Game Theory

PROF. *Leonardo Badia*, *Elvina Gindullina*
UNIVERSITY OF PADOVA

WRITTEN BY: *Francesco Manzali*

ACADEMIC YEAR 2020/21

Compiled on January 1, 2021

Contents

Foreword

If you find any mistake/typo/missing thing, please write at *francesco.manzali@studenti.unipd.it* (or even just for feedback).

Francesco Manzali, 31/12/2020

Acknowledgments

Introduction

0.1 Introduction

(Lesson 1 of 28/09/2020) Compiled: January 1, 2021

Game theory is the field that studies how to model **interactions** between **agents**. In this framework, all players are supposed to follow some **fixed rules** and are interested in certain outcomes, which however depend not only on individual actions, but also on the choices of other agents. Game theory seeks the optimal decision plan in such an environment, taking into account the interactions with the other players: this is why it is also sometimes called *Interaction Decision Theory*. This is quite different from other mathematical models of agents, such as in *Classical Decision Theory*, where problems are **single-person**, i.e. such that success solely depends on the individual's ability, and every uncertainty derives either from stochastic variables or approximations.

Historically, the first formal theorem in the theory of games was proved by E. Zermelo in 1913, and originally involved the game of Chess. One of the core books was "Theory of Games and of economic behavior", published in 1944 by von Neumann and Morgensten, and developed the framework of **zero-sum** games, where the interests of different players are strictly opposite (the advantage of one is the disadvantage of the other, like in chess). Then, some other fundamental results were derived by John Nash, concluding the basics of the theory.

The first application of game theory was in economical settings, in all levels ranging from micro (trading), intermediate (markets) and macro (countries and monetary systems). Nowadays, game theoretical models are applied in various fields, such as psychology (peoples are interacting agents), social sciences (laws are like rules of a game), political sciences (choices of parties), biology (competition of species), computational intelligence and network systems (distributed systems, multi-agent algorithms).

0.2 Decision problems

The first kind of problems we will consider are **decision problems**, where we want to choose the best action to achieve a certain goal.

To formalize this, we consider three elements:

- **Actions** (A)
- **Outcomes** of each action
- **Preferences**: a ranking of outcomes, describing which are preferable than the others

For example, suppose the goal is to arrive at university. To do this, we could go by foot, by bike or by bus. Each of these three actions (belonging to the set A) has consequences, namely:

- going by foot is free, but takes more time, risking arriving late at the lecture;
- going by bus is faster, but costs money (for the ticket);
- going by bike is sufficiently fast and it is free.

Here we prefer arriving early and not spending money. With this formulation, it is clear that in this case the best choice is taking the bike.

Formally, if A is a set of alternatives, with $|A| \geq 2$ (otherwise there is no choice to be made), we define a **preference** to be a binary relationship $\succcurlyeq$ on A. More precisely:

- If $a, b \in A$ are two elements of A, writing $a \succcurlyeq b$ means that a is ranked **above** b
- The relation $\succcurlyeq$ is reflexive ($a \succcurlyeq a$) and antisymmetric ($a \succcurlyeq b \Leftrightarrow b \preccurlyeq a$).

A preference is said to be **complete** if it ranks **all** elements of A, i.e. $\forall a, b \in A$, then either $a \succcurlyeq b$ or $b \succcurlyeq a$ (or both).
It is **transitive** if, for every three actions $\forall a, b, c \in A$, if $a \succcurlyeq b$ and $b \succcurlyeq c$ then $a \succcurlyeq c$. This property implies that there can be no *loops* in the ranking, and that there is a clear *best* choice (or a set of choices that are equivalent and better than all the others) in A.
If $\succcurlyeq$ is both **complete** and **transitive**, then it is a *total order relation*, and we say that the player adopting such preferences is **rational**.

Utility functions

Utility functions can be used to map inputs q (e.g. action or goods) to a certain payoff ("measure of goodness") $u(q) \in \mathbb{R}$. Note that, in this way, a rational preference ranking $\succcurlyeq$ on q is mapped in the order relation of real numbers $\leq$. In other words, if we prefer q over q', then $u(q) > u(q')$.
If $q \in \mathbb{R}$ indicates a countable good and $u(q)$ is differentiable, then usually $u(q)$ is an increasing function (having more goods is better) and mathematically $u'(q) \geq 0$ and $u''(q) \leq 0$.
Since $\leq$ is a total order relation on $\mathbb{R}$, using utility functions can work only for rational players: it is needed to have a complete transitive ranking of actions to be able to have utility functions! Formally, we say that u **represents** $\succcurlyeq$ if:

$$a \succcurlyeq b \Leftrightarrow u(a) \geq u(b) \qquad \forall a, b \in A$$

And then, we can state that the preferences that can be represented are only the ones of rational players:

Theorem 0.2.1. *On a finite set A, a preference $\succcurlyeq$ can be represented by a utility function u if and only if it is rational.*

Proof. $\Rightarrow$: If $\succcurlyeq$ can be represented, then it must be rational, since $\geq$ defines a total order relation on $\mathbb{R}$, and $a \succcurlyeq b \Leftrightarrow u(a) \geq u(b) \quad \forall a, b \in A$ implies that $\succcurlyeq$ must be a total order relation also on A, and thus rational. Conversely ($\Leftarrow$), given a rational preference, a suitable utility function can be defined as:

$$u(a) = |\{b \in A \colon a \succcurlyeq b\}|$$

In other words, the utility of $a \in A$ is the number of other elements $b \in A$ that are less preferable than a. Intuitively, this simply means to put all the elements of a in an ordered sequence $(a, b, c, \dots)$ using $\succcurlyeq$ (which is possible since by hypothesis $\succcurlyeq$ is rational) and then index them with the natural numbers: $a \to 0$, $b \to 1$, $c \to 2$. The indexing function u that maps each element to its position in the ordered list is the required utility function.

□

□

Rationality is a very strong requirement: in practice it means that all players act for their own good only (otherwise they could take "bad" outcomes out of altruism), and are aware of all consequences of their acts (if not, they could make things worse by error), and also of all the possible actions that can be taken.
It can be argued that assuming rationality is foolish: obviously humans are not rational, they can act crazy, make mistakes or simply be generous. While the point effectively becomes irrelevant in all applications that do not involve humans (e.g. algorithms, autonomous agents, etc.), the real actual problem is, at the end, the **accuracy** of a model. So, we can pragmatically assume rationality to prove results, and then tweak the models so that they correctly account for not rational behaviors (for example by inserting more parameters in the utility function) and are closer to reality.

One possible way to graphically represent decision problems is through **decision trees**. Basically, we start with the player at the **root** of the tree. Each possible decision is represented as a different **branch** that can be taken, which leads to an outcome, represented by its **utility** (on the **leaves**).

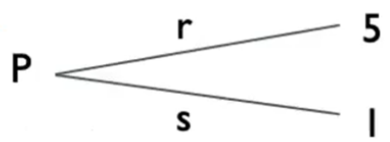

Figure (1) – Example of a decision tree. Here the goal is to have a nice dinner. For the first dish, the player P can choose between ravioli (r) or spaghetti (s), with respectively the utilities of $u(r) = 5$ and $u(1) = 1$ (they prefer ravioli).

Successive decisions form deeper *layers* in the decision tree:

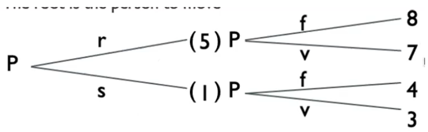

Figure (2) – Here we add another possible choice: fish (f) or vegetables (v)

But at any time we could always *group* each possible path in only one level, by simply listing every combination:

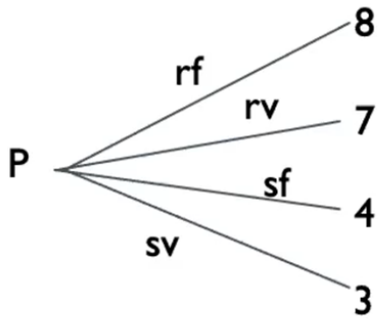

Figure (3) – Consecutive choices can always be collapsed in tree with a single layer.

In this simple setup, it is easy to choose the best possible choice. Here, however, we are dealing only with one player: things get much more interesting when we add multiple agents!

0.2.1 More players

(Lesson 2 of 01/10/2020) Compiled: January 1, 2021

⚠ *This section is just a draft!*

Let's add more players to the game. In general, each player does not know the **strategy** being used by other ones, and so must be able to make decisions under uncertainty.
A way to account for that is through the framework of **lotteries**, a generalization of decision problems where outcomes are stochastic in nature. For example, consider an investor who can choose how much money to invest in a certain market, and, depending on the ... probabilities
Let's formalize this by denoting with $X = \{x_1, x_2, \dots, x_n\}$ the finite set of outcomes. Then we define P to be the set of all **probability distributions** $p\colon X \to [0, 1]$ on X, which are normalized:

$$\sum_{x \in X} p(x) = 1$$

We call P the set of **lotteries**.

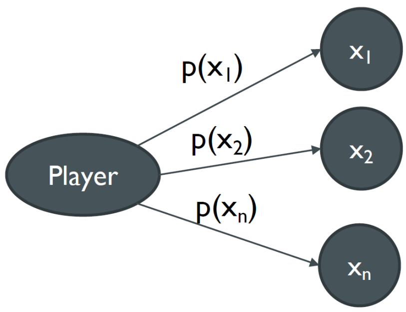

Figure (4)

Probabilities are **conditional** on actions $a \in A$, since outcomes are influenced on the actions that a player takes. Note that lotteries are a *generalization* of deterministic decision problems, which can be in fact modelled as **degenerate lotteries**, since here $p(x_i|a) = 1$ for a given i, and 0 for all the others, i.e. there is no uncertainty.

In the language of game theory, all randomness is represented as the consequences of the choices of another player, called **nature**.

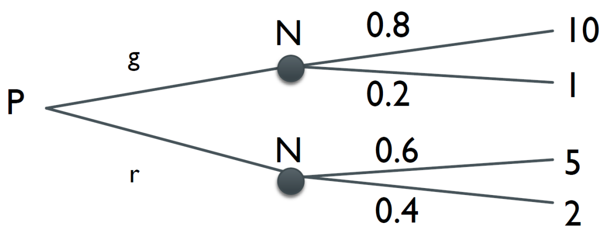

Figure (5) – Consider an investor who can choose to invest in two different fields g or r. The final outcome depends on external factors that are **not** under their control, and which can be represented as "nature's choices", which are hidden to the player.

Lotteries can be extended to a **continuous** space of events, by simply replacing discrete probabilities with cumulative distribution functions.
Suppose, for example, that choices are represented by real numbers, e.g. an amount of money to be invested. Then $X = [x_0, x_1]$ is a continuous interval and the lottery is a CDF $F: X \to [0,1]$, where $F(\hat{x}) = \mathbb{P}[x \leq \hat{x}]$.

How can we define **rationality** in such a framework? Consider, for example, the following case:

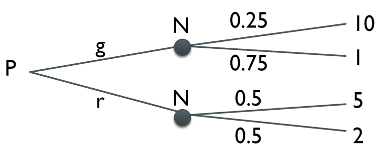

Figure (6)

Here, taking r could lead to the highest reward, but also to the lowest one (risk is high). On the other hand, g leads to more "moderate" outcomes, but also a lower maximum gain.

One simple way to decide which branch is best is to take **expectation values**. This was first proposed in the "Expected utility theory" developed by von Neumann and Morgenstern (1944). We define the expected payoff from lottery p over $X = \{x_1, x_2, \dots, x_n\}$ such that $p_i = \mathbb{P}[x = x_i]$ as follows:

$$\mathbb{E}[u(x)|p] = \sum_{x \in X} p(x)u(x)$$

So the idea is to choose the actions that maximize the expected outcome. Formally, we extend the correspondence between preference relation $\succcurlyeq$ and utility functions to the stochastic case by inserting expected values. Specifically, we say that $\succcurlyeq$ on P is said to be **represented** by a von Neumann-Morgenstern utility function $u\colon X \to \mathbb{R}$ if and only if:

$$p \succcurlyeq q \Leftrightarrow u(p) \equiv \sum_{x \in X} u(x)p(x) \geq \sum_{x \in X} u(x)q(x) \equiv u(q) \qquad \forall p, q \in P$$

Again, this can only be done if $\succcurlyeq$ is both complete and transitive (it has no loops) — and this condition is the **axiom of rationality**.
However, we need two more hypotheses: **independence**. Basically, two gambles mixed with an irrelevant third one will maintain the same order of preference as when the two are presented independently of the third one. Mathematically:

$$\forall p, q, r \in P(A),\ a \in (0, 1]$$

Nature can have more than one subsequent choice. In this case, all the above machinery can be naturally extended by computing *compound expectations*.

Example 1 (Continous investment)**:**

Suppose we can decide the amount of investment, i.e. $a \ni A = [0, 50]$. The expected outcomes are $X = [0, 100]$, and depend on our choice:

$$x|a \sim \mathcal{U}[0, 2a]$$

In other words, once we decide a, the final outcome will be some value between 0 and $2a$, uniformly extracted.

We can then compute the expected utility:

$$\nu(a) = \mathbb{E}[u(x)|a] = \frac{1}{2a} \int_0^{2a} u(x)\, \mathrm{d}x$$

Let's fix $u(x) = 18\sqrt{x}$. We substitute in the above, and select the value of a that maximizes the result:

$$\text{Choose a:} \max_{a \in [0,50]} \frac{1}{2a} \int_0^{2a} 18\sqrt{x}\, \mathrm{d}x - 2a$$

This can be done with some calculus, by computing the first derivative and setting it to 0:

$$12\sqrt{2a} - 2a \overset{!}{=} 0$$

And we find $a = 18$, with the expected payoff being $\nu(18) = 36$.

Note an important difference with the deterministic case: here the **absolute** value of utilities matters! First, only the order was important, but since we are now dealing with probabilities, we need also to understand *how much* an outcome is preferred than another.
In other words, changing the utility function (but maintaining the same preference ranking) does not change the best course of action in the deterministic case, but it does so in the stochastic one (in fact u enters explicitly the computations)! So we distinguish between **ordinal preferences**, where only order matters, and **cardinal ranking**, where both order and absolute value matter.

Different lotteries can have the same expected utility. For example, consider $X = (0, 1, 20)$, and $p_A = (0, 1, 0)$, $p_B = (0.95, 0, 0.05)$. The expected utility is 1 in both cases. However, A is a degenerate lottery: we are *certain* we will get 1 every time. B, instead, has a very high probability of giving no payoff ($p = .95$), and a small one of giving a high payoff (20 with $p = .05$).
A player could consider A and B equivalent, basing their choice only on the expected outcomes. This is the **risk neutral** case.
If a player prefers not taking risks, they will be called **risk adverse**, and prefer A over B, since A is deterministic. Otherwise, a **risk loving** player would do the opposite.

The choice of utility function can *encode* the **risk attitude** of a player. In essence, all monotonic utilities (such as $u(x) = x, x^2, \log x$), do not change the order of preference of a user, but change their risk attitude. In fact, a linear utility, evaluates *equally* both high outcomes with low probability and low outcomes with high probability. However, if u is concave, then
So, a player facing a decision problem with a **payoff function** u over outcomes is said to be **rational** if he chooses an action $a \in A$ that maximizes his expected payoff, i.e. chooses $a^* \in A$ such that:

$$\nu(a^*) = \mathbb{E}[u(x)|a^*] \geq \mathbb{E}[u(x)|a] = \nu(a) \quad \forall a \in A$$

Actions of player and nature may **alternate** over time. In this case, to decide which action is best we proceed by **Backward Induction** (or **Dynamic Programming**). Consider, for example, the following situation:

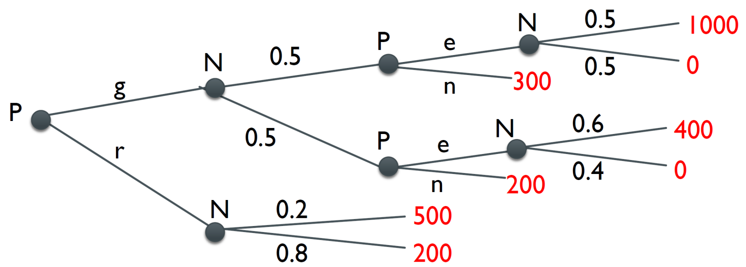

Figure (7) – A player can choose between two investments g and r. If they choose g, then, after a move from nature, he can choose again between two other actions e and n. Then nature plays again.

The idea is to classify all player's actions in groups:

- Group 1 contains all "final" actions, after which no other input from the player is needed.
- Group k contains all actions that are followed by group $k-1$ actions. So, for example, group 2 actions are the ones that precede the *final* actions.

In other words, start from the *end* of the tree, and work backwards.
In the example in fig. 7 we have just two groups:

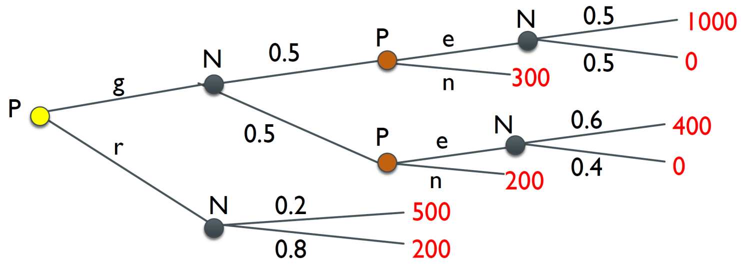

Figure (8) – Actions in group 1 are in **orange**, while group 2 are in **yellow**.

Then we start with group 1, compute the expected utilities and select the best choices:

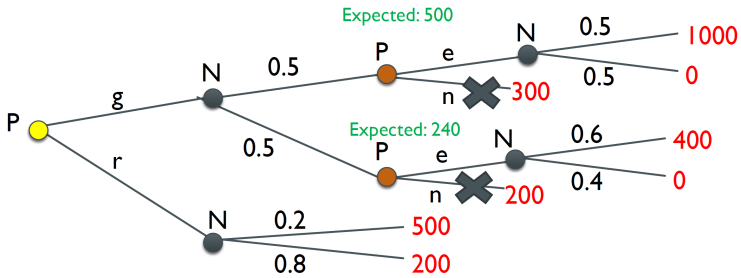

Figure (9) – Here in both cases we should select action e, since it leads to the highest expected utility.

Then we *prune* the tree, removing the groups already processed:

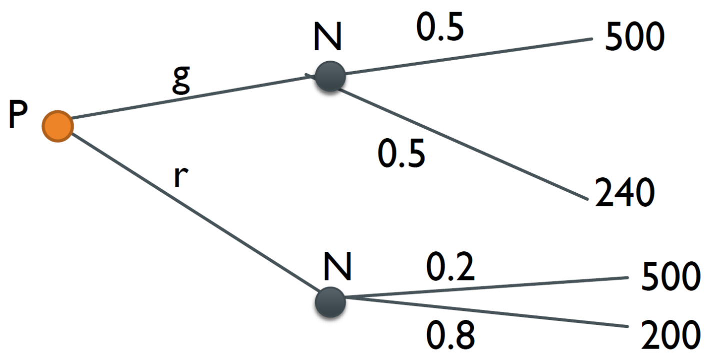

Figure (10) – Only group 2 actions are left.

Now we have a simple tree which can be finally solved:

$$\mathbb{E}[u|g] = 370 \qquad \mathbb{E}[u|r] = 260$$

So the best (rational) course of action is to first choose g and then e.

Since we are dealing with *future* decisions, we could also consider the fact that future outcomes are "discounted" (because the value of money lowers over time). In this case we multiply utilities by a discount factor $\delta < 1$:

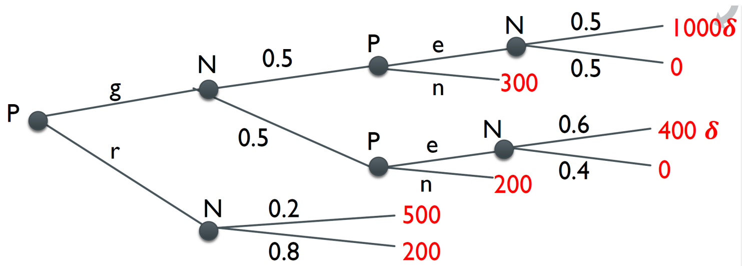

Figure (11)

The exact value of δ depends on the *greediness* of the player: a low δ means that *immediate gains*, i.e. high outcomes that are closer to the present, are valued more than the same outcomes that are further away in the future. A high δ, instead, indicates a *patient* player that is interested in passing over some higher immediate gains to come out on top at the end ("plays the long game").

In any case, knowing in advance which choice the **nature** player will take matters a lot. With decision trees we can actually *quantify* how much this information is valued. Basically, we compare the scenarios were nature's choice is known with the ones where it's not.
For example, consider the following case:

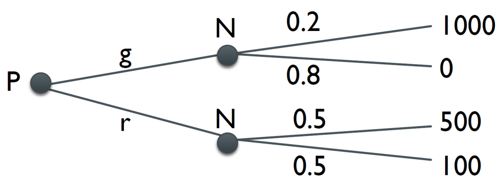

Figure (12)

If nature's choice is not known, the expected payoff for the best action (r) would be 300.

However, if we know how nature will play, we can plan the actions in advance, basically playing "in response" to nature:

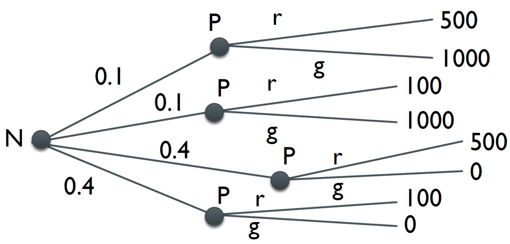

Figure (13)

Once nature's choice is made, the best action of the player is determined with certainty. We can then compute the expected utility, which comes out to 440 in this case.

The **difference** between the two cases is the **value** of information about nature's choice:

$$440 - 300 = 140$$

Chapter 1

Static Games of Complete Information

1.1 Definitions

(Lesson 3 of 05/10/20) Compiled: January 1, 2021

We are now ready to introduce *more players* to the game. There are several ways to accomplish this:

- Treat other players as **random variables**, with the same role nature had in the previous section. In this view, other players' actions only introduce more *randomness* in the system.
- Model other players as **agents**, with their own payoff function.

The second way is *more efficient*, since it allows for *strategizing*. If we assume that other players are **rational**, then they will try to maximize their own payoffs. Thus, by getting insight in their decision process, we can better **anticipate** their actions, which is surely advantageous.

As a start, we consider **static** games of **complete** information:

- **Static** means that all players *move together*, without taking turns. Moves need not happen simultaneously, but players must act **without knowledge** of everybody else's move (and so without communication). For instance, *rock/paper/scissors* is a static game, while chess is not.
- **Complete** information means that everybody knows all payoff functions. This means that *objectives are not hidden*, such as in the game of *Risk*. One example of a complete information game is *chess*, since both players compete to checkmate the other.

Examples of static games of complete information are *guessing games* (rock/paper/scissors, matching pennies).

The **static** requirement means that players act **independently**, each choosing an action from its own set A_i. Only after all n players have acted, choosing a vector of actions $(a_1, a_2, \dots, a_n)$, the outcome is determined, and players get their payoff: $u_i(a_1, \dots, a_n)$.

In this framework, it is useful to think in terms of **strategies**, which are just *plans of action*. A strategy is a mapping between certain *requirements* and an *action*. *Strategies*

For example, one possible strategy in rock/paper/scissors is to change the sign if the opponent has guessed right in the previous turn. A strategy can involve randomness (**mixed strategy**): if two or more actions are possible given the same requirement, one could be preferred (i.e. taken with a higher probability) than the others.

However, for now we limit ourselves to **pure strategies**, which are defined to be **deterministic**.

Then, we reformulate the problem by replacing actions with strategies. During each game, all players simultaneously choose their **strategy** $s_i \in S_i$, which can be gathered in a vector $(s_1, s_2, \dots, s_n)$. Then, each player i gets a payoff $u_i(s_1, \dots, s_n) \in \mathbb{R}$.

This notation allows specifying completely the **game**. In fact, the setting is entirely defined by specifying the set of strategies available to each player (S_i) and their utility function u_i. We denote this collection as the game's *normal form*:

$$G = \{S_1, \dots, S_n; u_1, \dots, u_n\}$$

We say that E is **common knowledge** if: *Common knowledge*

- All players know E
- All players know that all players know E
- And so on, *ad infinitum*.

This is *obvious* in technical settings, for example different machines sharing exactly the same protocol. However, in human groups, saying that something is common knowledge also implies that *everybody* is considering it in their decisions, which could not be the case (people can forget, or have different priorities).

We can now formalize **complete information** as saying that all that follows is *common knowledge*:

- all possible actions of all players
- all outcomes of these actions
- the individual preferences of all players about these outcomes

We will assume that also player's **rationality** is *common knowledge*.

An n player game can be represented as a function $f\colon S_1 \times S_2 \times \cdots \times S_n \to \mathbb{R}^n$, mapping a set of strategies $(s_1, \dots, s_n)$ to a vector of utilities $(u_1, \dots, u_n) \in \mathbb{R}^n$. If all S_i are discrete sets, with sizes $|S_1|, \dots, |S_n|$, then f can be represented as a $|S_1| \times \cdots \times |S_n|$ matrix, where each cell represents a choice of the n strategies, and contains a vector in $\mathbb{R}^n$ of the utilities. Note that this is an n-dimensional

matrix, since each player's strategies are contained in a *direction* orthogonal to the others.
In this course, we will focus on the $n = 2$ case (bi-matrices), since working with bi-dimensional matrices is easier.

