k \to \infty$, *and second,* π *is the limit of* π^k, *obtained by applying Bayes formula with* τ_1^k, τ_2^k.
In our case, it suffice to choose

$$\alpha_k = \frac{1}{k},\ \beta_k = \frac{0.5}{k},\ \epsilon_k = \frac{1}{k}.$$

[2] Here, $\tau_2 \notin \times_{s\in S_2} \Delta^0 D_s$. This, however, does not impact the beliefs of the players.

2. The equilibrium $([Rr],[Hh])$.

This situation is similar to the one above, and can be solved using the same approach. The problem lies in defining the beliefs at state 2.c*, which is the node reached when a candidate (unexpectedly) studies instead of relaxing. Intuitively, our strategy would then be rational under the condition that our beliefs would be "the candidate is then likely to have a high profile".*

Thus, assume player 1 *plays* $\tau_1^{\alpha,\beta} = (\alpha[S] + 1 - \alpha[R], \beta[s] + 1 - \beta[r])$*, and let* $\tau = (\tau_1^{\alpha,\beta}, \tau_2)$*. Note that as* $\alpha, \beta \to 0$*,* τ *converges to a behavioural representation of* $([Rr],[Hh])$*. For all* $\alpha, \beta \in [0,1]$*, we can compute*

$$\pi_{2.c}(X_2) = \frac{P(X_2|\tau)}{P(X_2|\tau) + P(X_9|\tau)} = \frac{\alpha}{\alpha+\beta}.$$

We may then compare the sequential values of the moves H *versus* N*, which should satisfy the inequality*

$$U_2(N|2.c,\pi) = \frac{0(\alpha+\beta)}{\alpha+\beta} \leq \frac{(-2\alpha+4\beta)}{\alpha+\beta} = U_2(H|2.c,\pi),$$

implying $\beta \geq \alpha/2$*. We are now able to show that the behavioural strategy* $([R]+[r],[H]+[h])$*, coupled with the strongly consistent belief vector generated e.g. by the sequence of strategies*

$$\tau_1^k = (\alpha_k[S] + 1 - \alpha_k[R], \beta_k[s] + 1 - \beta_k[r]), \tau_2^k = (\epsilon_k[N] + 1 - \epsilon_k[H], \epsilon_k[h] + 1 - \epsilon_k[n]),$$

where

$$\alpha_k = \frac{1}{k},\ \beta_k = \frac{0.5}{k},\ \epsilon_k = \frac{1}{k},$$

forms a sequential equilibrium.

3. The equilibrium $([Rr],[Nh])$.

We can show that the above corresponds to a sequential equilibria, with a belief vector that reflects that for Player 2, in the unlikely scenario that the candidate did study, it is more likely that the candidate has a low profile.

Chapter 5

Refinements of Equilibria

"Have no fear for perfection, you will never reach it." — Salvador Dali.
Chapter based on pages 213 to 232 of the book "Game theory - Analysis of conflict" by R. Myerson.

So far, our goal has been to understand the behaviour of rational and intelligent agents in a finite game, where the players take actions, without negotiation, with the goal of maximizing their payoffs. Decision theory (Chapter 1) and the Utility Maximization Theorem brought valuable insights to that end, that lead to the development of the Nash Equilibrium (Chapter 3). This chapter is the last one to be situated in this context.
At the end, Nash Equilibria have drawbacks in that they don't alone, capture the rational behaviour of players. But they are computed from the strategic form, a compact representation of a game. Sequential Equilibria (Chapter 4) capture the sequential rationality of players (a rational player wants to maximize its payoff at every move, not all Nash equilibria allow for this) are computed from the extensive form of a game.
In this chapter, we investigate other definitions of equilibria. Namely, we will describe the Sub-game perfect equilibrium, the Perfect Equilibrium. Ending this chapter, we situate these definitions relative to one another, providing a unifying picture.

5.1 The Sub-game perfect equilibrium

When playing rock-paper-scissors, the one and only Nash equilibrium of the game is to play each move with probability 1/3. If we played the game 10 times in a row, what should we do on the 10th game? The answer is easy: regardless of the outcomes of the previous 9 games, the 10th game being nothing else than a rock-paper-scissors game, we should again play each move with probability 1/3.
In this section, we tackle situations where a player, at some point in the game, has to make decisions knowing that the payoffs he will get from these decisions is actually independent of the past decisions made in the game (which comprise both his and his opponent's actions). To do so, we define the concept of a *subgame* of a game in strategic form. Note that we here make use of the formalism of Chapter 4 describing extensive form trees (Definition 4.4). An example is shown afterwards.

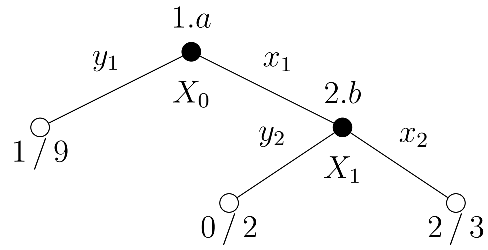
Definition 5.1: Subroots and subgames

Consider a game in extensive form Γ^e. For any node $x \in X$, let $F(x) \subseteq X$ be the set of nodes in the subtree rooted at x.
A *subroot* of the game is a node $x \in X$ such that, for any information state $s \in S$, we have

$$\text{either } Y_s \cap F(x) = \emptyset, \text{ or } Y_s \cap F(x) = Y_s.$$

Given a subroot $x \in X$ of Γ^e, a *subgame* rooted at x is the game obtained by removing from Γ^e all nodes not in $F(x)$.

Example 5.2 *Consider the following game, from Example 4.2:*



Both the nodes X_0 and X_1 are subroot of the game. The big idea is that at both of these nodes, the payoffs depend only on future actions.

This allows us to define the concept of subgame perfect equilibrium.

Definition 5.3

A *subgame perfect equilibrium* of an extensive form game Γ^e is any Nash equilibrium of Γ^e such that the restriction of this equilibrium to any subgame is a Nash equilibrium of that subgame.

Theorem 5.4

The set of subgame perfect equilibria is non-empty.

At a subgame perfect equilibrium, a player will play rationally at every subgame. If a subgame perfect equilibrium is a Nash equilibrium, all Nash equilibria are not subgame perfect.

Example 5.5 *The strategic form of the game of Example 5.2 is*

	x_2	y_2
x_1	*2, 3*	*0, 2*
y_1	*1, 9*	*1, 9*

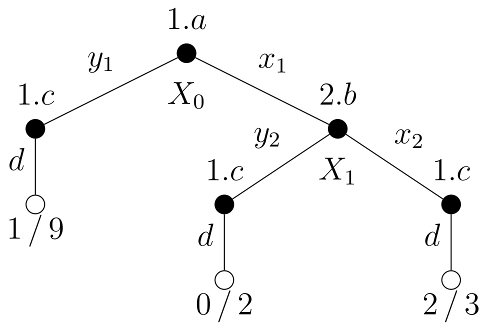
There are two Nash equilibria to that game, which are $([x_1],[x_2])$ *and* $([y_1],[y_2])$*. However, only the first one is subgame perfect. Indeed, if we consider the subgame rooted at node* X_1*, whose strategic form is*

X_1 *subgame*	x_2	y_2
·	*2, 3*	*0, 2*

we see that the payoffs depend only on the move of Player 2. Here, x_2 *dominates* y_2*. Thus, the restriction of the equilibrium* $([x_1],[x_2])$ *to the subgame rooted at* X_1 *is an equilibrium at that subgame.*

One of the main flaw in the definition of subgame-perfect equilibria is that they are highly sensitive to the modelling of the game, as shown next.

Example 5.6 *Consider this variant of the game of Example 5.2.*



The game now has only 1 subroot (which is X_0*), thus both the equilibria* $([x_1,d],[x_2])$ *and* $([y_1,d],[y_2])$ *are subgame perfect.*

This shows that simply by adding one more step to the game (with no effects at all on the outcomes), subgame-perfect equilibria may no longer help us detect irrational Nash Equilibria.

Even if they are sensitive to the modelization, subgame-perfect equilibria remain particularly relevant when addressing *Repeated Games* (see Chapter 8).

5.2 The Perfect equilibrium

Perfect equilibria, as we will explain, are very robust by definition. They are also known as *Trembling Hand perfect equilibria.*

Definition 5.7

A perfect equilibrium for a game in strategic form $\Gamma(N,C,u)$ is a strategy profile $\sigma \in \times_{i\in N}\Delta(C_i)$ such that there is a sequence of strategy profiles $(\sigma^k)_{k=1}^{\infty}$ satisfying

- $\sigma^k \in \times_{i\in N}\Delta^0(C_i)$ (non-zero probability of playing any strategy),

- $\lim_{k\to\infty} \sigma^k = \sigma$ (sequence converges to equilibrium),
- $\sigma_i \in \arg\max_{\tau_i \in \Delta(C_i)} u_i(\sigma^k_{-i}, \tau_i), \forall i \in N, k = 1, 2, \cdots$ (equilibrium is best response to all strategies in the sequence).

The concept shows some resemblance with the one of sequential equilibrium (Definitions 4.27 and 4.29). A sequential equilibrium requires a fully consistent belief vector. The two concepts embed within their definition the idea that the equilibrium strategy is obtained as the limit of strategies which put non-zero probabilities on every move. The difference here lies in the fact that the perfect equilibrium strategy must *additionally* be a best response to all the strategies in the sequence.

Example 5.8 *Consider the following game:*

	x_2	y_2
x_1	*2, 2*	*-1, 2*
y_1	*2, -1*	*0, 0*

For both players, the strategy x is weakly dominated by the strategy y. Nevertheless, amongst the two Nash equilibria of the game $([x_1],[x_2])$ and $([y_1],[y_2])$, the first uses weakly dominated strategies.
Let us now verify if these equilibria are perfect. *Take any randomized strategy of Player 2, say $\sigma_2 = (\alpha[x_2] + 1 - \alpha[y_2])$. If $\alpha \in]0,1[$, $\sigma_2 \in \Delta^0(C_2)$. For Player 1, the best response to σ_2 is always y_1, which grants him a payoff of 2α (compared to $3\alpha - 1 < 2\alpha$ for $\alpha \leq 1$). Since the game is symmetric and that best response of Player 1 to any strategy in $\Delta^0(C_2)$ is y_1, we conclude that $([y_1],[y_2])$ is a perfect equilibrium, but not $([x_1],[x_2])$.*

The previous example illustrates one of the drawbacks of perfect equilibria, in that players may no longer consider equilibria in weakly dominated strategies. Nevertheless, in many games, doing so makes a lot of sense!
Perfect equilibria, however, do come with a good news.

Theorem 5.9

For any *extensive form* game Γ^e, if the behavioural strategy profile σ is a perfect equilibrium for the multi-agent representation of Γ^e, then there exists a strongly consistent belief vector π such that (σ, π) is a sequential equilibrium.

Moreover, the existence of a perfect equilibria is guaranteed

Theorem 5.10

Given a game in strategic form, the set of all perfect equilibria is non-empty.

Example 5.11 *We go back to the signalling game of Example 4.32. We saw that the game had 3 Nash equilibria, which all match sequential equilibria, but two of them were composed of weakly dominated strategies.*
Consequently, amongst the 3, only one can correspond to a perfect equilibrium. Interestingly, it is the one that corresponds to always choosing the relax option for Player 1, and always choosing the hire option for Player 2...

5.2.1 Refinement - proper equilibrium

In the above, we learnt that a perfect equilibria of a multi-agent representation corresponds to a sequential equilibria. The result is actually quite surprising: we can find a *sequentially rational* set of moves without even looking at the extensive form of a game! We still need to consider the multi-agent form though.

The concept of perfect equilibrium can be further refined.

Definition 5.12

A *proper* equilibrium for a game in strategic form $\Gamma(N, C, u)$ is a strategy profile $\sigma \in \times_{i\in N}\Delta(C_i)$ such that there is a sequence of strategy profiles $(\sigma^k)_{k=1}^{\infty}$, and a sequence $(\epsilon_k)_{k=1}^{\infty}$ satisfying

- $\sigma^k \in \times_{i\in N}\Delta^0(C_i)$ (non-zero probability of playing any strategy),
- $\lim_{k\to\infty} \sigma^k = \sigma$ (sequence converges to equilibrium),
- $\sigma_i \in \arg\max_{\tau_i\in\Delta(C_i)} u_i(\sigma^k_{-i}, \tau_i), \forall i \in N, k = 1, 2, \cdots$ (equilibrium is best response to any strategy in the sequence),

and in addition,

- $\forall k, \forall i \in N, \forall c_i, e_i \in C_i : u_i(\sigma^k_{-i}, [c_i]) < u_i(\sigma^k_{-i}, [e_i]) \Rightarrow \sigma^k_i(c_i) \le \epsilon_k\sigma^k_i(e_i)$ (penalize bad moves),
- $\lim_{k\to\infty} \epsilon_k = 0$ (penalty vanishes at equilibrium).

The fun thing with proper equilibria is that we can now detect sequential equilibria of a game *from its strategic form!*

Theorem 5.13

Consider a game Γ^e in extensive form with a normal representation $\Gamma(N, C, u)$. If σ is a proper equilibrium of Γ, then there is a behavioural representation τ of σ and a belief vector π such that (τ, π) is a sequential equilibrium of Γ^e.

This has actually deep repercussions. One of the problems with the strategic form is that it may represent several games in extensive form. Consequently: an equilibrium for a game Γ can be perfect if and only if there is an extensive form game, whose strategic form is Γ, where a behavioural representation of the equilibrium is sequential.

5.3 Equilibria in non-cooperative games: unifying picture

Given a game in extensive form (its more general representation), we are interested in figuring out the moves that could be played by rational and intelligent player. Based on

the axioms of decision theory, and the normal form representation, the concept of Nash equilibrium was presented first.
Any game always has at least one *Nash Equilibrium*, and rational players need to play at Nash equilibria. However, the reverse is not true: not all Nash Equilibrium correspond to strategies rational players may play. The problem with Nash Equilibria is that players make several sequential decisions in a game, and seek to maximize their payoffs at every stage. In this context, *Subgame perfect* equilibria appear naturally. They correspond to Nash Equilibria whose restriction to any subgame is also a Nash Equilibria. Any game always has at least one Subgame perfect equilibrium! One step further is the Sequential Equilibrium. Here, players not only make rational moves at the beginning of the game (Nash) or at the beginning of each subgame (Subgame perfect), but at *every information state*. Again, any game always has at least one Sequential Equilibria. These equilibria are described in terms of behavioural strategies, and their computation relies on the extensive form of a game. Sequential equilibria are perhaps the solution concepts that capture the notion of rationality the best. At a sequential equilibria, player motivate their actions through their beliefs vectors: from their point of view, their actions are rational at every information state. Finally, we discussed perfect equilibria and proper equilibria. They offer cautious definitions for equilibria, where a player should play a strategy which is a best response to a sequence of randomized strategies, that put positive probabilities on all actions. They have some interesting theoretical properties: both are guaranteed to exists, a perfect equilibrium of the multi-agent representation corresponds to a sequential equilibrium, and so does a proper equilibrium of strategic form.

The relation between the above mentioned types of equilibria is shown on Figure 5.1.

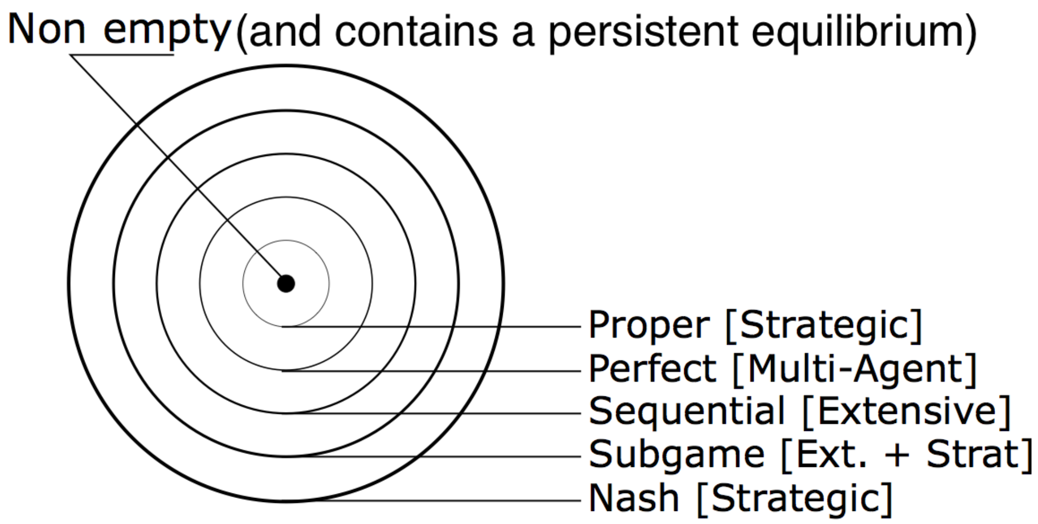


Figure 5.1: Inclusion relations between different types of equilibria.

Chapter 6

Games with communication

"Two monologues do not make a dialogue" — Jeff Daly.
Chapter based on (Myerson, 1991, pages 244 to 263).

Communication is a central part of our everyday life. In fact, in many strategic interactions it is common that players are able to communicate with each other. The goal of this chapter is to explore the possibilities arising from such communications.

6.1 Correlated strategies and mediators

We consider that rational and intelligent players will communicate in order to be able to attain a more profitable outcome out of their interaction. We are particularly interested by the following question: if we are given a game in strategic form, what are the payoffs achievable when enabling communication between the players?
To answer this question, we first need to define what "communication" means. A very broad (and rich) definition suggests that players can talk, send messages to each other, lie about their plans or their private information, etc. These can be refered to as *speech acts*.
This direct definition leads to a very natural view of communication in game theory: we may define, for each player, a set of available *speech acts*, and then we would form a game where a strategy consists of picking both speech acts and moves. Traditional notions of equilibria can then apply. You can however imagine that the above task may be daunting (as we need first to understand what constitute a speech act, and how it may affect the game...).

The good news is that we will be able to analyze games with communication in a general way by relying on a simple concept called a *mediator*.

Definition 6.1: Mediator (informal)

A mediator is a non-player agent whose role is to gather the information and preferences of each player in a game and provide recommendation about which action to take to each player.

The importance of the mediator is revealed through the *revelation principle* (Myerson, 1991, page 257).

Definition 6.2: The revelation principle (informal)

The set of equilibria of a game with communication coincides with the sets of recommandations, made by a mediator, that the players are willing to follow.

The next example serves to illustrate how communication can be important in strategic situations as well as the concept of mediator.

Example 6.3 (Crossroads (a.k.a. "chicken")) *Two drivers are arriving at high speed at a crossroad (Figure 6.1). The drivers can either wait (w) and stop at the crossroad, or*

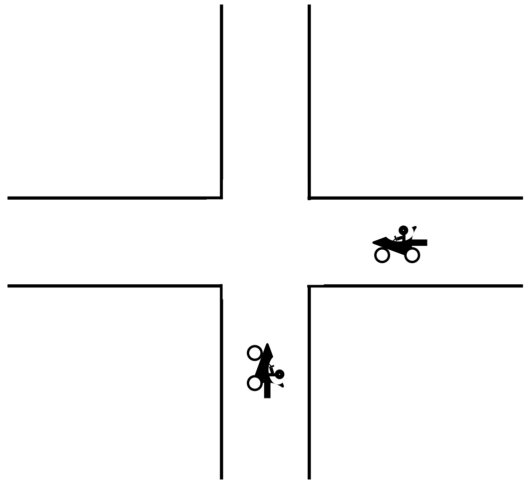


Figure 6.1: Who goes first?

go on (g). We will represent this as a 2 player game, with the following strategic form

	g	*w*
G	*-10, -10*	*2, 0*
W	*0, 2*	*-1, -1*

.

There are four outcomes for this game: players can have an accident (which matches the choice of strategies ([G],[g])), or Player 1 waits ([W],[g]), or Player 2 waits ([G],[w]), or both wait ([W],[w]).
The game has 3 Nash Equlibria: $([G],[w])$, $([W],[g])$, *and* $(\frac{3}{13}[G]+\frac{10}{13}[W], \frac{3}{13}[g]+\frac{10}{13}[w])$. *We know that, since the players are rational, they are* going to play at a Nash equilibrium. *What does this mean? It means that for this games, Player 1 may pick [G], [W] or 0.5[G] + 0.5[W] depending on what he* believes *Player 2 is going to do. The same occurs for Player 2. However, it may very well be the case that Player 1 thinks Player 2 is going to stop, and the reverse holds for Player 2. As a consequence, the players may crash rationally even if neither one played the randomized equilibrium.*

Of course, in practice, we know several ways to solve the crossroad problem. In Belgium, the law *states that we should always let the one arriving from the right go first. If we break the law, we face dire consequences (first, we may have an accident, and second, if we get caught, we can be heavily reprimanded).*

Another way to solve this problem is by installing a traffic light *at the crossroad. At every instant, the red light actually suggests to the players an outcome for the game ([G],[w]) or ([W],[g]). Most of the time, players follow this recommendation since if we receive a signal not to go, we know the other one is signalled to go. Thus, if we do not follow recommendations, we actually chose to have an accident.*

Yet another way to solve this problem would be through communication. We could imagine a situation where the players are able to send signals to each others, and decide (perhaps by tossing a coin) on a suitable outcome for the game. A reasonable strategy for the players would be to pick the outcome ([G],[w]) or ([W],[g]), each with probability $1/2$.

In the following, we will be considering games in strategic form $\Gamma = (N, C, u)$. Recall that in this case, the *outcomes of the games* are associated with pure strategies (hence, we can refer to C as the set of outcomes). Our players will try to find an agreement about which outcome should occur. As presented in the example, they may agree to randomize on these outcomes.

Definition 6.4: Correlated strategy

Given a game in strategic form $\Gamma = (N, C, u)$, the set of *correlated strategies* is

$$\Delta(C) = \Delta(\times_{i \in N} C_i).$$

As always, our players' decision will be driven by their utility functions.