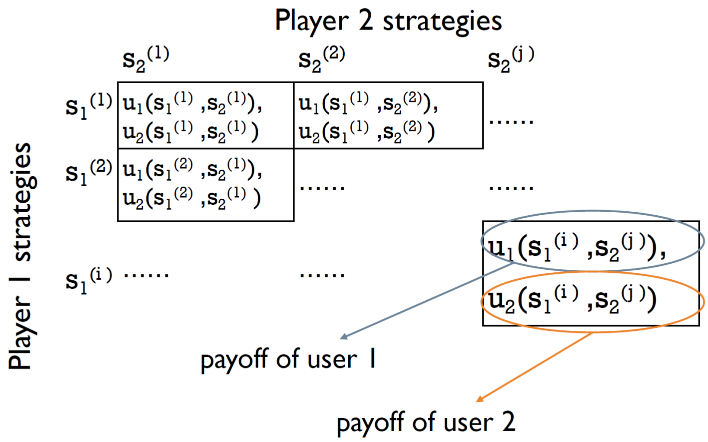

Figure (1.1) – Example of matrix representation of an $n = 2$ players game. Each cell contains the payoffs of both players.

player A		player B: L	player B: R
	U	8, 0	0, 5
	M	1, 0	4, 3
	D	0, 7	2, 0

Figure (1.2) – Example of utilities in a 2-player game. For example, if player A chooses the strategy M, and B chooses C, then player A will get 0, and B will get 5 as payoffs.

1.2 Examples

One explicit example is from the odd/even game, where two players pick a preferred outcome (odd/even) and a number. Then they both show simultaneously their numbers, and the player that correctly predicted the parity of the sum of the two numbers will win.

Odd		Even: 0	Even: 1
	0	-4, 4	4, -4
	1	4, -4	-4, 4

Figure (1.3) – Bi-matrix for the odd/even game, with a bet of 4 euros.

Another example is rock/paper/scissors:

		player B		
		R	P	S
player A	R	0, 0	-4, 4	4, -4
	P	4, -4	0, 0	-4, 4
	S	-4, 4	4, -4	0, 0

Figure (1.4) – Bi-matrix for the rock/paper/scissors game, with a bet of 4 euros. If both players choose the same action (diagonal) nobody wins, otherwise rock R beats scissors S, scissors S beat paper P and paper P beats rock R.

In all of these examples there is no deterministic best strategy: a rational player needs to use instead a *mixed* strategy, involving randomness.
Perhaps one of the most studied 2 players game of complete information is the **Prisoners' Dilemma**.
Alice and Bob have committed a crime together, and have been caught by the police, but the evidence against them is insufficient. They are given a choice to confess or not.
If both *do not confess*, they will both spend 1 month in jail (since the full charges could not be confirmed). If one *confesses* and the other not, then the first will be free (as thanks for their collaboration), while the second will spend 9 months in prison. Instead, if both confess, they will both be sentenced for 6 months in jail.

		Bob	
		M	F
Al	M	-1, -1	-9, 0
	F	0, -9	-6, -6

Figure (1.5) – Bi-matrix for the prisoners' dilemma, with M = not confess, and F = confess.

Assuming that both players act *independently*, without any way to communicate, and they both act *rationally*, i.e. trying to maximize their own utility function, which is the best course of action to take?

To tackle this problem, we first define a way to compare *joint strategies*, i.e. vectors $\boldsymbol{s}$ of all the players' strategies.
We say that $\boldsymbol{s}$ is **Pareto dominated** by $\boldsymbol{s'}$, if the latter *rises* the utility of some player i while *not lowering* the utilities of all other players:

$$u_i(\boldsymbol{s'}) \geq u_i(\boldsymbol{s}) \qquad \forall \text{ players}$$
$$u_i(\boldsymbol{s'}) > u_i(\boldsymbol{s}) \qquad \text{for some player } i$$

A joint strategy $\boldsymbol{s}$ that is not Pareto dominated by any other joint strategy is said to be **Pareto efficient**: it is "the best" possible outcome. Clearly there could be many distinct Pareto efficient strategies, each favoring a different player.

In the case of fig. 1.5, (F, F) is the Pareto dominated by (M, M), since the latter *rises* the utility of both players. Then, note that neither (F, M) nor (M, F) are Pareto dominated by (M, M) or (F, F), since in all cases the utility of a player *lowers*. For example, when going from (M, F) to (M, M), Alice rises by 8, but Bob lowers by 1.
Thus, in this case we have 3 Pareto efficient joint strategies: (M, M), (M, F) and (F, M). The first is "the best for all", while the others are the "best favoring a specific player".

1.3 Solving Games

1.3.1 IESDS: an intuitive algorithm

However, players can choose only their strategy, not the joint one! So, we reframe the same argument in terms of *single strategies*. Formally, consider a game $G = \{S_1, \dots, S_n; u_1, \dots, u_n\}$. We say that a strategy $s_i \in S_i$ for player i is **strictly dominated** by another strategy $s_i' \in S_i$ if the latter leads to a greater payoff independently of the actions taken by the other players:

$$u_i(s_1, \dots, s_i', \dots, s_n) > u_i(s_1, \dots, s_i, \dots, s_n) \qquad \forall(s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n) \in S_i \times \cdots \times S_{i-1} \times S_{i+1} \times \cdots$$

A rational player will **never** choose a strategy that it is strictly dominated by another, since it will perform worse.

player A \ player B	L	R
U	6, 0	0, 5
M	1, 0	4, 3
D	0, 7	2, 0

Figure (1.6) – Consider player A (first number), and compare the strategies D and M. For M, the utilities of player A are $\boldsymbol{u}_M = (1, 4)$, while for D are $\boldsymbol{u}_D = (0, 2)$. Note that the entries of $\boldsymbol{u}_A$ are strictly greater than the ones of $\boldsymbol{u}_D$, meaning that M strictly dominates D. So we can remove D from the table: a rational player won't ever choose such a bad strategy!

player A \ player B	L	R
U	6, 0	0, 5
M	1, 0	4, 3
D	0, 7	2, 0

Figure (1.7) – Similarly, we can compare strategies available to player B (after having removed D). We have $\boldsymbol{u}_L = (0, 0)$ and $\boldsymbol{u}_R = (5, 3)$. All entries of $\boldsymbol{u}_R$ are greater than the ones in $\boldsymbol{u}_L$, and so R strictly dominates L. Thus, B will choose strategy R (the only one left). Now, for player A strategy M dominates strategy U ($4 > 0$), and so they will choose M. So, if A and B are rational, they will play (M, R), with result $(4, 3)$.

This procedure is called *iterated elimination of strictly dominated strategies* (IESDS), and can be used to shrink a game to a "smaller" one by relying on common knowledge (in this case, that *both players are rational*). However, this will lead to a solution only in the simplest cases.

If we apply this reasoning to prisoner's dilemma, then paradoxically (F, F) will be the rational choice. In fact:

		Bob: M	Bob: F
Al	M	-1, -1	-9, 0
	F	0, -9	-6, -6

Figure (1.8) – Alice compares $\boldsymbol{u}_F = (0, -6)$ with $\boldsymbol{u}_M = (-1, -9)$, and notices that F strictly dominates M. Bob makes the same argument, and so the rational choice will be (F, F), resulting in $(6, 6)$.

This means that **rational players** choosing the strategy that *strictly dominates* the others, will produce a joint strategy that is not Pareto efficient, i.e. it is not "the best, most efficient one"!

However, note that IESDS is very limited. For example, in the following no strategy strictly dominates the others:

		player B: L	player B: C	player B: R
player A	U	0, 5	4, 0	7, 3
	M	4, 0	0, 5	7, 3
	D	3, 7	3, 7	9, 9

Figure (1.9) – There are no dominated strategy that can be eliminated from the table.

Game I

10, 10	3,15
15,3	5,5

Game II

8, 8	2, 7
7, 2	0, 0

Figure (1.10) – Which game to play? The first one has payoffs that are always better than the ones in the second. However, rational players in game 1 will always obtain $(5, 5)$, while in game 2 will always reach $(8, 8)$, which is better.

1.3.2 Formalization

(Lesson 4 of 08/10/20) Compiled: January 1, 2021

A game in normal form is given by:

$$\mathcal{G} = (M, (S_i)_{i\in M}, (u_i)_{i\in M})$$

where M is the set of players, S_i is the set of strategies available to player i, and u_i is the utility function of player i.

Then, a strategy profile s is the set of strategies chosen by the players, so $s \in S \equiv S_1 \times \cdots \times S_M \equiv \times_{i\in M} S_i$. Each utility function u_i maps a common strategy s to the payoff received by player i: $u_i \colon \times_{i\in M} S_i \to \mathbb{R}$.
For a two player game, where each player has access to 2 strategies, we can gather all the payoffs in the entries of a 2×2 matrix:

$$\mathbf{P} = \begin{array}{cc} & \text{Player 2} \\ & \begin{array}{cc} L & R \end{array} \\ \text{Player 1} \begin{array}{c} T \\ B \end{array} & \left\| \begin{array}{cc} a, b & c, d \\ e, f & g, h \end{array} \right\| \end{array}$$

This kind of simple matrices can be constructed only for **toy games**, and are useful to *understand* the basic principles that can be then applied to more complex cases.

A **solution Φ** of a class Γ of games $\mathcal{G}$ is a way to map Γ into a set of strategies: $\Phi \colon \Gamma \rightrightarrows \cup_{G\in\Gamma} S_G$. Then the solution of a specific game G is a subset of this general *class* solutions: $\Phi(\mathcal{G}) \subseteq S_G$. Note that this is not a function, since it maps input values not in a single point, but also in *sets of points*.

There are several ways to choose Φ, for example Nash equilibrium, or strict dominance (as seen by the *iterative elimination* method in the previous section). Solutions may not exist for all games, and if a solution exists it may not be unique.

Assumptions

In the following, we will assume that all players are **rational**, they *understand* all the elements of the game, and these two facts are **common knowledge** (everybody knows them, and everybody knows that everybody knows them).

Strict dominance

Recall that given two strategies $s_i, s_i' \in S_i$, we say that s_i **strictly dominates** s_i', and write $s_i \succ s_i'$, if, independently of the strategies $s_{-i} \equiv s \setminus \{s_i\}$ actuated by *all other players*, the utility function u_i for the player i will be better if they choose s_i instead of s_i':

$$u_i(s_i, s_{-i}) > u_i(s_i', s_{-i}) \qquad \forall s_{-i} \in S_{-i}$$

The notation s_{-i} with the *negative index* indicates the vector obtained by removing from s the i-th element. In this case, s is the vector of all players' strategies, s_i is the strategy of player i, and s_{-i} is the vector $(s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_M)$ of all strategies of players that are not i. Similarly, we denote with S_{-i} the set $S_1 \times \cdots \times S_{i-1} \times S_{i+1} \times \cdots \times S_M$.

Note that here we are focusing on the strategies of player i, basically ignoring the actions of other player. However, as we have seen in the previous section, not all games admit strictly dominant strategies. For example:

$$\mathbf{P} = \begin{array}{cc|cc|} & & \multicolumn{2}{c}{\text{Player 2}} \\ & & L & R \\ \multirow{2}{*}{\rotatebox{90}{Player 1}} & T & 2,1 & 0,0 \\ & B & 0,0 & 1,2 \end{array}$$

Best Response

In cases where there are no strictly dominating strategies, we need a **weaker** condition for the solution. One idea is that, *given* the choice of one player, it is clear *what* is the best strategy for the other. For example, if player 2 plays L (*left*), then player 1 should choose T (*top*). Conversely, if player 2 plays R (*right*), then player 1 should **reply** with B (*bottom*).

Reply

Let's formalize this notion. For player i, the strategy $s_i \in S_i$ is a **best reply** *against* the strategy $s_{-i} \in S_{-i}$ of the other players, if it leads to an equal or better utility than any other available strategy s_i':

$$u_i(s_i, s_{-i}) \geq u_i(s_i', s_{-i}) \qquad \forall s_i' \in S_i$$

To find s_i we just need to maximize u_i:

$$u_i(s_i, s_{-i}) \in \arg\max_{s_i \in S_i} u_i(s_i, s_{-i})$$

Now, note that a **dominated** strategy $s_i \in S_i$ is **never** a best reply.
Proof. Suppose that $s_i \in S_i$ is dominated, i.e. there exists a $s_i' \in S_i$ such that:

Best replies and strict domination

$$u_i(s_i', s_{-i}) > u_i(s_i, s_{-i}) \qquad \forall s_{-i} \in S_{-i}$$

This means that:

$$\nexists s_{-i} \in S_{-i} \text{ s.t. } u_i(s_i, s_{-i}) \geq u_i(s_i', s_{-i})$$

and so s_i can never be a best reply.

Nash Equilibrium

We can use the idea of **best replies** to find a solution for games where strictly dominance fails, leading to the concept of Nash Equilibrium, also known as Cournot equilibrium (who used the same result previously, unbeknownst to Nash, who however better formalized it).

A strategy profile s^* is a Nash Equilibrium if s_i^* is a best response to s_{-i}^* $\forall i \in M$, i.e. if every player has chosen a *best response* against the others. Thus:

$$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i', s_{-i}^*) \qquad \forall s_i' \in S_i$$

In a Nash Equilibrium, players have no incentive to *change* their decisions, since doing so would inevitably *decrease* their utilities.

For example, consider the following 2-player game:

$$\mathbf{P} = \begin{array}{cc} & \text{Player 2} \\ & \begin{array}{ccc} L & C & R \end{array} \\ \text{Player 1} \; \begin{array}{c} T \\ M \\ B \end{array} & \begin{Vmatrix} 4,3 & 5,1 & 6,2 \\ 2,1 & 8,4 & 3,6 \\ 3,0 & 9,6 & 2,8 \end{Vmatrix} \end{array}$$

Here $S_1 = \{T, M, B\}$ and $S_2 = \{L, C, R\}$. The Nash Equilibrium s^* is a strategy profile, i.e. $s \in S_1 \times S_2$, such that each player *does not want* to change their strategy.
To find it, we proceed as follows.

Finding a ***Nash Equilibrium***

Let's start with the first strategy available to player 2, which is L. In this case, the best response for player 1 is T. Similarly, if 2 plays C, then 1 responds with B, and if 2 plays R, 1 chooses T. In other words, for each column in the matrix, pick the *maximum* first entry (i.e. maximum payoff for player 1).

$$\mathbf{P} = \begin{array}{cc} & \text{Player 2} \\ & \begin{array}{ccc} L & C & R \end{array} \\ \text{Player 1} \; \begin{array}{c} T \\ M \\ B \end{array} & \begin{Vmatrix} 4,3 & 5,1 & 6,2 \\ 2,1 & 8,4 & 3,6 \\ 3,0 & 9,6 & 2,8 \end{Vmatrix} \end{array}$$

Then we do the same from the point of view of player 2. Practically, look at each row (which is a choice for 1) and select the entry with the maximum value of the second number (i.e. maximum payoff for player 2, meaning that this is the best reply).

$$\mathbf{P} = \begin{array}{cc} & \text{Player 2} \\ & \begin{array}{ccc} L & C & R \end{array} \\ \text{Player 1} \; \begin{array}{c} T \\ M \\ B \end{array} & \begin{Vmatrix} \boxed{4,3} & 5,1 & 6,2 \\ 2,1 & 8,4 & 3,6 \\ 3,0 & 9,6 & 2,8 \end{Vmatrix} \end{array}$$

The entries for which both players are choosing a best reply (i.e. the ones with *both* payoffs colored) correspond by definition to the game's Nash Equilibria. In this case, there is only one such entry: $s^* = \{T, L\}$. Note that *properly* the equilibrium is a set of strategies, not payoff, so $s^* \neq (4, 3)$!

Weak Dominance

However, Nash Equilibrium is not the only way to extend *strict dominance*. Another *natural* possibility is the following.

Let $s_i, s'_i \in S_i$. We say that $s_i \succcurlyeq s'_i$ (s_i *weakly dominates* s'_i) if choosing s_i does not decrease the utility u_i *independently* of all other players' choices:

$$u_i(s_i, s_{-i}) \geq u_i(s'_i, s_{-i}) \qquad \forall s_{-i} \in S_{-i}$$

and it *strictly improves* it for at least some other players' actions:

$$u_i(s_i, s_{-i}) > u(s'_i, s_{-i}) \qquad \text{for some } s_{-i} \in S_{-i}$$

For example:

Weak dominance ***elimination*** *method*

$$\mathbf{P} = \begin{array}{cc} & \begin{array}{cc} \multicolumn{2}{c}{\text{Player 2}} \\ L & R \end{array} \\ \text{Player 1} \begin{array}{c} T \\ M \\ B \end{array} & \left\| \begin{array}{cc} 2,3 & 1,3 \\ 1,0 & 0,1 \\ 0,1 & 1,0 \end{array} \right\| \end{array}$$

Note that M is strictly dominated by T, and so we can eliminate it:

$$\mathbf{P} = \begin{array}{cc} & \begin{array}{cc} \multicolumn{2}{c}{\text{Player 2}} \\ L & R \end{array} \\ \text{Player 1} \begin{array}{c} T \\ \cancel{M} \\ B \end{array} & \left\| \begin{array}{cc} 2,3 & 1,3 \\ \cancel{1,0} & \cancel{0,1} \\ 0,1 & 1,0 \end{array} \right\| \end{array}$$

If we now examine the strategies of player 2, we can see that L does not *strictly dominates* R, since in the first row they have both equal payoffs (3) for player 2. However, since L is *strictly* better than R if player 1 chooses B, we can say that it is **weakly dominant** over R:

$$\mathbf{P} = \begin{array}{cc} & \begin{array}{cc} \multicolumn{2}{c}{\text{Player 2}} \\ L & \cancel{R} \end{array} \\ \text{Player 1} \begin{array}{c} T \\ \cancel{M} \\ B \end{array} & \left\| \begin{array}{cc} 2,3 & \cancel{1,3} \\ \cancel{1,0} & \cancel{0,1} \\ 0,1 & \cancel{1,0} \end{array} \right\| \end{array}$$

Finally, we can see that T dominates over B, and so the final outcome will be $(2,3)$, given by $s = \{T, L\}$.

However, the result of this procedure unfortunately *depends* on the *order of operations*. In fact, let's consider again the same matrix, but this time we eliminate first the row B, since it is strictly dominated by T:

$$\mathbf{P} = \begin{array}{cc} & \begin{array}{cc} \multicolumn{2}{c}{\text{Player 2}} \\ L & R \end{array} \\ \text{Player 1} \begin{array}{c} T \\ M \\ \cancel{B} \end{array} & \left\| \begin{array}{cc} 2,3 & 1,3 \\ 1,0 & 0,1 \\ \cancel{0,1} & \cancel{1,0} \end{array} \right\| \end{array}$$

Now R *weakly dominates* L for player 2:

$$\mathbf{P} = \begin{array}{cc} & \text{Player 2} \\ & \begin{array}{cc} \cancel{L} & R \end{array} \\ \text{Player 1}\ \begin{array}{c} T \\ M \\ \cancel{B} \end{array} & \left\| \begin{array}{cc} \cancel{2,3} & 1,3 \\ \cancel{1,0} & 0,1 \\ \cancel{0,1} & \cancel{1,0} \end{array} \right\| \end{array}$$

Since T dominates over M, we get as final result $(1,3)$ given by $s = \{T, R\}$, which is not the same as before!

Strictly dominance is *not affected* by order of elimination, but *weak dominance* is. In this case, in fact, we have even *another* possible solution, that can be obtained if we note at the start that both M and B are strictly dominated by T:

$$\mathbf{P} = \begin{array}{cc} & \text{Player 2} \\ & \begin{array}{cc} L & R \end{array} \\ \text{Player 1}\ \begin{array}{c} T \\ \cancel{M} \\ \cancel{B} \end{array} & \left\| \begin{array}{cc} 2,3 & 1,3 \\ \cancel{1,0} & \cancel{0,1} \\ \cancel{0,1} & \cancel{1,0} \end{array} \right\| \end{array}$$

Here the "winning" strategy profiles are **two**: both $\{T, L\}$ and $\{T, R\}$.

Let's compare these results with the Nash Equilibrium of the same matrix (if it exists).

$$\mathbf{P} = \begin{array}{cc} & \text{Player 2} \\ & \begin{array}{cc} L & R \end{array} \\ \text{Player 1}\ \begin{array}{c} T \\ M \\ B \end{array} & \left\| \begin{array}{cc} \boxed{2,3} & \boxed{1,3} \\ 1,0 & 0,1 \\ 0,1 & 1,0 \end{array} \right\| \end{array}$$

The Nash Equilibria are $\{T, L\}$ and $\{T, R\}$, i.e. the same found before, suggesting a **connection** between the two concepts. We formalize this in the following proposition.

Proposition 1.3.1. *Let $\mathcal{G} = (M, (S_i)_{i \in M}, (u_i)_{i \in M})$, and let $j \in M$. Consider a strategy $\hat{s}_j \in S_j$ that is **weakly dominated**. We denote with $\hat{\mathcal{G}}$ the restriction of $\mathcal{G}$ obtained by removing $\hat{s}_j$, i.e. $\hat{S}_j = S_j \setminus \{\hat{s}_j\}$.*
Then, every Nash Equilibrium of $\hat{\mathcal{G}}$ is a Nash Equilibrium of $\mathcal{G}$, i.e. by iterating the weakly dominance elimination we will select one Nash Equilibrium.

*Weak dominance elimination always **selects** a Nash Equilibrium*

Proof. In the restricted game, the new set of strategies of player i is given by:

$$\hat{S}_i = \begin{cases} S_i & i \neq j \\ S_j \setminus \{\hat{s}_j\} & i = j \end{cases}$$

since the elimination step affects only player j.

Let s^* be a Nash Equilibrium of $\hat{\mathcal{G}}$. We want to show that it is a Nash Equilibrium also for $\mathcal{G}$. By definition, we have:

$$u_i(s^*) \geq u_i(s_i, s^*_{-i}) \qquad \forall s_i \in S_i$$

And for $i = j$:

$$u_j(s^*) \geq u_j(s_j, s^*_{-j}) \qquad \forall s_j \in \hat{S}_j \tag{1.1}$$

To show that s^* is a Nash Equilibrium in $\mathcal{G}$, we need to show that there are no possible *profitable* deviations in the strategies, i.e. that any different choice would lead to lower utilities for some players. For $i \neq j$ this can be immediately shown, since their strategies are *unaffected* by the elimination process: $S_i = \hat{S}_i$. In other words, players that are not j still act as if they were playing the full game $\mathcal{G}$.
Player j, instead, cannot choose $\hat{s}_j$ anymore in $\hat{\mathcal{G}}$. So, if we *extend back* to $\mathcal{G}$, we need to check if j can *switch their choice* to $\hat{s}_j$ and increase their utility. In other words, if we replace s^*_j with $\hat{s}_j$ in the Nash Equilibrium s^*, does the utility of j increase or decrease?

$$u_j(s^*_j, s^*_{-j}) \overset{?}{\geq} u_j(\hat{s}_j, s^*_{-j})$$

(Note that *for sure* switching to another state in $\hat{S}_j$ does not increase u_j, due to (1.1). However $\hat{s}_j \notin \hat{S}_j$).

The idea is to use the fact that $\hat{s}_j$ is weakly dominated. Thus, there is another state $t_j \neq \hat{s}_j$ that *dominates* $\hat{s}_j$, and since now $t_j \in \hat{S}_j$, we can use (1.1) to show that the utility of switching $s^*_j \to \hat{s}_j$ cannot increase.

More precisely, from the definition of **weakly dominance** we have that there exists a strategy $t_j \in S_j$ such that:

$$u_j(t_j, s_{-j}) \geq u_j(\hat{s}_j, s_{-j}) \qquad \forall s_{-j} \in S_{-j}$$

Since s_{-j} can be anything, we set $s_{-j} = s^*_{-j}$, obtaining:

$$u_j(t_j, s^*_{-j}) \geq u_j(\hat{s}_j, s^*_{-j}) \tag{1.2}$$

Since $t_j \neq s_j$, it is not removed by the elimination, and so $t_j \in \hat{S}_j$. So we can use (1.1) and set here $s_j = t_j$, leading to:

$$u_j(s^*) \geq u_j(t_j, s^*_{-j}) \tag{1.3}$$

Chaining (1.3) and (1.2) together we arrive at the result:

$$u_j(s^*) \geq u_j(\hat{s}_j, s^*_j)$$

proving that switching $s^*_j \to \hat{s}_j$ does not increase u_j, and so even j has no incentive to change their mind, meaning that s^* is indeed a Nash Equilibrium also for $\mathcal{G}$. $\square$

1.3.3 Prisoners' Dilemma

The **prisoner's dilemma** is a game consisting of a single Nash Equilibrium that is not Pareto's efficient. In other words, all players can be in a situation where nobody has any incentive to change their mind, but that it is not the "best" solution for all.

$$\mathbf{P} = \begin{array}{ccc} & & \text{Player 2} \\ & & \begin{array}{cc} L & R \end{array} \\ \text{Player 1} & \begin{array}{c} T \\ B \end{array} & \left\| \begin{array}{cc} 3,3 & 0,5 \\ 5,0 & \boxed{1,1} \end{array} \right\| \end{array}$$

The Nash Equilibrium is $\{B, R\}$, inducing the outcome $(1, 1)$. However, $\{T, L\}$ should be a *better* strategy profile, since it has a better outcome for everyone $(3, 3)$. So there is a **Pareto improvement** in passing from $(1, 1) \to (3, 3)$, but $\{T, L\}$ cannot be sustained: in this case there is *incentive* for a player to change their mind, since they can improve $3 \to 5$ their utility (but lowering the outcome for the opponent).

Moreover, if the Iterated Elimination of Dominated Strategies gives a unique solution, then the Nash Equilibrium is unique.

However, **multiple** Nash Equilibria may exist:

Non-uniqueness

$$\mathbf{P} = \begin{array}{ccc} & & \text{Player 2} \\ & & \begin{array}{cc} L & R \end{array} \\ \text{Player 1} & \begin{array}{c} T \\ B \end{array} & \left\| \begin{array}{cc} \boxed{2,1} & 0,0 \\ 0,0 & \boxed{1,2} \end{array} \right\| \end{array}$$

and sometimes a Nash Equilibrium may **not exist** at all:

Non-existance

$$\mathbf{P} = \begin{array}{ccc} & & \text{Player 2} \\ & & \begin{array}{cc} L & R \end{array} \\ \text{Player 1} & \begin{array}{c} T \\ B \end{array} & \left\| \begin{array}{cc} 1,-1 & -1,1 \\ -1,1 & 1,-1 \end{array} \right\| \end{array}$$

These cases may be solved by using instead *mixed strategies*, i.e. strategy profiles involving **probabilities**.