Definition 6.5: Expected payoff for correlated strategy

Given a game $\Gamma = (N, C, (u_i)_{i \in N})$ the payoff of player $i \in N$ for a correlated strategy $\mu \in \Delta(C)$

$$U_i(\mu) = \sum_{c \in C} \mu(c) \cdot u_i(c).$$

Before moving further, let us give more explanations about *why* do we call these μ *correlated strategies*? Any given a randomized strategy profile $\sigma \in \times_{i \in N} \Delta(C_i)$ corresponds to a unique correlated strategy μ, where

$$\forall c = (c_i)_{i \in N} \in C, \mu(c) = \prod_{i \in N} \sigma_i(c_i).$$

However, the reverse does not hold: when choosing their strategies *independently* of one another, there are some correlated strategy that cannot be reached by the players (this is shown in the next example). Thus the name: in correlated equilibria, the probability distribution of the actions of the different players are allowed to be correlated with each other, which enables much more possibilities than in classical standard games.

Example 6.6 *Assume, in the crossroads example, that the player want to implement the* correlated strategy $0.5([G], [w]) + 0.5([W], [g])$. *To do so, Player 1 can for example toss a coin, and communicate the result to Player 2 (assuming Player 1 is truthful). Alternatively, they can make use of a red-light (if present).*

Something they can not do, however, is to play according to a randomized strategy - say $\alpha_1[G] + (1-\alpha_1)[W]$ for the first player, and $\alpha_2[g] + (1-\alpha_2)[w]$ for the second. Indeed, this requires

$$\mu([G],[g]) = \alpha_1\alpha_2 = 0; \mu([G],[w]) = \alpha_1(1-\alpha_2) = 0.5, \mu([W],[g]) = (1-\alpha_1)(\alpha_2) = 0.5,$$

which is impossible.

We can already conclude that considering correlated strategies opens new horizons to the players, because this allows for outcomes that are impossible to reach when players chose their moves independently of each other.

Naturally, one may ask the question "how are the players able to implement a correlated strategy"? This is where the mediator comes in to play. It has a role similar to that of the traffic light in Example 6.3. The role of a mediator is to:

1. Describe a correlated strategy to the players.
2. Privately sample an outcome of the game following the correlated strategy.
3. Secretely inform each player of their own strategies (but not of that of the others).

6.2 Correlated equilibria

We now investigate the notion of *correlated equilibria*. To give a intuitive meaning to the concept and its shadow zones, we consider that the ultimate choice made by a player is up to that player. As seen in Figure 6.2, this implies that we need to study an augmented game where, on top of their own moves, player have the possibility to follow or not the recommendations of the mediator.

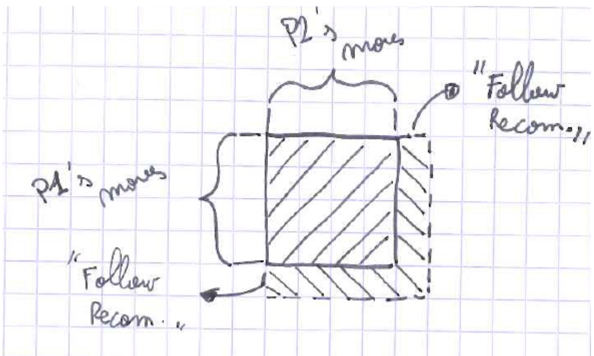


Figure 6.2: Players must chose rationally if they follow the recommendations of the mediator.

A correlated equilibrium occurs when it is a Nash equilibrium for players to follow the recommendations. It is therefore natural to study what happens when one wishes to not follow recommendations. In this context, we first study the concept of binding contract.

6.2.1 Binding contracts

In this context, we assume that the players have the options to sign a binding contract with the mediator. The recommendations of the mediator will vary according to which players signed the contract. The goal of the mediator is to make the contract attractive to everybody.

Definition 6.7: Contract

Given a game $\Gamma = (N, C, u)$, a contract τ is a function

$$\tau : S \subseteq N \to \Delta \times_{i \in S} (C_i),$$

that assigns to each subset of players S the correlated strategy they will play if S signs the contract and $N \backslash S$ does not.

In the definition above, $\tau(N)$ is a correlated strategy as defined in the previous section. For any other $S \subset N$, $\tau(S)$ is also a correlated strategy, but which is going to be implemented only by the members of the set S that sign the contract. The players who do not sign the contract are free to choose an action in their action set.

The following follows directly from the definition of Nash Equilibrium.

Proposition 6.8: Correlated equilibrium with binding contract

Consider a game $\Gamma = (N, C, u)$, and a contract τ for this game. The correlated strategy $\tau(N)$ is a correlated equilibrium if and only if

$$\max_{c_i \in C_i} \sum_{c_{-i} \in C_{-i}} \tau(N \backslash i, c_{-i}) u_i(c_i, c_{-i}) \leq \sum_{c \in C} \tau(N, c) u_i(c),$$

where for $S \subseteq N$, $c \in \times_{i \in S} C_i$, $\tau(S, c)$ is the probability that the players in S having signed the contract play the strategy c.

The above simply states that one won't sign a contract if doing so allows him to obtain an higher payoff.

Assume now that, as a mediator, you wish to propose a contract where $\tau(N)$ is a correlated equilibrium. A player won't sign the contract unless he has some incentive to do so. One way to do so is by using a threat, by choosing a $\tau(N \backslash i)$ that would guarantee that i would prefer to sign. In order to do this, we may use its *minimax value*.

Definition 6.9: Minimax

The *minimax value* for player i is defined as

$$v_i = \min_{\tau_{-i} \in \Delta(C_{-i})} \max_{c_i \in C_i} \sum_{c_{-i} \in C_{-i}} \tau_{-i}(c_{-i}) \cdot u_i(c_{-i}, c_i).$$

Example 6.10 *Alice and Bob plan to meet in the evening. There are two options available: either to go to the cinema, or to the ballet. Alice prefers to go to the ballet, and Bob to the cinema. They find it hard to decide and decide to ask Charles to mediate their dispute by providing them with a binding contract and a correlated equilibrium.*

Their payoffs and moves are summarized in the table below (Alice is Player 1, Bob is Player 2).

	c	*b*
C	*2, 3*	*-1, -1*
B	*1, 1*	*3, 2*

.

We begin by computing the minimax value v_1. Assume that Bob plays c with probability β, and b with probability $1-\beta$. The payoff of Alice are given by

$$u_1(C, \beta[c] + (1-\beta)[b]) = 3\beta - 1, \qquad u_1(B, \beta[c] + (1-\beta)[b]) = -2\beta + 3.$$

The minimax strategy for Bob is to pick $\beta = 4/5$, and the minimax value for Alice is $v_1 = 7/5$.

For v_2, we apply the same technique: if Alice plays C with probability α, we have

$$u_2(\alpha[C] + (1-\alpha)[B], c) = 2\alpha + 1, \qquad u_2(\alpha[C] + (1-\alpha)[B], b) = -3\alpha + 2.$$

The minimax strategy for Alice is to pick $\alpha = 1/5$, and the minimax value for Bob is $v_2 = 7/5$.

In conclusion, Alice and Bob will sign any contract proposed by Charles as long as their expected payoff is at least $7/5$. In particular, Charles could tell them to both go to the cinema, or to both go to the ballet. We discuss in Chapter 7 tools allowing Charles to select an appropriate correlated strategies among those agreeable to Alice and Bob (see Example 7.8).

The next result allow us the characterize all the correlated strategies for which there is a contract making them a correlated equilibria.

Theorem 6.11

Consider a correlated strategy μ. There exists a contract τ with $\tau(N) = \mu$ in which all players signing is an equilibrium if and only if $U_i(\mu) \geq v_i$ for all $i \in N$, where v_i is the minimax value for player i.

6.2.2 Correlated equilibria for games in strategic form

Let us assume now that the players do not have to sign any contract, so they are free to follow the plan of the mediator or to deviate from it during the game. To ensure that the latter case does not occur, a correlated equilibrium for such a game in *strategic form* must satisfy the following strategic incentive constraints:

Definition 6.12: Strategic incentive constraints

A correlated strategy μ satisfies to the *strategic incentive constraints* if

$$U_i(\mu) \geq \sum_{c \in C} \mu(c)\, u_i(c_{-i},\ \delta(c_i)), \quad \forall i \in N,\ \forall \delta : C_i \to C_i,$$

where $c = (c_{-i}, c_i)$, c_{-i} being the strategy of the players other than i, and where $\delta(c_i)$ is a cheating function that replaces any strategy $c_i \in C_i$ suggested by the mediator by a strategy $\delta(c_i) \in C_i$.
A correlated strategy satisfying the strategic incentive constraints is said to be *incentive compatible.*

These constraints require that the players cannot improve their payoff by deviating from the correlated strategy μ suggested by the mediator. Note that cheating would be perfectly rational if it was the case. Indeed, when a mediator dictates to a player what he should do to follow the correlated strategy, the player may infer a probability distribution on the other players moves. More precisely, if a mediator implementing μ tells i to play c_i, then the other players are going to play according to

$$\sigma_{-i}(c_{-i}) = \frac{\mu(c_i, c_{-i})}{\sum_{e_{-i} \in C_{-i}} \mu(c_i, e_{-i})}.$$

Given this knowledge, a rational player would naturally compute his best response to $\sigma_{-i}(c_{-i})$ and, if this brings him a better payoff than the one obtained by following the recommendation, he will adopt the best response.

Example 6.13 *For the crossroads example, the correlated strategy* $\mu = 0.5([G], [w]) + 0.5([W], [g])$ *is incentive compatible. Indeed, if a player is told to slow down, this player can infer that the other player has been told to keep going. Thus, if the first player decided to cheat and kept moving, an accident would surely happen.*
On the other hand, the correlated strategy $\mu = ([W], [w])$ *cannot be incentive compatible. Indeed, since the other player is being told to slow down, the best response is just to keep moving.*

Note that, because of the cheating function, the formulation of Definition 6.12 describes a polytope with an exponential number of inequalities. To avoid this problem, we can equivalently rewrite the conditions with a polynomial number of inequalities as follows:

$$\sum_{c_{-i} \in C_{-i}} \mu(c) \Big(u_i(c_{-i},\ c_i) - u_i(c_{-i},\ e_i) \Big) \geq 0, \quad \forall i \in N,\ \forall c_i \in C_i,\ \forall e_i \in C_i.$$

Finally, note that we may have several possible incentive compatible correlated strategies. The mediator still needs to pick one. One way to do this is by maximizing some objective function. For example, one may try to maximize $U_1(\mu) + U_2(\mu)$ (total amount of utility). The good news is that such objective functions are *linear* in the variables $\mu(c)$, $c \in C$, and thus for instance, *the problem of finding an incentive compatible mechanism that maximizes total payoff is a linear program.*

6.3 Correlated equilibria for Bayesian games

People are able to cheat when it is in their interest, but what about lying? In the context of Bayesian games (see Section 2.4.1), players now have access to private information able to shape the game. In order to make sound recommendations, the mediator needs access to this information. But maybe giving a false statement may steer the recommendations in your favor...

We now consider *Bayesian games with communication*. Recall that a Bayesian game is defined as $\Gamma = (N, C, T, p, u)$, where the $T = (T_i)_{i \in N}$ are the *types* of each players, $p = (p_i)_{i \in N}$ where $p_i : T_i \to \Delta(T_{-i})$ and the payoffs are defined as $u = (u_i)_{i \in N}$ with

$$u_i : T \times C \to \mathbb{R}.$$

That is, the payoffs of the players now depend on their types!

The mediator now needs to prepare recommendations for any combination of player times.

Definition 6.14: (Bayesian) correlated strategy

Given a bayesian game $\Gamma = (N, C, T, p, u)$, a Bayesian correlated strategy is a function of the form

$$\mu : T \to \Delta(C).$$

For a correlated strategy μ, the payoff of player i given his type t_i, is given by:

$$U_i(\mu \mid t_i) = \sum_{t_{-i} \in T_{-i}} \sum_{c \in C} p_i(t_{-i} \mid t_i)\, \mu(c \mid t)\, u_i(c,\ t),$$

where $t = (t_{-i}, t_i)$ and T_{-i} are the set of possible combinations of types for the players other than i.

The task of a mediator can now be described as

1. Propose to the players a set of correlated strategies $\mu(\cdot|t)$, one per combination of types $t \in T$.
2. Collect the types of the players.
3. Privately compute the outcome of the correlated strategy.
4. Secretly communicate his recommendations to each players.

The problem is that, to guarantee a desired outcome, the mediator needs not only propose an incentive compatible strategy for each combination of types, be he must also be sure that he knows the correct types for the players for doing the recommendations.

Example 6.15 *Back at our crossroads! Consider the following situation. You are an engineer working on a* "smart crossroad" *project. The idea is the following: some drivers (type V for VIP) are more important than others (type R for Regular). For example, those driving an ambulance should have priority on those who are just doing their groceries.*

The cars of your city have been equipped with rudimentary devices, that broadcast whether they are VIP drivers or not. We would like to allow the VIP to pass whenever they are in the presence of regular drivers.
Assume the following bayesian model for the game:

R vs R	g	w	R vs V	g	w
G	-10 / -10	2 / 0	G	-30 / -10	2 / -1
W	0 / 2	-1 / -1	W	1 / 5	-1 / -5
V vs R	g	w	V vs V	g	w
G	-10 / -30	5 / 1	G	-10 / -10	2 / 0
W	-1 / 2	-5 / -1	W	0 / 2	-1 / -1

Regarding the beliefs functions p, we may assume $p(V|R) = p(V|V) = p(V) = 0.1$, $p(R|R) = p(R|V) = p(R) = 0.9$ *as objective probabilities.*

The problem is that most resident of the city are talented engineers, and they can hack the broadcast device with ease. Could you build a policy, using red-lights, that allows VIP drivers to have priority over regular ones, but that would also discourage all drivers to hack their devices?

A naive approach would be to consider each of the four cases separately, and prepare one incentive compatible mechanism for each which, for example, would maximize total payoff. In our case, this yields

- $\mu(\cdot|RR) = 0.5([G],[w]) + 0.5([W],[g])$,
- $\mu(\cdot|RV) = ([W],[g])$,
- $\mu(\cdot|VR) = ([G],[w])$,
- $\mu(\cdot|VV) = 0.5([G],[w]) + 0.5([W],[g])$,

Assume now that, as a regular driver, you were aware that the above was the strategy implemented at the crossroads. Would you hack your broadcast device to appear as a VIP?

Assume you broadcast as a regular driver. Your payoff is then going to be

$$U_1(\mu|R) = p(R)u_1(\mu(\cdot|RR)|RR) + p(V)u_1(\mu(\cdot|RV)|RV) = 0.9 \cdot 1 + 0.1 \cdot 1 = 1.$$

Indeed, with probability $p(R)$, *the other player is of type R, so the recommendations are then given by* $\mu(\cdot|RR)$. *With probability* $p(V)$, *the other player is of type* V, *so we receive recommendations according to* $\mu(\cdot|RV)$. *If, instead we decided to broadcast your type as VIP, your payoff would become*

$$p(R)u_1(\mu(\cdot|VR)|RR) + p(V)u_1(\mu(\cdot|VV)|RV) = 0.9 \cdot 2 + 0.1 \cdot 1.5 = 2.85.$$

It is important to understand that we would receive the payoffs of a regular player, but the recommendations of a VIP driver. Thus, we get priority over regular drivers, and equal chances to go or wait versus a VIP driver.

In conclusion, yes, it would be rational to hack the device. The design of the proposed strategy did not take this into account.

Definition 6.16: Incentive Compatibility

A bayesian correlated strategy $\mu : T \to \Delta C$ is incentive compatible if it satisfies to the constraints

$$U_i(\mu \mid t_i) \geq U_i^*(\mu,\ \delta_i,\ s_i \mid t_i), \quad \forall i \in N,\ \forall t_i \in T_i,\ \forall s_i \in T_i,\ \forall \delta_i : C_i \to C_i,$$

with

$$U_i^*(\mu,\ \delta_i,\ s_i \mid t_i) = \sum_{t_{-i} \in T_{-i}} \sum_{c \in C} p_i(t_{-i} \mid t_i)\, \mu(c \mid t_{-i}, s_i)\, u_i\Big(\ (c_{-i},\ \delta_i(c_i))\ ,\ t\Big),$$

where t_i is the type of i, s_i is the type that i reveals to the mediator and δ_i is the function chosen by player i that, for any action c_i that can be suggested by the mediator, chooses an action $\delta_i(c_i)$ to play instead.

These constraints impose that the players neither have interest to lie on their type, nor to deviate from the correlated strategy μ suggested by the mediator.

Example 6.17 *We now conclude our crossroad example by presenting a mechanism satisfying to the incentive constraints for the bayesian game.*

One may verify that the following strategy is incentive compatible

- $\mu(\cdot|RR) = 0.5([G], [w]) + 0.5([W], [g])$,
- $\mu(\cdot|RV) = ([W], [g])$,
- $\mu(\cdot|VR) = ([G], [w])$,
- $\mu(\cdot|VV) = 1/3([G], [g]) + 1/3([G], [w]) + 1/3([W], [g])$.

The approach is quite dangerous: in order to deter a regular player from lying, we tell him that VIPs tend to crash into VIPs. Since crashing into a VIP has a very high cost for a regular driver, he then has no interest into lying.

6.3.1 Collective choice problems

In this section, we mention a special type of Bayesian games with communications, known as *Collective choices problems.*
These problems refer to situations where we seek to make a decision based on the type of players in a population. A typical example of this are *voting problems*, where each citizen's type corresponds to preferences in candidates, and the goal is to chose which candidate to elect (see e.g. (Shoham and Leyton-Brown, 2008, Chapter 9)). In this case, we which to establish rules under which it is in each citizen's interest to report his preferences truthfully.

With this in mind, we assume that a mediator is going to gather the types of the players and decide on the outcome of a game. In order for the player to be incentivized to divulge their true types, the mediator must pick an incentive compatible strategy in the following sense:

Definition 6.18: (Collective choice) incentive compatibility

For a collective choice problem, a Bayesian correlated strategy $\mu : T \to \Delta C$ is incentive compatible if it satisfies to the constraints

$$U_i(\mu \mid t_i) \geq U_i^*(\mu,\ s_i \mid t_i), \quad \forall i \in N,\ \forall t_i \in T_i,\ \forall s_i \in T_i,$$

with

$$U_i^*(\mu,\ s_i \mid t_i) = \sum_{t_{-i} \in T_{-i}} \sum_{c \in C} p_i(t_{-i} \mid t_i)\, \mu(c \mid t_{-i}, s_i)\, u_i\Big(\ (c_{-i},\ c_i)\ ,\ t\Big),$$

where t_i is the type of i, s_i is the type that i reveals to the mediator.

6.4 Moral hazards and adverse selection

The problem faced by mediators having to give incentive for player to be truthful about their types and to follow recommendations is ubiquitous in the insurance industry. The following definition is extracted from (Myerson, 1991, p. 263).

Definition 6.19: Moral hazard and adverse selection

Moral hazard is the need to give players an incentive to implement recommended actions. *Adverse selection* is the need to give players an incentive to report information honestly.

Overall, the contents of this chapter can be split into four categories:

	With moral hazard	Without moral hazard
With adverse selection	Section 6.3	Subsection 6.3.1
Without adverse selection	Subsection 6.2.2	Subsection 6.2.1

Chapter 7

Bargaining and Coalitions

"The whole is greater than the sum of its parts" — Aristotle.

Chapter based on (Myerson, 1991, pages 370 - 390 and 417 - 444) and (Shoham and Leyton-Brown, 2008, Chapter 12).

The question of *how can rational players cooperate?* is particularly intriguing. There are several reasons for this. First, of course, negotiations play a central role in our everyday interactions, should they be at the personal or corporate level. The second reason is deeper: our theoretical framework assumes that agents seek solely to improve their own payoff, but here our goal is to find the "fairest outcome" of a game, which is somehow contradictory with the fact that agents do not care about fairness in game theory, but only in optimizing their payoff. It is not surprising that we will have to postulate new axioms in order to formalize what we mean by "fairness". Negotiation is often tied to communication (arguably, communication is necessary for negotiation). In Chapter 6 we give a first glimpse at games where agents may communicate. We focused on highlighting those strategies that agents *may accept to play* in this context.In the current chapter, we discuss the process by which agents chose, among all the mechanisms they may accept to play, the one that seems the fairest for all.

The chapter is divided into two parts:

- 2-player games: The Nash Bargaining Problem. The two player case allows us to explore several key concepts, such as the notions of utilitarian and egalitarian solutions. We introduce the solution concept of the Nash-Bargaining solution, relying on an axiomatic approach.
- N-player games: Coallitions. When more than two players communicate, we face additional problems in trying to capture the phenomena behind negotiation. In particular, rational players may wish to make deals with a subset of the players, and this behavior may prevent them to reach agreements. We study the setting, and provide solution concepts for this case.

7.1 The two-players bargaining problem

We define (informally) a bargaining problem as a situation where a set of players must decide on a single outcome for a game such that "everyone benefits from the outcome".

We will shortly provide a mathematical formalism for its definition, and begin with Example 7.1.

Example 7.1 *Alice and Bob win a 100€ prize to be shared between them. They now discuss the matter of how they should split the money. The difficulty is that they each have very different plans about how to use their money, and need to take them into account.*

- *Bob says he* really *needs the money because he has a debt to pay to a friend. He says he would get 1.5* utils[1] *for each euro he gets for the first 40 euros, and 0.5* util *for each euro after this. Hence, Bob's utility is given by the formula*

$$u_B(x) = \begin{cases} 1.5x & \textit{if } x \leq 40, \\ 60 + 0.5(x - 40) & \textit{if } x > 40. \end{cases}$$

- *Alice declares that she would get 3 utils for each euro under 10 euro. After this, she will obtain one* util *for each euro, until 80 euro. She does not wish for more than 80 euros, and gains no more utility for additional money past this. Hence, Alice's utility is given by the formula*

$$u_A(x) = \begin{cases} 3x & \textit{if } x \leq 10, \\ 30 + (x - 10) & \textit{if } 10 < x \leq 80, \\ 100 & \textit{if } 80 < x. \end{cases}$$

- *If they can't agree on splitting the money, they decide to give away the prize to a charity. This gives a payoff of* 60 *to Alice, and* 25 *to Bob.*

Alice and Bob sit down to reflect on the situation. The first question they ask is what are all the possible way they can share the money? *They represent this at Figure 7.1. There, the scales show* utility *instead of the actual sum of money allocated.*
On the same figure, in red is shown the set of all allocations that make them at least as happy as giving the money to charity. Clearly, this set is of great importance: Alice and Bob being rational and intelligent, they would never pick an allocation outside the set. The point f appears particularly interesting: at that point, Alice and Bob would each get the same amount of utility above the point v, which is 24.