1.4 Preferences of Groups

(Lesson 5 of 12/10/20) Compiled: January 1, 2021

Until now, we have been considering the preferences of **each** player in the game. Assuming rational players, these utilities are both **complete** (each outcome can be ranked) and **transitive** (there are no *cycles*, but each player has well-defined preferred outcomes).
However, we would like a way to **aggregate** these preferences. For example, consider the following:

Filippo	Daniel	Egli
P	H	C
C	P	H
H	C	P

Figure (1.11) – Preference profiles (i.e. rankings) for 3 players, regarding different activities: *picnic* (P), *cinema* (C) and *hiking* (H).

Given these preference profiles, can we construct a ranking for the entire group?

First, let's define some notation. Let A be a set of outcomes. For a rational player i, a **preference profile** $\succcurlyeq_1$ is a **transitive** ranking of **all** (complete) the outcomes in A. For example, if $A = \{a, b, c\}$, a profile could be $a \succ b \succ c$. The ranking could be *non strict*, i.e. $a \succcurlyeq b \succcurlyeq c$, since a player may be *indifferent* about two outcomes that they see as *equivalent*.
Consider a set $N = \{1, \ldots, n\}$ of n players, and let $R(A)$ be the set of all preference profiles for a rational player. Then, all players preferences are elements in the space $R^n(A)$, which forms the *domain* of a preference-aggregating function.

Social choice function

The *simplest* example is the **social choice function**, which just selects a global outcome based on the rankings:

$$f \colon R^n(A) \to A$$

But in general, we could be interested in generating a new *aggregated* ranking of the outcomes A, i.e. a new "global" preference profile:

Social welfare function / Constitution

$$f \colon R^n(A) \to R(A)$$

Such a function f is called a **constitution** (or a **social welfare function**).

a. Pareto efficiency of a constitution

We expect a *good* aggregated ranking to be "good for everybody", i.e. representative of *global values*. This is formalized by the concept of Pareto efficiency.
A **constitution** f is **Pareto efficient** if the aggregated ranking it produces *respects* global agreements of the players. In other words, if all preference profiles $\{\succcurlyeq_i\}$ agree on the ranking of two outcomes $a \succcurlyeq_i b \; \forall i \in N$, then the aggregated ranking $\succcurlyeq_f$ must reproduce that ordering: $a \succcurlyeq_f b$.
In essence, the following must hold:

$$\forall a, b \in A, \qquad a \succcurlyeq_i b \quad \forall i \in N \Rightarrow a \succcurlyeq_f b$$

b. Independence of irrelevant alternatives

Another expected property for a *good* ranking is the so-called **independence of irrelevant alternatives**. The idea is that if we add or remove elements to A, this should not change the *relative* ranking of other pairs of outcomes.

To formalize this concept, we first need some notation. Consider a ranking $\succcurlyeq_i$ over a set of outcomes A. We define the **restriction** $\succcurlyeq_i |_B$, where $B \subseteq A$, as the ranking that $\succcurlyeq_i$ *induces* on the subset B. For example, if $A = \{a, b, c\}$, and $\succcurlyeq_i$ is:

$$a \succcurlyeq b \succcurlyeq c$$

Then $\succcurlyeq_i |_{a,c}$ is given by $a \succcurlyeq c$. Basically, it is obtained by removing the elements of $A \setminus B$ from $\succcurlyeq_i$.

Now we are ready to define the **independence** property. Consider two sets of profiles $\{\succcurlyeq_{(i)}\}_{i\in N}$ and $\{\succcurlyeq'_{(i)}\}_{i\in N}$. We say that f is **independent of irrelevant alternatives** if the following holds $\forall a, b \in A$:

$$\text{if } \forall i, (\succcurlyeq_i |_{\{a,b\}}) = (\succcurlyeq'_i |_{\{a,b\}}) \Rightarrow f(\succcurlyeq_{(i)})|_{\{a,b\}} = f(\succcurlyeq'_{(i)})|_{\{a,b\}}$$

To understand this (*cryptic*) notation, consider the following example. Suppose we have 3 players, and 4 outcomes $\{a, b, c, d\}$. We define the first set of rankings as follows:

$$\begin{aligned}
\succcurlyeq_1 &: a \succcurlyeq d \succcurlyeq b \succcurlyeq c \\
\succcurlyeq_2 &: c \succcurlyeq b \succcurlyeq a \succcurlyeq d \\
\succcurlyeq_3 &: a \succcurlyeq d \succcurlyeq c \succcurlyeq b
\end{aligned}$$

For the second set of rankings, we *shift* the position of c and d, leaving the *relative* ranking of a and b the same:

$$\begin{aligned}
\succcurlyeq'_1 &: d \succcurlyeq a \succcurlyeq b \succcurlyeq c \\
\succcurlyeq'_2 &: c \succcurlyeq b \succcurlyeq d \succcurlyeq a \\
\succcurlyeq'_3 &: c \succcurlyeq d \succcurlyeq a \succcurlyeq b
\end{aligned}$$

In fact:

$$\begin{aligned}
a \succcurlyeq_1 b &\text{ and } a \succcurlyeq'_1 b \\
b \succcurlyeq_2 a &\text{ and } b \succcurlyeq'_2 a \\
a \succcurlyeq_3 b &\text{ and } a \succcurlyeq'_3 b
\end{aligned}$$

This is what we mean when writing $\forall i, (\succcurlyeq_i |_{\{a,b\}}) = (\succcurlyeq'_i |_{\{a,b\}})$.
Then, if f is **independent of irrelevant alternatives**, we expect that it will produce the same raking of a and b for both sets. For example, denoting $f(\succcurlyeq_{(i)}) \equiv \succcurlyeq_f$ and $f(\succcurlyeq'_{(i)}) \equiv \succcurlyeq'_f$, if $a \succcurlyeq_f b$, then it must be $a \succcurlyeq'_f b$.

One of the simplest possible choices for f is that of a **dictatorship**, where we take $f(\succcurlyeq_{(i)}) \equiv \succcurlyeq_f = \succcurlyeq_i$ for an arbitrary player i. This means that the *aggregate* ranking simply reflects that of player i, effectively disregarding all the others' opinions: *Dictatorship*

$$a \succcurlyeq b \Rightarrow a \succcurlyeq_f b$$

Clearly, we would like for a *good* f to **not** be a dictatorship, since this would defeat the purpose of an aggregating function.

One last required property is **monotonicity**. This requires that, if an individual modifies their preference by ranking an outcome a to some *higher* position, then the *aggregating ranking* should either place a at the same place as before or to a higher rank, but certainly not to a lower one. In other words, an individual should not be able to *hurt* an option in the global ranking by ranking it higher in their personal ranking!

Unfortunately, all these requirements are too strong. In fact, in 1951 Kenneth Arrow proved the following theorem:

Theorem 1.4.1. *Any social welfare function f over three or more alternatives ($|A| \geq 3$) that is both Pareto efficient, monotonic and independent of Irrelevant Alternatives is* ***dictatorial****.*

As we will see in the next section, this has strong implications when discussing procedures where we need to *aggregate* preferences, such as **elections**.

1.5 Elections

In elections, players express their preferences in form of *votes*, so that a common outcome may be decided. There are many ways to do that, for example:

- The simplest idea is that of *single* voting, where each player expresses their preferred outcome.
- **Cumulative** voting: each player has a number of votes which are *distributed* over the outcomes.
- **Approval** voting: each player can vote multiple outcomes.

Suppose we have 3 voters and 2 candidates, and we obtain the following preferences:

voter	1	2	3
best	A	A	B
worst	B	B	A

Figure (1.12)

The majority has voted for A, and so by **plurality rule** we should prefer A. Let's add a third candidate:

voter	1	2	3
best	A	A	B
	B	C	C
worst	C	B	A

Figure (1.13)

Here A beats B, B beats C and A beats C. Since A has surpassed all the other candidates, the final choice should be for A.
However, these are lucky cases. If we have instead:

voter	1	2	3
best	A	C	B
	B	A	C
worst	C	B	A

Figure (1.14)

We see $A > B$, $B > C$ and $C > A$, forming a **cycle**. Thus, in this case there is no best candidate!

The *basic rules* we have been using in the past few samples define the concept of **Condorcet winner**, i.e. the outcome that is preferred to every other candidate in a pairwise majority-rule comparison (the one that "beats" the most competitors). As we have seen, it does not always exist, because there could be **Condorcet cycles**:

Condorcet winner

Condorcet cycles

$$A \succ B,\ B \succ C,\ C \succ A$$

Note that fig. 1.14 originates from the case with two candidates in fig. 1.12. Depending on the *position* of the new candidate, we can make A or C the winner, or originate a cycle. The probability of cycles actually *grows* with the number of candidates, and for $n \to +\infty$ cycles will *surely* occur.

1.5.1 Voting systems and Paradoxes

The **order** in which to compare the voters is called the **agenda**. Depending on the voting system, this can **affect or not** the result!
For example:

	1-4 (4 voters)	5-7 (3 voters)	8-9 (2 voters)
best	A	C	B
	B	A	C
	C	B	A
worst	D	D	D

Figure (1.15)

Here $A > B > C > A$ form a Condorcet cycle, and D is the worst candidate. Suppose the voting system consists of semifinals and a final: in this case the *order* of matchup will determine the winner! In fact, suppose A goes against D (A wins), and B against C (B wins). At the finals we have A vs B, and A wins. If instead we make B go against D (B wins), and A against C (C wins), the finals will be between B and C, with B winning! This is because whoever goes against the worst opponent will surely win, and the only other candidate that may defeat them will lose on the other branch of the semifinals.

Plurality voting

Each voter sorts the candidates in order of personal preference, and the winner is the one who appears in most *first places* in these rankings.

For example:

	1-4 (4 voters)	5-7 (3 voters)	8-9 (2 voters)
best	A	B	C
	B	C	B
worst	C	A	A

Figure (1.16)

Here A wins, because it has the most *top* votes (first line). However, if we look at the complete rankings we see that $B > A$ is preferred by 5 vs 4, and so is $C > A$. Since also $B > C$ is preferred by most (7 vs 2), B is the Condorcet winner.

Two-phase run-off

Voting happens in two phases: in the first one the *top* two candidates are selected, and only they compete in a second final round.
For example:

	1-4 (4 voters)	5-7 (3 voters)	8-9 (2 voters)
best	A	B	C
	C	C	B
worst	B	A	A

Figure (1.17)

By counting the *top* votes, A wins with 4, and B is the runner-up with 3 votes. Again, if we look at the complete rankings, we see that $C > A$ (5 vs 4), and also $C > B$ (6 vs 3), meaning that C is the Condorcet winner, but does not even make it to the ballot!

Borda voting

If we have M candidates, the voter ranks them. According to their position, each candidate receives a score: $M - 1$ for the *top* one, and 0 for the *worst* one.
For example:

		1-5 (5 voters)	6-8 (3 voters)	9
Best	2	A	B	C
	1	B	C	B
Worst	0	C	A	A

Figure (1.18)

A obtains $5 \cdot 2 = 10$ points, B gets $5 \cdot 1 + 3 \cdot 2 + 1 \cdot 1 = 12$ points, and C $3 \cdot 1 + 1 \cdot 2 = 5$, meaning that B wins. Still, $A > B$ (5 vs 4) and $A > C$ (5 vs 4), making A the Condorcet winner.

Since the scoring *depends* on the number of candidates, if one of them is removed from the list, the final ranking may significantly changing (also be reversed).

Approval voting

Each voter can give more than one preference (up to M), which counts as 1 point.
Here, depending on the number N of votes available to each player, the result may change. If $N = 1$, we recover the plurality voting system.

Instant run-off

Look at the top preferences and remove the candidate with the least number of votes. This procedure is iteratively repeated until the winner is found. The voters who put the removed candidate as first preference are kept, and their second preference is counted as votes for the other candidates.

Cheating

The previous paradoxes can get worse if we allow *cheating*, i.e. voters that choose rankings that do not correspond directly to their true preferences, but can *skew* the results of the election so that, at the end, their utility will be maximized.
For example, one can imagine a situation where a candidate is *winning* with a good margin. Voters of the least preferred candidate may *switch* their votes to support instead the *runner-up*, making them the winner (if this is the preferred situation).
The situation can get even worse if all the players *know* that cheating is happening. In this case, voters may cheat *in response* to the cheating of others! At the end, the outcome will depend not only on the order of comparisons (agenda) but also on the order of cheating.

A social function f may be made **strategy-proof**, i.e. non manipulable, but in this case it can be proven (Gibberard-Satterthwaite theorem 1973) that any strategy-proof constitution that does not forbid anyone to win **must** be a dictatorship.

1.6 Applications of Nash Equilibrium

(Lesson 6 of 15/10/20) Compiled: January 1, 2021

We will now see several examples of application of Nash Equilibrium to economics. The aim is twofold:

- First, these examples are very well known in literature, so knowing them is useful for understanding references
- Second, they serve as *inspiration* for finding applications in other fields, such as engineering. However, note that this won't be possible in all cases, and it may require first some *translation* work.

One way to *use* Nash Equilibrium is as a sort of **prediction**: we expect rational players to play a Nash Equilibrium. However, this can lead to some problems. In fact, some problems *look like* they have **no** Nash Equilibrium. However, this

can be solved by *generalizing* the definition, so that any game can have at least a NE.
In other problems there are instead **multiple** Nash Equilibria, and we would like a way to *choose* only one of them as a "preferred solution". Often, in fact, they are not *completely equivalent*, and so there should be a way to choose "the best one".
A bigger problem is that a Nash Equilibrium is not necessary Pareto efficient, i.e. players driven by egoism may not choose the *best* possible solution for the entire society.

1.6.1 Duopoly

A classic situation in economy is one where there are two competing *firms* that want to gain as many customers as possible.
A historic model for a duopoly is that of **Cournot** (1838), who *anticipated* the results of Nash. This is actually a *toy* model: it is not representative of reality nor accurate, but its simplicity allows exact computations, meaning that we can easily *understand* its inner workings.

Cournot duopoly

In Cournot model we consider two firms 1 and 2 producing a good in quantities q_1 and q_2, leading to a total $Q = q_1 + q_2$. We assume that the *cost* to produce q units of goods is the *same* for both firms, and it is given by a simple proportionality relation: $C(q) = c\,q$, with some arbitrary constant c.
When the good is sold on the market, its **price** is $P(Q)$ given by:

$$P(Q) = (a - Q)h(a - Q) \tag{1.4}$$

where h is the Heaviside function, needed to make $P(Q) \geq 0$. Clearly *demand and supply* are not that simple, and so this is a point where the model is unrealistic. For example, actually firms do not fix the quantity of produced goods q, but can set the price.
However, formula (1.4) captures the *basic concept* that if a firm sells *more goods*, the price will be lower, and viceversa (*the more you produce, the lower the price*). Note that the *price* of the goods depends on the total amount of goods produced Q, i.e. on both the firms.

Suppose the two firms choose q_1 and q_2 simultaneously and unbeknownst to each other. What is the Nash Equilibrium?

Let's start by formalizing the game. Each player $i \in \{1, 2\}$ decide their move $q_i \in S_i = [0, +\infty)$. Since producing *too much*, i.e. $q_i > a$, would mean that $P = 0$ (giving the goods away for free), which is clearly pointless, we can restrict S_i to $[0, a)$ without any loss.
The payoff of a firm is simple its profit, i.e. the revenue minus the cost:

$$u_i(q_i, q_j) = q_i[P(\underbrace{q_i + q_j}_{Q}) - c] \underset{(1.4)}{=} q_i(a - q_i - q_j - c)$$

with $(i, j) \in \{(1, 2), (2, 1)\}$. Then, we want to find (q_1^*, q_2^*) corresponding to the Nash Equilibrium. Since we are dealing with a *continuous* set of actions, we cannot write a matrix for describing all the payoffs.

Still, we can solve the problem. By definition, in a Nash Equilibrium both players are playing their best responses to each other. If player j plays q_j, the best reply available to player i is the one that maximizes their profit:

$$q_i = \arg\max_{q_i} u_i(q_i, q_j)$$

At the Nash Equilibrium, both player are using the best reply, and so the following must hold:

$$q_i^* = \arg\max_{q_i} u_i(q_i, q_j^*)$$

In other words, we need to solve the following system of equations:

$$\begin{cases} q_1^* = \arg\max_{q_1} u_1(q_1, q_2^*) \\ q_2^* = \arg\max_{q_2} u_2(q_1^*, q_2) \end{cases}$$

Given the symmetry, we just need to solve one of the equations, which can be done by differentiation:

$$q_i^* = \arg\max_{q_i} q_i(a - q_i - q_j^* - c) \Rightarrow (a - 2q_i^* - q_j^* - c) \stackrel{!}{=} 0$$

leading to the solution:

$$q_1^* = q_2^* = \frac{a-c}{3}$$

In this case, the profit is the same for both firms (as expected by symmetry):

$$u_1^* = u_2^* = \frac{(a-c)^2}{9}$$

1.6.2 Monopoly

Suppose one of the two firms from above *disappears*, leaving the other able to *control* the entire market. This can be obtained by setting (for example) $q_2^* = 0$, which leads to:

$$q_m = \arg\max_{q_1} q_1(a - q_1 - \cancel{q_2} - c) \Rightarrow q_m = \frac{a-c}{2} \Rightarrow u_m = \frac{(a-c)^2}{4} > u_1^* \quad (1.5)$$

Note that $(q_m, 0)$ is not a Nash Equilibrium, since clearly $q_2 = 0$ is not the best response for player 2. In this case, having no competition proves really advantageous for player 1, who is able to make a higher profit.
However, the *total amount of goods* produced will be lower:

Duopoly vs Monopoly

$$Q_m = q_m = \frac{1}{2}(a-c) < \frac{2}{3}(a-c) = q_1^* + q_2^* = Q^*$$

Thus a monopoly has incentive to produce *a bit less* to gain more profit. So this simple model shows that competition is good for the consumers, but bad for the producers. Quantitatively, we can compare the total utilities in both cases:

$$u_{\text{tot}} = u_m = \frac{1}{4}(a-c)^2 > \frac{2}{9}(a-c)^2 = u_1^* + u_2^* = u_{\text{tot}}^*$$

1.6.3 Trust

If we allow the two firms to communicate, the result can *change*. In this case, in fact, they can agree to "pretend to be a monopoly" by *combining their moves* and splitting the revenues. This is advantageous, since *half of the monopoly's revenue* is still more than the outcome of each player in a duopoly.
So, they each produce $q_m/2$, and get $u_m/2$, which is:

$$\frac{u_m}{2} = \frac{(a-c)^2}{8} > \frac{(a-c)^2}{9} = u_1^* = u_2^*$$

This situation would be *best* for both players, but it is not an equilibrium: any player has the *incentive* to break the agreement to earn more.
Note the similarity with the prisoners' dilemma. In both games, players reach an equilibrium solution which is not the "best" globally. In technical terms, the *combined monopoly* solution is Pareto dominating over the Nash Equilibrium.

1.6.4 Bertrand duopoly

Set the price, not the goods

Bertrand (1883) argued against Cournot model that firms choose prices, not the produced quantities q_j. This completely changes the game: now the strategies are the prices $p_i \in S_i = [0, \infty)$, and the sold quantities are determined by the prices. Specifically, customers will buy only the cheaper product, i.e. the one with the lowest p_i:

$$q_i = a - p_i$$

And for the other player, if $p_j > q_j$, $q_j = 0$ (nothing is sold). However, if both players *agree* on the price, then they will *share* the production of goods.
Let's assume the cost of producing goods to be the same as before, i.e. $C(q) = c\,q$, with $a > c$.

We can see immediately that, at the Nash Equilibrium, both prices must be the same, otherwise there is one player who is losing money and would like to change their action. Moreover, prices must be the *minimum* allowable, otherwise one player can *lower* their price and gain the entire market (since the other won't sell anything anymore). So, since producing a unit of goods costs c, we have:

$$p_1^* = p_2^* = c$$

Nash Equilibrium

Going below these prices would lead to *negative* profits, which are clearly not sustainable.

Note that this kind of Nash Equilibrium is definitely not the best outcome for the firms. As before, they could *agree* on a higher price and share the market, effectively forming a *monopoly*. In this case, the prices can go up to $(a+c)/2 > c$. However, while this solution Pareto dominates over the Nash Equilibrium, it is not an equilibrium: each player has an incentive to break the agreement (by lowering their price) to conquer the market.

Equivalent best replies in a continuum

As a final remark, note that if a player *believes* that the other will set the price at $p_2^* = c$, then *any choice they make* will lead to a profit of 0: they can sell goods at production cost, or set a higher price and don't sell at all. So, *any* $p_1 \geq 0$ is a best response to $p_2^* = c$. However, $(c + \epsilon, c)$ is **not** a Nash Equilibrium, since now c is *not* the best reply for player 2 to $p_1 = c + \epsilon$.

a. Imperfect substitutes

In reality, however, customers may sometimes buy a more expensive product. This is because goods produced by different firms are **imperfect substitutes**: they are not *completely equivalent*, and may have some different qualities that are preferred by certain customers. This can be modelled by changing the rules of the game:

$$q_i = a - p_i + b\,p_j \qquad b < 2$$

This means that the amount of goods sold by i is proportional to the price of the alternative: if p_j is high, customers will mostly buy goods from i, but if $p_j \approx p_i$, they will buy from both. In this equation, b acts as a sort of *exchange rate* between the different goods.
It can be shown that, in this case, the Nash Equilibrium becomes:

$$p_1^* = p_2^* = \frac{a + c}{2 - b}$$

b. Different production costs

What happens if the production costs are not the same for the two firms? For example, suppose that $c_1 = 1$ and $c_2 = 2$. Suppose also, for simplicity, that prices are discretized, and can be changed in steps of $\epsilon = 0.01$.
According to the Bertrand duopoly model, there is no way for player 2 to win, because player 1 can set any price *lower* than c_2, cutting out 2 and becoming a monopolist. Thus, in this case the Nash Equilibrium is $(1.99, 2.00)$.

Discontinuities

However, note that as $\epsilon \to 0$, the continuous space of actions leads to a problem. In fact, player 1 has an incentive to set a price that is *as much close as possible* to c_2, but still lower. If $\epsilon \to 0$, this leads to a **discontinuity** in u_1, which is that of a monopoly for any $p_1 < c_2$, and that of a duopoly for $p_1 = c_2$. Clearly $(2, 2)$ cannot be the Nash Equilibrium, since p_1 can achieve a better result by lowering their price. But if they lower by ϵ, they could do better by instead lowering *less*, i.e. by $\epsilon_1 < \epsilon$, and so on.
Practically, we can just consider a discretization to solve this kind of problems: this is also realistic, since there is a minimum amount of money.

1.6.5 Hotelling model

Another famous duopoly model is the one proposed by Hotelling (1929).
Here we consider two firms that sell *perfect substitutes* for the same goods along a street. For example, consider two ice-cream vendors at a seaside boulevard, assumed to be 1 km long.
Customers will buy ice-cream from the **nearest** vendor, and are uniformly distributed along the street. For simplicity, suppose there are 101 possible locations for the ice-cream stands (one each 10 m). What is the best choice for the stands' location?

For example, if A chooses position $x_A = 22$, and B chooses $x_B = 35$, then A will attract all people in $[0, 28]$, and B the rest. So, x_A is not optimal for A: they should move instead to the right, more specifically to 36, so that now A gets $[36, 101]$ and B the rest. Similarly, now B would like to move *to the right* of A again. This process continues until the equilibrium is reached at the **middle-point**, with $x_A^* = x_B^* = 50$ (one player gets the left side, one all the right one).

This result shows a sort of **convergence**. The same can happen with politicians from **two** groups trying to attract voters from different sides, who often "gather at the middle", claiming very similar things.

However, if we allow *more than two groups*, this convergence is broken, and the street will be divided equally.

1.6.6 Tragedy of the commons

Many political philosophers and economists, since at least Hume (1739) have understood that, if moved only by private incentives, citizens tend to *misuse* public resources (e.g. by throwing trash in the streets or polluting the environment). This problem is commonly referred to as the **tragedy of commons**.

This situation can be formalized in the context of game theory in several ways. The **classic version** of the problem is that of Hardin (1968), in which n farmers, each owning g_i goats, all use a common green area for foraging. In total, the number of goats is:

Hardin model

$$G = g_1 + g_2 + \cdots + g_n$$

Each goat costs c in caring expenses. The value of *foraging in the common green* is $v(G)$, which is a decreasing function of G (more goats using the same area means that each goat has access to *less* grass). Mathematically, it is a **concave** function:

$$v(G_{\max}) = 0 \qquad v'(G) < 0 \qquad v''(G) < 0$$

Generalization. Note that this framework can be generalized to many situations. For example, consider n users of a WiFi hotspot, each using g_i processes. Clearly, the more users connect at the same time, the less bandwidth will be available for each of them, meaning that the *value* of accessing the common connection $v(G)$ decreases with G.

Each farmer can choose g_i, i.e. how many goats to bring to the common green. The payoff is the difference between benefits and maintenance costs:

$$u_i = g_i(v(G) - c)$$

To find the Nash Equilibrium, we search for the strategy profile consisting of all best replies. To simplify notation, denote with $\boldsymbol{g_{-i}}$ the profile strategy vector without the i-th entry:

$$\boldsymbol{g_{-i}} \equiv (g_1, \ldots, g_{i-1}, g_{i+1}, \ldots, g_n)$$

and with G_{-i} the *sum* of all entries of $\boldsymbol{g_{-i}}$.
Then, the Nash Equilibrium is the solution of the following n equations:

$$g_i^* = \arg\max_{g_i} g_i(v(g_i + G_{-i}^*) - c)$$

Note that *increasing* g_i increases the multiplier of benefit, but decreases each benefit.