Definition 7.2: Bargaining problem

A two-player bargaining problem is a pair

$$(F, v)$$

where F is a closed convex set of $\mathbb{R}^2$ which represents the set of possible payoffs and $v = (v_1, v_2) \in \mathbb{R}^2$ is the *disagreement point* and represents the payoffs that both players would receive in the event of failure of the negotiation.

[1] Utility is not necessarily equal to the amount of money received. Hence, we introduce a new unit, *util*, to denote the amount of utility one gets!

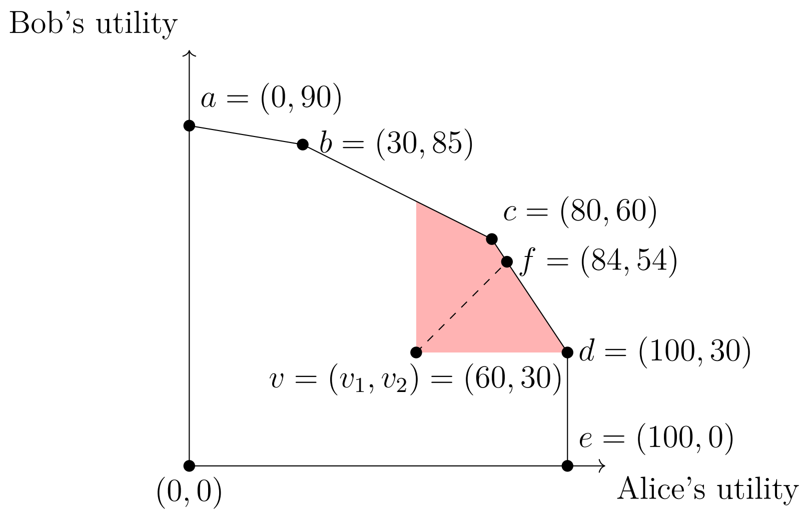


Figure 7.1: Possible utility allocations for Alice and Bob. If they decide to give the money away to charity, they obtain the allocation v. Clearly, they should not chose any allocations outside of the red set.

To better interpret these structures, note that Bargaining problems arise naturally in the context of Chapter 6. Consider players strategic form game $\Gamma = (\{1, 2\}, C, u)$. For this game, we can have

$$F = \Big\{ \Big(u_1(\mu),\, u_2(\mu) \Big) \mid \mu \in \Delta(C) \Big\}, \quad \text{where } u_i(\mu) \;=\; \sum_{c \in C} \mu(c) u_i(c),$$

where $\mu \in \Delta(C)$ is a correlated strategy satisfying, depending on the context, to the conditions of Theorem 6.11 (payoff greater than minimax for binding contracts), or strategic incentive constraints (see Definitions 6.12 or 6.16). Observe that in all these cases, the set of possible payoff allocations is convex (why?).
The disagreement point can be defined in several ways depending on the problem studied. We discuss three typical definitions in Section 7.1.3.
Convexity of F can be justified by the possibility to *randomize* among allocations by the players.

Definition 7.3: Essential

The Bargaining Problem (F, v) is *essential* if there exists at least one allocation y in F such that $y > v$.[a]

[a] For two vectors $x, y \in \mathbb{R}^n$, $x \geq y$ if $x_i \geq y_i$ for all entries, and $x > y$ if the inequality is strict everywhere.

Definition 7.4: Efficiency, Egalitarian, Utilitarian

Consider a Bargaining Problem (F, v) and a point $x \in F$. We say that x is

- *strongly (Pareto) efficient* if $\nexists y \in F : y \geq x$ and $y_i > x_i$ for at least one i;
- *weakly (Pareto) efficient* if $\nexists y \in F : y > x$;
- *rational* if $x \geq v$;
- *egalitarian* if weakly efficient and $\forall i, j : x_i - v_i = x_j - v_j$;
- *utilitarian* if $x \in \arg\max \{\sum_i x_i : x \in F\}$.

Example 7.5 (Example 7.1 Continued) *Figure 7.1 represents a* Bargaining problem (F, v). *F is the polytope whose vertices are the origin and the points a, b, c, d and e. The disagreement point v is represented in F. The following hold:*

- (F, v) *is essential.*
- *All points in the red area are rational.*
- *All points on the segments* $[a, b]$, $[b, c]$ *and* $[c, d]$ *are* strongly Pareto efficient.
- *Point f is* egalitarian, *at coordinates* $(84, 54)$. *Observe that since we have* $v = (60, 30)$, *we get* $x_1 - v_1 = x_2 - v_2 = 24$ *at this point.*
- *Point c is* utilitarian, *with* $x_1 + x_2 = 140$. *Observe that, on the figure, no other points achieve the same sum of payoffs.*

Given a bargaining problem, we wish to find a point that will satisfy all players. This is clearly a hard task, as we can easily oppose different views regarding what is a "good solution". First of all, players may not share a same utility scale, rendering the comparison of payoffs hazardous. Second, some may prone the usage of an *egalitarian solution*, appearing *fair to everyone*, while others may view this as a waste, preferring a *utilitarian solution*.

In the next section, we will see how Nash was able to propose a meaningful solution to this question. We will also see in Section 7.1.2 how Nash's solution relates to the egalitarian and utilitarian solutions.

7.1.1 Nash's Bargaining solution

Like Von Neumann and Morgenstern before him, John Nash uses an axiomatic approach to derive his solution to the two-player bargaining problem. Thereby, he tries to objectify how rational players should negotiate while hopefully ensuring the existence and unicity of a solution.

Notation 7.6: Solution of a Bargaining Problem

Given a Bargaining problem (F, v), we denote by

$$\phi(F, v)$$

the *solution* of the problem.

The axioms in question are the following.

Axiom 7.1 (Efficiency.)

1. *Strong: $\phi(F, v)$ is an allocation in F and, for all x in F, if $x \geq \phi(F, v)$, then $x = \phi(F, v)$.*

2. *Weak: $\phi(F, v) \in F$ and there does not exist any y in F such that $y > \phi(F, v)$.*

Axiom 7.2 (Individual rationality.) *$\phi(F, v) \in F$ is rational, i.e. $\phi(F, v) \geq v$.*

Axiom 7.3 (Scale covariance.) *For all $\lambda_1 > 0, \lambda_2 > 0, \gamma_1, \gamma_2$, let*

- $G = \{\, (\, \lambda_1 x_1 + \gamma_1,\ \lambda_2 x_2 + \gamma_2 \,) \mid (x_1, x_2) \in F \,\}$,
- $w = (\, \lambda_1 v_1 + \gamma_1,\ \lambda_2 v_2 + \gamma_2 \,)$.

Then $\phi(G, w) = (\, \lambda_1\, \phi_1(F, v) + \gamma_1,\ \lambda_2\, \phi_2(F, v) + \gamma_2 \,)$.

The scale covariance axiom is the wierdest looking one, but is very natural when viewed in the context of Decision Theory. we will comment on it Section 7.1.2.

Axiom 7.4 (Independence of irrational alternatives.) *For any closed convex set G, if $G \subseteq F$ and $\phi(F, v) \in G$, then $\phi(G, v) = \phi(F, v)$.*

Axiom 7.5 (Symmetry.) *If $v_1 = v_2$ and $\{\, (x_2, x_1) \mid (x_1, x_2) \in F \,\} = F$, then $\phi_1(F, v) = \phi_2(F, v)$.*

From these axioms results a single solution, as stated by Nash's theorem for bargaining.

Theorem 7.7: Nash's bargaining solution

There is a unique solution function $\phi(\cdot, \cdot)$ that satisfies Axioms 7.1-7.5 above. This solution function satisfies, for every two-person bargaining problem (F, v):

$$\phi(F, v) \in \operatorname*{argmax}_{x \in F,\, x \geq v} (x_1 - v_1)(x_2 - v_2). \tag{7.1}$$

Example 7.8 (Example 6.10 continued.) *In Example 6.10, Charles needs to pick an allocation for Alice and Bob, who are playing the game*

	c	b
C	2, 3	-1, -1
B	1, 1	3, 2

.

That allocation must be so that $u_i \geq 7/5$ for both players. In this case, it is easy to see that $u_1 + u_3 \leq 5$. Moreover, Charles can propose the allocation $u_i = 2.5$ for both players by picking the correlated strategy $0.5([C],[c]) + 0.5([B],[b])$. Since the allocation is both utilitarian and egalitarian, it is the Nash Bargaining Solution.

Example 7.9 (Example 7.5 continued.) We compute the Nash Bargaining Solution for the game of Figure 7.1. In (7.1), we see that we can focus on the set of all points that are rational. Hence, as shown in Figure 7.2, we can focus on this set for computing the solution.

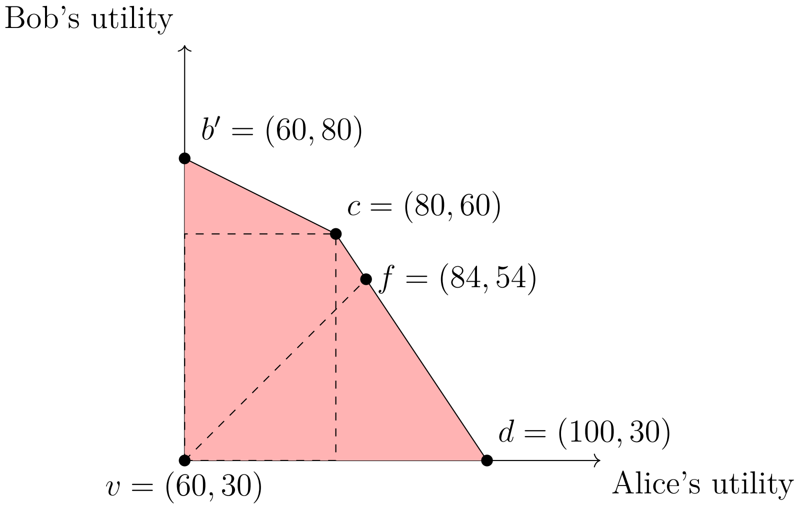


Figure 7.2: The set of rational points in Figure 7.1.

The solution in (7.1) is obtained by solving a quadratic program with (linear) constraints. To compute it, in this case, we may rely on our intuition. We know for a fact that the solution is Strongly Efficient (Axiom 7.1). Hence, it lies on one of the two lines forming the boundary of the domain. We may just find the maximum along each line, compare these maxima, and pick the best one as our solution.

- Let us first investigate the line segment $[b', c]$. It corresponds by a line with equation

$$y = \frac{80 - 60}{60 - 80}x + 80 - \frac{80 - 60}{60 - 80}60 = -x + 140.$$

When restricted on this line, our program becomes

$$\max_{x \in [60,80]} (x - 60)(-x + 110).$$

To solve this, we compute the derivative of the objective $(x - 60)(-x + 110)$ which gives $-2x + 170$ which vanishes at $x = 85$. The solution $x = 85$ is outside the range $[60, 80]$, however since our objective function to be maximized is concave, we conclude that the solution here is $x = 80$.

- *Second, we consider the line segment* $[c, d]$. *It corresponds to a line with equation*

$$y = -1.5x + 180.$$

Repeating our procedure, we want to solve the quadratic program

$$\max_{x \in [80,100]} (x - 60)(-1.5x + 150).$$

The derivate of the objective is $-3x + 240$ *which vanishes at* $x = 80$.

In conclusion, our Nash Bargaining Solution here is the one that gives a payoff of 80 *to Alice and* 60 *to Bob. For our original problem of Example 7.1, this corresponds to giving* 60 € *to Alice, and* 40 € *to Bob.*

7.1.2 Interpersonal comparison of weighted utility

As mentioned above, in real bargaining situations, the players often reason by comparing their respective utilities. They usually do so in two different ways:

- The "equal gains" principle: *"You should do that for me because I do more for you."*

  In agreement with this principle, we define the $\lambda-$*egalitarian solution* of a bargaining problem (F, v) as the unique point x of F that is weakly efficient in F and that satisfies the following condition:

$$\lambda_1(x_1 - v_1) = \lambda_2(x_2 - v_2).$$

- The "greater good" principle: *"You should do this for me because it helps me more than it harms you."*

  The $\lambda-$*utilitarian solution* of a bargaining problem (F, v) that derives from this principle is any solution function that yields $x \in F$ such that:

$$\lambda_1 x_1 + \lambda_2 x_2 = \max_{y \in F} (\lambda_1 y_1 + \lambda_2 y_2).$$

The following theorem expresses the fact that the Nash bargaining solution is a natural synthesis of the equal gains and greater good principles.

Theorem 7.10

Let (F, v) be an essential bargaining problem with two players and let x be an allocation vector such that $x \in F$ and $x \geq v$. Then x is the Nash bargaining solution for (F, v) iff there exist $\lambda_1 > 0$ and $\lambda_2 > 0$ such that:

$$\begin{aligned} \lambda_1 x_1 - \lambda_1 v_1 &= \lambda_2 x_2 - \lambda_2 v_2 \\ \text{and} \qquad \lambda_1 x_1 + \lambda_2 x_2 &= \max_{y \in F} (\lambda_1 y_1 + \lambda_2 y_2). \end{aligned}$$

7.1.3 About the disagreement point

Depending on the problem studied, the disagreement point v can be defined in at least three ways: we may choose the minimax solution, use a Nash equilibrium or make rational threats.

Minimax. In this case, players assume that in case of a disagreement, they will get the value that they can always ensure, that is, their minimax value which is given by:

$$v_i = \min_{\sigma_{-i} \in \Delta(C_{-i})} \max_{\sigma_i \in \Delta(C_i)} u_i(\sigma_i, \sigma_{-i}), \quad \text{with } i \in \{1, 2\}.$$

This choice of v can be useful to model unpredictable, irrational or risk averse behaviors for instance. Playing the minimax strategy can also make sense in a situation where the failure of a negotiation would be the fault of some particular player who could then choose to play safely because he cannot be threatened of having a payoff smaller than his minimax strategy.

Nash equilibrium. If σ is a Nash equilibrium of the game, then we may define v such that $v_i = u_i(\sigma)$.

Rational threats. The idea of rational threats is for each player to choose (and announce) his disagreement strategy in order to maximize his value with the Nash bargaining solution. Doing so, the players of course assume for an eventual agreement. More precisely, suppose that in case of a disagreement, the players announce that they will play τ_1 and τ_2 respectively. We call τ_1 and τ_2 *threats*. If those actions were to happen, the payoffs of the players would be $u(\tau_1, \tau_2)$ and the Nash bargaining solution would be $\phi(F, u(\tau_1, \tau_2))$. We say that the threats τ_1 and τ_2 are *rational* iff

$$\begin{aligned} &\phi_1(F, u(\tau_1, \tau_2)) \geq \phi_1(F, u(\sigma_1, \tau_2)) \quad \text{for all } \sigma_1 \in \Delta(C_1) \\ \text{and} \quad &\phi_2(F, u(\tau_1, \tau_2)) \geq \phi_2(F, u(\tau_1, \sigma_2)) \quad \text{for all } \sigma_2 \in \Delta(C_2) \end{aligned}$$

and in that case, we define v as $u(\tau_1, \tau_2)$. In other words, changing τ_i to any other threat would deteriorate the Nash bargaining solution for player i. This definition of v is perhaps the most meaningful of the three but it may also be too hard to compute... except in the case of transferable utility. Note that in the framework of Rational Threats, we assume that the players commit to play their disagreement strategy in case the negotiation fails (via some sort of binding contract). This assumption may not hold in some practical cases.

7.1.4 The case of transferable utility

Often in bargaining, it makes sense for a player to be willing to share some of his payoff with the other player in order to achieve a better solution for both. We then say that the utility is *transferable*[2]. Another way of interpreting transferable utility is to imagine

[2]To be precise, we should add that the utility beeing transferable also assumes that the players can throw away their payoffs.

that the players are playing as a team aiming for the highest total reward and then that they share this reward among them, according to each player's contribution.

In the case where we have transferable utility, the set of feasible payoffs is simply given by

$$F = \{ \ y \in \mathbb{R}^2 \mid y_1 + y_2 \leq v_{12} \ \}, \quad \text{where} \quad v_{12} = \max_{\mu \in \Delta(C)} u_1(\mu) + u_2(\mu).$$

Here v_{12} corresponds to the largest total payoff that the players can achieve together. In that case, the Nash bargaining solution can be explicitly computed by

$$\phi_1 = \frac{v_{12} + v_1 - v_2}{2} \quad \text{and} \quad \phi_2 = \frac{v_{12} - v_1 + v_2}{2}.$$

We end this section with a theorem that shows that rational threats can be easily computed in the case of transferable utility.

Theorem 7.11

The solution of the two-player Nash bargaining problem with transferable utility when the players are using the threats τ_1 and τ_2 respectively is given by:

$$\phi_1(F, u(\tau_1, \tau_2)) = \frac{v_{12} + u_1(\tau_1, \tau_2) - u_2(\tau_1, \tau_2)}{2} \tag{7.2}$$

$$\text{and} \quad \phi_2(F, u(\tau_1, \tau_2)) = \frac{v_{12} - u_1(\tau_1, \tau_2) + u_2(\tau_1, \tau_2)}{2}. \tag{7.3}$$

Therefore, the threats τ_1 and τ_2 are rational iff they are an equilibrium of the two-player zero-sum game given by $\Gamma^* = (\{1, 2\}, \Delta(C_1), \Delta(C_2), u_1 - u_2, u_2 - u_1)$.

The last part of the theorem comes from the fact that the players aim to maximize expression (7.2) and (7.3) respectively, which boils down to maximizing $u_1(\tau_1, \tau_2) - u_2(\tau_1, \tau_2)$ for Player 1 and $u_2(\tau_1, \tau_2) - u_1(\tau_1, \tau_2)$ for Player 2. We remind that the equilibria of a two-player zero-sum game can be easily computed.

7.1.5 Game of Alternative offers

We end the topic of two-player bargaining problems by discussing Alternating-Offer Bargaining Games. This refers to a situation where both players negociate on a prize, offering a price one after another, until either they agree to exchange the item at this prize, or they cease negotiations.

Let (F, v) be a regular two-person bargaining problem, and let $0 < p_1, p_2 < 1$ be two numbers that can be interpreted as the *patience* of each player.

The game plays as follows.

- At every odd numbered round, Player 1 makes an *offer* $x \in F$. If Player 2 accepts the offer, the game ends, and the players receive the payoff corresponding to the offer x. If the player rejects the offer, then there are two possibilities. Either the game ends with a disagreement, which occurs with probability p_1 where p_1 represents the probability that Player 1 gets impatient. In this case the players receive a payoff (v_1, v_2). Otherwise, the next round begins.

- At every even numbered round, Player 2 makes an *offer* $x \in F$. If Player 1 accepts the offer, the game ends. If the player rejects the offer, then there are two possibilities. The game may end with probability p_2 and the players have a payoff (v_1, v_2). Otherwise, the next round begins.

The game appears complex at first sight, but it turns out the equilibrium strategies of the players are well understood.

Theorem 7.12

The alternative offer game has a *unique* subgame-perfect equilibrium where

- Player 1's strategy is to offer a strongly efficient allocation $\bar{x}$, which will be accepted at round 1;
- Player 2's strategy is to offer a strongly efficient allocation $\bar{y}$,
- and the offers satisfy

$$\bar{y}_1 = (1 - p_2)(\bar{x}_1 - v_1) + v_1,\ \bar{x}_2 = (1 - p_1)(\bar{y}_2 - v_2) + v_2.$$

Furthermore, player 1 would accept any offer giving him at least $\bar{y}_1$, and player 2 would accept any offer giving him at least $\bar{x}_2$.

7.2 The bargaining problem with more than two players

We have dealt with the case where two players seek to find an agreement through the means of negociations. The natural next step is to consider what happens when three or more players, each with their respective objectives, find themselves in a similar scenario. Surprisingly, we will see that the solution concept of the previous section, the Nash Bargaining Solution, becomes inadequate when considering $N \geq 3$ players. One of the main root of this is that, as more players enter the picture, we now have to deal with an additional phenomenon: *players may now form coalitions, where a part of the group band together in order to maximize their own objectives.* Example 7.13 illustrates the fact that a direct generalization of the tools of Section 7.1 may be inadequate.

Example 7.13 *Alice, Bob, and Jean-Henry are awarded 300€ to be shared between them. They want to do the following. Let x_1, x_2 and x_3 be the shares of Alice, Bob and Jean-Henry respectively.*
They seek to find an allocation $x = (x_1, x_2, x_3)$, *where $x_i \geq 0$ for each player, and* $x_1 + x_2 + x_3 \leq 300$.
An allocation is accepted *if* 2 *players out of* 3 *vote to accept it. They are of course allowed to discuss before the vote.*

A natural way to tackle the issue is by a direct extension of the tools of Section 7.1. We form a 3-player bargaining problem (F, v) where

$$F = \{(x_1, x_2, x_3 \mid x_i \geq 0, \sum_i x_i \leq 300\},$$

and where we pick, for the time being, $v = (0, 0, 0)$. *With this setup, the Nash-Bargaining solution for this problem is given by* $x = (100, 100, 100)$ *(this is easy to see simply from the Symmetry Axiom 7.5).*

Let us add a new information to the discussion. Our three players are fervent believers in democracy, and that the will of a majority should always decide. Consequently, we will see that the $x = (100, 100, 100)$ *allocation should* not *be accepted by the players. Indeed, Bob and Jean-Henry can team up and see that the allocation* $x = (0, 150, 150)$ *is better for both of them, and that they can obtain it if they both vote for it.*
Alice can then propose a deal to Bob: the allocation $x = (50, 250, 0)$ *is clearly better for both of them (she would have got nothing in the previous deal).*
Jean-Henry could then approach Alice with a new deal she could not refuse: the allocation $x = (150, 0, 150)$ *would be better for both of them compared to the previous one!*

As the reader may guess, the above reasoning will not lead to a solution. Every time a group of players discuss an allocation, another group can be formed for which a better allocation exist. It is critical to study when these dynamics may occur.

We focus on games with *transferable utility*. That is, we assume that after the game happened, the players are allowed to redistribute their payoffs among the players in the coalition. In this setting, a coalition game is completely described by a function that returns the value earned by each coalition (that is, the sum of the payoffs earned by the members of the coalition).

Our main tool to describe a coalition game with transferable utility is the *characteristic function*.

Definition 7.14: Characteristic function

Given a coalition game with N players, a *characteristic function* is a function of the form

$$v : 2^N \to \mathbb{R} : S \mapsto v(S),$$

with $v(\emptyset) = 0$, that assigns to each coalition $S \subseteq N$ the amount of transferable utility $v(S)$ it can achieve. We call it the *worth* of the coalition.

The characteristic function is the central concept we use here to define coalitional games.

Definition 7.15

A coalitional game is a pair (N, v), where v is a characteristic function.

Definition 7.16: (Super-Additivity and Convexity

A coalition game (N, v) is:

- *super-additive* if, for any pair $S, T \subseteq N$, with $S \cap T = \emptyset$,

$$v(S \cup T) \geq v(S) + v(T),$$

- *convex* (or *super-modular*) if, for any pair $S, T \subseteq N$,

$$v(S \cup T) \geq v(S) + v(T) - v(S \cap T).$$

With a super-additive characteristic function, the payoff that all players get together is greater than what they would get if the group was divided into smaller teams. Unless stated otherwise, we make the assumption that super-additivity holds. Note that there are easy techniques in order to obtain an equivalent super-additive characteristic function from an arbitrary characteristic function. Hence, this assumption is not conservative. Convexity is stronger than super-additivity when the characteristic function is positive. These games may appear particular, but are not so rare in practice (Shoham and Leyton-Brown (2008)). In fact, we call it 'convexity' because it is the natural equivalent to the notion of convexity in functional analysis. Hence, it is not surprising that this property 1) is naturally appearing in many applications, and 2) often enhances our ability to solve problems in an efficient way.

Example 7.17 *(Airport Game, (Shoham and Leyton-Brown (2008))) A number N of cities need airport capacity. A first solution is for each city to build its own airport. The cost of an airport for city i is y_i, that scales with the size of the runway that city needs. A group $S \subseteq N$ of cities can also decide to collaborate and build a regional airport. Then, they need to share the cost $\max_{i \in S} y_i$, in order to build the largest runway that is needed. The worth of a coalition S is given by*

$$v(S) = \sum_{i \in S} y_i - \max_{i \in S} y_i,$$

which is the benefit made by the coalition if it allocates its resources to building its regional airport rather than spending it on individual airports.

We have here a convex coalition game (N, v). Indeed,

$$\begin{aligned} v(S \cup T) &= \left(\sum_{i \in S} y_i + \sum_{i \in T} y_i - \sum_{i \in S \cap T} y_i \right) - \left(\max_{i \in S \cup T} y_i \right) \\ &\geq \left(\sum_{i \in S} y_i + \sum_{i \in T} y_i - \sum_{i \in S \cap T} y_i \right) - \left(\max_{i \in S} y_i + \max_{i \in T} y_i + \max_{i \in S \cap T} y_i \right). \end{aligned}$$

The rest of this section is divided into two parts. First, we will linger a bit more on the concept of characteristic function, and on how we can construct it from a game in strategic form. Second, we will discuss different ways one can obtain, from a characteristic function, payoff allocations for each player (which is, in fact, our main interest here).

7.2.1 Constructing characteristic functions from a game in strategic form

We are given a game in strategic form $\Gamma = (N, C, u)$ with transferable utility (recall that this means that, first, all players share the same utility scale, and that second, players have the possibility to give away part of their utility payoffs to other players).

We discuss three ways to define a characteristic function for this game, that are similar in fact to what we use to define the disagreement points in two-players bargaining problems 7.1.3.

As a common grounds among the three approaches, for each $S \subseteq N$, we consider a game with two *meta-players* corresponding to the coalition $S \subseteq N$ on the one hand and the coalition $N \backslash S$ on the other hand. We refer to this as the coalition game.

Minimax. In the minimax representation, we have

$$\forall S \subseteq N : v(S) = \min_{\sigma_{N\backslash S} \in \Delta(C_{N\backslash S})} \max_{\sigma_S \in \Delta(C_S)} \sum_{i \in S} u_i(\sigma_S, \sigma_{N\backslash S}).$$

That is, we assume that the coalition S wants to maximize his total amount of utility while the coalition $N\backslash S$ wants to deny S the most utility as possible. In the above, $C_S = \times_{s \in S} C_s$, the set $\Delta(C_S)$ is the set of all *correlated strategies* available for the players within the coalition S, and

$$u_i(\sigma_S, \sigma_{N\backslash S}) = \sum_{c_S \in C_S} \sum_{c_{N\backslash S} \in C_{N\backslash S}} \sigma_S(c_S)\sigma_{N\backslash S}(c_{N\backslash S}) u_i(c_S, c_{N\backslash S}).$$

The interpretation of the minimax representation is straightforward. It assumes that everyone would like to take part in the negotiations, and to achieve this, they take an aggressive posture towards those within the coalition S. Remark that this somehow violates the assumption that players are rationals, because attacking the coalition S may hurt the group $N\backslash S$ way more.

Nash Equilibrium. In the Nash-Equilibrium representation, the coalitions S and $N\backslash S$ pick their strategies such that they are a Nash-Equilibrium of the coalition game:

$$\bar{\sigma}_S \in \arg\max_{\sigma_S \in \Delta(C_S)} \sum_{i \in S} u_i(\sigma_S, \bar{\sigma}_{N\backslash S}),\ \bar{\sigma}_{N\backslash S} \in \arg\max_{\sigma_{N\backslash S} \in \Delta(C_{N\backslash S})} \sum_{i \in N\backslash S} u_i(\bar{\sigma}_S, \sigma_{N\backslash S}).$$

The characteristic function is given by

$$v(S) = \sum_{i \in S} u_i(\bar{\sigma}_S, \bar{\sigma}_{N\backslash S}).$$

The interpretation here is more in line with our rationality assumption: for every choice of S, we assume that both S and $N\backslash S$ seek only to maximize their payoff. However, there could be several different Nash Equilibria in the coalition game, and so this model leaves some arbitrary choice, which is a drawback.

Rational threats. The rational threat representation is somehow *in between* the two others. The focus here is one making a coalition S *more appealing* than a coalition $N\backslash S$. The strategies of S and $N\backslash S$ are chosen as follows:

$$\bar{\sigma}_S \in \arg\max_{\sigma_S \in \Delta(C_S)} \left(\sum_{i \in S} u_i(\sigma_S, \bar{\sigma_{N\backslash S}}) - \sum_{j \in N\backslash S} u_j(\sigma_S, \bar{\sigma_{N\backslash S}}) \right), \tag{7.4}$$

$$\bar{\sigma}_{N\backslash S} \in \arg\max_{\sigma_{N\backslash S} \in \Delta(C_{N\backslash S})} \left(\sum_{j \in N\backslash S} u_j(\bar{\sigma}_S, \sigma_{N\backslash S}) - \sum_{i \in S} u_i(\bar{\sigma}_S, \sigma_{N\backslash S}) \right), \tag{7.5}$$

$$v(S) = \sum_{i \in S} u_i(\bar{\sigma}_S, \bar{\sigma}_{N\backslash S}). \tag{7.6}$$

When the three representations lead to the same characteristic function, we say the game has *orthogonal coalitions* (see (Myerson, 1991, p. 426)).

Example 7.18 ((Myerson, 1991, pages 425 - 426)) *Consider the following three players game in strategic form.*

	a_2		b_2	
	a_3	b_3	a_3	b_3
a_1	$4,4,4$	$2,2,5$	$2,5,2$	$0,3,3$
b_1	$5,2,2$	$3,0,3$	$3,3,0$	$1,1,1$

- *With the* minimax representation*, we obtain the following characteristic function:*

$$v(\{1\}) = v(\{2\}) = v(\{3\}) = 1, \textit{ where } S \textit{ and } N\backslash S \textit{ play } b,$$
$$v(\{1,2\}) = v(\{1,3\}) = v(\{2,3\}) = 4, \textit{ where } S \textit{ plays } a \textit{ and } N\backslash S \textit{ plays } b,$$
$$v(\{1,2,3\}) = 12.$$

- *With the Nash-Equilibrium representation, we obtain:*

$$v(\{1\}) = v(\{2\}) = v(\{3\}) = 5, \textit{ where } S \textit{ plays } b \textit{ and } N\backslash S \textit{ plays } a,$$
$$v(\{1,2\}) = v(\{1,3\}) = v(\{2,3\}) = 4, \textit{ where } S \textit{ plays } a \textit{ and } N\backslash S \textit{ plays } b,$$
$$v(\{1,2,3\}) = 12.$$

- *With the Rational threats representation, we obtain:*

$$v(\{1\}) = v(\{2\}) = v(\{3\}) = 1, \textit{ where } S \textit{ and } N\backslash S \textit{ play } b,$$
$$v(\{1,2\}) = v(\{1,3\}) = v(\{2,3\}) = 2, \textit{ where } S \textit{ and } N\backslash S \textit{ play } b,$$
$$v(\{1,2,3\}) = 12.$$

7.2.2 The Core

Until now, we have provided a modeling framework for a coalitional game. We are now in position to propose a *solution concept*. That is, we seek for a payoff allocation for each player that reflects their ability to generate payoffs for the group. Since we are restricting ourself to super-additive games, we can start our analysis by assuming that the overall payoff generated by the full set of agent is the payoff corresponding to N, which we call the *grand* coalition. Indeed, it would not be rational to generate less payoff, as $v(N)$ is the maximal possible overall payoff. Our goal (our "solution concept") boils thus down to answering the question: 'how should we split the payoff of $v(N)$ into the different players?'.

The core is the first solution concept that we propose for a game in coalitional form. It comes naturally from our intuition: we want to allocate payoffs to the players such that they can never secure a better payoff by forming a smaller coalition.

Definition 7.19: Feasibility, Improvability, and the Core

Consider a characteristic function $v = (v(S))_{S \subseteq N}$ for a coalition game and a payoff allocation $x \in \mathbb{R}^N$.

For a coalition $S \subseteq N$; x is *feasible* for S if

$$\sum_{i \in S} x_i \leq v(S).$$

If x is feasible for S and $\sum_{i \in S} x_i < v(S)$, then we say that S can *improve on x*. The *core* of the game is the set

$$\{x \in \mathbb{R}^n \mid \sum_{i \in N} x_i = v(N), \forall S \subseteq N : \sum_{i \in S} x_i \geq v(S)\},$$

i.e. the set of all allocations feasible for the coalition N and which cannot be improved upon by any smaller coalitions.

The concept of the core is perfectly in line with the notion of rationality: the only way players will play together in a big coalition N is that they cannot improve their payoffs by forming smaller coalitions (including the possibility of playing alone). Remark that the core describes a set of feasible allocations that players should accept, but does not produce a unique solution. To remedy this, one can add a criterion that needs to be optimized, e.g. minimize the difference between allocations to make them more "egalitarian". However the function to optimize might not be agreed upon by all the players, and this constitutes thus a first weakness for this solution concept.

Surprisingly, while very intuitive with respect to our goals, the core has another weakness of even greater importance, in that there are some games for which it is *empty*.

Example 7.20 (Example 7.13 continued) *In Example 7.13, Alice, Bob and Jean-Henry wanted to share 300€ between them. The twist was that an allocation would be accepted if* 2 *of the three accepted it. Let us build a characteristic function for the game. We chose the minimax representation here (however, one can show that the three representations coincide here). We have the following characteristic functions:*

$$\begin{aligned} &v(\{1\}) = v(\{2\}) = v(\{3\}) = 0, \ \textit{a single player always gets bypassed,} \\ &v(\{1,2\}) = v(\{1,3\}) = v(\{2,3\}) = 300, \\ &v(\{1,2,3\}) = 300. \end{aligned}$$

For a point to be in the core, it needs to satisfy the following constraints,

$$x_1 + x_2 \geq 300, \ x_1 + x_3 \geq 300, \ x_2 + x_3 \geq 300, \ x_1 + x_2 + x_3 = 300,$$

which is impossible.

In other case, the core may not be empty, but the allocations in the core may seem *unfair*, or are *highly sensitive to disturbances*; we show this third weakness in the following example.

Example 7.21 (Matching gloves) *There are two groups of n_1 and n_2 people. The first group manufactures* left gloves, *and the second group manufactures* right gloves. *While a single glove can't be sold, a* pair *is sold* 1 € *on the market. Therefore, members of the two groups must find an arrangement if they hope to make a profit.*

In this case, we denote a coalition by a pair (a, b), where $0 \leq a \leq n_1$ are the number of members of the first group within, and $0 \leq b \leq n_2$ the number of members of the second group. We have

$$v(a, b) = \min(a, b).$$

We consider three cases:

- *$n_1 > n_2$. In this case, we can see that the only possible allocation in the core is to give a payoff of 1 to each member of the second group, and a payoff of 0 to each member of the first group. To see why, assume that we gave a payoff of $\epsilon > 0$ to some player in the first group, and zero to all the others. Then, members of the second group could offer $\epsilon/(n_1 - 1)$ to the others and form a coalition with them.*
- *$n_1 = n_2$. In this case, there is a continuum of possible allocations.*
- *$n_1 < n_2$. This is similar to the first case.*

Hence, the allocation depends only ('non-continuously') on the number of players in each group.

The core is therefore a very elegant concept and comes with its flaws. Several extensions have been studied to remedy these flaws.

Perhaps one of the more natural is the $\epsilon-$core.

Definition 7.22: the ϵ - Core

Consider a characteristic function $v = (v(S))_{S \subseteq N}$ and a number $\epsilon \geq 0$.
The $\epsilon-$*core* of the game is the set

$$\{x \in \mathbb{R}^n \mid \sum_{i \in N} x_i = v(N), \forall S \subseteq N : \sum_{i \in S} x_i \geq v(S) - \epsilon|S|\},$$

i.e. the set of all allocations feasible for the coalition N and for which no other coalitions can provide ϵ more to everyone of its members.

The 0-core therefore coincides with the core.

Additionally, there exists several results that allow us to predict the emptiness of the core depending on the structure of the game ((Shoham and Leyton-Brown, 2008, Section 12.2)).

Theorem 7.23

Every convex game has a non-empty core.

7.2.3 The Shapley value

In view of the failure of the core, it is natural to take a step back and ask ourselves what would be a *fair* division of the worth $v(N)$ among a set of players? Like Nash did with its Nash Bargaining Solution, Shapley's solution is built from reasonable axioms. Given a coalition game (N, v), we are looking for an $x = \phi(N, v)$ that satisfy the following axioms.

Axiom 7.6 (Symmetry) *In a game (N, v), agents i and j are interchangeable if, for any $S \subset N$ that does not contain i or j, $v(S \cup i) = v(S \cup j)$.*
For any v, if i and j are interchangeable, then we have $\phi_i(N, v) = \phi_j(N, v)$.

Axiom 7.7 (Support) *In a game (N, v), agent i is a* dummy *if $\forall S \subset N$ where $i \notin S$, $v(S \cup i) = v(S) + v(\{i\})$.*
For any v, if i is a dummy player, then $\phi_i(N, v) = v(\{i\})$.

Axiom 7.8 (Linearity) *For any two characteristic functions v_1 and v_2, we have that*

$$\phi_i(N, v_1 + v_2) = \phi_i(N, v_1) + \phi_i(N, v_2),$$

where $(N, v_1 + v_2)$ is the game where $(v_1 + v_2)(S) = v_1(S) + v_2(S)$.

Theorem 7.24: Shapley Value

Given any characteristic function v, there exists a *unique* function $\phi(N, v)$ satisfying the Symmetry, Support and Linearity axioms which is given by

$$\phi_i(N, v) = \sum_{S \subseteq N - i} \frac{|S|!(|N| - |S| - 1)!}{|N|!}(v(S \cup i) - v(S)).$$

The above formula finds an intuitive explanation in the following idea. Suppose that all the players were entering the game one by one. Each time a new player, i, enters the game, he brings some additional value $v(S \cup i) - v(S)$ to the coalition S already there, which he takes for himself. Then we consider every possible order in which the players can enter the game (there are $|N|!$ such orders) and it is easy to see that the Shapley Value is exactly this 'expected earning', when the ordering in which agents enter the coalition is choosen randomly.

Let us now comment on the link between the Shapley Value and the core. At first sight, the Shapley value is not guaranteed to be in the core. This means that, in theory, rational agents may not agree on the Shapley value allocation. In practice, the Shapley value is very appealing since it provides us a unique and always well-defined solution concept satisfying such axioms.

There are some cases where the Shapley value is indeed in the core.

Theorem 7.25

If a game is convex, then the Shapley value is in the core.

Example 7.26 (Example 7.17 continued) *Consider a set of $N = 3$ towns, called a, b, and c, that require airport capacities at Figure 7.3.*

In order to compute the Shapley value, we can construct a table such as in Figure 7.4. In this table, for each permutation of the towns, we make the towns enter the coalitions one by one and compute their respective contribution to the worth of the coalition. For example, for the order a, b, c, we write $v(\{a\}) = 0$ for town a, $v(\{a, b\}) - v(\{b\}) = 3$ for town b, and $v(\{a, b, c\}) - v(\{a, b\}) = 22 - 12 - 3 = 7$ for town c.

The Shapley values is then obtained by taking the average along each columns:

$$\phi(N, v) = (4, 2, 4).$$

Town	Airport Cost
a	7
b	3
c	12

Figure 7.3: Required airport capacities

N (in all possible orders)	a's contribution	b's contribution	c's contribution
abc	0	3	7
acb	0	3	7
cab	7	3	0
cba	7	3	0
bca	7	0	3
bac	3	0	7
Expectancy	4	2	4

Figure 7.4: Computing the Shapley Values

Chapter 8

Repeated Games

"Insanity: Doing the same thing over and over again and expecting different results." — Albert Einstein.

Chapter based on (Myerson, 1991, pages 308 - 331), (Shoham and Leyton-Brown, 2008, Section 6.1). and on the Stanford/Coursera MOOC on Game Theory (for the section on learning in repeated games).

A repeated game is a strategic situation where a set of players play a same game several times in a row. These situations occur quite often in real life, and in fact, most of our everyday interactions can be seen as part of a repeated game.
Interestingly, the behavior of a player in a repeated game may be different than that of the same player when playing the game only once. For example, while we know that the only Nash Equilibrium of the Prisoner's Dilemma game is for both players to betray the other, cooperation may occur rationally in the repeated game setting.

The chapter is split in three sections, presenting matters from the more particular to the more general. Section 8.1 discusses finitely repeated game, that can be studied directly with the analysis tools of previous chapters. Section 8.2 discusses infinitely repeated games. There, we start observing interesting phenomena, in particular concerning the Prisoner's Dilemma. Finally, in Section 8.3, we present a more general and involved model for repeated games, and discuss the notions of equilibria in more details.

8.1 Finitely repeated games

In a *finitely repeated game*, players play the *same game* Γ repeatedly for a finite number of times K. The number K is known to the players at the beginning of the game. This situation is represented in Figure 8.1.

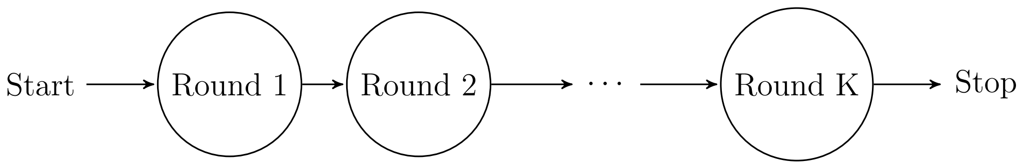


Figure 8.1: Finitely Repeated Games.

After each round, the players collect their payoffs and then move on to the next.

A finitely repeated game can be written as a classical game, where players have to decide before the first round on a strategy to play at each round. The next example illustrates this on the Prisoner's Dilemma.