To solve this problem, note that by **symmetry**, all farmers are equivalent, and so we expect that $g_i^* \equiv g^*$ are all the same. So, the total number of goats is $G^* = ng^*$. Substituting in the above:

$$g^* = \arg\max_{g^*} g^*(v(G^*) - c)$$

We can now differentiate with respect to g^* and set the derivative to 0 to find the maximum:

$$\frac{\partial}{\partial g^*} g^*(v(G^*) - c) \overset{!}{=} 0 \Rightarrow v(G^*) - c + \underbrace{g^*}_{G^*/n} v'(G^*) = 0 \tag{1.6}$$

We cannot find an explicit solution without knowing the full expression for v. However, we can compare (1.6) with that where *all* farmers agree on "acting like a monopoly" (i.e. the "best possible scenario" for *all* the farmers). In this case, it is as if there was a *single* farmer, and so we need to maximize just one condition:

$$G_m = \arg\max_G G(v(G) - c) \Rightarrow v(G_m) + G_m v'(G_m) - c \overset{!}{=} 0 \tag{1.7}$$

Note that both (1.6) and (1.7) are set to 0, and so we can equate them:

$$v(G_m) + G_m v'(G_m) = v(G^*) + \frac{G^* v'(G^*)}{n}$$

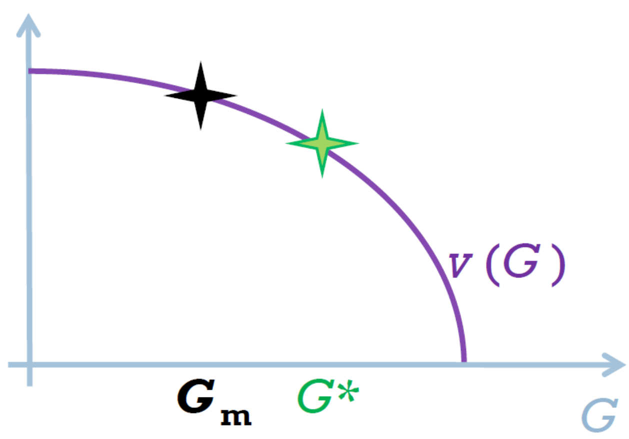

Figure (1.19) – Comparison between the Nash Equilibrium and the *monopoly* solution for the Hardin model. Note how $G^* > G_m$, meaning that the benefit for each individual $v(G^*) < v(G_m)$, since v is a decreasing function. In other words, *selfishness* of the players leads to a non-optimal utilization of the common resources.

Let's try to understand this situation. Suppose a user with g_i goats considers an *increment* h, i.e. thinks about changing their move $g_i \to g_i + h$.
Looking at the payoff function:

$$u_i = g_i(v(G) - c)$$

At first order, increasing g_i leads to a benefit gain:

$$\frac{(g_i + h)v(G) - v(G)}{h} = v(G)$$

However, this is counteracted by:

- An **increase** in the cost of possessions of $ch/h = c$.
- A **decrease** in the value of each possession $[v(G+h) - v(G)]/h \approx v'(G)$, for a total of $v'(G)g_i < 0$ (recall that $v'(G) < 0$).

So, the Nash Equilibrium is exactly the situation where the benefit gain is *exactly cancelled* by the other two losses:

$$v(G^*) + v'(G^*)\frac{G^*}{n} - c = 0$$

Note that here we are considering the losses of a **single** user (see the $1/n$ factor).

In the *monopoly* scenario, the equation is the same, but without the $1/n$ factor. So, we are considering the loss of **all** users, which is $v'(G_m)G_m$. This is because we are not playing the game at the level of users, but at the one of the entire society.

This explains the *tragedy of commons*: when all players have access to a common ground, losses are shared, and so they matter less for each individual (they are divided by n), meaning that they do not feel a strong need to avoid them!

1.6.7 Selfish routing

Another model related to the tragedy of commons, but this time set as an engineering situation, is that of **selfish routing**.

Pigou (1920) considered a scenario with two **paths**, red R and blue B, going from a source s to a destination d. Let x_R and x_B be the fractions of *traffic* on each road, and suppose that travelling through a path incurs in some cost, which is x for the red one (i.e. proportional to the *congestion*) and fixed to 1 for the blue one.

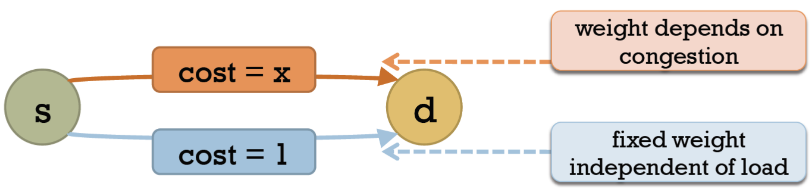

Figure (1.20) – Diagram of the two paths in the selfish routing model.

Clearly, traversing the red path is a **dominating strategy**, since depending on congestion there is the possibility of paying a cost $x < 1$.

Since all players can see that, they will all go through the red path, making $x = 1$, and the two roads equivalent. Note that, even in this situation, there is still *no unilateral incentive* for one user to go instead through the blue path, which is now empty.

Clearly, this is not very efficient! For example, if we *force* the traffic to split **equally** over the two paths, we will have a cost $1/2$ for the red road, and of 1 for the blue one, leading to an average:

$$\text{Average cost} = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot 1 = \frac{3}{4}$$

It can be shown that this is the globally **optimal** situation, i.e. the one with the lowest average cost, and so this is the Pareto efficient strategy. However, it is not an equilibrium: there is an incentive for the users moving through the blue path to choose instead the red path.

The ratio between the Nash Equilibrium average cost and that of the Pareto efficient strategy is the **price of anarchy**, which in this case is $1/(3/4) = 4/3$. This measures the *relative loss* of "global efficiency" produced by the selfish behavior of players.

The situation becomes even worse if the cost for the red path is non-linear, i.e. x^a with $a > 1$. In this case, the "sacrifice" of only few users going through the blue path would allow *almost all* to pay a very low cost. In fact, for $a \to +\infty$, the average cost of the Pareto efficient solution goes to 0, meaning that the price of anarchy diverges to infinity.

Counterintuitively, a "better" network can have a "worse" Nash Equilibrium. For example, consider the following scenario:

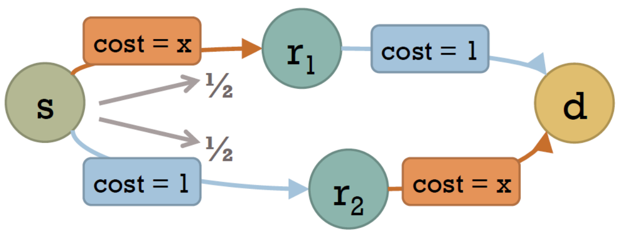

Figure (1.21) – Two *balanced* but independent paths.

In this case, the average cost at equilibrium is $3/2$.

Suppose we add a new *costless* connection between the nodes r_1 and r_2, allowing re-routing of the traffic. We could expect the average cost to decrease, or at least remain the same:

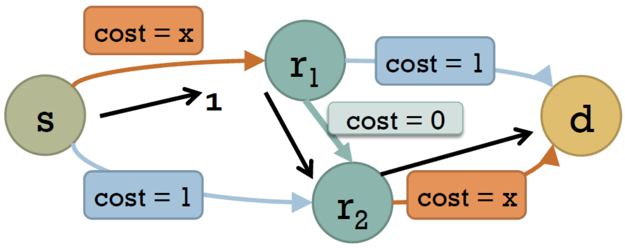

Figure (1.22) – A *costless* route is added between r_1 and r_2.

However, now the equilibrium is reached with *all players* going through the orange path, incurring in a higher average cost of 2!

This is an example of the so-called **Braess paradox**.

Generalization to graph with linear costs

Regarding this kind of problem, a more general result can be proven. In *any* network where latencies (costs) are linear in the occupancy (congestion), the **price of anarchy** is always 4/3. Intuitively, this is because when confronted with two choices, all selfish users take the better one, leading to an overload of paths. So, the price of anarchy is an intrinsic property of the problem, it does not depend on the graph's topology.

1.7 Generalized Nash Equilibrium

(Lesson 7 of 19/10/20) Compiled: January 1, 2021

As it was discussed previously, there are games that do not seem to have a Nash Equilibrium.
For example, consider the Odds & Evens game. Here each player *bets* on a result (Even or Odd), and we denote the player with their bet. Then, they can play an odd number (0) or an even one (1). The parity of the sum of the two numbers determines the winner, who *gains* 4 points, while the other *loses* 4 points.

$$
\mathbf{P} = \begin{array}{cc} & \text{Even} \\ & \begin{array}{cc} 0 & 1 \end{array} \\ \text{Odd} \begin{array}{c} 0 \\ 1 \end{array} & \begin{Vmatrix} -4, 4 & 4, -4 \\ 4, -4 & -4, 4 \end{Vmatrix} \end{array}
$$

Here, there is no strategy profile in which all players are choosing a best reply, because one will certainly lose and regret their action.

Note that the problem lies in the *uncertainty* intrinsic in this kind of game, which effectively *simulates* a coin toss. So, to extend the concept of Nash Equilibrium also to these situations (which is desirable if we want to use it as a sort of *predictive* tool), we need to allow strategies involving randomness, i.e. **mixed strategies**.

As a start, let's *extend* the game, allowing each player to make an "intermediate" move 1/2, resulting in an *intermediate* outcome between the two alternatives:

$$
\mathbf{P} = \begin{array}{cc|ccc|}
 & & & \text{Even} & \\
 & & 0 & 1/2 & 1 \\
\hline
 & 0 & -4,4 & 0,0 & 4,-4 \\
\text{Odd} & 1/2 & 0,0 & 0,0 & 0,0 \\
 & 1 & -4,0 & 0,0 & -4,4
\end{array} \tag{1.8}
$$

In this *extended* game, there is a Nash Equilibrium, which is $(1/2, 1/2)$.

Let's try to formalize this. First, given a non-empty discrete set A, a **probability distribution** over A is a function $p\colon A \to [0,1]$ satisfying the normalization constraint:

$$
\sum_{x \in A} p(x) = 1 \tag{1.9}
$$

A specific probability distribution can be represented as a vector $\boldsymbol{p} = (p(a_1), \ldots, p(a_n))$, where $a_1, \ldots, a_n$ are all the elements of A. Since the entries of $\boldsymbol{p}$ sum to 1 (due to (1.9)), $\boldsymbol{p}$ belong geometrically to a **simplex**, which is denoted by $A \in \Delta A$.

Mixed strategy definition

Then, for a normal form game $(S_1, \ldots, S_n; u_1, \ldots, u_n)$, a **mixed strategy** for player i is defined to be a **probability distribution** $m_i\colon S_i \to [0,1]$ over the set S_i.
Expanding the notation, this means that i chooses strategies in $S_i = (s_i^{(1)}, \ldots, s_i^{(n)})$ with probabilities $(m_i(s_i^{(1)}), \ldots, m_i(s_i^{(n)}))$. In other words, a player "makes use of a random number generator" for choosing their strategy[1].

Expected utility

The **expected** utility for player i is then a function over *all the probability distribution of all players*, i.e. $\Delta S_1 \times \cdots \times \Delta S_n$, mapping a vector of distributions $(m_1, \ldots, m_n)$ to a number:

$$
u_i(m_1, \ldots, m_n) = \sum_{s \in S} m_1(s_1) \cdot m_2(s_2) \cdot \cdots \cdot m_n(s_n) \cdot u_i(s)
$$

where $u_i\colon S \equiv S_1 \times \cdots \times S_n$ is the utility function for player i. In other words, we are summing the utilities for *all possible strategy profiles*, weighted by the likelihood that the players actually play them.
Note that here we are assuming all choices to be **independent** (as it happens in a static game of complete information).

Now, let's consider again the Odds & Evens game, allowing players to play mixed strategies. Suppose that Odd plays 0 with probability q, and Even plays 0 with probability r. Then the *weighted* payoffs are:

$$
\mathbf{P} = \begin{array}{cc|cc|}
 & & \text{Even} & \\
 & & \begin{matrix}0\\(\text{prob } r)\end{matrix} & \begin{matrix}1\\(\text{prob } 1-r)\end{matrix} \\
\hline
\text{Odd} & \begin{matrix}0\\(\text{prob } q)\end{matrix} & -4qr, 4qr & 4q(1-r), -4q(1-r) \\
 & \begin{matrix}1\\(\text{p. } 1-q)\end{matrix} & 4(1-q)r, -4(1-q)r & -4(1-q)(1-r), 4(1-q)(1-r)
\end{array}
$$

[1] ∧How can this be *rational*? Imagine like this: a player makes a choice by drawing a card from a deck (e.g. in Magic the Gathering). They cannot *choose* which card to draw, but they (rationally) constructed the deck so that the probability distribution of the drawn card is *determined* by them.

This is the representation of a **single** strategy profile $\boldsymbol{m} = (m_1, m_2) = (q, r)$ (it is not the normal form!). The expected payoff for Odd is:

$$\langle u_{\text{Odd}} \rangle = -16qr + 8q + 8r - 4 = -4(2q-1)(2r-1)$$

For even, observe that all payoffs are the same but with opposite sign, and so $\langle u_{\text{Even}} \rangle = -\langle u_{\text{Odd}} \rangle$.

The *normal form* for the game would be an *infinite* matrix with the expected payoffs for all values of (q, r) ranging from $(0, 0)$ to $(1, 1)$. Basically it is the same thing as (1.8), with *all other possible intermediate states.*

Pure strategies as degenerate mixed strategies

Degenerate mixed strategies (i.e. with the probability of one strategy being 1, and all the others 0) are **pure strategies**.
Formally, we define the **support** of a mixed strategy $m_i \in \Delta S_i$ as the set of strategies with non-zero probability: $\{s_i \in S_i \colon m_i(s_i) > 0\}$. A pure strategy is just a mixed strategy with *singleton* support. So, $s_i \in S_i$ can be identified with m_i such that $p(s_i) = 1$ and $p(s_j) = 0$ for all $s_j \neq s_i$ (degenerate probability distribution).

We are now ready to **generalize** all the previous definitions.
Consider a game $\mathcal{G} = \{S_1, \dots, S_n; u_1, \dots, u_n\}$.

Strict/Weak domination

- If $m_i', m_i \in \Delta S_i$, m_i' **strictly dominates** m_i if:

$$u_i(m_i', \boldsymbol{m_{-i}}) > u_i(m_i, \boldsymbol{m_{-i}}) \qquad \forall \boldsymbol{m_{-i}}$$

- We say that m_i' **weakly dominates** m_i if:

$$\begin{aligned} u_i(m_i', \boldsymbol{m_{-i}}) &\geq u_i(m_i, \boldsymbol{m_{-i}}) \qquad \forall \boldsymbol{m_{-i}} \\ \exists \boldsymbol{m_{-i}} \text{ s.t. } u_i(m_i', \boldsymbol{m_{-i}}) &> u_i(m_i, \boldsymbol{m_{-i}}) \end{aligned}$$

So, generalization is immediate: we need just to replace pure strategies with distributions (mixed strategies), and $S_i \to \Delta S_i$.
However, verifying these inequalities becomes much more complex, since now:

$$\boldsymbol{m_{-i}} \in \Delta S_1 \times \cdots \times \Delta S_{i-1} \times \Delta S_{i+1} \times \cdots \times \Delta S_n$$

which is a continuous set, and so the possible values to check are infinite!

Fortunately, we can *restrict* $\boldsymbol{m_{-i}}$ to just pure strategies, achieving the same results:

- If $m_i', m_i \in \Delta S_i$, m_i' **strictly dominates** m_i if:

$$u_i(m_i', \boldsymbol{s_{-i}}) > u_i(m_i, \boldsymbol{s_{-i}}) \qquad \forall \boldsymbol{s_{-i}} \in S_{-i}$$

- We say that m_i' **weakly dominates** m_i if:

$$\begin{aligned} u_i(m_i', \boldsymbol{s_{-i}}) &\geq u_i(m_i, \boldsymbol{s_{-i}}) \qquad \forall \boldsymbol{s_{-i}} \in S_{-i} \\ \exists \boldsymbol{s_{-i}} \in S_{-i} \text{ s.t. } u_i(m_i', \boldsymbol{s_{-i}}) &> u_i(m_i, \boldsymbol{s_{-i}}) \end{aligned}$$

In other words, we can limit our search to only pure strategies of the opponents. Intuitively, this is possible because u_i is a linear combination of the utilities computed over pure strategies, since it is an *expected* value. So, if an equality holds for all pure strategies, it holds for all mixed strategies too.

Generalized Nash Equilibrium

Then, a joint mixed strategy $\boldsymbol{m} \in \Delta S_1 \times \cdots \times \Delta S_n$ is said to be a Nash Equilibrium if all players are choosing best replies:

$$\forall i,\ u_i(\boldsymbol{m}) \geq u_i(m_i', \boldsymbol{m_{-i}}) \qquad \forall m_i' \in \Delta S_i$$

Applying this definition to the Odd/Even game, recall that the payoff for Odd is $-4(2q-1)(2r-1)$, and the opposite for Even. If $q = 1/2$ or $r = 1/2$, **both** players have payoff of 0. If they *both* play 1/2, then no one has an incentive to change the action, because any unilateral change leads to the *same* payoff of 0 for both.

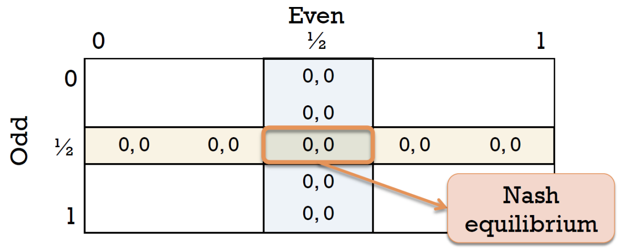

Figure (1.23) – Nash Equilibrium for the Odd/Even game.

Exercise 1.7.1:

> Prove that $(1/2, 1/2)$ is the **only** Nash Equilibrium of the Odd/Even game. *Hint*: consider three cases where the payoff of player Odd is < 0, > 0 or $= 0$, but the joint strategy is not $(1/2, 1/2)$. Show that in all of them there is a player having an incentive in changing strategy.

1.8 IESDS and Mixed Strategies

In practice, we have shown that Iterated Elimination of Strictly Dominated Strategies *selects* a Nash Equilibrium of a game. However, if there are no strictly dominated strategies, this method can not be applied.
For example, consider the following game:

player A \ player B	L	C	R
T	7, 4	5, 0	8, 1
D	6, 0	3, 4	9, 1

Figure (1.24) – No strategy strictly dominates another, so no elimination can happen in IESDS.

Fortunately, mixed strategies can solve these problems.
In this case, focus on strategy R. According to the previous definitions, it is not dominated by L or C, since there the payoff could be $0 < 1$. However, it seems *inefficient* to play R, since the expected payoff given by L and C is bigger. For example, we can consider the strategy $m = 1/2L + 1/2C$, for which the expected gain $u_B = 2$ regardless of A's move, which is *strictly more* than 1, the gain obtained by playing R.
Thus, if we allow mixed strategies, R is strictly dominated by m, and so we can eliminate it:

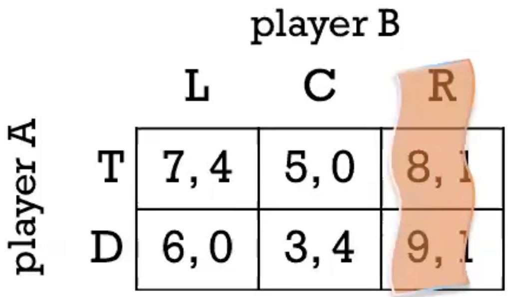

Figure (1.25) – R is dominated by $m = 1/2L + 1/2C$.

From there, we can make further eliminations. First, T now dominates D, and so D can be eliminated. Then, L dominates C. The only remaining strategy profile is the Nash Equilibrium:

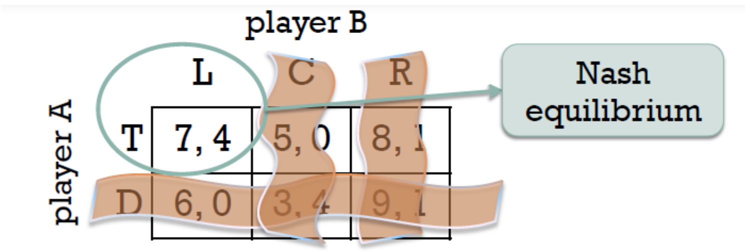

Figure (1.26) – Nash Equilibrium with mixed strategies.

In general, we can extend the theorems about IESDS to IESDSm (IESDS + mixed strategies). In particular:

- **Theorem**. Nash equilibria survive IESDSm, i.e. they are *selected* by the method.

- **Theorem**. The order of eliminations in IESDSm is irrelevant.

Note that here we are always using **strict** dominance, not **weak**, which guarantees the order independence. In fact, weakly dominated strategies can be Nash equilibria or part of the support of a Nash equilibrium.

An interesting result is that, if *all players* do not have any incentive to unilaterally change their play to another **pure** strategy, then they are playing a Nash Equilibrium (and clearly viceversa).

Theorem 1.8.1. *Given a game $\mathcal{G} = \{S_1, \ldots, S_n; u_1, \ldots, u_n\}$, and a joint mixed strategy m, the following statements are equivalent:*

- *The joint mixed strategy $\boldsymbol{m}$ is a Nash equilibrium*
- *For each i:*

$$\begin{aligned} u_i(\boldsymbol{m}) &= u_i(s_i, \boldsymbol{m_{-i}}) \qquad \forall s_i \in \text{support}(m_i) \\ u_i(\boldsymbol{m}) &\geq u_i(s_i, \boldsymbol{m_{-i}}) \qquad \forall s_i \notin \text{support}(m_i) \end{aligned}$$

In particular, if a player deviates from Nash Equilibrium to a pure strategy, their payoff does not change (and so they have no incentive to deviate in the first place), and in particular it does not get worse. This theorem explains that the situation of fig. 1.23, with the row and column full of zeros merging in the middle, is actually general, not a coincidence.

Again, the intuition comes from the fact that $u_i(\boldsymbol{m})$ is a linear combination of the payoffs given by the pure strategies in the *support* of $\boldsymbol{m}$:

$$\mathbb{E}_{\boldsymbol{m}}[u_i] = \sum_{a_k \in S_i} \underbrace{\mathbb{E}[u_i | s_i = s_1]}_{u_i(s_i, \boldsymbol{m_{-i}})} \mathbb{P}[s_i = a_k]$$

Now, the *weight* $\mathbb{P}[s_i = a_k]$ are controlled by player i, who seeks to maximize $\mathbb{E}[u_i]$. Assuming they play a best reply m_i, then we expect that any *unilateral* change does not increase $\mathbb{E}[u_i]$. Suppose that there is a unilateral change that *decreases* $\mathbb{E}[u_i]$. This means that, only by changing the weights, we can make u_i lower. So, there must be a *worse* pure strategy for i, $u_i(s_i, \boldsymbol{m_{-i}}) < u_i$, which we are giving more weight. However, probability is conserved: all weights must some to 1. So, the weight we are giving to this worse strategy, is weight that we are subtracting from some *better* strategy. But this means that we are not playing the best possible reply!
In other words, if a perturbation of m_i can decrease u_i, then there is an *opposite* perturbation that can increase it, and viceversa. So, if u_i cannot increase given a perturbation, it cannot decrease too, meaning that it must remain constant.
For example, suppose $s_i \in \{1, 2\}$, and:

$$u_i = w_1 \cdot 1 + w_2 \cdot 10 \qquad w_1 + w_2 \stackrel{!}{=} 1$$

Starting with $(w_1, w_2) = (1/2, 1/2)$, we can see that going to $(3/4, 1/4)$ lowers u_i, and conversely $(1/4, 3/4)$ rises it. But if we assume that u_i cannot be made bigger, the only way is to have something like this:

$$u_i = w_1 \cdot 5 + w_2 \cdot 5$$

This theorem shows also that the Nash Equilibria found from the previous definition (pure strategies) are *compatible* with this new definition (mixed strategy). In other words, if s^* is a NE, also the corresponding (degenerate) m^* will be a NE.

1.8.1 Examples

Recall the game of the "battle of sexes":

$$
\mathbf{P} = \begin{array}{cc|c} & & \text{Brian} \\ & & \begin{array}{cc} \text{R} & \text{S} \end{array} \\ \text{Ann} & \begin{array}{c} \text{R} \\ \text{S} \end{array} & \left\| \begin{array}{cc} \boxed{2,1} & 0,0 \\ 0,0 & \boxed{1,2} \end{array} \right\| \end{array}
$$

This game has two *pure* NEs: (R, R) and (S, S).
We can show that there is also a mixed NE. Suppose Ann (or Brian) plays R with probability q (or r). Then their expected payoffs are:

$$u_A(q, r) = 2qr + (1-q)(1-r) \qquad u_B(q, r) = qr + 2(1-q)(1-r) \tag{1.10}$$

Assume (a, b) is a mixed NE. Since it is not pure, the strategies are combinations of both R and S, meaning that the support of both a and b is $\{R, S\}$. Then, from the theorem:

$$u_A(a, b) = u_A(0, b) = u_A(1, b)$$

Inserting (1.10) we get:

$$2ab + (1-a)(1-b) = 1 - b = 2b$$

leading to $b = 1/3$. From there, or by using $u_B(a, 0) = u_B(a, 1)$, we can then find $a = 2/3$. The asymmetry is given by the fact that they each have a different *preference* for the positive outcomes.

1.9 Nash theorem

Do mixed strategy allow finding *at least* one Nash equilibrium in any game? The answer is affirmative, as it was proved by Nash (1950).

Theorem 1.9.1. *Every game with finite S_i has* ***at least one*** *Nash equilibrium (possibly involving mixes strategies).*

(But actually finding it/them could be difficult).

Before *proving* this theorem, it is worth gaining some more understanding on mixed strategies. Are they just a *mathematical postulate* needed for the nice property of having Nash equilibria, or is there some physical meaning? Surely we do not expect rational players to *follow* the toss of a coin!
Fortunately, there are several possible interpretations:

- **Iterated games**. Suppose a game is repeated M times, forgetting everything after each iteration (*memoryless* process). Here, playing a mixed strategy means *fixing* the frequency of each pure strategy during the whole run. So, at each specific game the choice is randomized, but in the long run there is a very clear trend.