Example 8.1 (Twice Repeated Prisoner's Dilemma) *Consider the Prisoner's Dilemma game* Γ *below*

Prisoner's Dilemma	c	d
C	$1,1$	$-1,2$
D	$2,-1$	$0,0$

The only Nash Equilibrium is the situation where both players defect (D, d)*. However, it has been observed that, in many social interactions, players have a natural tendency to cooperate. Furthermore, it makes sense, intuitively, that if one is going to interact with another on several occasions, it would be rational to show ourselves as agreeable and cooperative in order to foster a prosperous relation.*

Let us see what happens then when the game is repeated twice. The players are now faced with the situation represented in Figure 8.2. The key to analyze repeated games

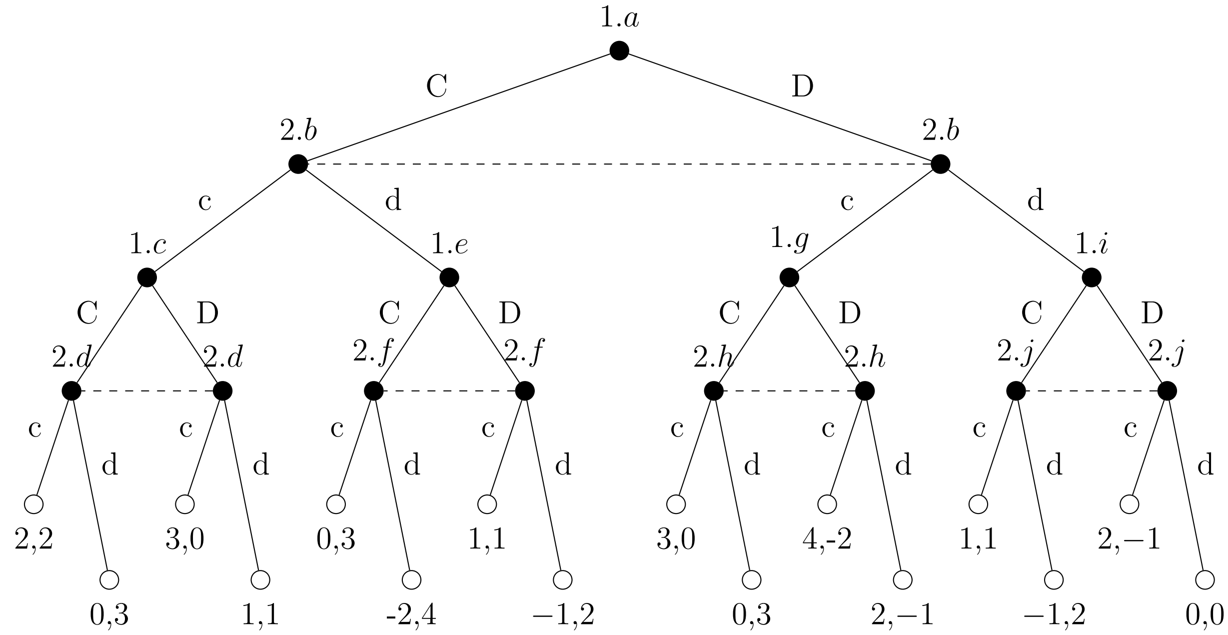


Figure 8.2: The twice repeated Prisoner's Dilemma, in extensive form. The payoffs at the terminal nodes are the sum of those at both rounds.

is the backward analysis. *For this, we begin by studying how the players should act at the second round. Clearly, they only want to maximize their contributions to their own payoffs. Hence, in the second round, they should select* $([D], [d])$ *(the Nash Equilibrium). Knowing this, we can infer (by recurrence) that the players should pick* $([D], [d])$ *as well in the first round.*

In conclusion, when playing a finitely repeated version of the Prisoner's Dilemma, players should always defect at every round. We did the exercise for when the game is repeated twice, but this holds for any fixed number K *of repetitions.*

The study of repeated games calls for the definition of an appropriate notion of *strategy*. The following definition, though informal, has the advantage to represent our intuition and it is valid for both the finitely and infinitely repeated case.

Definition 8.2: Strategy in a repeated game (informal)

In a repeated game with N players, a strategy for player i defines, at each round $k = 1, 2, \ldots$ a (randomized) strategy to play for the game at round k given the information available at the beginning of round k.

Let us linger a bit longer on what constitutes *information* for a player in round k. A player should at least be aware of *the history of his moves and payoffs* so far. The players may also know other players' moves or their payoffs depending on the scenario.

Definition 8.3: Stationary Strategy (informal)

In a repeated game, a *stationary strategy* is a strategy that depends only on the current round.

Proposition 8.4: Equilibrium in Finitely Repeated Prisoner's Dilemma

In a finitely repeated Prisoner's Dilemma, the unique equilibrium strategy is stationary, where each player defect at every round.

This can be easily proven by using the *backward analysis*, see Ch. 4, Sect. 4.1. Remarkably, in *infinitely repeated games*, the players will have access to a richer array of actions and payoffs. They will not be confined to playing Nash Equilibria at each round, but they will now be able to learn and adapt their moves in accordance to the history of the game. In particular, we can observe the rise of cooperation in the Prisoner's Dilemma in the infinitely repeated setting.

8.2 Infinitely Repeated Games

In an infinitely repeated game, all the players involved know that they will have to play the game an infinite amount of time. At each round, the players gather their payoffs for the current round before playing the next one (see Figure 8.3).

Notation 8.5: Time superscript

Given a sequence r in time, we use the superscript $r^{[t]}$ to denote the element of the sequence r at time t.

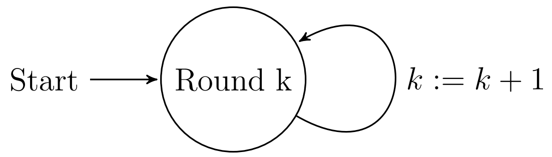


Figure 8.3: Infinitely Repeated Games.

Definition 8.6: Standard Repeated Game

A *standard repeated game* is an infinitely repeated game where a same game in strategic form Γ is repeated infinitely and where at each round k, every player knows all the players past moves.

When playing a repeated game, players needs to select some strategy to be played instead of others. To do so, they need to be able to express preferences, as prescribed by Decision Theory. The following definitions provide two classical criteria for ranking payoffs in repeated games.

Definition 8.7: Average and Discounted Payoffs

Given a sequence of payoffs $r_i^{[1]}, r_i^{[2]}, \ldots$ for player i in an infinitely repeated game,

- the *limit average payoff* of i is

$$\lim_{k\to\infty} \frac{\sum_{j=1}^{k} r_i^{[j]}}{k}. \tag{8.1}$$

  Note that this limit may not exist (but more sophisticated limit definitions can be used to circumvent this),

- for a *discount factor* $0 \leq \delta < 1$, the $\delta-$*discounted payoff* is

$$\sum_{j=1}^{\infty} \delta^{j-1} r_i^{[j]}. \tag{8.2}$$

- for a *discount factor* $0 \leq \delta < 1$, the $\delta-$*discounted average payoff* is

$$(1-\delta)\sum_{j=1}^{\infty} \delta^{j-1} r_i^{[j]}. \tag{8.3}$$

Other criteria exist, see (Myerson, 1991, page 315).

The *discounted payoff* can be interpreted in a number of ways.
The *discount factor* δ is commonly understood as a measure of the player's *patience*: the closer to 0, the more the player aims at obtaining immediate payoffs, and the closer to 1,

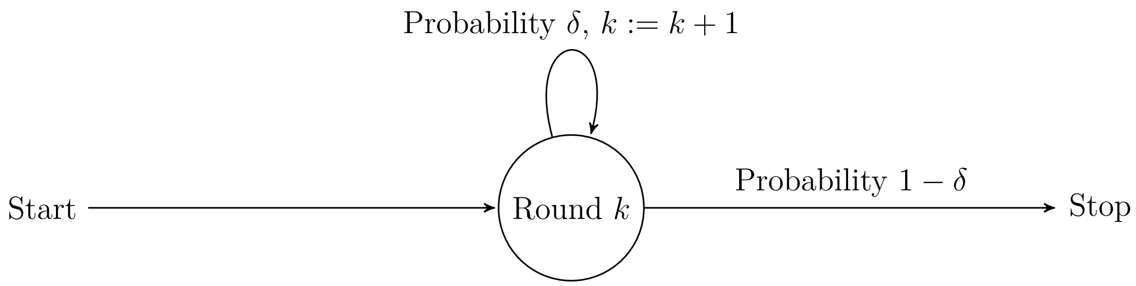


Figure 8.4: Interpretation of discounted payoff: the game ends with probability $1 - \delta$.

the more the player is willing to wait for future payoffs.
There is another way to interpret the discount factor, which is useful to model situations where the game will actually not be played an infinite number of time, but will end at some moment, *in a probabilistic way.* Indeed, suppose that at every round, the game may terminate with probability $1 - \delta$, or go on for at least one more round with probability δ (see Figure 8.4). The reader can observe that the equations for the expected payoff are exactly the same than in the discounted setting. Indeed, if the game may stop at every round with probability $1 - \delta$, in order to obtain the payoff $r_i^{(j)}$ the game must be repeated $j - 1$ times, which occurs with probability δ^{j-1}. The discounted payoff is then the expected payoff of the game under this interpretation. Finally, as the reader has guessed, there is a third interpretation/reason to use the discounted payoff: it represents situation where inflation occurs, that is, the actual value of a gain at some time t is preferred to the same gain later in time.

The *discounted average payoff* is a normalized version of the discounted payoff, where if $r_i^{[t]} = 1$ for all t,

$$(1 - \delta) \sum_{j=1}^{\infty} \delta^{j-1} 1 = 1.$$

Definition 8.8: Equilibrium for Repeated Games (informal)

An equilibrium of a repeated game is a set of strategies (one per player) where each strategy is a best response to the others, in that it maximizes the selected payoff criterion given the other players' strategies.

Example 8.9 *In an infinitely repeated version of the Prisoner's Dilemma, cooperation can appear as a rational strategy to be played at each round. Consider a discount factor δ and the game of Example 8.1. We define the strategy of the player as follows: each player is supposed to start the game by cooperating. Then, if at any round a player defects, the other reacts by defecting for the rest of the game. The δ-discounted payoff for the strategy is*

$$\sum_{j=1}^{\infty} (\delta)^{j-1} 1 = \frac{1}{1 - \delta}$$

where players cooperate at every round. If a player defects at the first round, he gets a payoff of 2 and then always gets a payoff of 0 because the other player reacts by always *defecting. His total payoff is then simply 2, independently of δ. Consequently, if*

$$\frac{1}{1-\delta} \geq 2 \Leftrightarrow \delta \geq 0.5,$$

the strategy is an equilibrium, leading to a cooperative outcome for the game.

In the example above, the strategy leading to the cooperative equilibrium infinitely punishes a player for uncooperative behaviors. This is called the **grim** (or **trigger**) strategy. Other classical strategies, leading to cooperative equilibria, are as follows

- **Grim**: the grim strategy is that of Example 8.9. We begin the game with cooperation and continue cooperating at each round unless the other player defects. From this point on, we defect. See the analysis in (Myerson, 1991, pp. 308-310) for a detailed analysis of this example.
- **tit-for-tat**: we begin by cooperating and then, at every round, we choose the move of our opponent at the last round,
- **getting-even**: we begin by cooperating, and continue cooperating *unless* the opponent has defected more time that we did. If this is the case, we defect.

8.2.1 Folk Theorems and Cooperation

Infinitely repeated games offer opportunities for cooperative behaviors, which can be linked to the theory of cooperative game theory. The following is taken from (Shoham and Leyton-Brown, 2008, Theorem 6.1.5).

Theorem 8.10: Folk Theorem

Consider any N-player normal-form game $\Gamma = (N, C, u)$, and a payoff profile $r = (r_i)_{i \in N}$ with $r_i \in \mathbb{R}$. Assume that

- the profile r satisfies $r_i \geq v_i$, where

$$v_i = \min_{c_{-i} \in C_{-i}} \max_{c_i \in C_i} u_i(c_{-i}, c_i),$$

- and there is a correlated strategy $\tau \in \Delta C$, where $\tau(c) \in \mathbb{Q}$ for each $c \in C$ where $\mathbb{Q}$ is the set of rational numbers, such that

$$r_i = \sum_{c \in C} \tau(c) u_i(c).$$

Then r is the payoff obtained by the players for some Nash Equilibrium of the infinitely repeated game with *limit average rewards.*

There is an interesting parallel here with Chapter 6 and in particular Section 6.2.1. As a reminder, the topic there was finding rules under which a group of players is willing to sign a contract, forcing each of them to follow the recommendations of a mediator. The mediator is a proxy to represent communication processes between the players. The following example illustrates the results.

Example 8.11 (Example 6.10 revisited.) *Alice and Bob wish to hold an old promise: once a month, for the rest of their lives (which we assume to be infinite for the exercise), they will meet either at the cinema or at the ballet. Alice prefers to go to the ballet, and Bob to the cinema. Their payoffs and moves are summarized in the table below (Alice is player 1, Bob is Player 2).*

$$\begin{array}{c|cc} & c & b \\ \hline C & 2,3 & -1,-1 \\ B & 1,1 & 3,2 \end{array}.$$

We computed in Example 6.10 the minimax values $v = (7/5, 7/5)$ *for this game, and saw in Example 7.8 that a very good correlated strategy here is to play* $\tau([C],[c]) = 0.5$, $\tau([B],[b]) = 0.5$, *with expected payoffs* $(2.5, 2.5)$.

For the sake of the exercise, let us assume that Alice and Bob don't plan ahead during these activities: they just enjoy the moment, and have no other means of communication. They decide their moves based on the history of the game (here, it is easy to find out where was the other if they do not meet).

It is not difficult to construct strategies for Alice and Bob that grant them an average payoff *of* 2.5 *in the infinitely repeated game.*

Let us consider the sequence of outcomes $a^{(t)} = ([C],[c])$ *if* t *is even, and* $a^{(t)} = ([B],[b])$ *if* t *is odd. Consider the following strategy for each player: begin by playing with* $a^{(1)}$ *at round 1, and then at round* k

- *if the outcomes at all previous rounds are those of the sequence* a*, then play according to* $a^{(k)}$*,*
- *else, play the minimax strategy for the other player.*

Clearly, if both player follow this strategy, they get an average payoff of 2.5*. Moreover, if one deviates from this strategy, his expected payoff drops to* 7/5*, because the other plays is minimax for eternity. Hence, the strategy is a Nash Equilibrium.*

8.3 General Model of Repeated Games

So far, we have focused on finitely repeated games and standard repeated games. We now discuss a more general model, with the associated computational tools for formally defining strategies and equilibria.

To represent a repeated game, we need a number of elements. First, of course, we need a set of players N and a set of actions $D = \times_{i \in N} D_i$ that the players can play at each round.
Additionally, we may consider that in each round, the game is in a different *state*. For example we can understand the situation Figure 8.4 as one infinitely repeated game with *two* states: the game is either running or stopped, and the finitely repeated game of Figure 8.1 as infinitely repeated games with $K + 1$ states: K rounds and the stop state. In all generality, we let Θ be the set of *states of the world*.
The actions of players lead to rewards at the end of each round, depending on the state of nature, through a utility function $u_i : D \times \Theta \to \mathbb{R}$ for each $i \in N$.

As mentioned before, players make decision based on currently available information, gathered through past rounds. This is something we represent through a set of *signals* $S = \times_{i \in N} S_i$. Signals are flexible in their nature: the set of signal S_i for player i could contain D, meaning he obtain at each round information about the moves of all players on the past round; or it may contain Θ, meaning the player has information on the state of the world, etc... This flexibility can be used for modeling purpose.
Additionally, we need to define how the game *starts*. At the beginning of the game, each player must receive a signal (which can be as simple as "the game as started"), and a state of world has to be selected. This is represented by an *initial distribution* $q \in \Delta(S \times \Theta)$. Finally, as the game evolves, players will be fed signals and the state of the world will change. In all generality, we represent this by a *transition function* $p : D \times \Theta \to \Delta(S \times \Theta)$.

Definition 8.12: General Repeated Game

A general repeated game is a structure of the form

$$\Gamma^r = (N, \Theta, (D_i, S_i, u_i)_{i \in N}, q, p),$$

where

- N is the set of *players*,
- Θ is the set of *state of the worlds*,
- D_i is the set of *moves* available to player i,
- S_i is the set of *signals* that player i may receive,
- $q \in \Delta(S \times \Theta)$ is the *initial distribution*,
- $p : D \times \Theta \to \Delta(S \times \Theta)$ is the *transition function*.

Definition 8.13: Strategy

Consider a repeated game $\Gamma^r = (N, \Theta, (D_i, S_i, u_i)_{i \in N}, q, p)$. The set of all *pure strategies* is $C = \times_{i \in N} C_i$,

$$C_i = \{c_i = (c_i^{[k]})_{k=1}^{\infty} \mid \forall k, c_i^{[k]} : (S_i)^{\times k} \to D_i\}.$$

The set of all *behavioral strategies* is $B = \times_{i \in N} B_i$,

$$B_i = \{\sigma_i = (\sigma_i^{[k]})_{k=1}^{\infty} \mid \forall k, \sigma_i^{[k]} : (S_i)^{\times k} \to \Delta D_i\}.$$

In the above, the set $(S_i)^{\times k}$ is the k-fold Cartesian product of the set S_i.
The next concept will allow us to compute the payoffs of a given set of strategies.

Definition 8.14: Probability of playing move and reaching state

Given a repeated game $\Gamma^r = (N, \Theta, (D_i, S_i, u_i)_{i\in N}, q, p)$, a behavioral strategy $\sigma \in B$, moves $d \in D$ and a state of the world $\theta \in \Theta$, we let

$$P^{[k]}(d, \theta \mid \sigma)$$

be the probability that, at round k, the state of the world is θ and the players' moves are d if they play according to σ.

These probabilities can be computed as follows. Let

$$Q^{[k]}(s, \theta \mid \sigma)$$

be the probability that, at time k, the players are in the state θ and have received the signal $s \in (S)^{\times k}$. By definition, we have

$$P^{[k]}(d, \theta \mid \sigma) = \sum_{s\in(S)^{\times k}} Q^{[k]}(s, \theta \mid \sigma) \prod_{i\in N} \sigma_i^{[k]}(d_i \mid s).$$

The probabilities $Q^{[k]}$ can be computed recursively:

$$Q^{[1]}(s^{[1]}, \theta^{[1]} \mid \sigma) = q(s^{[1]}, \theta^{[1]}),$$
$$Q^{[k+1]}(s^{[1]} \dots s^{[k+1]}, \theta^{[k+1]} \mid \sigma) =$$
$$\sum_{\theta^{[k]}\in\Theta} \sum_{d\in D} Q^{[k]}(s^{[1]} \dots s^{[k]}, \theta^{[k]} \mid \sigma) \left(\prod_{i\in N} \sigma_i^{[k]}(d_i \mid s^{[1]} \dots s^{[k]}) \right) p(s^{[k+1]}, \theta^{[k+1]} | d, \theta^{[k]}).$$

The payoff of a player $i \in N$ at round k when following the strategies in σ is therefore given by

$$r_i^{[k]}(\sigma) = \sum_{d\in D} \sum_{\theta\in\Theta} P^{[k]}(d, \theta|\sigma) u_i(d, \theta).$$

Definition 8.15: Equilibrium

The behavioral strategy $\sigma = (\sigma_i, \sigma_{-i})$ is an equilibrium for

- the δ - discounted payoff criterion if

$$\forall i \in N, \forall \hat{\sigma}_i \in B_i : \sum_{k=1}^{\infty} \delta^{k-1} r_i^{[k]}(\sigma) \geq \sum_{k=1}^{\infty} \delta^{k-1} r_i^{[k]}(\hat{\sigma}_i, \sigma_{-i});$$

- the average payoff criterion if

$$\forall i \in N, \forall \hat{\sigma}_i \in B_i : \lim_{k\to\infty} \frac{\sum_{j=1}^{k} r_i^{[j]}(\sigma)}{k} \geq \lim_{k\to\infty} \frac{\sum_{j=1}^{k} r_i^{[j]}(\hat{\sigma}_i, \sigma_{-i})}{k}.$$

The formalism above highlights one of the big challenges of the analysis of repeated games: the set of behavioral strategies is infinite, and so checking if a strategy is an equilibrium can be very hard! In practice one often focuses on strategies relying only on limited information (for example, use only the ℓ last signal received to define the strategy). This can be understood as a relaxation of the concept of rationality and intelligence in favor of pragmatism: in reality, players may not have access to a perfect memory or be able to reflect perfectly about the game. Think for example about the following situation: what if rational players, instead of choosing a move in each round by themselves, have to program a computer with a limited amount of memory and computational power to do so? We refer to this as the *bounded rationality* principle, which is discussed in e.g. Rubinstein (1998), (Shoham and Leyton-Brown, 2008, Section 6.1.3), (Osborne and Rubinstein, 1994, Chapter 9).

8.3.1 Games with complete state information and discounting

We now narrow our scope to a particular (yet expressive) class of games.

Definition 8.16: Games with Complete state information

A repeated game $\Gamma^r = (N, \Theta, (D_i, S_i, u_i)_{i \in N}, q, p)$ has *complete state information* if, at every round, every player knows the current state of nature. That is, there is a function $w_i : S_i \to \Theta$ known by each player such that

$$\forall s \in S, \forall \theta, \theta' \in \Theta, \forall d \in D : p(s, \theta' | d, \theta) = 0 \text{ if } \theta' \neq w_i(s_i).$$

Definition 8.17: Stationary strategy

In a game with complete state information, a behavioral strategy is *stationary* if the move of each player at each time *depends only on the information state of the game at that time.* That is, there is a function $\tau_i : \Theta \to \Delta(D_i)$ such that

$$\forall k, \forall s \in (S)^{\times k} : \sigma_i^{[k]}(\cdot \mid s) = \tau_i(\cdot \mid w_i(s_i^{[k]})).$$

Example 8.18 *Consider the Prisoner's Dilemma game of Example 8.1 in a setting where players make this decision based only on the last moves of their adversary. Note that this goes against the idea that players are rational and intelligent, since they use only a part of the information at their disposal (bounded rationality). Nevertheless, many intuitive strategies can be analyzed in such a setting. It is for example the case for the Tit-for-Tat strategy discussed in Section 8.2.*
A natural way to model the situation is as a game with complete state information. Here, we have 4 states, each one assigned to the way the player moved on the last round:

$$\Theta = \{([C], [c]),\ ([C], [d]),\ ([D], [c]),\ ([D], [d])\}.$$

Regarding the signals, we let $S_i = \Theta$ and both players see the state of the world at every round.

For the initial distribution, we can assume that $q(([C],[c])) = 1$ *in the beginning. The transition function is straightforward:*

$$p(s, \theta \mid d, \theta') = 1 \text{ if } d = \theta \text{ and } s = (\theta, \theta),$$

and it is equal to 0 *for any other* s, θ.
If both players play according to the Tit-for-Tat (stationary here) strategy, the execution of the game can be summarized at Figure 8.5.

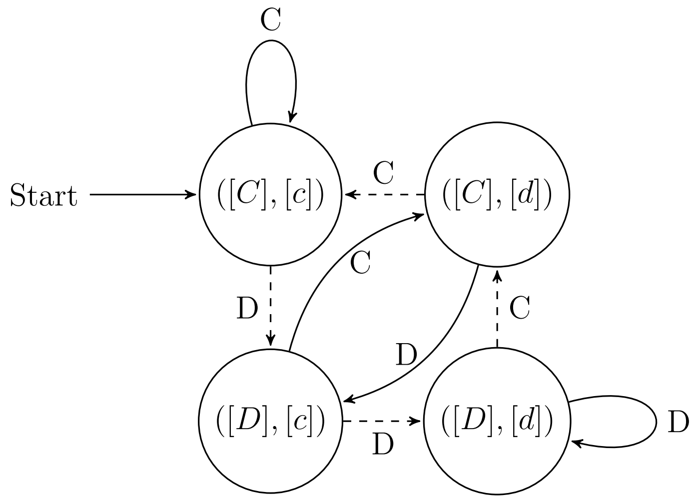


Figure 8.5: Prisonner's dilemma when playing Tit-for-Tat. Assuming Player 2 plays Tit-for-Tat, the plain arrows correspond to when Player 1 plays Tit-for-Tat as well, and the dashed arrows are what would happen should he decide to do the opposite.