  The problem in this interpretation is that, in reality, iterated games *do involve* memory, and so strategies should take into account past iterations.

- **Fuzzy values**. Mixed strategies encode *lack of decision*, i.e. the need to avoid *sticking to just one choice* because there is some intrinsic randomness in the game.

- **Beliefs**. In modern game theory, the uncertainty of mixed strategies reflects the uncertainty the opponent has about the choice of the other player. For example, in the Odd/Even game, at any moment each player *knows* which (pure) move they are going to make. However, they act so that their opponent *cannot anticipate* their beliefs. In other words, there is no way for player B to predict what A will play, and *viceversa*, even though they both have in mind some particular, specific and concrete move. Nonetheless, they both know that they *don't know*, which is sufficient to find a strategy to reach equilibrium.

(Lesson 8 of 23/10/20) Compiled: January 1, 2021

Different definitions for strategies

In summary, at the start we interpreted *strategies* as just *actions*, which was effectively a bit redundant. Truly, strategies should be interpreted as *plans of actions*. The distinction becomes clear for mixed strategies: an action can't be random, yet a strategy can involve randomness! But while I know that my opponent will choose an action (a *pure strategy*), if I do not have any way to predict it perfectly, then I may still form some *belief* about its action, which is a probability distribution. So, players *actually perform* pure strategies, but *assume* that each other are playing *mixed strategies*.

This *change* of definition may appear unintuitive, and confusing. However, it was intended from the start. The idea is similar to that of real numbers. Values with infinite non-repeating digits are not intuitive, but they are needed to find solutions for certain equations, and that's why we discuss them. Yet, to *ease* their introduction, we start by explaining *integers*, which are a "special case" of real numbers which well agrees with our innate numerical intuition. We can think about *mixed strategies* as strategies "in-between" pure strategies, as real numbers lie "in-between" integers.

Introduction

Before *proving* Nash theorem, we recall the previous definitions of beliefs and best-replies, which are the "ingredients" for finding Nash equilibria, and explicitly *restate* them in the general case with mixed strategies. Then, we will compute a *mixed* Nash equilibrium in an example, and introduce a visualization that will give some intuition about the proof.

Belief

First, a **belief** of player i is a possible profile of their opponents' strategies, i.e. an element of the set ΔS_{-i} (all strategy profiles, *except* that of the current player i).

Generalized best responses

A **best-response** is a *correspondence*, i.e. a "mapping that can be one to many", $\text{BR}\colon \Delta S_{-i} \to \rho(\Delta S_i)$, where $\rho(\Delta S_i)$ is the set of *parts* of ΔS_i, i.e. its elements are *subsets* of ΔS_i. Thus, BR associates an opponents' strategy $\boldsymbol{m_{-i}} \in \Delta S_{-i}$ to a subset $\text{BR}(\Delta S_{-i}) \subseteq \Delta S_i$ of the i-th player's strategies ΔS_i. If this were a *function*, i.e. a one-to-one mapping, then $\text{BR}(\Delta S_{-i})$ would always be a singleton, i.e. a set containing just one element.
This correspondence is such that each $m_i \in \text{BR}(\boldsymbol{m_{-i}})$ is a *best response* to $\boldsymbol{m_{-i}}$, that is a "good counter-action" to $\boldsymbol{m_{-i}}$.

For example, consider the following game:

$$\mathbf{P} = \begin{array}{cc} & \text{Bea} \\ & \begin{array}{cc} F\,(r) & G\,(1-r) \end{array} \\ \text{Art} \begin{array}{c} U\,(q) \\ D\,(1-q) \end{array} & \left\| \begin{array}{cc} 6,1 & 0,4 \\ 2,5 & 4,0 \end{array} \right\| \end{array} \tag{1.11}$$

Here there is no clear "best strategy" to choose. Then, Bea ignores what Art will play, but being rational she assumes he will play U with probability q. Similarly, Art thinks Bea will play F with probability r.

Suppose that, for some reason, Bea is known for always playing F, meaning that $r = 1$. In this case, Art's best response is clearly U, which means $q = 1$. Let's generalize this to a generic fixed r, to see what is the best response available for Art. His expected utility for option D is:

Example: mixed Nash equilibrium

$$\underbrace{u_A(D, r)}_{\substack{\text{Utility for Art} \\ \text{playing } D}} = \underbrace{2r}_{\text{Bea plays } F} + \underbrace{4(1-r)}_{\text{Bea plays } G}$$

And similarly, for U we have:

$$u_A(U, r) = 6r$$

We can compare these two quantities:

$$U_A(D, r) = -2r + 4 > 6r = u_A(U, r) \Leftrightarrow r < \frac{1}{2}$$

So, if $r < 1/2$, D is the best response, and if instead $r > 1/2$, U is the best response. If $r = 1/2$, the two are equivalent.
We have found that Art's best response *depends* on the parameter r, meaning that the probability q that Art will play U, depends on r, and we denote it with $q^*(r)$. From the previous analysis:

$$q^*(r) = \begin{cases} 0 & r < 1/2 \\ 1 & r > 1/2 \end{cases}$$

Similarly, this same computation can be done from Bea's point of view, obtaining:

$$u_B(q, F) = q + 5(1-q) \qquad u_B(q, G) = 4q \Rightarrow r^*(q) = \begin{cases} 1 & q < 5/8 \\ 0 & q > 5/8 \end{cases}$$

Now, a Nash Equilibrium is a situation in which each player has chosen the best response to the other player's choice. So, it is the value (q, r) that belongs to **both** graphs $q^*(r)$ and $r^*(q)$, i.e. to their intersection. Visually, if we choose q and r as axes, $q^*(r)$ and $r^*(q)$ will appear as:

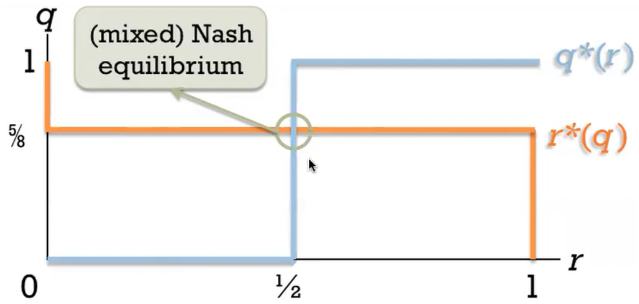

Figure (1.27) – Mixed Nash equilibrium as the intersection of the best response functions of the two players.

In particular, note that *all points* at $r = 1/2$ (or at $q = 5/8$) are best responses for Bea (or Art).

Figure 1.27 gives an intuition about one way to *prove* Nash theorem! The idea is that each graph *must* change from 0 to 1, otherwise one of the two strategies is a constant 0, and so it is strictly dominated and not viable. But then, there must be an intersection between the two, since they are *continuous lines*, meaning that they *traverse* all points in $[0, 1]$. *Proof: intuition*

Note that there could be more complex situations, where the lines go "up and down" multiple times. In these cases there are multiple intersections, and so multiple Nash equilibria. For example:

$$\mathbf{P} = \begin{array}{cc} & \text{Brian} \\ & \begin{array}{cc} R & S \end{array} \\ \text{Ann} \begin{array}{c} R \\ S \end{array} & \left\| \begin{array}{cc} 2,1 & 0,0 \\ 0,0 & 1,2 \end{array} \right\| \end{array}$$

Here:

$$u_A(R, r) = 2r,\ u_A(S, r) = 1 - r \Rightarrow q^*(r) = 1 - h(r - 1/3)$$
$$u_B(q, R) = q,\ u_B(q, S) = 2(1 - q) \Rightarrow r^*(q) = 1 - h(q - 2/3)$$

and the graphs look like this:

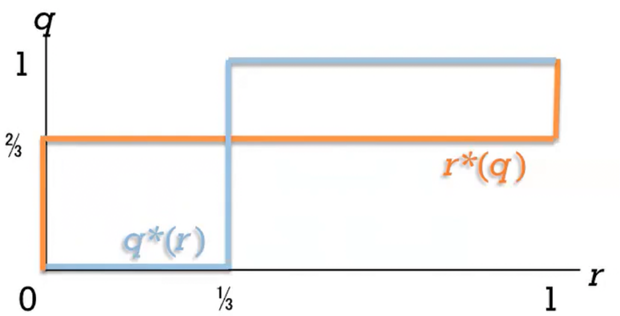

Figure (1.28)

There are 3 intersections (Nash equilibria): $(r, s) = (0, 0)$, $(1/3, 2/3)$ and $(1, 1)$.

Proof. We are now ready to provide a formal proof. Consider a game G written in normal form $G = \{S_1, \ldots, S_n; u_1, \ldots, u_n\}$. The *best response* of player i is the correspondence $\mathrm{BR}_i\colon \Delta S_1 \times \cdots \times \Delta S_{i-1} \times \Delta S_{i+1} \times \cdots \times \Delta S_n \to \rho\Delta S_i$, where $\rho\Delta S_i$ is the set of parts (subsets) of ΔS_i, and $\mathrm{BR}_i(\boldsymbol{m_{-i}}) = \{m_i \in \Delta S_i\colon u_i(m_i, \boldsymbol{m_{-i}})$ is maximal$\}$. For example in (1.11) $\mathrm{BR}_{\mathrm{Art}}(r = 2/3) = \{q^*(2/3)\} = \{1\} \equiv \{U\}$.

Note that BR_i are *sets*. So, if we want to pick a best response *for each player*, we can consider a vector $\boldsymbol{m} = (m_1, \ldots, m_n)$ where $m_i \in \mathrm{BR}_i$. In general, each BR_i can have more than 1 element, and so *all possible combinations* of elements *chosen from each set* are contained in the *cartesian product* of these sets. We denote it as $\mathbf{BR}\colon \Delta S \to \rho\Delta S$, defined by:

$$\mathbf{BR}(\boldsymbol{m}) = \mathrm{BR}_1(\boldsymbol{m_{-1}}) \times \cdots \times \mathrm{BR}_n(\boldsymbol{m_{-n}})$$

In other words, given a strategy $\boldsymbol{m}$ for all players, the elements of $\mathbf{BR}(\boldsymbol{m})$ are global strategies in which each player i is choosing their best response to the opponents actions $\boldsymbol{m_{-i}}$. But if everybody is already playing a best response $\boldsymbol{m}$, i.e. they are in a Nash equilibrium, then $\boldsymbol{m} \in \mathbf{BR}(\boldsymbol{m})$. Thus:

$$\boldsymbol{m} \text{ is a Nash Equilibrium } \Leftrightarrow \boldsymbol{m} \in \mathbf{BR}(\boldsymbol{m})$$

Following again the example from (1.11), recall that at equilibrium $(r^*, q^*) = (1/2, 5/8)$. Note that, if $r = 1/2$, then Art can play *any move* (or mixed strategy) and his expected utility will not change. The same happens for Bea when $q = 5/8$. Thus, in this case $\mathbf{BR}((r^*, q^*))$ consists of *all* the possible mixed strategies available by both players! This, of course, includes (r^*, q^*).

We can think as $\mathbf{BR}(\boldsymbol{m})$ as a "transformation", in which each player is *improving* their action. Then, the Nash Equilibrium is a **fixed point** of $\mathbf{BR}$: in this case there is no way for any single player to "improve" their action anymore, and so the *best possibility* is to *keep playing the same move.*

Now, any player has at least a best response, and so $\mathrm{BR}_i(\boldsymbol{m_{-i}})$ is always non-empty. Moreover, it must contain at least a pure strategy. In fact, we have seen before that mixed strategies must be "intermediate" between (i.e. linear combination of) two best responses.

An important existence theorem regarding fixed points is **Brouwer's Fixed Point Theorem**. In its simplest form, it states that if $f(x)$ is a continuous function from a closed real interval $\mathcal{I}$ to itself ($\mathcal{I} \to \mathcal{I}$), then there exists a fixed point: $\exists x^* \in \mathcal{I}$ s.t. $f(x^*) = x^*$.
This can be seen as a generalization of Bolzano-Weierstrass theorem. For example, consider the interval $\mathcal{I} = [0, 1]$. Searching for a x^* s.t. $f(x^*) = x^*$ is the same as searching for an intersection with the line $y = x$. Since $f(x) \in [0, 1]$ (because $f\colon \mathcal{I} \to \mathcal{I}$), there are two possibilities:

- $f(0) = 0$ or $f(1) = 1$ (or both). Immediately we have $x^* = 0$ or $x^* = 1$.
- If the previous option does not hold, then $f(0) > 0$, and $f(1) < 1$. Subtracting x, this means that $f(0) - 0 > 0$, and $f(1) - 1 < 0$. So, the

function $f(x) - x$ is positive at $x = 0$, and negative at $x = 1$. Since it is continuous, by the Bolzano-Weierstrass theorem, there must be a $x^* \in (0, 1)$ s.t. $f(x^*) - x^* = 0 \Rightarrow f(x^*) = x^*$, which confirms the theorem.

However, we can't directly apply this result to our case: we are not working on real numbers, but on strategy sets, and we do not have functions, but correspondences (i.e. one-to-many). So, we need a *generalization* of this theorem, which goes under the name of **Kakutani's Fixed Point Theorem**. It states that:

- *If* A is a non-empty, compact, convex subset of $\mathbb{R}^n$
- *If* the correspondence $F\colon A \rightrightarrows A$ is such that:
  - For all $\boldsymbol{x} \in A$, $F(\boldsymbol{x})$ is non-empty and convex
  - If $\{\boldsymbol{x_i}\}$, $\{\boldsymbol{y_i}\}$ are sequences in $\mathbb{R}^n$ converging to $\boldsymbol{x}$ and $\boldsymbol{y}$ respectively, then $\boldsymbol{y_i} \in F(\boldsymbol{x_i}) \Rightarrow \boldsymbol{y} \in F(\boldsymbol{x})$ (i.e. F has a *closed graph*, meaning that it *contains its limit points*)

then there exists a fixed point $\boldsymbol{x}^* \in A$ such that $\boldsymbol{x}^* \in F(\boldsymbol{x}^*)$. (Note how we say $\in$ and not $=$, because we are dealing with *correspondences*!)

We will not prove this, but just observe that all the hypotheses hold for our case. This follows naturally from the fact that strategies are really probability distributions, i.e. sets of numbers that must sum to 1. For example, any combined strategy $\boldsymbol{m}$ in example (1.11) could be uniquely mapped to a tuple (q, r), with $0 \leq q, r \leq 1$. This forms a set $A \subset \mathbb{R}^n$ which is clearly non-empty, compact and convex (it is a *square*). Then $\mathrm{BR}_i(\boldsymbol{m_{-i}})$ is non-empty (there is always a best response), and convex (if it contains two pure strategies, then all the *mixed strategies in-between* are also best responses). The closed graph property is satisfied by observing that *all* points in the curve 1.27 belong to the graph. For example, consider a sequence ($\boldsymbol{x}$) $\{r_i\}$ converging to a $\bar{r} < 1/2$. The corresponding $\boldsymbol{y}$ is a sequence of 0s ($\mathrm{BR}_{\mathrm{Art}}(r_i) = \{0\}$) which clearly converges to $0 \ni \mathrm{BR}_{\mathrm{Art}}(\bar{r}) = \{0\}$. The same happens for a $\bar{r} > 1/2$, but now with $\mathrm{BR}_{\mathrm{Art}}(\bar{r}) = \{1\}$. The only interesting part to check is for a limit $\bar{r} = 1/2$. In this case, the r_i can be either from the left or from the right, leading to $y_i = \mathrm{BR}_{\mathrm{Art}}(r_i)$ that are either all 0s or all 1s. However, note that $\mathrm{BR}_{\mathrm{Art}}(1/2) = [0, 1]$ (i.e. 0, 1 and *all* the in-between values[2]), which contains both 0 and 1. So, the convergence property is satisfied.
Of course, the last two properties can be generalized to all players, i.e. to the full **BR**, satisfying the theorem's requirements. Then, this guarantees the existence of a fixed point for **BR**, which is nothing else than a Nash equilibrium. □

1.10 Finding Nash Equilibria with time

How to find Nash Equilibria in practice? One way is to add some kind of "time evolution" to the static game, which converges to the desired equilibrium.

[2] ∧Think of this part of the graph as an actual line. This is in fact permitted by BR (i.e. $q^*(r)$ in the example) not being a function, but a correspondence, meaning that it *can* have multiple values! If it were a function, we would have needed a *discontinuous jump*.

This is the idea of **fictitious play** (G.W. Brown, 1951), where we *allow players to change their moves*. In practice, we start with a global strategy $\boldsymbol{m}$, then choose a player i at random, who gets a chance to *play again*, performing a best response to all the other players' moves $\boldsymbol{m_{-i}}$, regarded as fixed. *Fictitious play*
We then proceed iteratively, *evolving* $\boldsymbol{m}$ by making random players *move again*, until a stationary state is achieved.

Note that this procedure *denies* "full rationality" of players. If everyone already has the *correct* beliefs about the others, then they will play the best possible response, and have no regrets, i.e. no incentive to change their move.

In any case, we can see that Nash equilibria are **absorbing states**. That is, at a Nash equilibrium $\boldsymbol{m}^*$, all players have no regrets, and so nobody will change their move, basically *ending* the evolution.

We can then ask if, by starting from *any* $\boldsymbol{m}$, we can end up in a Nash equilibrium.

This (unfortunately) does not happen. For example, there are games (such as Rock/Paper/scissors) in which the evolution of $\boldsymbol{m}$ is *cyclical*, and it reaches no fixed state.

However, fictitious play *does converge* in some cases:

- If the game can be solved by IESDS
- **Potential games**, which will explored in the next section.
- Other cases, such as $2 \times N$ games with generic payoffs.

The advantage of fictitious play is that it is a **distributed algorithm**, i.e. it does not require to synchronize all players' moves. It is then a good way to *implement* a program to find a Nash equilibrium.

1.10.1 Potential Games

Potential Games are a special class of static games of complete information that are "solvable" through fictitious play.
To define them, we first need some notation. Let's start with a game in normal form $G = \{S_1, \dots, S_n; u_1, \dots, u_n\}$, and denote $S \equiv S_1 \times \cdots \times S_n$. A function $\Omega \colon S \to \mathbb{R}$ is an (exact) **potential** for G if:

$$\Omega(s_i', \boldsymbol{s_{-i}}) - \Omega(s_i, \boldsymbol{s_{-i}}) = u_i(s_i', \boldsymbol{s_{-i}}) - u_i(s_i, \boldsymbol{s_{-i}}) = \Delta u_i$$

In other words, if player i changes their move from s_i to s_i', then they will have a difference in (expected) utility Δu_i of $\Omega(s_i', \boldsymbol{s_{-i}}) - \Omega(s_i, \boldsymbol{s_{-i}})$. Note that Ω is *the same* for every player, which is quite a strong requirement.
In practice, it suffices that the difference in potential is *proportional* to Δu_i, with a coefficient $w_i > 0$ that is player-specific. In this case, we call $\Omega \colon S \to \mathbb{R}$ a **weighted potential**:

$$\Omega(s_i', \boldsymbol{s_{-i}}) - \Omega(s_i, \boldsymbol{s_{-i}}) = w_i \Delta u_i \qquad \boldsymbol{w} = \{w_i > 0\}$$

Even more in general, we can consider a **ordinal potential** $\Omega\colon S \to \mathbb{R}$ for G, which encodes just the *ordering* of utilities, i.e. such that $\Delta\Omega$ is a monotonic difference of Δu_i, but not necessarily linear:

$$\Omega(s_i', \boldsymbol{s_{-i}}) > \Omega(s_i, \boldsymbol{s_{-i}}) \Leftrightarrow u_i(s_i', \boldsymbol{s_{-i}}) > u_i(s_i, \boldsymbol{s_{-i}})$$

A **potential game** is a game G that admits a **potential**. Local maxima of the potential are *equilibria* for the game.

So, the main idea is that if $G = \{S_1, \dots, S_n; u_1, \dots, u_n\}$ has an ordinal potential Ω, it is immediate that its set of Nash equilibria is the same of $G' = \{S_1, \dots, S_n; \Omega, \dots, \Omega\}$. In other words, all the players want to maximize the *same* potential, reducing a multi-utility optimization to a single-goal optimization.

As an example, note that the Prisoner's Dilemma has an **exact potential**:

$$\mathbf{P} = \begin{array}{cc} & \text{Bob} \\ & \begin{array}{cc} M & F \end{array} \\ \text{Alice} \begin{array}{c} M \\ F \end{array} & \left\| \begin{array}{cc} -1,-1 & -9,0 \\ 0,-9 & -6,-6 \end{array} \right\| \end{array} \Rightarrow \textbf{Potential} = \begin{array}{cc} & \text{Bob} \\ & \begin{array}{cc} M & F \end{array} \\ \text{Alice} \begin{array}{c} M \\ F \end{array} & \left\| \begin{array}{cc} 0 & 1 \\ 0 & 4 \end{array} \right\| \end{array} \tag{1.12}$$

Another example is Cournot oligopoly, for which there is an ordinal potential function:

$$u_i(q_i, q_j) = q_i(a - q_i - q_j - c) \Rightarrow \Omega(q_i, q_j) = q_i q_j (a - q_i - q_j - c)$$

Theorem 1.10.1. *Every finite ordinal potential game has (at least) a* ***pure*** *strategy Nash equilibrium (which can be found deterministically).* *Potential games and pure Nash equilibria*

Proof. The idea is to use fictitious play. At each move, players maximize their utility, and so the potential. Since the strategies are finitely many, a local maximum will be reached, which is a (pure) Nash equilibrium. □

Other kinds of games that allow a potential are **congestion games**, where we try to *minimize congestion*, or (anti)**coordination games**, where higher payoffs are assigned to players that make (or not) the same choice (so, there is some sort of *common goal*).

In fact, it can be shown that any potential game is a sum of a **pure coordination** game and a **dummy game**. Here, a **dummy** (or pure externality) game is such that for all $\boldsymbol{s_{-i}}$, $u_i(s_i, \boldsymbol{s_{-i}}) = u_i(s_i', \boldsymbol{s_{-i}})$, i.e. where the payoff of each player i only depends on the strategies $\boldsymbol{s_{-i}}$ of the opponents.
For example, for the Prisoners' Dilemma:

$$\mathbf{P} = \begin{array}{cc} & \text{Bob} \\ & \begin{array}{cc} M & F \end{array} \\ \text{Alice} \begin{array}{c} M \\ F \end{array} & \left\| \begin{array}{cc} -1,-1 & -9,0 \\ 0,-9 & -6,-6 \end{array} \right\| \end{array} = \underbrace{\begin{array}{cc} & \text{Bob} \\ & \begin{array}{cc} M & F \end{array} \\ \text{Alice} \begin{array}{c} M \\ F \end{array} & \left\| \begin{array}{cc} -1,-1 & 0,0 \\ 0,0 & 3,3 \end{array} \right\| \end{array}}_{\text{Coordination}} + \underbrace{\begin{array}{cc} & \text{Bob} \\ & \begin{array}{cc} M & F \end{array} \\ \text{Alice} \begin{array}{c} M \\ F \end{array} & \left\| \begin{array}{cc} 0,0 & -9,0 \\ 0,-9 & -9,-9 \end{array} \right\| \end{array}}_{\text{Dummy}}$$

In this case, players would want to collaborate (coordination part), but a player confessing gives a huge penalty (-9) to the opponent (dummy part). This *shifts* the Nash Equilibrium to the "suboptimal" (i.e. Pareto inefficient) (M, M).

1.10.2 Computational Complexity

While at least a Nash Equilibrium must exist, actually finding it could be difficult.
There are specific cases where efficient algorithms (such as fictitious play) can be used. However, in general, it is a **hard problem**, as demonstrated by Papadimitriou et al.
More precisely, it belongs to a complexity class called PPAD, which is somehow intermediate between P and NP.

1.11 Exercises

(Lesson 9 of 26/10/2020) Compiled: January 1, 2021

Exercise 1.11.1 (Find Nash Equilibria)**:**

Consider the following static games of complete information played by A and B, where the normal-form representation of the game is given. For all of them, find the entire set of Nash equilibria.

We start from:

$$\mathbf{P} = \begin{array}{cc} & \text{B} \\ & \begin{array}{cc} f & g \end{array} \\ \text{A} \begin{array}{c} F \\ G \end{array} & \begin{Vmatrix} 2,4 & 0,1 \\ 1,6 & 3,5 \end{Vmatrix} \end{array} \tag{1.13}$$

Note that g is a strictly dominated strategy for player B, because $4 > 1$ and $6 > 5$, and so it can be removed from the game. Then we proceed with IESDS, removing G for player A. The only remaining strategy, i.e. (F, f), is the Nash equilibrium.

Consider now:

$$\mathbf{P} = \begin{array}{cc} & \text{B} \\ & \begin{array}{cc} f & g \end{array} \\ \text{A} \begin{array}{c} F \\ G \end{array} & \begin{Vmatrix} 0,4 & 3,0 \\ 6,0 & 0,5 \end{Vmatrix} \end{array} \tag{1.14}$$

Here IESDS can't be applied. In (F, f), A would want to play G, but here B would want to play g, and now A wants to change to F again, and then B goes back to f, cycling forever. This is similar to what happens in the odds/evens game, and in fact this is in the same category of **discoordination games**. So, there is no Nash equilibrium in pure strategies. However, the must be a Nash equilibrium in mixed strategies.