We now provide characterization of equilibrium in stationary strategies for games with complete state information and the δ-discounted average payoff criterion, see Eq. (8.3), recalled below: given a sequence of payoffs $r^{[1]}, r^{[2]}, \ldots$ and $0 \leq \delta < 1$, the δ-discounted average of the sequence is $(1-\delta)\sum_{j=1}^{\infty} \delta^{j-1} r^{[j-1]}$.

Observe that this satisfies the following recursion: letting

$$w(k) = (1-\delta)\sum_{j=1}^{\infty} \delta^{j-1} r^{[j+k-1]},$$

which is the average from time k, we have

$$w(1) = (1-\delta) r^{[1]} + \delta w(2). \tag{8.4}$$

We are now in position to analyze games. To do so, we first let

$$\nu_i(\theta, \tau)$$

be the *expected δ-discounted average* of player i when beginning from the state $\theta \in \Theta$ when everyone plays according to τ. The following holds from (8.4):

$$\nu_i(\theta, \tau) = \sum_{d_i \in D_i} \tau_i(d_i|\theta) \sum_{d_{-i} \in D_{-i}} \tau_{-i}(d_{-i}|\theta) \left((1-\delta) u_i(d_i, \theta) + \delta \sum_{\theta' \in \Theta} p(\theta'|d, \theta) \nu_i(\theta', \tau) \right). \tag{8.5}$$

The above highlights an interesting term:

$$Y_i(\tau, d_i, \nu_i, \theta, \delta) = \sum_{d_{-i} \in D_{-i}} \tau_{-i}(d_{-i}|\theta) \left((1-\delta) u_i(d_i, \theta) + \delta \sum_{\theta' \in \Theta} p(\theta'|d, \theta) \nu_i(\theta', \tau) \right),$$

which can be interpreted as the gain of player i if he played the move d_i *once* if at state θ while everyone else follows τ (himself included in the future).

We end this section with a characterization of stationary equilibria for games with complete state information when considering the discounted average payoff:

Theorem 8.19

Given a repeated game with complete state information and bounded payoffs, and given a profile of stationary strategies τ. If there exists a bounded vector

$$\nu = (\nu_i(\theta))_{\theta \in \Theta, i \in N}$$

such that

$$\nu_i(\theta) = \sum_{d_i \in D_i} \tau_i(d_i|\theta) Y_i(\tau, d_i, \nu_i, \theta, \delta), \tag{8.6}$$

$$\nu_i(\theta) = \max_{d_i \in D_i} Y_i(\tau, d_i, \nu_i, \theta, \delta). \tag{8.7}$$

the strategy profile τ is an equilibrium of the repeated game.
In this equilibrium, $\nu_i(\theta)$ is the expected $\delta-$discounted average payoff for player i in the repeated game if the initial state is θ.

In general the equations of Theorem 8.19 can be quite hard to solve. In the following, we give an example applied to our Prisoner Dilemma game where we analyze the Tit-for-Tat and Grim strategies introduced in Section 8.2.

Example 8.20 (Example 8.18 continued) *Our goal is now to find the values* $\nu_i(\theta)$ *for the prisoner dilemma at Example 8.18. We saw in Example 8.9 that the Grim strategy was an equilibrium for the game with the* $\delta-$*discounted payoff when* $\delta \geq 0.5$. *Let us now consider the Tit-for-Tat strategies for both players and we focus on Player 1. For a given* $0 \leq \delta < 1$, *applying* (8.5) *to solve* (8.6) *(it's the same...), we have the following:*

$$\begin{aligned}
\nu_1([C],[c]) &= (1-\delta) \cdot u_1([C],[c]) + \delta \cdot \nu_1([C],[c]) \\
&= (1-\delta) \cdot 1 + \delta \cdot \nu_1([C],[c]), \\
\nu_1([D],[c]) &= (1-\delta) \cdot u_1([C],[d]) + \delta \cdot \nu_1([C],[d]) \\
&= (1-\delta) \cdot (-1) + \delta \cdot \nu_1([C],[d]), \\
\nu_1([C],[d]) &= (1-\delta) \cdot u_1([D],[c]) + \delta \cdot \nu_1([D],[c]) \\
&= (1-\delta) \cdot 2 + \delta \cdot \nu_1([D],[c]), \\
\nu_1([D],[d]) &= (1-\delta) \cdot u_1([D],[d]) + \delta \cdot \nu_1([D],[d]) \\
&= (1-\delta) \cdot 0 + \delta \cdot \nu_1([D],[d]).
\end{aligned}$$

The solutions are

$$\begin{aligned}
\nu_1([C],[c]) &= 1,\\
\nu_1([D],[c]) &= \frac{2\delta - 1}{1+\delta},\\
\nu_1([C],[d]) &= \frac{2-\delta}{1+\delta},\\
\nu_1([D],[d]) &= 0.
\end{aligned}$$

Then, for each state, we need to verify (8.7)*. In our present case, this boils down to checking, at any state $\theta \in \Theta$, if we would not have an higher payoff by deviating from the prescribed action at this state. Hence, we would be at equilibrium whenever the following inequalities are satisfied.*

$$\begin{aligned}
&\nu_1([C],[c]) \geq (1-\delta)\cdot 2 + \delta\nu_1([D],[c]) = \nu_1([C],[d]), \ \textit{doing } D \textit{ instead of } C,\\
&\nu_1([D],[c]) \geq (1-\delta)\cdot 0 + \delta\nu_1([D],[d]) = 0, \ \textit{doing } D \textit{ instead of } C\ ,\\
&\nu_1([C],[d]) \geq (1-\delta)\cdot 1 + \delta\nu_1([C],[c]) = 1, \ \textit{doing } C \textit{ instead of } D,\\
&\nu_1([D],[d]) \geq (1-\delta)\cdot(-1) + \delta\nu_1([C],[d]) = \nu_1([D],[c]). \ \textit{doing } C \textit{ instead of } D.
\end{aligned}$$

So we see that, for the equations to hold, we need to have exactly $\delta = 0.5$. If it is not the case, then:

- *For $\delta < 0.5$, the two first inequalities are invalid, but not the two second. This means that we should always do D. Hence, we revert to our original behavior when playing the non-repeated Prisoner's Dilemma. This fits the interpretation of a low δ being associated to the search for an immediate reward.*
- *For $\delta > 0.5$, the two second inequalities are invalid, but not the two first. This means that if we know the second player is going for Tit-for-Tat, we should always cooperate when δ is big enough.*

8.4 Learning in repeated games

In repeated games, *learning* strategies can be defined as follows.

Definition 8.21: Learning strategy

A learning strategy is a strategy that uses the information collected by a player during the previous rounds to improve its behavior in future rounds.

This corresponds to the way we see learning in our everyday life: using past inputs to improve itself in the future. In addition, learning strategies are supposed to draw interesting inferences or use accumulated experience in an interesting way. This will be illustrated with the convergence of "fictitious play" below.

There exists a lot of learning strategies. Two will be discussed in this section *fictitious play* and *no-regret playing.*

	l	r
U	3,3	0,0
D	4,0	1,1

Table 8.1: Normal form game used in the Fictitious play example. The only Nash Equilibrium of this game is (D, r).

8.4.1 Fictitious play

Fictitious play is an algorithm introduced by BROWN at the RAND corporation in 1949. Originally, the objective of this algorithm was to provide an explanation for the origin of Nash equilibria as well a heuristic to compute Nash equilibria. One can imagine this algorithm as a player playing against itself multiple times, trying to devise the best strategy (i.e. a Nash equilibrium) for when the time to play against a real opponent will come.

More rigorously, fictitious play is an instance of *model-based* learning where learners maintain beliefs about opponent's behaviors. The way the opponent's model is build is really simple and relies on a single assumption: the opponent is playing a mixed strategy given by the empirical distribution of its past moves.

Procedure 8.22: Fictitious play

Repeat:

1. Play a best response to the assessed strategy of the opponent (if first round, play an arbitrary strategy)
2. Observe the opponent's actual play and update beliefs accordingly

Note that this algorithm is often given with an additional beliefs initialization step. As a consequence, step 1 always consists in playing a best reponse (and no longer an arbitrary strategy at the first round). However, this version does not correspond to the original algorithm proposed by BROWN.

Example 8.23 (Fictitious play) *Let's consider the normal form game given in table 8.1 and two players using the fictitious play algorithm.*

Let's first define some notations

- ν_i^t*: belief of player* i *at turn* t*,*
- μ_i^t*: expected play of player* i *by the other players at turn* t*,*
- σ_i^t*: strategy of player* i *at turn* $t+1$*.*

At the first round, each player plays an arbitrary strategy. Let's assume then that Player 1 plays U and Player 2 plays l initially.

At the second round the belief vectors of each player are the following: $\nu_1^2 = (1,0), \nu_2^2 = (1,0)$*. The expected move of the two players are thus the following: Player 1 expects*

Player 2 to play $\mu_1^2 = l$ and player 2 expects Player 1 to play $\mu_2^2 = U$. The best response strategy of each player is thus going to be the following: Player 1 is going to play $\sigma_1^2 = D$ and Player 2 is going to play $\sigma_2^2 = l$.

At the third round the belief vectors are $\nu_1^3 = (2, 0)$ and $\nu_2^3 = (1, 1)$. The expected strategy of each player is thus $\mu_1^3 = l$ and $\mu_2^3 = \frac{1}{2}U + \frac{1}{2}D$. The best response strategy for each player is thus $\sigma_1^3 = D$ and $\sigma_2^3 = l$: in fact for Player 2 the expected payoff if he plays l is going to be equal to $\frac{1}{2} \cdot 3 + \frac{1}{2} \cdot 0 = 1.5$ and the payoff for playing right is $\frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 1 = 0.5$, thus playing l is the best response against the supposed mixed strategy of the opponent.

At the fourth round the belief vectors are $\nu_1^4 = (3, 0)$ and $\nu_2^4 = (1, 2)$. The expected strategy of each player is thus $\mu_1^4 = l$ and $\mu_2^4 = \frac{1}{3}U + \frac{2}{3}D$. The best response strategy for each player is thus $\sigma_1^4 = D$ and $\sigma_2^4 = l$: the expected payoff for Player 2 is equal to $\frac{1}{3} \cdot 3 + \frac{2}{3} \cdot 0 = 1$ if he plays l and $\frac{1}{3} \cdot 0 + \frac{2}{3} \cdot 1 = \frac{2}{3}$ if he plays right.

At the fifth round the belief vectors are $\nu_1^5 = (4, 0)$ and $\nu_2^5 = (1, 3)$. The expected strategy of each player is thus $\mu_1^5 = l$ and $\mu_2^5 = \frac{1}{4}U + \frac{3}{4}D$. The best response strategy for each player is thus $\sigma_1^5 = D$ and σ_2^5 can be either equal to l or to r, in fact, the expected payoffs for Player 2 are the same: $\frac{1}{4} \cdot 3 + \frac{3}{4} \cdot 0 = 0.75$ if he plays l and $\frac{1}{4} \cdot 0 + \frac{3}{4} \cdot 1 = 0.75$ if he plays right.

Finally at the sixth round, the belief vectors are $\nu_1^6 = (5, 0)$ and $\nu_2^6 = (1, 4)$. The expected strategy of each player is thus $\mu_1^6 = l$ and $\mu_2^6 = \frac{1}{5}U + \frac{4}{5}D$. The best response strategy for each player is thus $\sigma_1^4 = D$ and $\sigma_2^4 = r$: in fact, the expected payoff for Player 2 is equal to $\frac{1}{5} \cdot 3 + \frac{4}{5} \cdot 0 = 0.6$ if he plays l and $\frac{1}{5} \cdot 0 + \frac{4}{5} \cdot 1 = 0.8$ if he plays right. Since the strategy (D, r) is a Nash equilibrium (and a strongly dominating strategy) of the game, the strategy of each player won't change anymore.

This example suggests that fictitious play might converge. In particular, in might converge to a Nash's Equilibrium. Two questions remain, however: does it always converge? Does it always converge to a Nash's Equilibrium? The two following theorems answer these two questions.

Theorem 8.24

If the empirical distribution of each player's strategies converges in fictitious play, then *it converges to a Nash equilibrium.*

Theorem 8.25

Each of the following are a sufficient conditions for the empirical frequencies of play to converge in fictitious play[a]

- the game is zero-sum;

- the game is solvable by iterated elimination of strictly dominated strategies

[a] Additionnal conditions that are out of the scope of this course exist but are not mentionned here.

8.4.2 No-regret playing

The second algorithm we are going to study in this section is *no-regret playing*. To be able to define the concept of *no-regret playing*, we first need to define the concept of *regret*.

Definition 8.26: Regret

Let α^t be the average per-period reward the agent received up to time t. Let $\alpha^t(d)$ be the average per-period reward the agent would have received by playing strategy d. The regret an agent experiences at time t for not having played d is defined as

$$R^t(d) = \alpha^t(d) - \alpha^t.$$

Once we have defined the concept of regret, we can define *no-regret playing*:

Definition 8.27: No-regret playing

A learning rule exhibits no regret if it guarantees with high probability that the agent will not experience any positive regret asymptotically, for all strategy d

$$P([\liminf_{t\to\infty} R^t(d)] \leq 0) = 1.$$

Thus, No-regret is an enviable property for any learning strategy. But how to achieve it? Is it possible at all? One example of a no-regret playing algorithm is *regret matching*.

The procedure for the regret matching algorithm is given below.

Procedure 8.28: Regret matching

Repeat:
At each time step each action is chosen with probability proportional to its positive regret:

$$\sigma_i^{t+1}(d) = \frac{R^{t,+}(d)}{\sum_{d' \in D_i} R^{t,+}(d')}$$

where $R^{t,+}(.) = \max(R^t(.), 0)$.

Similarly to fictitious play it has been shown that in zero-sum games if the two players act with no regrets then the empirical frequency of play converges to a Nash-equilibrium.

Moreover, under a stronger requirement on the asymptotic regret of each player convergence to a correlated equilibria in general games can be guaranteed.

Chapter 9

Auctions

"3... 2... 1... Meins!" — eBay.

Chapter based based on the book Game theory for applied economists *(1992) by R. Gibbons (Sec. 9.3 and 9.6) and the script* Multiagent systems: Algorithmic, Game-Theoretic, and Logical Foundations *(2009) by Y. Shoham and K. Leyton-Brown (Sec. 9.1, 9.2, 9.4 and 9.5).*

In this chapter we consider the problem of allocating resources among selfish agents by means of *auctions*. Auctions are widely used in real life. For instance by many people to trade goods on the internet, by governments to sell public resources (e.g., 4G frequencies bands), and very often also in computational settings, for instance to allocate bandwidth to users.

In this chapter, we focus on different types of simple auctions and how they provoke different kinds of behaviour among bidders. We consider the case of a *seller* auctioning one item to a set of *buyers*. The underlying assumption we make when modelling auctions is that each bidder has an intrinsic value (type) for the item being auctioned; she is willing to purchase the item for a price up to this value, but not for any higher price. We will also refer to this intrinsic value as the bidder's *true value* for the item.

9.1 Canonical Auction Families

Before digging into the theory, we start by presenting the most famous families: English, Japanese, Dutch, sealed-bid and double auctions.

English auction. The *English auction* is perhaps the best-known family of auctions, since it is used on most online consumer sites. The auctioneer sets a starting price for a good, and the agents then have the option to announce higher bids. The final bidder (the agent with the highest bid by the end of the auction) has to buy the good for the amount of his final bid.

Japanese auction. The *Japanese auction* is a variant of the English auction, where instead of the bidders the auctioneer increases the prices and each agent must declare if she is still "in" at that price or not. The Japanese auction is often easier to analyze than the English auction because the increments are well specified. In the English auction agents can place so-called *jump bids* which are much higher than the previous bid and complicate the analysis.

Dutch auction. In a *Dutch auction* it is the seller who makes iterative offers. He starts with a high enough price, and then proceeds by announcing successively lower prices. The first agent who bids wins the auction and must buy the good at the announced price. This auction is for instance used in the Amsterdam flower market.

Sealed-bid auction. So far we disused only open auctions where the agents call out their bid in public. In the family of *sealed-bid auctions*, each agent submits a secret, "sealed" bid to the auctioneer. The agent with the highest price must buy the good but the price depends on the type of the auction. In a k^{th}-price auction, the agent must pay the amount of the k^{th} highest bid.

Observe that the first-price auction and the Dutch auction are actually the *same* auction. In the first-price auction each bidder has to send a message to the auctioneer, in the Dutch auction only one bidder has to send one bit of information to the auctioneer, but the auctioneer has to broadcast several prices.

Double auction. In a *two-sided* or *double auction* there are many buyers and many sellers. A typical example is the stock market where people trade any given stock. Below we are going to study an example where two traders want to exchange one single good, for instance one single share of a company.

9.2 Definitions and Assumption

We will now formalize auctions as Bayesian games.

Recall that a *Bayesian game* is a tuple (N, X, T, p, u), where N is a finite set of n agents; X is a finite set of outcomes, $T = T_1 \times \cdots \times T_N$ a set of possible joint type vectors; p is a (common-prior) probability distribution on T; and $u = (u_1, \ldots, u_N)$, where $u_i \colon X \times \mathbb{R}^n \times T \to \mathbb{R}$ is the utility function for each player i (that also depends on the prices $p \in \mathbb{R}^n$ that the players have to pay). Here, X represents a set of nonmonetary outcomes, such as the allocation of the object to one of the bidders.

- In the following we will always consider auctions where a single good is allocated to one of the bidders, hence $X = [1, \ldots, N]$. An element of X corresponds to the winner of the auction. Moreover, we assume the following (quasilinear) structure of the utility function of each player i: $u_i(x, p, v) = v_i \cdot \delta_{x=i} - p_i$, where v_i is the value i associates to winning the auction, $\delta_{x=i} = 1$ if $x = i$. Thus, the utility of player i is $v_i - p_i$ if the item is allocated to her and she has to pay p_i, and just $-p_i$ if she has to make the payment p_i without obtaining the item.

  The utility functions imply that each agent has the same value-for-money. Consequently, we have *transferable utility.* Each agent can transfer any given amount of utility to another agent by giving her an appropriate sum of money (and the sum of the total utilities of all the players remains the same). Note further that agents with such utility functions are *risk neutral.* Suppose that an agent must decide whether she wants to participate in a fair lottery, that awards the amount x half of the time and the amount $-x$ the other half of the time. The expected utility is zero – thus the agent is indifferent between participating in the lottery or not participating. In real life the situation is often more complicated, as for instance *risk seeking* agents might prefer to participate in a lottery that cost \$ 1 and has a 0.01 % chance of paying off \$ 1000, whereas a risk neutral player is indifferent.

- The agents know the common distribution p from which the private values are drawn, but besides their own valuation they do not know the valuations of the other agents. One of the best-known and most extensively studied settings is the *independent private value* (IPV) setting. In this setting, all agents' valuations are drawn independently from the same (commonly known) distribution, i.e. $T = V \times \cdots \times V$ and $p = F \times \cdots \times F$ for a probability distribution F over V. An example where the IPV setting is appropriate is in auctions where the bidders buy a good simply for their own entertainment and the valuations depend only on their personal taste. In contrast, the IPV setting is not appropriate if the good provides some extra "utility" to the buyer (for instance by reselling on the market). Throughout this section, we will always assume the IPV setting.

An auction can be described as a mediator, like in Chapter 5. In Auctions theory however, we often talk of a mechanism. A *mechanism* (for a Bayesian game setting) is a triple $(A, \chi, \wp)$ where $A = A_1 \times \cdots, A_n$, where each A_i is the set of actions available to agent $i \in N$; $\chi \colon A \to \Delta(X)$ maps each action profile to a distribution over choices; and $\wp \colon A \to \mathbb{R}^n$ maps each action profile to a payment for each agent. The mechanism is *deterministic* if for every $a \in A$, there exists $x \in X$ such that $\chi(a)(x) = 1$; in this case we simply write $\chi(a) = x$.

An *auction* is a structured framework for negotiation. Each such negotiation has certain rules which must be specified: (*i*) bidding rules (how are offers made?); (*ii*) clearing rules (when does the trade occur or what are these trades, i.e. which player gets the good, and what are the prices for the players?); and (*iii*) information rules (who knows what about the state of negotiation?). Auctions can be formulated as Bayesian games by specifying the set of agents N, the set of outcomes X, the set of actions A_i available to each agent $i \in N$, the choice function χ, the payment function $\wp$, the utility functions u_i for each agent $i \in N$ and the common prior distribution F.

9.3 First-price and Dutch auction

In this section, we derive a Bayes-Nash equilibrium of a two agent sealed-bid first-price auction. The case of more than two agents is addressed at the end of Section 9.5.

> There are two agents (labeled $i = 1, 2$) who simultaneously submit a real valued positive bid. The higher bidder wins the good and pays the price she bid, the other bidder gets and pays nothing (in case of a tie, the winner is determined by a flip of a coin). The valuations of the agents are uniformly distributed on [0,1]. Let b_i denote the bid of player i and v_i its true valuation. If she wins, her payoff is $u_i = v_i - b_i$, if she loses, its $u_i = 0$.

Recall that a strategy is a function from player types (the valuations v_i) to actions (the bids $b_i \in [0, \infty)$). In a Bayes-Nash equilibrium, Player 1's strategy $b_1(v_1)$ is a best response to Player 2's strategy $b_2(v_2)$, and vice versa. Formally, for each $v_i \in [0,1]$, $b_i(v_i)$ solves

$$\max_{b_i} \left[(v_i - b_i) P\{b_i > b_j\} + \frac{1}{2}(v_i - b_i) P\{b_i = b_j\} \right] \tag{9.1}$$

The agents face the following dilemma when placing their bids: the higher the bid, the more likely they are to win, the lower the bid, the higher the payoff if the agent does win. We will derive the following result:

Theorem 9.1

In a sealed-bid first-price auction with two players where the valuations are drawn independently and uniformly from [0,1], $\left(\frac{1}{2}v_1, \frac{1}{2}v_2\right)$ is a Bayes-Nash equilibrium strategy profile.

Let us look at a sketch of the proof.

Proof 9.2 *As a reminder, we have two players, Player 1 and Player 2, who each have a valuation for an item that is being sold. We know that v_1 and v_2 are independent and identically distributed from a uniform distribution with parameters 0 and 1. We use b_1 and b_2 to denote the bid of Player 1 and 2, respectively. Also, we use u_1 and u_2 to denote the utility of Player 1 and 2, respectively. We have three cases*

1. *If $b_1 < b_2$, then $u_1 = 0$ and $u_2 = v_2 - b_2$;*
2. *If $b_1 > b_2$, then $u_1 = v_1 - b_1$ and $u_2 = 0$;*
3. *If $b_1 = b_2$, then $u_1 = \frac{1}{2} \cdot (v_1 - b_1)$ and $u_2 = \frac{1}{2} \cdot (v_2 - b_2)$.*

However, we know that the last case will happen with a probability zero.

First, let us show that $b_1(v_1) = \frac{1}{2}v_1$ is a best response for Player 1 to the strategy of Player 2. To be more precise, we need to check that for any fixed type v_1 for Player 1, $b_1(v_1) = \frac{1}{2}v_1$ is a best response to the strategy of Player 2, knowing that the strategy of Player 2 will be $b_2(v_2) = \frac{1}{2}v_2$. We need to compute

$$b_1^* = argmax \ (v_1 - b_1) \cdot \mathbb{P}[b_1 > B_2] + \frac{1}{2}(v_1 - b_1) \cdot \mathbb{P}[b_1 = B_2]$$

where B_2 is written with a capital letter, because it denotes a random variable. Indeed, in this part of the proof, we fix b_2 such that $b_2(v_2) = \frac{1}{2}v_2$. Since we know that $V_2 \sim Unif(0,1)$, we know that $B_2 \sim Unif\left(0, \frac{1}{2}\right)$. Now let us compute the two probabilities. We have

$$\mathbb{P}[b_1 = B_2] = 0,$$

because B_2 is a continuous random variable, and

$$\mathbb{P}[b_1 > B_2] = \mathbb{P}[B_2 \leq b_1] = F_{B_2}(b_1) = \frac{b_1 - 0}{\frac{1}{2} - 0} = 2b_1.$$

where F_{B_2} denotes the cumulative distribution function of B_2. As a reminder, if $X \sim Unif(a,b)$, then

$$F_X(x) = \begin{cases} 0 & \textit{for } x < a, \\ \dfrac{x-a}{b-a} & \textit{for } x \in [a,b), \\ 1 & \textit{for } x \geq b. \end{cases}$$

Now, let us make an important remark. At this point, we have an unknown value for b_1, hence it can be negative, it can be equal to 42, and so on. Hence, when we write that $\mathbb{P}[b_1 > B_2] = 2b_1$, we should be very careful. Indeed, by the axioms of probabilities, we know that $0 \leq \mathbb{P}[b_1 > B_2] \leq 1$. Hence, we should write

$$\mathbb{P}[b_1 > B_2] = \max\{0, \min\{1, 2b_1\}\}$$

Hence we have

$$b_1^* = \textit{argmax } (v_1 - b_1) \cdot \max\{0, \min\{1, 2b_1\}\} = \textit{argmax } f(b_1)$$

where the function f is defined like

$$f(b_1) = \begin{cases} 2b_1(v_1 - b_1) & 0 \leq b_1 \leq \frac{1}{2}, \\ 0 & b_1 < 0, \\ (v_1 - b_1) & b_1 > \frac{1}{2}. \end{cases}$$

In order to maximize f, we compute its derivative with respect to b_1. We find

$$f'(b_1) = \begin{cases} 2(v_1 - 2b_1) & 0 \leq b_1 \leq \frac{1}{2}, \\ 0 & b_1 < 0, \\ -1 & b_1 > \frac{1}{2}. \end{cases}$$

Hence we find the value of b_1^* by cancelling the derivative. We find

$$f'(b_1^*) = 0 \Leftrightarrow \begin{cases} \frac{1}{2}v_1 = b_1^* & 0 \leq b_1 \leq \frac{1}{2}, \\ 0 & b_1 < 0. \end{cases}$$

We know that $b_1 < 0$ is absurd. Hence we only consider the case $0 \leq b_1 \leq \frac{1}{2}$, and we conclude that $b_1^* = b_1(v_1) = \frac{1}{2}v_1$.

Note that one should still check the values of f at discontinuous points, which is left to the reader.

Now let us show that $b_2(v_2) = \frac{1}{2}v_2$ is a best response for Player 2 to the strategy of Player 1. For this, we simply use the symmetry of the problem. This concludes the proof.

It is not hard to check that $b_1(v_1) = \frac{1}{2}v_1$ is indeed a best response to $b_2(v_2) = \frac{1}{2}v_2$ and vice versa. However, in this exposition we are also going to *derive* the equilibrium strategies.

In the following, we are looking for a linear equilibrium, i.e. strategies of the form $b_1(v_1) = a_1 + c_1 v_1$, $b_2(v_2) = a_2 + c_2 v_2$. We are going to prove that the equilibrium we just found is the only linear equilibrium.

Suppose that player j adopts the strategy $b_j(v_j) = a_j + c_j v_j$. Since it makes no sense for player i to bid below player j's minimum bid or above player j's maximum bid, we can assume $a_j \leq b_i \leq a_j + c_j$, so

$$P\left\{b_i > a_j + c_j v_j\right\} = P\left\{v_j < \frac{b_i - a_j}{c_j}\right\} = \frac{b_i - a_j}{c_j}\,. \tag{9.2}$$

Therefore player i's best response as given in Equation (9.1) simplifies to (the event $\{b_i = b_j\}$ happens with probability zero):

$$b_i(v_i) = \arg\max_{b_i}\left[(v_i - b_i)\cdot\frac{b_i - a_j}{c_j}\right] = \begin{cases}(v_i + a_j)/2 & \text{if } v_i \geq a_j\,,\\ a_j & \text{if } v_i < a_j\,.\end{cases} \tag{9.3}$$

If $0 < a_j < 1$ then $b_i(v_i)$ is not linear (rather flat at the beginning and linearly increasing later). As we are looking for a linear equilibrium, we therefore rule out $0 < a_j < 1$ and focus on $a_j \geq 1$ and $a_j \leq 0$ instead. We can also rule out $c_j < 0$ because for a higher type $v_j' \geq v_j$ always better to bid as least as much for type v_j. But therefore also $a_j \geq 1$ cannot occur: it would imply $b_j(v_j) \geq 1 + c_j v_j \geq v_j(1 + c_j) \geq v_j$ which can clearly not be optimal (the best payoff would be zero). Thus, we must have $a_j \leq 0$, in which case $b_i(v_i) = (v_i + a_j)/2$.

By repeating the analysis for player j one equivalently deduces $b_j(v_j) = (v_j + a_i)/2$ and $a_i \leq 0$. Hence $a_i = a_j = 0$, and $b_1(v_1) = \frac{1}{2}v_1$ and $b_2(v_2) = \frac{1}{2}v_2$ as claimed earlier.

9.4 Second-price and Japanese auction

In this section, we are going to describe a Bayes-Nash equilibrium for the second-price sealed-bid auction.

> There are n agents who simultaneously submit a real valued positive bid. The highest bidder wins the good and pays the amount b_2 of the second highest bid, the other bidders get and pay nothing (in case of a tie, the winner is determined by a flip of a coin). If agent i wins, her utility is $u_i = v_i - b_2$, the utilities of all other agents are zero.

It turns out that the situation is much simpler than for the first-price auctions as there exists a dominant strategy for each player. Interestingly, the dominant strategy for each agent is to tell the truth, i.e. to reveal her private valuation. That is, the second-price auction is a *truthfull* mechanism.

Theorem 9.3

In a second-price auction truth telling is a dominant strategy.

Proof 9.4 *Assume that all bidders other than i bid in some arbitrary way, and consider i's best response. First, consider the case where i's valuation is larger than the highest of the other bidders' bids. In this case i would win and would pay the next-highest bid amount. Could i be better off by bidding dishonestly in this case? If she bids higher, she would still win and would still pay the same amount. If she bids lower, she would either still win and still pay the same amount or lose and pay zero. Since i gets nonnegative utility for receiving the good at a price less than or equal to her valuation, i cannot gain, and would sometimes lose by bidding dishonestly in this case.*

Now consider the other case, where i's valuation is less than at least one other bidder's bid. In this case i would lose and pay zero. If she bids less, she would still lose and pay zero. If she bids more, either she would still lose and pay zero or she would win and pay more than her valuation, achieving negative utility. Thus again, i cannot gain, and would sometimes lose by bidding dishonestly in this case.

We can now also identify a strong relationship between the second-price auction and the Japanese auction. In both cases the bidder must select a number (in the sealed-bid case the number is the one written down, and in the Japanese case it is the price at which the agent will drop out); the bidder with highest amount wins, and pays the amount selected by the second-highest bidder. The difference between the auctions is that information about other agents' bid amounts is disclosed in the Japanese auction. In the sealed-bid auction an agent's bid amount must be selected without knowing anything about the amounts selected by others, whereas in the Japanese auction the amount can be updated based on the prices at which lower bidders are observed to drop out. In general, this difference can be important; however, it makes no difference in the IPV case. Thus, Japanese auctions are also dominant-strategy truthful when agents have independent private values.

9.5 Revenue equivalence

The choice function implemented in the first-price auction and in the second-price auction is exactly the same: the item is awarded to the agent with the highest bid. But what about the payments? In this section we study the revenues, i.e. the payments that the auctioneer gets.

For two agents with valuations v_1 and v_2, we can readily express the revenue of the first-price auction as $\max\{v_1/2, v_2/2\}$, and the revenue of the second-price auction as $\min\{v_1, v_2\}$. Of course, each of these expressions can be higher than the other for different values of v_1 and v_2. However, if the valuations are drawn independently and uniformly from $[0, 1]$, the *expected* revenue is exactly $\frac{1}{3}$ in both cases. Thus, for the auctioneer it makes no difference if he chooses a first or second-price auction, his expected revenue remains the same. This does also hold in a more general setting.

Theorem 9.5: Revenue equivalence theorem

Assume that n agents have values $v_1, \ldots, v_n$ drawn independently from a common atomless distribution of support $[\underline{v}, \bar{v}]$ with a strictly increasing cumulative density function F. Then, all auction mechanisms which, at equilibrium, (i) always award the item to the agent with the highest value, and where (*ii*) the expected payment of bidders with valuation $\underline{v}$ is zero, generate the same expected revenue, and hence result in any bidder with valuation $v \in [\underline{v}, \bar{v}]$ making the same expected payment.

Proof 9.6 (Sketch).
Consider any mechanism for allocating the good. Let $u_i(v_i)$ be i's expected utility given true valuation v_i, assuming that all agents including i follow their equilibrium strategies. Let $P_i(v_i)$ be i's probability of being awarded the good given (i) that his true type is v_i; (ii) that he follows the equilibrium strategy for an agent with type v_i and (iii) that all other agents follow their equilibrium strategies.

$$u_i(v_i) = E\left[v_i \cdot \chi_{x=i} - p_i\right] = v_i P_i(v_i) - E\left[\textit{payment by type } v_i \textit{ of player } i\right] \tag{9.4}$$

Below will derive that $u_i(v_i)$ can be expressed as follows:

$$u_i(v_i) = u_i(\underline{v}) + \int_{x=\underline{v}}^{v_i} P_i(x)\mathrm{d}x\,. \tag{9.5}$$

Now consider any two mechanisms that satisfy the assumption of the theorem. A bidder with valuation $\underline{v}$ will never win (since the distribution F is atomless), so his expected utility $u_i(\underline{v}) = 0$. Because both mechanisms award the good to the agent with the highest valuation, every agent has the same $P_i(v_i)$ (his probability of winning given his type v_i). By Equation (9.5) each agent must therefore have the same expected utility u_i in both mechanisms. From Equation (9.4) this means that a player of any given type v_i must make the same expected payment in both mechanisms. Since this is true for all i, the auctioneer's expected revenue is also the same in both mechanism.

It remains to prove Equation (9.5). We will derive it from the incentive compatible equation. Consider an agent i with valuation v_i. Suppose that i follows the equilibrium strategy for a player with valuation $\hat{v}_i$ rather than for his true valuation v_i. Then i would make all the same payments and would win the good with the same probability as a player with valuation $\hat{v}_i$. However, whenever he wins the good, he values it $(v_i - \hat{v}_i)$ more than a player of type $\hat{v}_i$, hence his expected utility is $u_i(\hat{v}_i) + (v_i - \hat{v}_i)P_i(\hat{v}_i)$. In equilibrium, such a deviation must be unprofitable, hence

$$u_i(v_i) \geq u_i(\hat{v}_i) + (v_i - \hat{v}_i)P_i(\hat{v}_i)\,.$$

This inequality must also hold in particular for $\hat{v}_i$ of the form $\hat{v}_i = v_i + \epsilon$ or $\hat{v}_i = v_i - \epsilon$ for $\epsilon > 0$. Moreover, this inequality must also hold if the agent has type $v_i + \epsilon$ (or $v_i - \epsilon$) and announce as a type $\hat{v}_i = v_i$. This gives us two more equations. Summarizing, we have the four following equations:

$$\begin{aligned}
u_i(v_i) &\geq u_i(v_i + \epsilon) - \epsilon P_i(v_i + \epsilon)\,,\\
u_i(v_i) &\geq u_i(v_i - \epsilon) + \epsilon P_i(v_i - \epsilon)\,,\\
u_i(v_i + \epsilon) &\geq u_i(v_i) + \epsilon P_i(v_i)\,,\\
u_i(v_i - \epsilon) &\geq u_i(v_i) - \epsilon P_i(v_i)\,.
\end{aligned}$$

By combining the above four inequalities we obtain

$$\frac{u_i(v_i+\epsilon)-u_i(v_i-\epsilon)}{2\epsilon} \geq \frac{1}{2}\left(P_i(v_i-\epsilon)+P_i(v_i)\right),$$
$$\frac{1}{2}\left(P_i(v_i+\epsilon)+P_i(v_i)\right) \geq \frac{u_i(v_i+\epsilon)-u_i(v_i-\epsilon)}{2\epsilon},$$

and for $(\epsilon \to 0)$ *the derivative* $\frac{du_i}{dv_i} = P_i(v_i)$. *Integrating, we get* (9.5) *as claimed.*

Theorem 9.5 tells us that an auctioneer cannot increase the revenue of an auction without changing the allocation rule itself. If he is willing to modify the choice function he can certainly increase the revenue. For instance, consider two bidders with valuations distributed independently and uniformly in [0,1], and an auctioneer who puts a *reservation price* $x \in [0,1]$ and sells the item for a price that is the maximum of the second highest bid and x. If both agents bid below x, then none of them wins. A short calculation reveals that this mechanism is incentive compatible and that the expected revenue $r(x)$ of this auction is $r(x) = \frac{1}{3} + x^2 - \frac{4}{3}x^3$. This expression is maximized for $x = \frac{1}{2}$ and $r(\frac{1}{2}) = \frac{5}{12} > r(0) = \frac{1}{3}$, the expected revenue of the auction with no reservation price.

The revenue equivalence theorem can also be used to predict the bidding strategies of the players. Consider again two agents whose valuations are drawn independently and uniformly at random from [0,1]. Let $r_{1\text{st}}(v)$ and $r_{2\text{dn}}(v)$ denote the expected payment of an agent with type v in a first-price auction or second-price auction, respectively, where all players follow their equilibrium strategies. By the revenue equivalence theorem (both auctions award the item to the agent with highest valuation in equilibrium and an agent with valuation zero has to pay nothing), both payments must be equal: $r_{1\text{st}}(v) = r_{2\text{nd}}(v)$. In a second-price auction, the expected payment of an agent with valuation v can easily be computed: his probability of winning is exactly v, and the expected payment conditioned that he wins the auction is just $\frac{v}{2}$, hence $r_{2\text{nd}}(v) = \frac{v^2}{2}$. In a first-price auction the probability of winning remains exactly the same, however, in case of winning the player has to pay the value $b(v)$ of the bid he made. Hence it must hold $b(v) \cdot v = \frac{v^2}{2}$, and we recover the bidding strategy $b(v) = \frac{v}{2}$ for an agent in the first-price auction. This gives also an interpretation of the bidding strategy in the first-price auction: each player bids the expectation of the second-highest valuation, conditioned on the assumption that his own valuation is the highest. The same technique also allows to derive the bidding strategies in the first-price auctions when more than two agents are participating.

9.6 Double auctions

In the last section of this chapter we analyze a trading game called a double auction.

> Two players, a buyer and a seller, both have a private valuation of a good. The seller names a price p_s and the buyer simultaneously offers a price p_b. If he asks a price $p_b \geq p_s$, the trade will be executed at price $p = (p_b + p_s)/2$, otherwise no trade occurs. The private valuations v_b and v_s are drawn independently and uniformly at random from [0,1]. If the trade occurs at price p, the utility of the buyer is $v_b - p$, the utility of the seller $p - v_s$. If no trade occurs, the utility of both players is equal to zero.

In this game, a strategy for the buyer is a function $p_b(v_b)$ specifying the price that the buyer will offer depending on his valuation. Likewise, a strategy for the seller is a function $p_s(v_s)$. A pair of strategies $\{p_b, p_s\}$ is a Bayes-Nash equilibrium if the following two conditions hold. For each $v_b \in [0,1]$,

$$p_b(v_b) = \arg\max_{p_b} \left[v_b - \frac{p_b + E\left[p_s(v_s) \mid p_b \geq p_s(v_s)\right]}{2} \right] P\left\{p_b \geq p_s(v_s)\right\} , \tag{9.6}$$

where $E\left[p_s(v_s) \mid p_b \geq p_s(v_s)\right]$ is the expected price the seller will demand, conditional on its demand being less than the buyer's offer p_b. For each $v_s \in [0,1]$,

$$p_s(v_s) = \arg\max_{p_s} \left[\frac{p_s + E\left[p_b(v_b) \mid p_b(v_b) \geq p_s\right]}{2} - v_s \right] P\left\{p_b(v_b) \geq p_s\right\} , \tag{9.7}$$

where $E\left[p_b(v_b) \mid p_b(v_b) \geq p_s\right]$ is the expected price the buyer will offer, conditioned on its offer being greater than the seller's demand p_s.

There are many Bayes-Nash equilibria in this game. Two families of equilibria are described in the following theorem.

Theorem 9.7

1. For every $x \in [0,1]$, the strategies

$$p_b(v_b) = \begin{cases} x, & \text{if } v_b \geq x \\ 0, & \text{otherwise} \end{cases} \qquad\qquad p_s(v_s) = \begin{cases} x, & \text{if } v_s \leq x \\ 1, & \text{otherwise} \end{cases}$$

describe a Bayes-Nash equilibrium of the double auction.

2. The strategies

$$p_b(v_b) = \frac{2}{3}v_b + \frac{1}{12} \qquad\qquad p_s(v_s) = \frac{2}{3}v_s + \frac{1}{4} ,$$

describe a linear Bayes-Nash equilibrium of the double auction.

Proof 9.8 *It is not hard to see that the first two strategies are best responses to each other. Given the seller's strategy, the buyer's choices amount to trading at x or not trading at all. So the seller's strategy is a best response to the buyer's, and vice versa.*

Let us now consider the second pair of strategies. Suppose the seller's strategy is $p_s(v_s) = \frac{2}{3}v_s + \frac{1}{4}$. Then p_s is uniformly distributed on $[\frac{1}{4}, \frac{11}{12}]$. Clearly, it makes no sense for the bidder to bid more than $\frac{11}{12}$, hence it must hold $p_b(v_b) \leq \frac{11}{12}$ for all $v_b \in [0,1]$. In the case $p_b < \frac{1}{4}$, the bidder always looses, hence it makes only sense to bid strictly less than $\frac{1}{4}$ if also his valuation $v_b < \frac{1}{4}$. We will come back to this observation shortly.