To find it, we can use the *indifference theorem*, which says that in a mixed Nash equilibrium, changing the *combination* of pure strategies in its support must not change the payoff. Denote with α the probability that A plays F. Given that, B must be indifferent when answering with f or g, because the support of B's mixed strategy is $\{f, g\}$ (a proper mixed strategy in this game must involve both pure strategies). So, equating the expected payoffs for playing f or g:

$$\alpha \cdot 4 + (1-\alpha) \cdot 0 \stackrel{!}{=} \alpha \cdot 0 + (1-\alpha) \cdot 5 \Rightarrow \alpha = \frac{5}{9}$$

Similarly, let β be the probability of B playing f, and repeating the same reasoning from the point of view of A, we get:

$$\beta \cdot 0 + (1-\beta) \cdot 3 \stackrel{!}{=} \beta \cdot 6 + (1-\beta) \cdot 0 \Rightarrow \beta = \frac{1}{3}$$

Moving on to the next game, consider:

$$\mathbf{P} = \begin{array}{cc} & \text{B} \\ & \begin{array}{cc} f & g \end{array} \\ A \begin{array}{c} F \\ G \end{array} & \left\| \begin{array}{cc} 9,3 & 2,2 \\ 0,0 & 3,9 \end{array} \right\| \end{array} \tag{1.15}$$

This is a **coordination game** (similar to the Battle of Sexes one), since the highest payoffs are the ones where both players choose the same move. Here there are two Nash equilibria. To find them, let's *highlight* the best responses available to each player. A considers F if B plays f, and G if B plays g. Similarly, B chooses f if A plays F, and g if A plays G:

$$\mathbf{P} = \begin{array}{cc} & \color{blue}{B} \\ & \begin{array}{cc} f & g \end{array} \\ \color{red}{A} \begin{array}{c} F \\ G \end{array} & \left\| \begin{array}{cc} \color{red}{9},\color{blue}{3} & 2,2 \\ 0,0 & \color{red}{3},\color{blue}{9} \end{array} \right\| \end{array}$$

The strategies for which both players are playing a best response are, by definition, Nash equilibria.

Moreover, there is a mixed Nash equilibrium, which can be found by setting indifference for player B:

$$u_B(\alpha, f) \stackrel{!}{=} u_B(\alpha, g) \Rightarrow 3\alpha = 2\alpha + 9(1-\alpha) \Rightarrow \alpha = \frac{9}{10}$$

And the analogous computations for player A lead to $\beta = 1/10$.

Finally, we consider:

$$\mathbf{P} = \begin{array}{cc} & \text{B} \\ & \begin{array}{cc} f & g \end{array} \\ A \begin{array}{c} F \\ G \end{array} & \left\| \begin{array}{cc} 2,2 & 0,6 \\ 6,0 & 1,1 \end{array} \right\| \end{array} \tag{1.16}$$

which is an example of Prisoners' dilemma. By highlighting the best responses (or proceeding through IESDS) we find that $(1,1)$ is the only pure Nash equilibrium.

Just to see what would happen, let's search for a mixed Nash equilibrium (which we expect not to be there). By setting indifference for A:

$$2\beta = 6\beta + 1 - \beta \Rightarrow \beta = -\frac{1}{3}$$

which is not acceptable, since $\beta \in [0,1]$ because it is a probability.

Exercise 1.11.2 (Mixed Nash equilibria)**:**

Consider the following static game of complete information played by A and B, where the normal-form representation is given below:

$$\mathbf{P} = \;\lessdot\; \begin{array}{c} \\ \\ X \\ Y \\ Z \end{array} \begin{array}{c} \text{B} \\ \begin{array}{cccc} J & K & L & M \end{array} \\ \left\| \begin{array}{cccc} 6,7 & 5,5 & 3,8 & 8,1 \\ 4,9 & 9,2 & 0,4 & 2,3 \\ 8,4 & 2,8 & 4,2 & 3,6 \end{array} \right\| \end{array} \tag{1.17}$$

1. Prove that there is no Nash equilibrium in pure strategies.
2. Prove that these (m_A, m_B) are Nash equilibria in mixed strategies:
   - $m_A = (2/3, 0, 1/3)$, $m_B = (5/11, 4/11, 2/11, 0)$
   - $m_A = (0, 4/11, 7/11)$, $m_B = (7/11, 4/11, 0, 0)$
3. List all the joint pure strategies that are Pareto optimal.

Solution.

1. One possibility is to go through all 12 strategies, and see that each one of them leaves some player with regrets. A quicker way is to highlight the best responses.

$$\mathbf{P} = \;\lessdot\; \begin{array}{c} \\ \\ X \\ Y \\ Z \end{array} \begin{array}{c} \text{B} \\ \begin{array}{cccc} J & K & L & M \end{array} \\ \left\| \begin{array}{cccc} 6,7 & 5,5 & 3,\boxed{8} & \boxed{8},1 \\ 4,\boxed{9} & \boxed{9},2 & 0,4 & 2,3 \\ \boxed{8},4 & 2,\boxed{8} & \boxed{4},2 & 3,6 \end{array} \right\| \end{array} \tag{1.18}$$

Since there is no entry in the table with both payoffs highlighted, there is no Nash equilibrium in pure strategies.

2. Note that finding mixed Nash equilibria in a such a game is, in general, difficult, given it is bigger than 2×2. One possibility would be to use the indifference principle. However, this requires to *known* beforehand the support of the mixed state we are searching. In the 2×2 case, there is only one possibility, with both pure states belonging to the support. However, here there are *more possible combinations* (for example, A mixing X and Y, and B mixing K and M, and so on...), and each should be checked independently. In general, many of them will have *no solution*. Thus, the process becomes tedious very quickly.

   Fortunately, in this case we are just required to *prove* that certain mixed strategies are equilibria. So, we *do know* the support of these mixed strategies.

   - Here m_A has support $\{X, Z\}$, and so we impose that X and Z give the same payoff against m_B. We also need to check that Y gives a $\leq$ payoff than the others (otherwise it would be a meaningful unilateral deviation for A).

     $$u_A(X, m_B) = 6 \cdot \frac{5}{11} + 5 \cdot \frac{4}{11} + 3 \cdot \frac{2}{11} + 8 \cdot 0 = \frac{56}{11}$$
     $$u_A(Z, m_B) = 8 \cdot \frac{5}{11} + 2 \cdot \frac{4}{11} + 4 \cdot \frac{2}{11} + 3 \cdot 0 = \frac{56}{11}$$
     $$u_A(Y, m_B) = 4 \cdot \frac{5}{11} + 9 \cdot \frac{4}{11} + 0 \cdot \frac{2}{11} + 2 \cdot 0 = \frac{56}{11}$$

     Analogously, we see that m_B has support $\{J, K, L\}$, meaning that these pure strategies all give the same payoff against m_A:

     $$u_B(J, m_A) = 7 \cdot \frac{2}{3} + 9 \cdot 0 + 4 \cdot \frac{1}{3} = \frac{18}{3} = 6$$
     $$u_B(K, m_A) = 5 \cdot \frac{2}{3} + 2 \cdot 0 + 8 \cdot \frac{1}{3} = 6$$
     $$u_B(L, m_A) = 8 \cdot \frac{2}{3} + 4 \cdot 0 + 2 \cdot \frac{1}{3} = 6$$
     $$u_B(M, m_A) = 1 \cdot \frac{2}{3} + 3 \cdot 0 + 6 \cdot \frac{1}{3} = \frac{8}{3} < 6$$

     And this completes the proof.

   - Proceeding as before, for A:

     $$u_A(Y, m_B) = 4 \cdot \frac{7}{11} + 9 \cdot \frac{4}{11} = \frac{64}{11}$$
     $$u_A(Z, m_B) = 8 \cdot \frac{7}{11} + 4 \cdot \frac{2}{11} = \frac{64}{11}$$
     $$u_A(X, m_B) = 6 \cdot \frac{7}{11} + 5 \cdot \frac{4}{11} = \frac{62}{11} < \frac{69}{11}$$

     And for B:

     $$u_B(J, m_A) = 9 \cdot \frac{4}{11} + 4 \cdot \frac{7}{11} = \frac{64}{11}$$

$$u_B(K, m_A) = 2 \cdot \frac{4}{11} + 8 \cdot \frac{7}{11} = \frac{64}{11}$$
$$u_B(L, m_A) = 4 \cdot \frac{4}{11} + 2 \cdot \frac{7}{11} = \frac{30}{11} < \frac{64}{11}$$
$$u_B(M, m_A) = 3 \cdot \frac{4}{11} + 6 \cdot \frac{7}{11} = \frac{54}{11} < \frac{64}{11}$$

3. In this case, we need to proceed by *enumeration.* To find the Pareto efficient outcomes, we start from the maximum payoffs of A and B, which happen in (Y, G), giving $(4, 9)$ and (Y, K), giving $(9, 2)$. These are Pareto optimal, since there is no way for each player to get a higher payoff.

   We can then *eliminate* all strategies that are dominated by these ones, i.e. with both payoffs $\leq$ than $(4, 9)$ or $(9, 2)$. For example, (Y, L) gives $(0, 4) < (4, 9)$, and so can be removed. Similarly, $(3, 8)$, $(8, 1)$, $(2, 3)$, $(2, 8)$, $(4, 2)$ and $(3, 6)$. We are left with $(8, 4)$, $(6, 7)$ and $(5, 5)$. The latter is dominated by $(6, 7)$, and so can be removed. However, there is no way to *improve* on $(8, 4)$ or $(6, 7)$ without some loss for a player, and so (X, J) and (Z, J) are Pareto efficient too.

   In summary, we find the following Pareto optimal strategies:

   - (Y, J) giving $(4, 9)$
   - (Y, K) giving $(9, 2)$
   - (X, J) giving $(6, 7)$
   - (Z, J) giving $(8, 4)$

Exercise 1.11.3 (Normal Form)**:**

Two students, Charlotte (C) and Daniel (D), need to write their MS Thesis. They need to choose (independently and unbeknownst to each other) a supervising professor. Three professors are available for this role: **X**avier, **Y**uan, and **Z**ingberry. The utility of a student is given by the amount of help he/she receives from the supervisor, which is quantified as 40 for Xavier, 60 for Yuan, 50 for Zingberry. However, if the two students select the *same* professor as their supervisor, they only get 70% of the utility that they would get if the professor had only one of them to supervise.

1. Write the game in normal form.
2. Find all the Nash equilibria of the game in pure strategies.
3. Find all the Nash equilibria of the game.

Solution.

1. The normal form is usually expressed as a table. Also, it can be useful to write the payoffs *diagonally*, so that the payoffs of the same player

are *aligned*:

$$
\mathbf{P} = \begin{array}{cc} & \mathrm{D} \\ & \begin{array}{ccc} X & Y & Z \end{array} \\ \mathrm{C} \; \begin{array}{c} X \\ Y \\ Z \end{array} & \left\Vert \begin{array}{ccc} {}^{28}{}_{28} & {}^{40}{}_{60} & {}^{40}{}_{50} \\ {}^{60}{}_{40} & {}^{42}{}_{42} & {}^{60}{}_{50} \\ {}^{40}{}_{60} & {}^{50}{}_{60} & {}^{35}{}_{35} \end{array} \right\Vert \end{array} \tag{1.19}
$$

Note that $0.7 \cdot 40 = 28$, $0.7 \cdot 60 = 42$ and $0.7 \cdot 50 = 35$.

One way is to highlight the best responses available to each player:

$$
\mathbf{P} = \begin{array}{cc} & \color{blue}{\mathrm{D}} \\ & \begin{array}{ccc} X & Y & Z \end{array} \\ \color{red}{\mathrm{C}} \; \begin{array}{c} X \\ Y \\ Z \end{array} & \left\Vert \begin{array}{ccc} {}^{28}{}_{28} & {}^{40}{}_{\color{blue}{60}} & {}^{40}{}_{50} \\ {}^{\color{red}{60}}{}_{40} & {}^{42}{}_{42} & {}^{\color{red}{60}}{}_{\color{blue}{50}} \\ {}^{50}{}_{40} & {}^{\color{red}{50}}{}_{\color{blue}{60}} & {}^{35}{}_{35} \end{array} \right\Vert \end{array} \tag{1.20}
$$

So two pure Nash equilibria are (Y, Z) and (Z, Y).
Note that the strategy X *for both players* is strictly dominated by the others, and so can be effectively removed from the game. The game restricted to $\{Y, Z\}$ is then an **anti-correlation** game, in which players receive a higher payoff when they choose *different* strategies.

Noting that we are dealing with an anti-correlation game, we know that there must be a third Nash equilibrium, this time belonging to mixed strategies. Its support must be $\{Y, Z\}$, since X is strictly dominated.

It can be found through the principle of indifference, by denoting with p the probability that C plays Y. For **symmetry**, this must be *equal* to the probability that D plays Y (exchanging the players does not change the game!), which further simplifies the computations. So:

$$
u_D(p, Y) \stackrel{!}{=} u_D(p, Z) \Rightarrow 42 \cdot p + 60 \cdot (1 - p) = 50 \cdot p + 35 \cdot (1 - p) \Rightarrow p = \frac{25}{33}
$$

So the Nash equilibrium is (p, p).

Exercise 1.11.4 (Continuous strategies)**:**

A strategic interaction takes place between a taxpayer T and the tax inspector I. T is supposed to pay a share S of its income to have a net income after paying taxes equal to R. However, T is considering two alternatives: hide part of the taxes (H) to get an additional dishonest income of L (so the tax paid is $S - L$ and the net income is $R + L$) or pay all due taxes in full (P). Player I also has two choices: check T for tax evasion (C) or do not check T (D). Performing a check has a cost equal to E. If the tax inspector discovers that T did not pay the tax, then the taxpayer will be fined and will have to pay an additional amount equal to F, that goes into the inspector. The probability of being discovered by a tax inspector after a check is p. The purpose of the taxpayer is to get the maximum possible amount of money. The goal of the tax inspector is to maximize the collected amount (minus the cost). Formalize this conflict in the form of a static game of complete information and find its Nash equilibria.

Solution.
We start by writing the game in normal form, equating the players' utilities to the money they earn. First, if the taxpayer is honest (P) and the inspector does not check (D), we have the simplest case, in which T gets R, and I gets S. If I does not check (D) and T *hides* the taxes (H), then T gets $R+L$, and I will receive less taxes, i.e. $S-L$. If I checks, but T is honest, then I will lose money, getting $S-E$, and T will still get R.

$$\mathbf{P} = \begin{array}{cc} & \begin{array}{cc} & \text{I} \\ C & D \end{array} \\ \text{T}\begin{array}{c} H \\ P \end{array} & \left\| \begin{array}{cc} R+(1-p)L-pF,\ S-E+pF-(1-p)L & R+L,\ S-L \\ R,\ S-E & R,\ S \end{array} \right\| \end{array} \tag{1.21}$$

The difficult part is when T is dishonest (H) and I checks (C). Here we need to compute the *expected* earnings, given the probability p of detecting the fraud. T has to pay F with probability p, but gets L with probability $1-p$. So his expected reward is $R+(1-p)L-pF$.
The situation is *opposite* for the inspector, who gets $S-E$ by default (due to the expense of checking), plus an additional $pF-(1-p)L$ on average.

Note that all payoffs of I include a term S, and all payoffs of T a term R. Given the fact that we are interested in *ordinal* payoffs, we can *remove* these constants to simplify a bit the normal form:

$$\mathbf{P} = \begin{array}{cc} & \begin{array}{cc} & \text{I} \\ C & D \end{array} \\ \text{T}\begin{array}{c} H \\ P \end{array} & \left\| \begin{array}{cc} (1-p)L-pF,\ -E+pF-(1-p)L & L,\ -L \\ 0,\ -E & 0,\ 0 \end{array} \right\| \end{array} \tag{1.22}$$

In other words, setting $R=S=0$ has *no* impact on Nash equilibria.

Depending on the values of the parameters, we have 3 possible *positions* for the Nash equilibrium.

1. If the effort of checking taxes is *too high* (precisely, $E > p(F+L)$), then there is no incentive for I to do it. As a consequence, T has the incentive to *hide taxes*, and the NE is (H,D). The inequality comes by imposing the strategy C to be strictly dominated by D. We already have $-E<0$, and so the remaining condition is:

$$-E+pF-(1-p)L < -L \Leftrightarrow E > p(F+L)$$

2. Suppose that the effort of checking is sufficiently low ($E < p(F+L)$). There are two possibilities:

- The probability of T getting caught is so low that the expected payoff of playing H is higher than that of playing P, regardless of the response. In other words, we impose that H strictly dominates P. $L > 0$ by hypothesis, and the other condition requires:

$$(1-p)L - pF > 0 \Rightarrow p < \frac{L}{L+F}$$

  In this case, the Nash equilibrium is (H, C): the taxpayer is always dishonest, and the inspector always checks.

- The probability of getting caught is sufficiently high ($p > L/(L+F)$), and the cost of checking is low ($E < p(F+L)$). In this case, the situation is that of a **discoordination game**, for which no pure Nash Equilibrium exists. Instead, we need to search for a mixed Nash equilibrium, in which T and I play H and C with respective probabilities:

$$\alpha = \frac{L}{p(L+F)}; \qquad \beta = \frac{E}{p(L+F)}$$

  Note that $\alpha, \beta \in [0,1]$, i.e. they are *meaningful probabilities*, only if the above conditions hold.

Chapter 2

Dynamic Games

(Lesson 11 of 03/11/2020) Compiled: January 1, 2021

A **dynamic game** involves time-dependent situations, in which players take **turns** to play.
As before, we will assume **complete information**, i.e. that everyone knows all the payoffs, and everybody knows that they know, and so on. In other words, the *rules of the game* are available to everyone, and all players act assuming that other players know those.
However, the time dimension adds another *quality* to information. A player that moves *after* another knows what they did play. This is very different from the situation we analyzed in static games, where players moved simultaneously, *without knowing beforehand their opponents' moves.* This distinction is formalized with the following terms:

Perfect vs Imperfect information

- **Perfect information**, when a player can make a decision with full awareness of the opponents' moves (e.g. in chess).
- **Imperfect information**, when some moves are simultaneous. In these cases, certain players may not have full awareness about other moves. This situation appears also when there are *external factors* that are non-controllable by players, such as moves "done by Nature".

Note that this is *independent* of **complete** information. That is, there are games of *complete perfect* information, and games of *complete imperfect* information.
As a starting example, consider the framework of the *Battle of Sexes* game, where Ann and Brian agreed to meet either at the romance (R) of the sci-fi (S) movie (with lowercase letters for Brian).

$$\mathbf{P} = \begin{array}{cc} & \text{Brian} \\ & \begin{array}{cc} r & s \end{array} \\ \text{Ann} \begin{array}{c} R \\ S \end{array} & \begin{Vmatrix} 2,1 & 0,0 \\ 0,0 & 1,2 \end{Vmatrix} \end{array} \tag{2.1}$$

If we treat this as a static game, we are assuming that Ann and Brian act *unbeknownst* to each other, which is not very realistic.

So, let's suppose that the two players can interact. For instance, Ann decides which movie to see before Brian, and then calls him to communicate her proposal.

Ann knows (**complete** information) that Brian's best response is to go with her. Thus, since Ann prefers R over S, she chooses R, *anticipating* Brian's move.

Note that we need a way to *order in time* the actions. The normal form is then replaced by the game's **extensive form**, consisting of the following information:

Extensive form

1. **Normal form**: set of players, and all their payoff functions
2. **Time dimension**: order of their move turns, actions allowed to players *when* they can move (they may change at each turn), information they have when they can move.
3. **Complete information**: probability of external events, and the fact that all the above is *common knowledge*.

Graphically, this can be represented as a **decision tree**. The nodes are the player that move, and each branch denotes a possible action. The leaves are the payoffs for both players.

Graphical representation

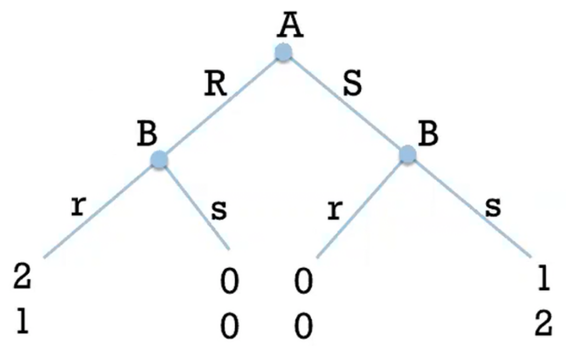

Figure (2.1) – Extensive-form representation for the Battle of Sexes *dynamic* game.

An extensive form game may omit some players, e.g. if they have a single actions. In this case, there are no nodes for a player, but they will still receive payoffs. For them, the game is a *dummy game*: what they get only depends on the action of others.

Dummy players

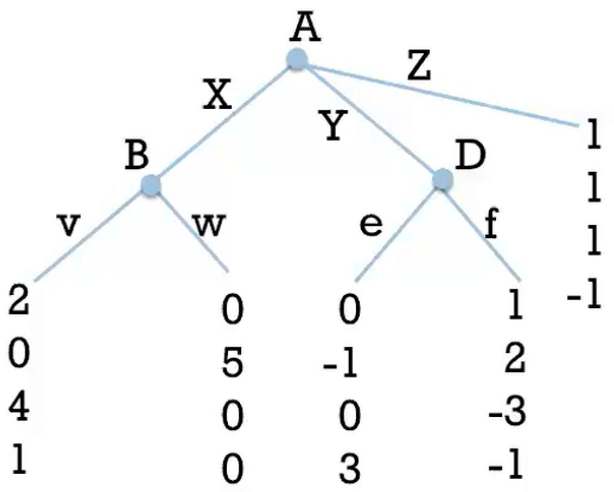

Figure (2.2) – Not all players need to have moved before payoffs are assigned. Here, for example, C never acts, but still receives some points at the end. We say that C is a **dummy player**. Also, some branches may end *before* others.

On different branches, the information available to the same player is generally different, and this can *change* the preferred action.

Information sets

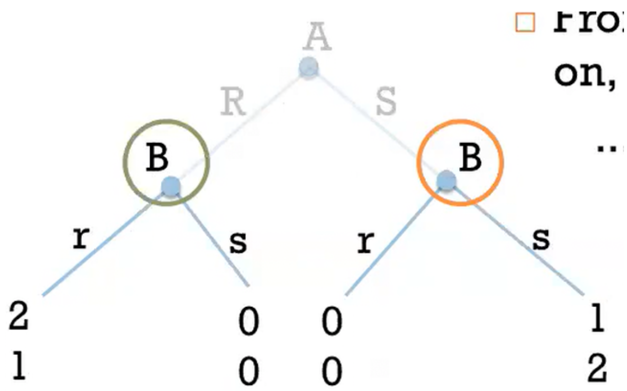

Figure (2.3) – The same player B acts differently depending on the prior choice of A, which is known when they decide the action (r or s) to take. This difference is graphically represented by having *two distinct nodes* for the same player, each in a separate branch of the tree.

Each node has *access* to all the *previous information*, which is encoded in its **parent nodes**. We can explicitly highlight this by giving a **unique** label to each node:

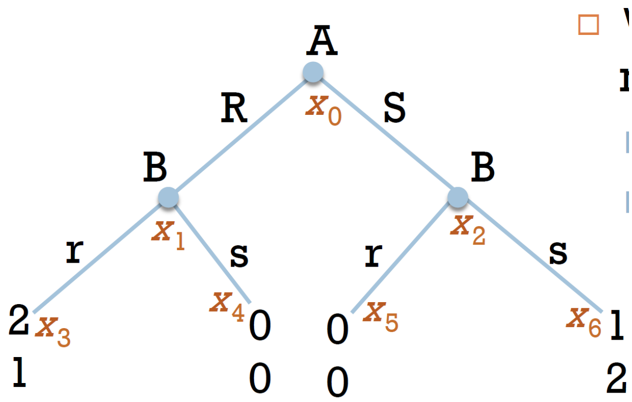

Figure (2.4) – Player B has two distinct nodes: x_1 and x_2. Depending on which node he finds himself in, his choice will be different. This *more detailed* notation makes this explicit, since labels are not repeated.

So, we can think of a node as *encoding* the **information set** h_i available to the player i that is to move. For example, if the information set of B is $\{x_2\}$, then they know that A has played S before.
More in general, if h_i is a *singleton*, then the player has full awareness of all previous moves.

If some choices are simultaneous, a player may *not know* at which node they are. For example, in the original Battle of Sexes, B does not know if they are at x_1 or x_2, and so their information set is $\{x_1, x_2\}$. This is graphically represented as follows:

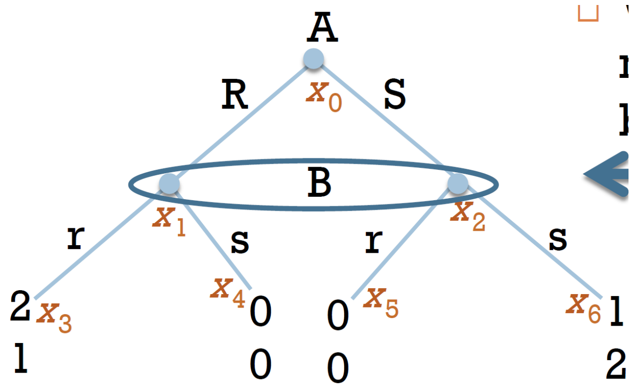

Figure (2.5) – If B does not know if he is at x_1 or x_2, then his information set is $\{x_1, x_2\}$, which is not a singleton.

Since there is no way for B to distinguish between x_1 and x_2 in this situation, their strategy must be the same in both cases.

We can map each node x_j to its information set $h_i(x_j) \in H_i$. In other words, this maps a node (which can be identified by a sequence of moves) to the *knowledge* that a player has in that node.
We assume that $x_j \in h_i(x_j)$, since the game is of complete information. This means that a player at a certain node x_j knows (for sure or with some probability) to be at x_j. Then there are two possibilities:

- $h_i(x_j) = \{x_j\}$. In this case the player i can move knowing for sure that they are at node x_j.

- $h_i(x_j)$ is not a singleton, i.e. there is another $x_k \in h_i(x_j)$ with $x_k \neq x_j$. In this case, player i does not know for certain their node (due to some *simultaneous* action), but knows they must be in some subset of the graph.

  Here, the available actions $A(x_j)$ at node x_j must be *the same* of all nodes in the information sets (so $A(x_j) = A(x_k)$). Otherwise, since the player is always aware of what they can do (complete information), they will have a way to distinguish the nodes, and *identify* their position in the graph, removing the original uncertainty.

  This requirement can be *enforced* (if necessary) by adding "forbidden actions", which have a strongly negative payoff for all players. Thus, they won't ever be taken, but they "symmetrize" the graph.