Assume now that $p_b \in [\frac{1}{4}, \frac{11}{12}]$, then the probability to win for the bidder becomes $P\left\{p_b \geq p_s(v_s)\right\} = \frac{3}{2}\left(p_b - \frac{1}{4}\right) = \frac{3}{2}p_b - \frac{3}{8}$; the expected price the seller will demand if there is a trade equals $E\left[p_s(v_s) \mid p_b \geq p_s(v_s)\right] = \frac{1}{2}\left(\frac{1}{4} + p_b\right)$; hence (9.6) becomes

$$\max_{p_b} \left[v_b - \frac{3}{4}p_b - \frac{1}{16} \right] \left(\frac{3}{2}p_b - \frac{3}{8} \right) = \max_{p_b} \left[-\frac{9}{8}p_b^2 + \frac{3}{2}p_b v_b + \frac{6}{32}p_b - \frac{3}{8}v_b + \frac{3}{128} \right] .$$

The first-order condition $-\frac{9}{4}p_b + \frac{3}{2}v_b + \frac{6}{32} = 0$ *yields* $p_b(v_b) = \frac{2}{3}v_b + \frac{1}{12}$ *as best response. Note that we assumed* $p_b \in [\frac{1}{4}, \frac{11}{12}]$, *which implies that* $p_b(v_b) = \frac{2}{3}v_b + \frac{1}{12}$ *is a best response for all* $v_b \in [\frac{1}{4}, 1]$. *We already noted that in case* $v_b < \frac{1}{4}$, *the actual bid of the bidder does not matter as long as it stays below* $\frac{1}{4}$. *Hence, there are* arbitrary *many strategies for the bidder how to play if* $v_b \in [0, \frac{1}{4})$. *The* linear *strategy given in the theorem is just* one *of the many many possible choices for the bidder.*

Analogously, suppose that the buyer's strategy is $p_b(v_b) = \frac{2}{3}v_b + \frac{1}{12}$. *Then* p_b *is uniformly distributed on* $[\frac{1}{12}, \frac{3}{4}]$. *Clearly, the seller should always ask for at least* $p_s \geq \frac{1}{12}$, *and in case he asks for more than* $\frac{3}{4}$ *the buyers will never agree. Hence, if* $v_s > \frac{3}{4}$ *any bid strictly above* $\frac{3}{4}$ *does not alter the outcome of the auction. The linear strategy suggested by the theorem is just one of the many possible solutions.*

Assume now that $p_s \in [\frac{1}{12}, \frac{3}{4}]$. *Then the probability to win for the bidder becomes* $P\left\{p_b(v_b) \geq p_s\right\} = 1 - \frac{3}{2}\left(p_s - \frac{1}{12}\right) = \frac{9}{8} - \frac{3}{2}p_s$ *and the expected price the buyer will offer if there is a trade equals* $E\left[p_b(v_b) \mid p_b(v_b) \geq v_s\right] = \frac{1}{2}\left(p_s + \frac{3}{4}\right)$, *hence* (9.7) *becomes*

$$\max_{p_s}\left[\frac{3}{4}p_s + \frac{3}{16} - v_s\right]\left(\frac{9}{8} - \frac{3}{2}p_s\right) = \max_{p_s}\left[-\frac{9}{8}p_s^2 + \frac{3}{2}p_s v_s + \frac{9}{16}p_s - \frac{9}{8}v_s + \frac{27}{128}\right],$$

the first order condition $-\frac{9}{4}p_s + \frac{3}{2}v_s + \frac{9}{16} \overset{!}{=} 0$ *for which yields* $p_s(v_s) = \frac{2}{3}v_s + \frac{1}{4}$. *This finishes the proof.*

In the first family of Bayes-Nash equilibria, the trade occurs only for the pairs $(v_b, v_s) \in [x, 1] \times [0, x]$, see Figure 9.1. However, the trade would be efficient (beneficial for both players) for all pairs such that $v_b \geq v_s$. For the two linear strategies, the trade occurs only if $v_b \geq v_s + \frac{1}{4}$, see Figure 9.1. In both cases, the most valuable possible trade (namely $v_s = 0$, $v_b = 1$) does occur. But the one-price equilibria miss some valuable trades (such as $v_s = 0$ and $v_b = x - \epsilon$, where $\epsilon > 0$ is small) and achieves some trades that are worth to almost nothing (such as $v_s = x - \epsilon$, and $v_b = x + \epsilon$). The linear equilibrium, in contrast, misses all trades that are worth to nothing but achieves all trades worth at least $\frac{1}{4}$. This suggests that the linear equilibrium may dominate the fixed-price equilibria, in terms of the expected gains the player receive, but also raises the possibility that the players might do even better in an alternative equilibrium.

Myerson and Satterthwaite (1983) show that, for the uniform value distributions considered here, the linear equilibrium yields highest expected gains for the players than any other Bayes-Nash equilibrium of the double auction. This implies that there is no Bayes-Nash equilibrium of the double auction in which trade occurs if and only if it is efficient. Indeed, the expected gains for the players in this situation would be higher than the expected gains for the linear equilibrium, which is not possible.

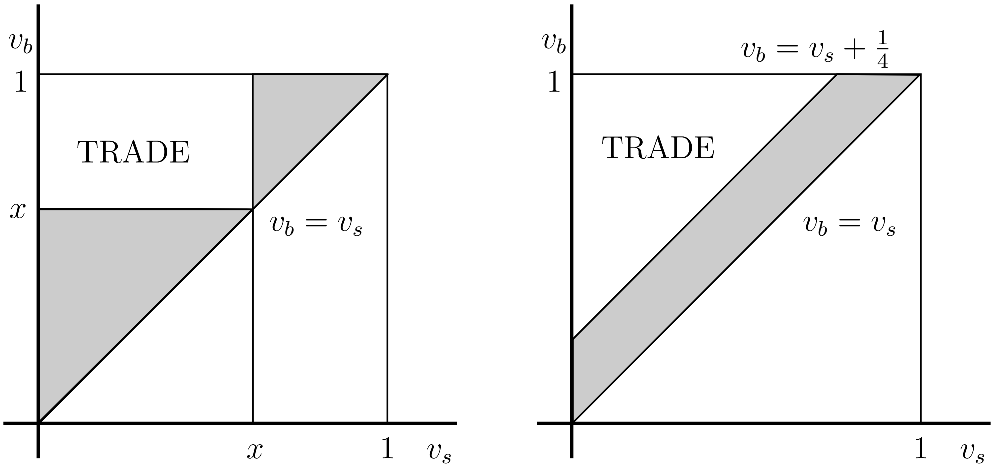


Figure 9.1: Comparison of the executed trades for the two equilibrium strategies of Theorem 9.7.

Chapter 10

Evolutionary Game Theory

"If you have an idea that was going to outrage society, would you keep it to yourself?" — Charles Darwin.

This chapter is inspired from Nowak (2006). Figures are taken from this book.

10.1 Introduction

In the seventies, people like William Hamilton and Robert Trivers proposed to apply ideas from mathematical ecology, or more generally life sciences, to Game Theory. This has led to an important subfield of Game Theory (and quite disconnected from the rest of it) called Evolutionnary Game Theory. Other main actors of the developments in Evolutionary Game Theory are John Maynard Smith Smith and Price (1973) (a geneticist and biologist from the UK) and Karl Sigmund (a mathematician from Vienna).

Evolutionary Game Theory analyses special types of games where *a large number of identical agents* are playing. The payoff of a player, here called *fitness*, (which depends on the strategies of the agent, and of the other agents playing, like in classical Game Theory) determines the *reproduction rates* of agents. One is interested in the different proportions of agents playing different strategies asymptotically. These proportions are to be interpreted as an Equilibrium, in the sense of Game Theory. Moreover, Evolutionary Game Theory does not assume any rationality of the agents, and these agents interact randomly, reproducing themselves (or disappearing) dependently of their types and the types of their fellow agents.

This idea is directly borrowed from classical concepts in Ecology. Let us see an example: on Figure 10.1 are represented two species; one of which having the ability to move, while the other one is static. When there are not many agents in the pool, the agent with the ability to move takes a great advantage of it, and its fitness is high. However, when many agents are present, the agents are bound to be static, whatever their initial capacities are, and the mobile agent looses its comparative advantage. It thus makes sense to model the fitness (the rate of reproduction) of the agents as depending on the relative 'concentrations' of the different agents.

Let us put formulas on this model, for two competing strategies A and B. Let x_A be the frequency of A and x_B be the frequency of B.The vector $\overrightarrow{x} = (x_A, x_B)$ defines the state of our dynamical system. In general, let the function $f_A(\overrightarrow{x})$ be the fitness of A

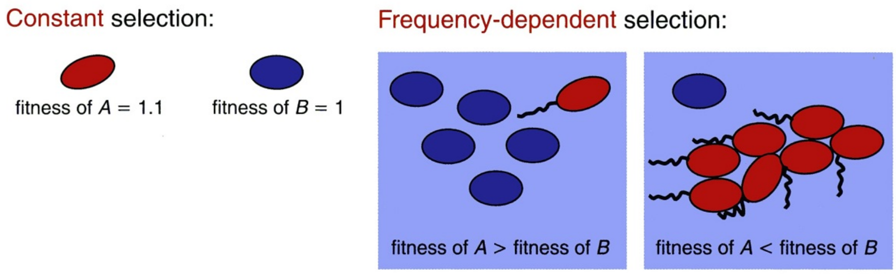


Figure 10.1: Constant selection VS Frequency-dependent selection

and $f_B(\overrightarrow{x})$ the fitness of B. We note $\phi = x_A f_A(\overrightarrow{x}) + x_B f_B(\overrightarrow{x})$ be the average fitness. The laws of dynamics selection then write:

$$\begin{cases} \dot{x}_A & = x_A \left[f_A(\overrightarrow{x}) - \phi \right] \\ \dot{x}_B & = x_B \left[f_B(\overrightarrow{x}) - \phi \right] \end{cases} \tag{10.1}$$

In what follows, we will only be interested in fitness functions depending on the *relative* concentrations of each species. Thus, we suppose that $x_A + x_B = 1$, and we can rewrite our dynamical system as a one-dimensional system with state-space variable x such that : $x_A = x$ et $x_B = 1 - x$. We have:

$$\dot{x} = x(1-x) \left[f_A(x) - f_B(x) \right] \tag{10.2}$$

The different equilibria of (10.2) are

- $x = 0$,
- $x = 1$ and
- all the values $x \in (0, 1)$ such that $f_A(x) = f_B(x)$.

The equilibrium $x = 0$ is stable if $f_A(0) < f_B(0)$ and the equilibrium $x = 1$ is stable if $f_A(1) > f_B(1)$. Any other equilibrium $x^\star$ is stable if the functions f_A et f_B satisfy $f'_A(x^\star) < f'_B(x^\star)$ (the functions are supposed to be twice differentiable). In general, there can be many equilibria in the interval $[0, 1]$. However, we will now restrict the shape of the functions $f_A(x)$, as they will be defined according to the game we are studying.

10.2 Model and equilibria analysis

How can we get inspiration from this biological example in order to introduce a solution concept for two-player games? We will consider our game from a new point of view: in evolutionary game theory, there is not a pair of players, playing one against each other, but rather, a large number of identical players, each of which having to pick a strategy in a common finite set. This choice is their 'DNA', and the impact of their choice is reflected in their ´reproduction rate', the probability that they will be able to reproduce themselve, given their DNA, and given the ecological system defined by the other players' choice. Just like in classical game theory, the choice of other players will impact the 'payoff' (here, the reproduction rate), of each player. Hence, the payoff matrix can only be symmetric, since it does not represent several different players, but the payoff that each of them will receive, when facing one of their fellow agents. As a consequence, in Evolutionary Game Theory, one represents the payoff matrix in the following simplified way:

$$\begin{array}{c|cc} & A & B \\ \hline A & a & b \\ B & c & d \end{array} \Longleftrightarrow \begin{array}{c|cc} & A & B \\ \hline A & a,a & b,c \\ B & c,b & d,d \end{array}. \tag{10.3}$$

The main idea of Evolutionary Game Theory is to consider the payoff as a fitness, in a dynamical system where agents reproduce proportionnally to their fitness. Knowing x_A, the frequency of agent A, and x_B, the frequency of agent B, the expected fitness of respectively, agents of type A and agents of type B will be:

$$\begin{cases} f_A = ax_A + bx_B \\ f_B = cx_A + dx_B \end{cases}. \tag{10.4}$$

These equations constitute a crucial point in our modeling procedure, where we interpret our payoff matrix as an *infinite population* game. Indeed, we made here the assumption that the expected fitness of an agent can be reinterpreted as being the fitness of all the agents in the population. This is a classical assumption in modeling, also sometimes called 'fluid model'. If we insert these equations (10.4) in our dynamical system (10.2), we obtain :

$$\dot{x} = x(1-x)\left[(a-b-c+d)x + b - d\right]. \tag{10.5}$$

Thus, we are now facing a particular case of Equation (10.2), where the dynamics are described by four parameters a, b, c, d. It turns out that in this case, only five situations are possible.

Theorem 10.1

In an evolutionary dynamics model derived from a two-player game as in Equation (10.5), one of the five following situations occur:

1. A dominates B if $a > c$ et $b > d$, and the whole population will turn into A-individuals.

2. B dominates A if $a < c$ et $b < d$, and the whole population will turn into B-individuals.

3. A et B are bistables if $a > c$ et $b < d$: depending of the initial proportion of A and B-individuals, the whole population will turn into individuals of the same type. The treshold between these two antagonistic situations is an unstable equilibrium, given by $x^* = (d-b)/(a-d-b+c)$.

4. A et B coexist at equilibrium. This happens if $a < c$ et $b > d$. In this case there is only one stable equilibrium, given by $x^* = (d-b)/(a-d-b+c)$.

5. A et B are neutral if $a = c$ et $b = d$. in this situation the population do not evolve in time, whatever the initial conditions are.

Proof 10.2 *Left for exercise.*

10.3 Evolutionary Stable Strategy (ESS)

This concept, Introduced by John Maynard Smith, aims at translating to Game Theory the concept of robustness from Systems and Control. More precisely, we consider now a large population of A-agents, and we will say that strategy A is *Evolutionary Stable* if it is robust against mutation of a small fraction of its population into B-individuals. In this section,, we are only interested in characterizing pure strategies.

10.3.1 Two-Player games

We begin with a Two-Player game for simplicity:

Definition 10.3

Consider the game (10.3). A strategy A is *Evolutionnary Stable* (in short, ESS), if a population constituted initially of only A-individuals is stable under the mutation of an ϵ fraction of this population to another type of individual.

Mathematically, let us consider a population composed of invading individuals B, so that $x_B = \epsilon$, in an initial population of A-agents, so that $x_A = 1 - \epsilon$. In this situation, the fitness of A is greater than the fitness of B if:

- $$\forall \epsilon > 0,\ a(1-\epsilon) + b\epsilon > c(1-\epsilon) + d\epsilon \tag{10.6}$$
  and thus:
  $$a > c \tag{10.7}$$
- or, if $a = c$, we have
  $$b > d. \tag{10.8}$$

We can thus conclude:

Proposition 10.4

A is an Evolutionary Stable Strategy if one of the following conditions is satisfied:

1. $a > c$
2. $a = c$ et $b > d$.

10.3.2 More than two strategies

Let us generalize our analysis to games with more than two strategies. In this case, the payoff for Strategy S_i against Strategy S_j is given by: $E(S_i, S_k)\ \forall i \neq k$. We have:

1. Strategy S_k is a strict nash equilibrium if:

$$E(S_k, S_k) > E(S_i, S_k) \quad \forall i \neq k$$

2. Strategy S_k is a Nash equilibrium if :

$$E(S_k, S_k) \geq E(S_i, S_k)\ \forall i$$

3. Strategy S_k is an ESS if $\forall i \neq k$ one has either :

$$E(S_k, S_k) > E(S_i, S_k)$$

or

$$E(S_k, S_k) = E(S_i, S_k) \text{ and } E(S_k, S_i) > E(S_i, S_i)$$

4. Strategy S_k is stable against invasion (or 'weak ESS') if $\forall i \neq k$ one has either:

$$E(S_k, S_k) > E(S_i, S_k)$$

or

$$E(S_k, S_k) = E(S_i, S_k) \text{ and } E(S_k, S_i) \geq E(S_i, S_i)$$

.

Point 4 represents the case where A will not be eliminated by B, but, on the other hand, A won't manage to eliminate a small fraction of mutant B agents. In this case both species will finally co-exist asymptotically.

Summarizing, it is easy to see the following implications between our different pure-strategies equilibria:

$$\text{strict Nash} \quad \Rightarrow \quad \text{ESS} \quad \Rightarrow \quad \text{weak ESS} \quad \Rightarrow \quad \text{Nash.} \tag{10.9}$$

10.4 Replicator Dynamics

Just like in game theory, we would like to mathematically represent the fact that sometimes a 'mix' of strategies is a good representation of the actual outcome of the game. To this purpose, we are going to analyse the System (10.1) in the whole state space, and not only in situations where the complete population (except an ϵ proportion) is composed of a single agent. Generalizing to a n-Player game, the expected payoff of stragegy i is given by $f_i = \sum_{j=1}^{n} x_j a_{ij}$. The average payoff among all the strategies is given by $\phi = \sum_{i=1}^{n} x_i f_i$. Then, we obtain the so-called *Replicator equation* (See Figure 10.2):

$$\dot{x}_i = x_i(f_i - \phi) \quad i = 1, \ldots, n. \tag{10.10}$$

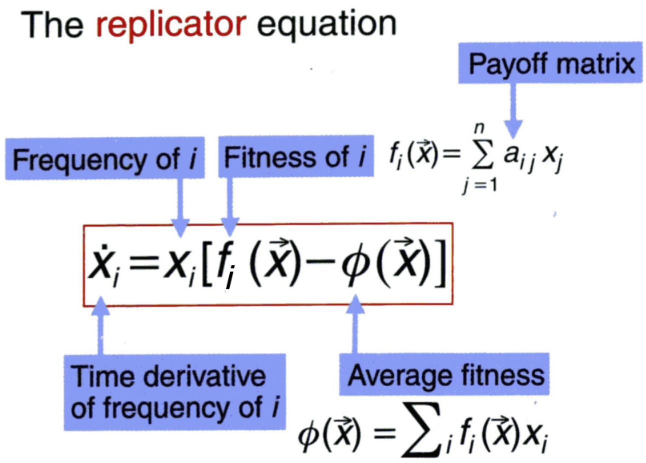


Figure 10.2: The replicator equation

Equation (10.10) is defined on the simplex S_n (i.e. $\{x : \sum_{i=1}^{n} x_i = 1\}$). It is easy to see that by design, the interior of the simplex is invariant: a trajectory can converge towards the boundary of the simplex (representing the fact that a population is dying), but it will never exactly hit the boundary. Let us now study how trajectories can behave. For low dimensional systems, it is possible to completely characterize the behavior of all possible trajectories, depending on the values of the parameters.

10.4.1 Two-strategies games

The 2-Player case has been studied above: one can have fixed points in the interior of the simplex only if $(a_{11} - a_{21})(a_{12} - a_{22}) < 0$. In that case, no population dominates the other one. This equilibrium is stable if $a_{11} < a_{12}$ and $a_{21} > a_{22}$. Obviously, in the neutral case ($a_{11} = a_{12}$ and $a_{21} = a_{22}$), all the points in the interval $[0, 1]$ are equilibrium points.

10.4.2 3-Player games

If $n = 3$, we have the following payoff matrix:

$$\begin{pmatrix} 0 & -a_2 & b_3 \\ b_1 & 0 & -a_3 \\ -a_1 & b_2 & 0 \end{pmatrix}$$

(Note that this matrix was obtained by subtracting from every column the corresponding diagonal element. This will not change the equilibria, since the Replicator Equation (10.10) is also conserved. One has thus the following possible cases:

- $a_1a_2a_3 < b_1b_2b_3$
- $a_1a_2a_3 > b_1b_2b_3$.

In the first case, there exists a unique interior equilibrium which is globally stable. In the second case, there exists a unique interior equilibrium which is unstable.

10.4.3 Games with several strategies

When $n > 3$, one obtains the equilibria by solving these equations:

$$f_1 = f_2 = ... = f_n$$

$$x_1 + x_2 + ... + x_n = 1.$$

This is a simple n-dimensional linear system. Thus, there can at most be one single isolated equilibrium.

10.4.4 Lotka-Volterra Equations

Predator-Prey models have been developed by Vito Volterra, an italian physicist, just after WW2. His goal was to understand a sudden rise in the concentration of sharks in the Adriatic sea following the war. His equations represented populations of preys x, and predators, y. The preys reproduce at rate a and are eaten at a rate b, predators dying at a rate c and reproducing at rate d.

$$\dot{x} = x(a - by)$$

$$\dot{y} = y(-c + dx)$$

First, if there are no predators, or no prey, the other populations evolve according to increasing (resp. decreasing) curves.

$$x(t) = x(0)e^{at}$$
$$y(t) = y(0)e^{-ct}$$

In the general case, we have the following interior stable equilibrium.

$$x^* = c/d$$

$$y^* = a/b$$

Now, if fishing decreases, the concentration of preys increases, which indirectly causes a rise in concentration of sharks.

Generalizing these equations to more than two species, one obtains the famous *Lotka-Volterra equations:*

$$\dot{y}_i = y_i(r_i + \sum_{j=1}^{n} b_{ij}y_j).$$

It is easy to show that these equations are actually perfectly equivalent to the *replicator equation* (10.10): Evolutionnary Game Theory is nothing more than Mathematical Ecology applied to Game Theory.

10.4.5 An example: Hawk or Dove

This paradigmatic game in Game Theory comes from Ecology: scientists wanted to understand and explain violence and domination in nature. Violence can indeed be beneficial for an individual, if it allows him to enhance reproduction of its own genes, compared with other species. Of course, it can also be harmful when two such individuals meet. We thus consider two species: hawks, and doves. Each species can chose (through long term mutations) to fight everytime they meet another bird, of to flee. In the following payoff matrix, b represents the payoff in case of victory, and c represents the cost of injuries after loosing a fight. Thus, the expected payoff when two hawks meet is $(b-c)/2$. If a hawk meets a dove, the hawk wins, and thus receives a payoff of b, while the dove doesn't engage, and hence leaves with a zero payoff. We thus have the following payoff matrix:

$$\begin{pmatrix} \frac{b-c}{2} & b \\ 0 & \frac{b}{2} \end{pmatrix}$$

If $c > b$, there is no pure-strategy Nash equilibrium. If every individual choses to play Hawk, playing dove is the best strategy, and vice versa. Hence, there is a mixed population equilibrium, with a rate of hawks-individuals equal to b/c.

10.5 Finite population

It is possible to rewrite all the conditions above, by supposing that the (large number of) identical agents is finite, instead of infinite. Mathematically, one must then model the birth-death process in a stochastic way (as in so-called *Moran* processes). By making use of similar notions of robustness to mutation-perturbations, one can define new notions of equilibria for a given (symmetric) payoff table. Interestingly, this finite population assumption can lead to significantly different numerical values of equilibria.

References

Myerson, R. B. (1991). *Game theory: Analysis of conflict.* Harvard university press.

Nowak, M. A. (2006). *Evolutionary Dynamics: Exploring the Equations of Life.* Harvard University Press.

Osborne, M. J. and Rubinstein, A. (1994). *A course in game theory.* MIT press.

Rubinstein, A. (1998). *Modeling bounded rationality.* MIT press.

Shoham, Y. and Leyton-Brown, K. (2008). *Multiagent systems: Algorithmic, game-theoretic, and logical foundations.* Cambridge University Press.

Smith, J. M. and Price, G. (1973). The logic of animal conflict. *Nature*, 246:15.