The cardinality of information sets *defines* the perfect/imperfect nature of the game:

- In a game of **perfect information**, all $h_i(x_j)$ are singletons. Moreover, all actions are decided by the players, i.e. there is not another "uncontrollable player" (Nature) affecting the game.

- In a game of **imperfect information**, $h_i(x_j)$ may contain multiple nodes (there is *endogenous uncertainty*, i.e. "from within", due to simultaneous moves), or there are external effects (choices of Nature, *exogeneous uncertainty*).

In dynamic games, **actions** need to account also the *history of play*, i.e. all the traversed information sets. In other words, strategies need to specify the *reactions* to all possibilities. Thus, a **pure strategy** in a dynamic game is a plan of action encompassing all actions to do in response to every possible situation. In a sense, it is a "big list of if/else statements".
Finally, we have a clear distinction between actions and strategies: a strategy consists of several actions, with conditions attached.

Actions and Strategies

Explicitly, for the Battle of Sexes with Ann moving first, both players choose a move within the set $A = \{R, S\}$. A strategy for Brian specifies which action a to play in response to Ann's action. So, it is a tuple (a_R, a_S), with 4 different possibilities:

Pure strategies

- (s, s): Brian goes to S *no matter what*
- (r, s): follows the choice of Ann
- (s, r): avoids Ann
- (r, r): Brian goes to R *no matter what*

Of course there are *better* and *worse* strategies, but note that while $|A| = 2$, the number of possible strategies is 4, which is higher. In more complex games, this number can quickly explode.

For example, consider the *static* Battle of Sexes repeated two times. In this case Ann makes a decision at the start, and then one at the second night. However, there are 4 possible configurations after the first night, and the second choice can be tailored to each one of them. Thus, a **pure strategy** for Ann involves 5 actions (one at the start, 4 in response to the first night). Since each action is chosen from a set of 2, the total number of available pure strategies is $2^5 = 32$. Note that some of these can never be played, either because they are very bad, or because they are straight out impossible. For example, if Ann chooses R at the start, the configuration (S, S) is impossible. Nonetheless, here we are just *defining* all the possibilities, and need to include all of them as *part of the game's definition*, even those which have 0 probability.

Mixed strategies can be extended to dynamic games as probability distributions over the whole set of pure strategies.
This picture may appear a bit strange: playing a *mixed strategy* would mean drawing a random *plan of action* and sticking to it until the end.

Mixed strategies

A more natural way of playing would be to draw probabilities *at each turn*, depending on the previous actions. Intuitively, we consider a probability distribution unique to the available actions at each position in the graph (or, more generally, at each element of an information set).
Behavioral strategies are the formalization of this idea, and they assign probability distributions to the actions $A_i(h_i)$ associated to each information set $h_i(x_j) \in H_i$. They are mappings $\sigma_i \colon H_i \to \Delta A_i(h_i)$, where $\Delta A_i(h_i)$ is a probability distribution over $A_i(h_i)$.

Behavioural strategies

Expanding the notation, we have that $\sigma_i(a_i|h_i)$ is the probability that i plays action $a_i \in A_i(h_i)$ given the information set $h_i = h_i(x_j)$, i.e. when he/she is at any x_j belonging to it.

Under *certain mild assumptions*, mixed strategies and behavioral strategies are equivalent. Let's see it in an example, again using the Battle of Sexes game.

Behavioural strategies ≡ Mixed Strategies (Example)

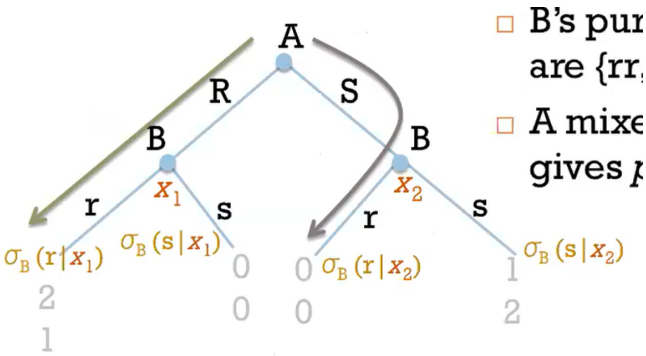

Figure (2.6)

The pure strategies available to B are $\{rr, rs, sr, ss\}$. For example, rs means: play r if Ann plays R (the first action), and s if Ann plays S (the second). So, a mixed strategy assigns a probability to each of them: $\{p_{rr}, p_{rs}, p_{sr}, p_{ss}\}$.

A behavioral strategy is fully defined by the $\sigma_i(a_i|h_i)$, i.e. the probability distributions over the available actions at each information set. In this case, B has only two possible information sets, either $\{x_1\}$ or $\{x_2\}$, which are both singletons (perfect information). So we need to specify: $\sigma_B(r|x_1)$, $\sigma_B(s|x_1)$, $\sigma_B(r|x_2)$ and $\sigma_B(s|x_2)$.

Now, in a mixed strategy, B picks a strategy at random at the start, and sticks to it. Suppose Ann plays R. Brian will play r if he has picked rr, but also if he has picked rs, because both of these strategies specify to respond to Ann's R with an r. Since they are *disjoint events*, their probabilities sum up: $\sigma_B(r|x_1) = p_{rr} + p_{rs}$. Similarly, $\sigma_B(r|x_2) = p_{rr} + p_{sr}$, since now B is responding with r to Ann's S. And the same can be done for $\sigma_B(s|x_1)$ and $\sigma_B(s|x_2)$. Together, these form 4 equations[1] in 4 unknowns, proving a 1 to 1 correspondence between the parameters p_{rr}, p_{rs}, p_{sr} and p_{ss} of the mixed strategy and those $\sigma_B(r|x_1)$, $\sigma_B(s|x_1)$, $\sigma_B(r|x_2)$, $\sigma_B(s|x_2)$ of the behavioral strategy.

Perfect recall

This reasoning can be generalized to more complex cases, such as the ones with non-singleton information sets. However, we are making an implicit (yet reasonable) **assumption**, which is **perfect recall**. This means that players *do not forget* information about the past. A player may be confused about the node in the graph they are occupying, but they know *which* turn they are playing, i.e. "at which level" the node must be.

No perfect recall: the absent-minded driver

In the following, we will always assume perfect recall to hold, since this applies

[1] ∧Actually we should count also the normalization conditions, since these are all probabilities, making the system over-determined.

to most realistic situations. However, it is useful to see an example of situation in which it does *not* hold.

Andrew is driving on the highway and is now close to home. He sees the sign of an exit from the road, and he knows that there are 3 possibilities:

- First exit: bad neighborhood, payoff 0
- Second exit: direct way home, payoff 4
- Third exit: need to go back, very long, payoff 1.

Andrew is tired, and when he passes an exit he is unsure of which is it. Graphically, the situation looks like this:

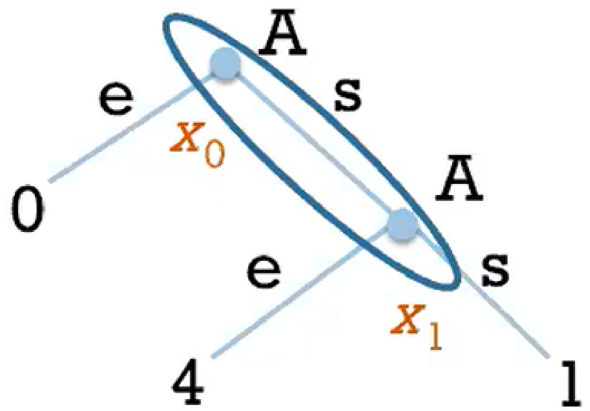

Figure (2.7)

Andrew does not know which node he is at, and so the information set is $\{x_0, x_1\}$. Note how it spans nodes at different *times*. Clearly, the only sensible strategy is e, since it must be the same in both cases.

Formally, the set of strategies is only $\{e, s\}$ (there is only one player and one information set). A mixed strategy is a distribution over it. Let's denote with p the probability of choosing e. Then we compute the expected payoff as follows:

$$\mathbb{E}[u_A] = \underbrace{0 \cdot p}_{\text{Exit immediately}} + \overbrace{4 \cdot (1-p)p}^{\text{Exit at the middle}} + \underbrace{1 \cdot (1-p)^2}_{\text{Exit at the end}} = -3p^2 + 2p + 1 \Rightarrow \text{Max at } p^* = \frac{1}{3}$$

The situation changes with a behavioral strategy. Let q be the probability that Andrew is at node x_0. The strategy defines $p = \sigma(e|\{x_0, x_1\})$, i.e. the probability of going out at a node. Now there are two possibilities: either A starts at x_0 (with probability q) or at x_1 (with $1-q$). So:

$$\begin{aligned}\mathbb{E}[u_A] &= q[0 \cdot p + 4 \cdot (1-p)p + 1 \cdot (1-p)^2] + (1-q)[4 \cdot p + 1 \cdot (1-p)] = \\ &= -3qp^2 - qp + 3p + 1\end{aligned}$$

This is the same as before only if $q = 1$. So, without perfect recall, behavioral strategies are *more general* than mixed strategies. The key point is that in a mixed strategy we always begin at the first turn. In a behavioral strategy, since choices are *delayed*, we can consider also some *prior knowledge* about *which turn we are in*.

After all this discussion, we can *adapt* the normal form also to a dynamic game. The idea is to extend it to *all pure strategies*:

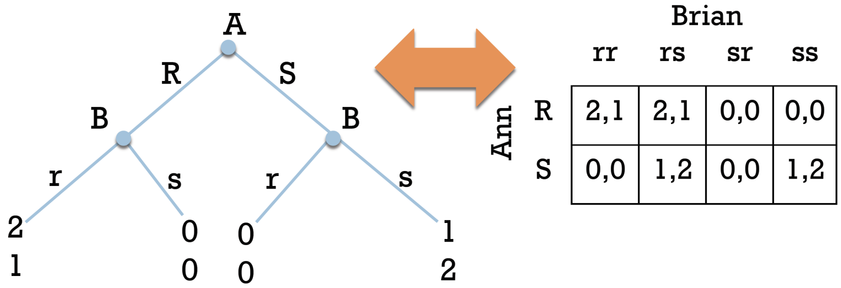

Figure (2.8)

This generalizes also to situations with simultaneous actions. The idea is that a strategy must treat the same all nodes in an information set.

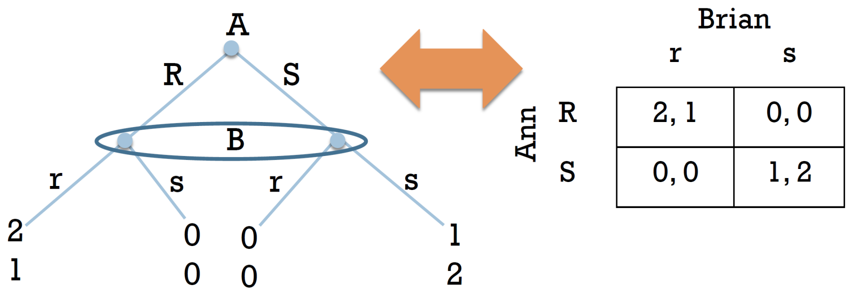

Figure (2.9)

Since games can involve dummy players, *different* extensive forms can be represented by the *same* normal form:

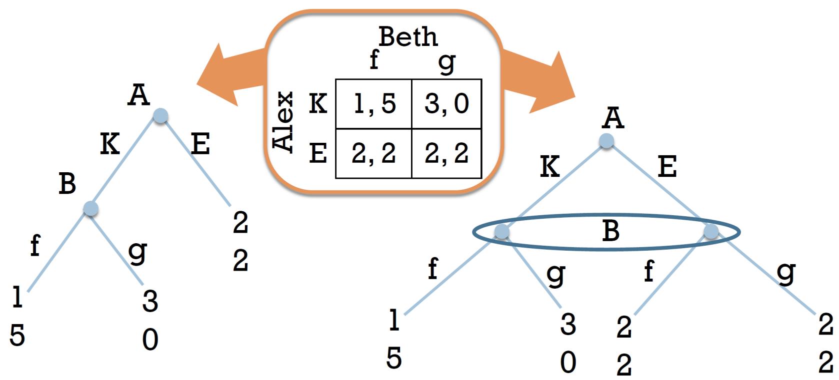

Figure (2.10)

2.1 Dynamic Nash Equilibria

(Lesson 12 of 06/11/2020) Compiled: January 1, 2021

While other concepts, such as pure/mixed strategies, generalize naturally to the dynamic case, Nash equilibria require more work.

To see that, let's start from the most natural way to define them in this new *dynamic* setting. We have seen that a game in extensive form can be *rewritten* in normal form, by just *enumerating* all possible available strategies. In a sense, we are dealing with "paranoid" players that *plan in advance* all countermoves to their opponents.

The *dynamic* Battle of Sexes, when expressed in normal form, looks like this:

$$\mathbf{P} = \begin{array}{cc} & \text{Brian} \\ & \begin{array}{cccc} rr & rs & sr & ss \end{array} \\ \text{Ann} \begin{array}{c} R \\ S \end{array} & \left\| \begin{array}{cccc} 2,1 & 2,1 & 0,0 & 0,0 \\ 0,0 & 1,2 & 0,0 & 1,2 \end{array} \right\| \end{array} \tag{2.2}$$

By highlighting all the best responses, we see that there are 3 Nash equilibria:

- (R, rr): Ann plays R, Brian *always plays* R
- (R, rs): Ann plays R, Brian *copies her move.*
- (S, ss): Ann plays S, Brian *always plays* S.

Note that (S, rs) is not a Nash equilibrium, because in this case Ann has an incentive to change her move to R, knowing that Brian will then follow her.

(R, rr) and (R, rs) are equivalent at equilibrium, since they only differ in Brian's planned response (r or s) to Ann's S, which is never played. Still, they cannot be *really* equivalent. If we are *not* at equilibrium, (R, rs), i.e. "Brian copies Ann's move", is the *rational* strategy, because Brian *wants* to be with Ann.

Then, consider (S, ss). While it is a NE, it is not rational for Ann to play S. She goes first, and she can get a higher payoff by choosing R, and she knows it. Also, she is *rational*, and so she should *maximize* her payoff.

Still, the possibility that Brian plays ss must be considered. This looks like a **threat**[2], forcing her to play S: otherwise, both get nothing. However, it cannot be sustained. The best move for Brian is to *always copy* Ann's move (because he wants to be with her), and so ss is **non-credible**: it is an *empty* threat. If executed, it will damage both players. So, while (S, ss) *exists*, it won't be chosen by rational players.

Note how, in the last few paragraphs, we considered plans for moves that never happen. These are just needed for evaluation purposes, to be *sure* that everything is considered. To be rational means to *think about all the possibilities*, not just the ones that seem "more likely". If we say that a strategy is *the best*, it must be so *in the whole graph*, not just one part of it!

[2]∧All of this reasoning happens in Ann's mind, since she is *very paranoid*, and must consider every possibility, as expected for a rational player.

In summary, we see that between the 3 NE, only one will be really played: (R, rs).

This leads us to introduce the concept of an **equilibrium path**. Given a joint profile of behavioral strategies $\boldsymbol{m}^* = (m_1^*, m_2^*, \dots, m_n^*)$ corresponding to a Nash Equilibrium, its *equilibrium path* contains the decision tree nodes that are reached with probability > 0.

Equilibrium path

Example 2 (Ultimatum game)**:**

Two players share 10 candies as follows:

- Player 1, the *proposer*, presents a split
- Player 2, the *responder*, decides whether to accept it.

If player 2 refuses, they both get nothing.

There are 11 ways to split candies, but for simplicity we will consider just 3: $A_P = \{D: (9,1), F: (5,5), G: (2-8)\}$. Since the proposer plays first, he has only 3 (pure) strategies, one for each move.
The responder has just 2 actions, $A_R = \{Y: \text{(accept)}, N: \text{(Refuse)}\}$. However, since he goes *second*, the total number of strategies will be $3 \cdot 2 = 6$, because each strategy must specify a *response* to one of the choices of P.

The game's extensive form is the following:

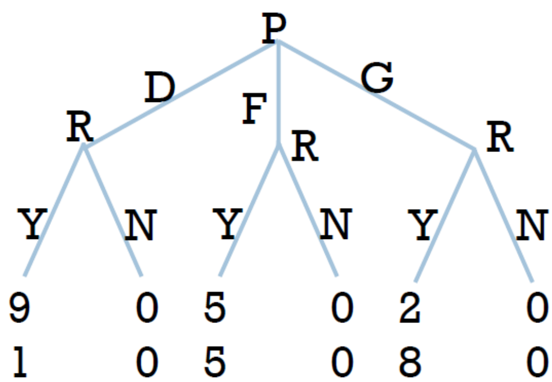

Figure (2.11)

To find the Nash equilibria, we can reason as follows. Suppose P plays F. Clearly, R's plan must respond Y to F (*accept* the proposal), otherwise he would get nothing. However, he needs to specify also the answers to *other proposals*. He must plan to say N to every other split that would be better for P, which, in this case, is D. In this way, P cannot unilaterally change his move to D and gain an advantage: doing so would get him a rejection, and so nothing. On the other hand, it does not matter which action R has planned to all the other possibilities with a *lower* payoff for P than the current one. So, he can say Y or N to G, and nothing will change, because P has no incentive to move from F to G.

From this reasoning, we see that P won't change its move. The same holds for R: if he changes the planned response (Y) to the current split, he will get nothing. Otherwise, if he unilaterally changes his response to another move, nothing will change, because that move has not been played. So, again, no incentive to make any different move from his part, which confirms the Nash Equilibrium.

In summary, a full list of NEs is as follows:

- (D, YYY), and all the combinations (D, YNN), (D, YNY), (D, YYN).
- (F, NYY) and (F, NYN)
- (G, NNY)

But, which of these equilibria is likely to get *really* played?

There is only one: (D, YYY). Basically, the idea is that R always prefers getting something rather than nothing, and P knows that. Thus, all strategies in which R *refuses* a split are *empty threats*: they are not rational moves, and so they are **non-credible**.

2.1.1 Rational Nash Equilibria

We introduced Nash Equilibria as a way to **solve** games, in the sense that we expect rational players to recognize and **play** a NE. However, the NEs found by analyzing the normal form of a dynamic game are not all really *playable*: several of them, as seen in the above examples, seem to be *irrational*. So, we need a more powerful definition.

First, we will see how to get a *good* (i.e. playable) solution in the first place. Then, we will formalize that procedure in the following section, *extending* the definition of a Nash equilibrium.

Let's start with a special case of dynamic games, called **sequential games**, which are just dynamic games with perfect information. So, all players take turns, there are no simultaneous decisions, all players know what happened before their turn, and all information sets are singletons. Moreover, all of this is *common knowledge*.

(a) Perfect information

This kind of games can be solved by means of **backward induction**.
To see that, consider a 2-players game.

Backward induction

1. Player 1 chooses action $a_1 \in A_1$.
2. Player 2 sees a_1, and chooses action $a_2 \in A_2$. The set A_2 *depends* on a_1, because a_1 could somewhat *limit* decisions of player 2. This is known by both players.
3. At the end, players receive payoffs $u_1(a_1, a_2)$ and $u_2(a_1, a_2)$.

Since both players are **rational**, they can always make optimal decisions. In particular, player 2 knows always where they are in the decision tree (**perfect information**) and so they can choose the optimal move here.

But player 1 has **complete** information, and so they can anticipate what player 2 will do, and then optimize their choice.

So, by starting from the end and *optimizing* at each step we can *solve* the entire game. In fact, this can be proved by a theorem:

Theorem 2.1.1 (Zermelo)**.** *Any dynamic game of* ***perfect*** *information has a* ***backward induction*** *solution that is sequentially* ***rational****. Moreover, if terminal payoffs are all different, the solution is unique.*

In practice, consider the following.

1. When it is their turn, Player 2 sees Player 1's move a_1^h and solves the optimization problem:

$$R_2(a_1^h) = \arg\max_{a_2 \in A_2} u_2(a_1^h, a_2)$$

2. Player 1 can anticipate that computation, and solve:

$$a_1^* = \arg\max_{a_1 \in A_1} u_1(a_1, R_2(a_1))$$

The outcome $(a_1^*, R_2(a_1^*))$ is a Nash Equilibrium in pure strategies.

For example, consider the following game:

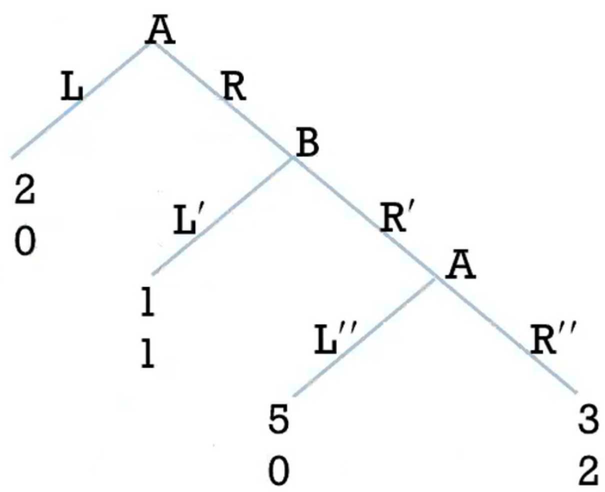

Figure (2.12)

At the last move, A chooses L'', because this leads to a better payoff $5 > 3$. So, we remove the R'' branch from the game.
Now B can choose between L', which ends the game with $(1, 1)$ or R', which again ends the game (since then B will play L'') with $(5, 0)$. The best choice is L', since $1 > 0$, and we can remove R'. Then, A prefers L, because $2 > 1$.

So, the final strategies are (LL'', L'). Note that we need to specify *all of them*, also the actions that never get played (since the game effectively ends after L). Moreover, this is not a Pareto efficient solution, which would be the one with payoff $(3, 2)$. However, a strategy (RR'', R') is non-credible: if A is given the opportunity to play the second move, they will play L'', not R'', to maximize their payoff. B knows that, and so, if given the opportunity to play, would choose L', not R'. But again A *knows* that, and so they won't give B the opportunity to play, and choose L at the start!

Credibility

Formally, if we know that a_1 implies a response $R_2(a_1)$, then any strategy involving "if a_1, then $a_2 \neq R_2(a_1)$" is classified as **non-credible**: it is an *empty treat*, it won't be chosen by rational players.

(b) Imperfect information

Backward induction can be generalized to the case of **imperfect information**, where some players act simultaneously. For instance:

1. Players 1 and 2 choose actions $a_1 \in A_1$ and $a_2 \in A_2$
2. Players 3 and 4 observe the outcome a_1 and a_2, and choose $a_3 \in A_3$ and $a_4 \in A_4$.
3. Payoffs are computed as $u_j(a_1, a_2, a_3, a_4)$ for $j = 1, 2, 3, 4$.

The idea is that simultaneous moves can be regarded as a *static game*, and we know that rational players will opt for a Nash Equilibrium. Specifically:

1. For every choice (a_1, a_2) of the first two players, players 3 and 4 play a Nash Equilibrium $(a_3^*(a_1, a_2), a_4^*(a_1, a_2))$
2. Players 1 and 2 can anticipate this (**complete** information), and compute their payoffs as $u_j(a_1, a_2, a_3^*(a_1, a_2), a_4^*(a_1, a_2))$. Since they are rational players, they will choose a NE a_1^*, a_2^*, leading to the outcome $(a_1^*, a_2^*, a_3^*(a_1^*, a_2^*), a_4^*(a_1^*, a_2^*))$.

We are sure that there is *at least* one NE, but there could be more than one, and so the results of this *backward induction* will generally be many.

2.1.2 Subgame-perfect Nash Equilibrium

From the previous section, we saw that a *dynamic* game can be *solved* by backward induction, which amounts to finding the NE *at the end*, and "propagating it back". In other words, we start from the *last turn*, and treat it as a *one-turn separate game* (a **subgame**), solve it, and substitute the result in the previous turn, which can now be regarded as a *one-turn separate game* too. Thus, the **solution** we seek has the property of being a NE *for all these subgames*, which suggests a possible extension of the NE definition.

First, we need to precisely define what we mean with *subgame*. A (proper) **subgame** G contains a single node of the (extensive-form) tree and all of its successor nodes, with the requirement of including *all nodes* belonging to the same information sets: $x_j \in G, x_k \in h_i(x_j) \Rightarrow x_k \in G$. Clearly, the whole game is a subgame of itself.

For example, consider the following game:

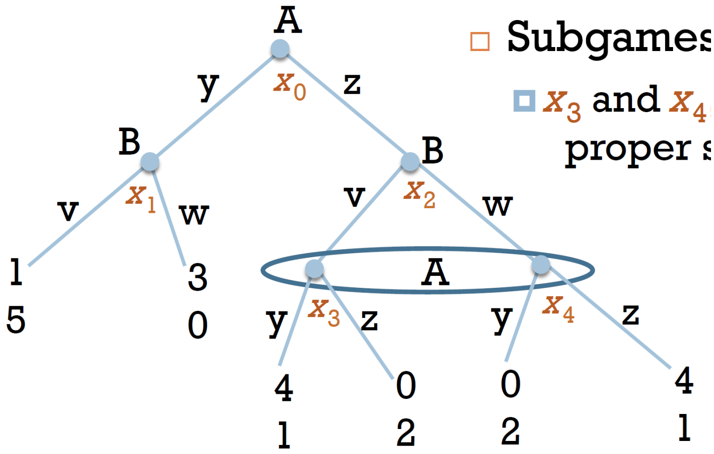

Figure (2.13)

The subtrees that start from x_0, x_1 and x_2 are all subgames. However, we can't pick only x_3 and its successors as a proper subgame, since x_3 and x_4 belong to the same information set, meaning that we would need to include also x_4 and its successors (and the same holds for x_4). In other words, subgames "can't break *bubbles* in the graph".

In this view, the choice of A, i.e. y or z, decides which *subgame* (x_1 or x_2) will be played by B.

Definition 1 (R. Selten)**.** A Nash equilibrium is **subgame-perfect** (SPE) if the strategies chosen by the players give a NE in **every** subgame.

Subgame-perfect NE

It can be shown that *every finite extensive form game* has an SPE.

2.2 Minimax

(Lesson 13 of 10/11/2020) Compiled: January 1, 2021

We now take a small parenthesis to introduce an important optimization method with many applications, also in Game Theory.

Consider a "two" player game, in which i battles against their opponents $-i$, who may or may not multiple players, but we treat them like a single group.

We define a function $f_i \colon S_i \to \mathbb{R}$ as follows:

$$f_i(s_i) = \min_{s_{-i} \in S_{-i}} u_i(s_i, s_{-i})$$

In other words, if i plays s_i, the *worst* payoff they could get (depending on the choices from the other players) is $f_i(s_i)$.

So, it is convenient for i to choose s_i^* that **maximizes** this worst outcome. This means that i will get *at least* $w_i \equiv f_i(s_i^*)$:

$$s_i^* = \arg\max_{s_i \in S_i} f_i(s_i)$$

Such s_i^* (which may be not unique) is called a **security strategy** and assures a payoff:

$$w_i = \max_{s_i \in S_i} \min_{s_{-i} \in S_{-i}} u_i(s_i, s_{-i})$$

called the **security payoff**, or **maximin** (*max*imum of the *min*imum payoffs).

A security strategy is a *conservative* approach, in which we expect opponents $-i$ to be very smart, and immediately make use of any misstep player i makes.

Similarly, suppose i plays *after* $-i$. If they play a best response, the payoff they get is given by the function $F_i \colon S_{-i} \to \mathbb{R}$ defined as follows:

$$F_i(s_{-i}) = \max_{s_i \in S_i} u_i(s_i, s_{-i})$$

Over all possible moves s_{-i}, the minimum payoff *guaranteed* to i is:

$$z_i = \min_{s_{-i} \in S_{-i}} F_i(s_{-i}) \max_{s_i \in S_i} u_i(s_i, s_{-i})$$

and z_i is called the **minimax** for player i (*min*imum of the *max*imum payoffs).

Note that, in both definitions, we only need the utility u_i of player i, and so we can ignore the utilities of the opponents.

As an example, consider the following game:

$$\mathbf{P} = \begin{array}{c|c} & \text{player B} \\ \text{player A} & \begin{array}{c|ccc|c} & L & C & R & f_{\min} \\ \hline T & 5,- & 3,- & 4,- & 3 \\ D & 2,- & 6,- & 1,- & 1 \\ \hline F_{\max} & 5 & 6 & 4 & \end{array} \end{array} \tag{2.3}$$

When computing $F_{\max}$, player A *knows* which move player B will play, and thus chooses their move with maximum payoff. For example, if B plays L, then A will respond with T, getting $F_{\max} = 5$. If A *does not know* what B will play, they still know that playing T will get, at the minimum, a payoff $f_{\min} = 3$. Similarly, playing D gets 1.
Then, the maximin is 3, since it is the maximum of the minimum row values. The minimax is instead 4, which is the minimum of the maximum payoffs.

In all cases, $\text{maximin}_i \leq \text{minimax}_i$, because in the latter case i is always playing best responses. Moreover, it can be shown that if a joint strategy $\boldsymbol{s}$ is a NE, then for every player i, $\text{minimax}_i \leq u_i(\boldsymbol{s})$. In fact, in a NE all players choose a best response. The minimax is the minimum payoff of all best responses, and so it must be $\leq$ than u_i computed at the NE $\boldsymbol{s}$.

To find the Nash Equilibria, we need to know also the payoff for the other player:

$$\mathbf{P} = \begin{array}{c|c} & \text{player B} \\ \text{player A} & \begin{array}{c|ccc|} & L & C & R \\ \hline T & 5,6 & 3,2 & 4,1 \\ D & 2,0 & 6,8 & 1,2 \end{array} \end{array} \tag{2.4}$$

There are two NE, which are (T, L) and (D, C). In both cases, $u_A > \text{minimax}_A$.

If there is only one NE, such as (D, L) in the following game:

$$\mathbf{P} = \begin{array}{cc} & \text{player B} \\ & \begin{array}{ccc} L & C & R \end{array} \\ \text{player A} \begin{array}{c} T \\ D \end{array} & \begin{Vmatrix} 3,4 & 5,0 & 3,1 \\ 5,4 & 6,2 & 7,2 \end{Vmatrix} \end{array} \tag{2.5}$$

then, for both players, the maximin must be payoff at the NE, and from the other properties we have:

$$\text{maximin}_i = \text{minimax}_i = u_i(\text{NE})$$

The converse is not true: $\text{maximin}_i = \text{minimax}_i$ does not imply that there is a unique NE with that payoff!

2.2.1 Zero-sum games

A stronger connection between minimax/maximin and NE can be obtained for a specific kind of games, called **zero-sum games**. Here $u_i(s) = -u_{-i}(s)$, that is i gaining an advantage happens at the expense of their opponent $-i$ (and viceversa).

For example:

$$\mathbf{P} = \begin{array}{cc} & \text{player B} \\ & \begin{array}{ccc} L & C & R \end{array} \\ \text{player A} \begin{array}{c} T \\ M \\ D \end{array} & \begin{Vmatrix} -9,9 & 8,-8 & -5,5 \\ -2,2 & -6,6 & 2,-2 \\ -1,1 & 3,-3 & 4,-4 \end{Vmatrix} \end{array} \tag{2.6}$$

Zero-sum games are a special case of **adversarial games**, in which when the utility u_i of a player i increases, then the utility of the opponents u_{-i} must decrease. In other words, there is always a winner and a loser.

If G is a zero-sum game with finitely many strategies, then (von Neumann, 1928):

1. G has a NE $\Leftrightarrow$ $\text{maximin}_i = \text{minimax}_i$ for each i. Actually, this needs to be checked only for a player i, since $u_{-i} = -u_i$, and so $\text{maximin}_i = \text{nimimax}_{-i}$, and similarly for minimax_i.

2. All NEs yield the same payoff, which is the maximin_i

3. NEs have the form (s_i^*, s_{-i}^*) with s_i^* being a security strategy.

The common value of $\text{maximin}_1 = \text{minimax}_1$ is called the **value** of the game (player 1 is taken *by convention*). A joint security strategy (if any), i.e. a NE, is called a **saddle point** of the game. To see why, consider the utility function $u_i(s_i, s_{-i})$. i wants to *maximize* it. Since $u_i = -u_{-i}$, and $-i$ wants to maximize $-i$, we see that $-i$ wants to *minimize* u_i. Thus, i will move along the

s_i direction[3] to maximize u_i, and they will choose a s_i^* that does so. Conversely, $-i$ will move along s_{-i} to minimize u_i, choosing s_{-i}^*. A point (s_i^*, s_{-i}^*) that is a maximum along one direction (that of s_i) and a minimum along another (that of s_{-i}) is exactly a **saddle point**, and in this case it is the choice that satisfies both players. Since they do not have any incentive for unilateral deviation, this is a Nash Equilibrium.

In zero-sum games, the utilities u_{-i} can be left implicit, thus using a regular matrix instead of a bi-matrix.

2.2.2 Mixed Minimax

We now extend the previous results to the general case of mixed strategies. So, consider the function $f_i\colon \Delta S_i \to \mathbb{R}$, where ΔS_i is the set of probability distributions over the pure strategies, defined by:

$$f_i(m_i) = \min_{m_{-i} \in \Delta S_{-i}} u_i(m_i, m_{-i})$$

A **mixed security strategy** for i is the one that maximizes f_i:

$$m_i^* = \arg\max_{m_i \in \Delta S_i} f_i(m_i)$$

And the **mixed security payoff** (in mixed strategies) is:

$$\text{maximin}_i^m = \max_{m_i \in \Delta S_i} \min_{m_{-i} \in \Delta S_{-i}} u_i(m_i, m_{-i})$$

As before, a mixed security strategy is the *conservative* mixed strategy guaranteeing the highest payoff for i in case of the worst mixed strategy by the opponents $-i$.

Similarly, we define $F_i\colon \Delta S_{-i} \to \mathbb{R}$ as:

$$F_i(m_{-i}) = \max_{m_i \in \Delta S_i} u_i(m_i, m_{-i})$$

And the minimum of the best response utilities is the minimax:

$$\text{minimax}_i^m = \min_{m_{-i} \in \Delta S_{-i}} F_i(m_{-i}) = \min_{m_{-i} \in \Delta S_{-i}} \max_{m_i \in \Delta S_i} u_i(m_i, m_{-i})$$

By exploiting the linearity of expected utilities, we can find $f_i(m_i)$ by minimizing $u_i(m_i, s_{-i})$, i.e. by limiting the opponents' strategies to *pure* strategies. Similarly, $F_i(m_{-i})$ can be found by maximizing $u_i(s_i, m_{-i})$. This is because if m_i is a best response, every pure strategy in the support of that mixed strategy must also be a best response.

Finally, maximin_i^m and maximin_i^m always exist, since the u_i are now continuous functions over compact domains, and their minima/maxima always exist.

Also in this case, if $\boldsymbol{m}$ is a NE, then for every player i:

$$\text{minimax}_i^m \leq u_i(\boldsymbol{m})$$

[3] ∧Imagine the possible strategies of player i as points along their axis i in a $2D$ plane (since we are dealing with 2 players).

As an example, consider the following game:

$$
\mathbf{P} = \begin{array}{cc} & \text{Joe} \\ & \begin{array}{cc} S & C \end{array} \\ \text{Jim} \begin{array}{c} T \\ M \end{array} & \left\| \begin{array}{cc} 3,- & 0,- \\ 1,- & 2,- \end{array} \right\| \end{array} \tag{2.7}
$$

For Jim, the best responses (in *pure* strategies) are 3 and 2, with a minimum (minimax) of 2. The minimum payoffs for each action are 0 and 1, with a maximum (maximin) of 1.

By using *mixed* strategies, Jim can increase its maximin by playing $m_i = 1/4\, T + 3/4\, M$, reaching maximin$^m = 1.5$.
In fact:

$$
\begin{aligned}
u_i(m_i, S) &= \frac{1}{4} \cdot 3 + \frac{3}{4} \cdot 1 = 1.5 \\
u_i(m_i, C) &= \frac{1}{4} \cdot 0 + \frac{3}{4} \cdot 2 = 1.5
\end{aligned}
$$

In other words, playing m_i makes Jim's payoff *indifferent* on Joe's choice.

Similarly, the *worst* strategy that Joe can play (from the point of view of Jim) is $m_{-i} = 1/2\, S + 1/2\, C$, with a minimax$^m = 1.5$. In fact:

$$
\begin{aligned}
u_i(T, m_{-i}) &= \frac{1}{2} \cdot 3 + \frac{1}{2} \cdot 0 = 1.5 \\
u_i(M, m_{-i}) &= \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 2 = 1.5
\end{aligned}
$$

Note that:

$$
\text{minimax}^m = \text{maximin}^m
$$

Let's see another example to understand why.

$$
\mathbf{P} = \begin{array}{cc} & \text{Bea} \\ & \begin{array}{cc} F & G \end{array} \\ \text{Art} \begin{array}{c} U \\ D \end{array} & \left\| \begin{array}{cc} 6,1 & 0,4 \\ 2,5 & 4,0 \end{array} \right\| \end{array} \tag{2.8}
$$

Let q be the probability of Art playing U, which we will use to parametrize A's mixed strategies. Then, we compute f_A explicitly from the definition:

$$
\begin{aligned}
f_A(q) \equiv \min_{m_B \in \Delta S_B} u_A(q, m_B) &= \min_{s_i \in S_{-i}} u_A(q, s_B) = \min_{s_B \in \{F,G\}} u_A(q, s_B) = \\
&= \min\{u_A(q,F), u_A(q,G)\} = \min\{6q + 2(1-q), 4(1-q)\} = \min\{2+4q, 4-4q\}
\end{aligned}
$$

Since $q \in [0,1]$, we can find graphically this minimum:

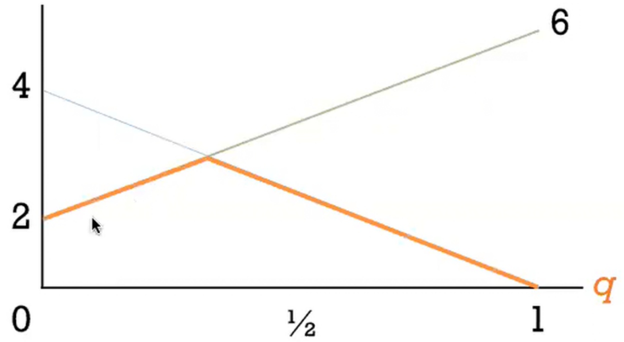

Figure (2.14) – The green line is $2+4q$, the blue one is $4-4q$. The minimum is highlighted in orange.

Now, the maximin is defined as:

$$\text{maximin}^m = \max_{m_A \in \Delta S_A} f_A(m_A) = \max_{q\in[0,1]} f_A(q) = \max_{q\in[0,1]} \min\{2+4q, 4-4q\} = 3$$

And graphically is the uppermost point of the orange line:

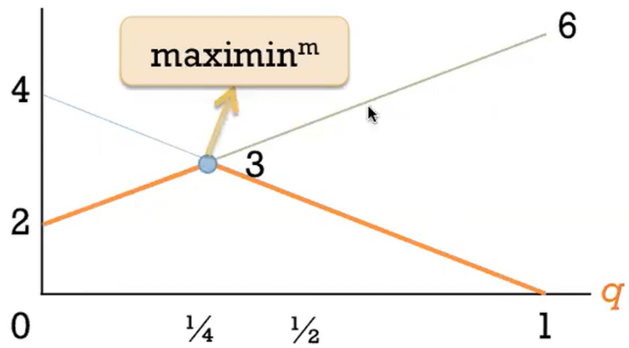

Figure (2.15)

So, moving in mixed strategies allows increasing the maximin of pure strategies.

Similarly, for the minimax we denote with p the probability that Bea plays F. Then we compute $F_A(p)$ from the definition:

$$\begin{aligned} F_A(p) &\equiv \max_{m_A \in \Delta S_A} u_A(m_A, p) = \max_{s_A \in S_A} u_A(s_A, p) = \max_{s_A \in \{U,D\}} u_A(s_A, p) = \\ &= \max\{u_A(U,p), u_A(D,p)\} = \max\{6p + 0(1-p), 2p + 4(1-p)\} = \max\{6p, 4-2p\} \end{aligned}$$

And for the minimax:

$$\text{minimax}^m = \min_{m_B \in \Delta S_B} F_A(m_B) = \min_{p\in[0,1]} F_A(p) = \min_{p\in[0,1]} \max\{6p, 4-2p\} = 3$$

Graphically, the situation is the same as before, but *reversed*. That is, moving in the mixed strategies *decreases* the maximin from that of pure strategies.

Since $\text{maximin}_i \le \text{minimax}_i$ always for pure strategies, and in general $\text{maximin}_i^m \ge \text{maximin}_i$ and $\text{minimax}_i^m \le \text{minimax}_i$, we expect the mixed minimax/maximin to "meet at the middle", when they *saturate* the inequality. This can in fact be verified:

$$\text{maximin}_i \le \text{maximin}_i^m = \text{minimax}_i^m \le \text{minimax}_i$$

Now, this does tell us something about Nash Equilibria **only** in zero-sum games. In particular, at the NE, $u_i(\boldsymbol{m}) = \text{minimax}_i^m = \text{maximin}_i^m$.

In the above example, we can see that $(q = 1/4, p = 1/2)$ is a NE, with an expected payoff of 3 for both players, which coincides with the maximin and minimax. However, that game is **not** a zero-sum game, and so this is just a coincidence!

As a counterexample, consider the Battle of Sexes:

$$
\mathbf{P} = \begin{array}{cc} & \text{Brian} \\ & \begin{array}{cc} R & S \end{array} \\ \text{Ann} \begin{array}{c} R \\ S \end{array} & \left\| \begin{array}{cc} 2,1 & 0,0 \\ 0,0 & 1,2 \end{array} \right\| \end{array} \tag{2.9}
$$

Here maximin = 0 and minimax = 1 for both players in pure strategies. Then $\text{maximin}^m = \text{minimax}^m = 2/3$. However, there are 3 NEs with payoffs 1, 2 and 1.67, which are **not** that of the minimax/maximin, since this game is **not** zero-sum.

Note that, since $\text{maximin}^m = \text{maximin}^m$ always in mixed strategies, this assures that zero-sum games have always a NE in mixed strategies, which is nothing else than Nash theorem.

In summary, we can now extend the previously stated theorem.

Theorem 2.2.1 (Zero-sum games)**.** *For a zero-sum game G with finitely many strategies:*

1. *For every player i, $\text{maximin}_i^m = \text{minimax}_i^m$, and thus G must have a Nash equilibrium (this is actually how to find it).*
2. *All Nash Equilibria in mixed strategies are security strategies for player i and yield a payoff to i equal to maximin_i^m.*

Linear programming

Optimizing zero-sum games can still be hard, due to the sheer number of strategies involved (e.g. in *chess*). In practice, it can be solved through **linear programming** as follows.

Suppose player 1 has pure strategies $\{A_1, A_2, \ldots, A_L\}$, and player 2 has $\{B_1, B_2, \ldots, B_M\}$. A mixed strategy $\boldsymbol{a} = \{a_j\}$ for player A is a linear combination:

$$
a_1 A_1 + \cdots + a_L A_L
$$

And similarly for B:

$$
b_1 B_1 + \cdots + b_M B_M
$$

Let u_1 be the utility of player 1. Since we are dealing with a zero-sum game, $u_2 = -u_1$. Then we consider the following:

$$
\sum_j a_j u_1(A_j, B_k) \geq W \qquad \forall k
$$

with $a_j \geq 0$ and $\sum_j a_j = 1$, since they are probabilities. The idea is to find the *maximum* W for which values of $\boldsymbol{a}$ can be found that satisfy the above constraint.

2.3 Stackelberg games

Stackelberg games are a special case of sequential games, i.e. dynamic games with perfect information, in which there are two players: the first, called the **leader**, and the second, the **follower**. They can be solved, as previously seen, by *backward induction*, which leads to the so-called **Stackelberg equilibrium**.

Stackelberg games can be represented in normal form as static games. For example, the *dynamic* Battle of Sexes in which Ann plays first is a Stackelberg game:

$$\mathbf{P} = \begin{array}{cc} & \text{Brian} \\ & \begin{array}{cc} R & S \end{array} \\ \text{Ann} \begin{array}{c} R \\ S \end{array} & \left\| \begin{array}{cc} 2,1 & 0,0 \\ 0,0 & 1,2 \end{array} \right\| \end{array} \tag{2.10}$$

The Stackelberg equilibrium is (R, R), as can be found through backward induction. Note that, since Brian *knows* the choice of Ann, and plays his best response to that, he achieves at least his minimax, which is 1 in this case: this follows directly by definition.

Mixed strategies are allowed, in the sense of a player "not revealing their hand". So, for example, in the Odds/Even game:

$$\mathbf{P} = \begin{array}{cc} & \text{Even} \\ & \begin{array}{cc} 0 & 1 \end{array} \\ \text{Odd} \begin{array}{c} 0 \\ 1 \end{array} & \left\| \begin{array}{cc} -4,4 & 4,-4 \\ 4,-4 & -4,4 \end{array} \right\| \end{array}$$

the Stackelberg equilibrium is the *same* as the Nash Equilibrium, i.e. the mixed strategy $1/2\,0 + 1/2\,1$ for both players, who achieve the minimax (0).

As another example, consider the following:

$$\mathbf{P} = \begin{array}{cc} & \text{Joe} \\ & \begin{array}{ccc} F & G & H \end{array} \\ \text{Carl} \begin{array}{c} R \\ S \\ T \end{array} & \left\| \begin{array}{ccc} 2,2 & 3,1 & 0,0 \\ 1,6 & 5,4 & 6,4 \\ 0,1 & 4,3 & 6,2 \end{array} \right\| \end{array} \tag{2.11}$$

If this is a static game, then (R, F) is a pure Nash Equilibrium. If instead Carl goes first, as in a Stackelberg game, then the equilibrium will be (T, G). The idea is to apply backward induction: Carl *knows* that Joe will respond to any of his actions with his best response (highlighted in blue). For example, if Carl plays S, then Joe will respond with F, resulting in a payoff of 1 for Carl. By looking at these payoffs (which are, in order, 2, 1, 4, i.e. the values *to the left* of the blue numbers), Carl chooses the *highest one* (4), and so he plays T.
Note that, in this case, Joe gets a payoff of 3, which is *better* than his minimax (2). This is not surprising, since the game is not zero-sum.

However, backward induction may not produce a unique solution if there are *repeated* payoffs, such as the two 6 for Carl. This becomes evident if we let Joe go first. Now, if Joe plays F, Carl with respond with R, if he plays G, Carl reacts with S, but if Joe chooses H, Carl can play either S or T. For Carl, they lead to the same payoff, but the situation for Joe is very different!

To resolve the stalemate, we **assume** that Carl is a *generous follower*: when indifferent about his payoff, he chooses the option that maximizes the opponent's payoff. In this case, if Joe goes with H, Carl reacts with S.

Thus, the possible outcomes for Joe (which are the values *to the right* of the red numbers) are $2, 4, 4$. Here we have another ambiguity: will Joe play G or H?

Again, we need a further **assumption**: Joe is a *generous leader*, and so he will resolve ties by favoring the follower (basically *answering* Carl's kindness in choosing S rather than T). So, Joe plays H, which gives Carls a better payoff than G ($6 > 5$). And so the Stackelberg equilibrium is (S, H).

Note that, in this situation, Carl obtains payoff 6, which is much higher than his minimax (2).

As a final remark, note that in Stackelberg games, or even in general in sequential games, the player that goes first has an *advantage* (**first-move advantage**): their payoff is always $\geq$ than that of the Nash equilibrium. This may appear *counterintuitive*: the second player knows *more*, because they have seen the first player move! However, recall that both players are rational, and the game has **complete** knowledge: player 1 can *anticipate* all of player 2's knowledge, and so they get more.

2.4 Consistency of time discounting

(Lesson 14 of 13/11/2020) Compiled: January 1, 2021

Consider a dynamic game that can terminate at different turns. Until now, we have treated all payoffs to be *of the same importance*. However, intuitively an *immediate* payoff should be preferable to a *delayed* one: it is best to take a reward now than tomorrow.

This can be modelled by applying a *discount* factor to future payoffs, lowering those who are very far into the future. As we will now see, there is really *only one way* to do this that is **consistent**.

As a start, consider a player who has a fixed resource budget $K = 1$ (e.g. *food*) to allocate over N subsequent time steps (e.g. *days*). Let's assume $N = 3$. Every time the player has access to x units of the allocated resource, they get a utility of $u(x)$. Since delayed rewards are worse than immediate rewards, at every time step the computed utilities are multiplied by a discount factor $\delta < 1$. So, x_1 on the first timestep leads to $u(x_1)$, but x_2 on the *second* timestep gives $\delta u(x_2) < u(x_2)$. The total payoff *visible* at the start is then:

$$v(x_1, x_2, x_3) = u(x_1) + \delta u(x_2) + \delta^2 u(x_3) \tag{2.12}$$

The player wants to maximize v, subject to the constraint $x_1 + x_2 + x_3 = K = 1$.

Note that this is a standard one-player optimization problem, which can be solved directly through calculus, without the need of any Game Theory method.

To solve it, we need to specify a utility function, which is:

$$u(x) = \log(1 + x)$$

Note that it is an increasing function, which is somewhat *saturating*, and in particular it is concave.

Then, from the constraint we have $x_1 = 1 - x_2 - x_3$, which can be substituted into $v(x_1, x_2, x_3) = v(1 - x_2 - x_3, x_2, x_3)$. Differentiating and setting the first derivatives to 0 we get:

$$x_1 = \frac{3 - \delta - \delta^2}{1 + \delta + \delta^2} \quad x_2 = \frac{-1 + 3\delta - \delta^2}{1 + \delta + \delta^2} \quad x_3 = \frac{-1 - \delta + 3\delta^2}{1 + \delta + \delta^2}$$

With $\delta = 1$, there is no discount in the future, and so we get $x_1 = x_2 = x_3$, i.e. an **equal split**. Otherwise, it is convenient to set x_1 higher. For $\delta = 0.8$, for example, $x_1 = 0.6393$, $x_2 = 0.3115$ and $x_3 = 0.0492$. Note that $\delta > (\sqrt{5} - 1)/2 = 0.618$ must hold, otherwise the constraint cannot be satisfied.

Now, for this choice to be **consistent**, it should remain invariant through time. That is, suppose we consider the player at the start of the second day, with $1 - x_1$ available resources, and consider the split between that day (the second) and the following one (the third). Will this lead to the same x_2 and x_3 previously computed, or will the play *regret* their previous decision?

So, let's consider:

$$w = u(x_2) + \delta u(x_3)$$

and maximize w with the constraint $x_2 + x_3 = 1 - x_1$, with x_1 constant. The result is:

$$x_2 = \frac{2 - x_1 - \delta}{1 + \delta} \qquad x_3 = \frac{-1 + 2\delta - \delta x_1}{1 + \delta}$$

From this, we see that x_2 and x_3 *agree* with the previously computed values! For example, with $\delta = 0.8$ and $x_1 = 0.6393$, we get again $x_2 = 0.3115$ and $x_3 = 0.0492$.

In fact, it can be shown that the exponential discount procedure (2.12) is the *only one* with this *self-similarity* property.

As a counter-example, consider the following *non-exponential* discounting:

$$v(x_1, x_2, x_3) = u(x_1) + \delta u(x_2) + \delta u(x_3)$$

With the same procedure of maximizing v we get:

$$x_1 = \frac{3 - 2\delta}{2\delta + 1} \quad x_2 = \frac{2\delta - 1}{2\delta + 1} \quad x_3 = \frac{2\delta - 1}{2\delta + 1}$$

Now we need $\delta > 0.5$. For $\delta = 1$ we still get an equal split, but with $\delta = 0.8$ we have $x_1 = 0.5385$, $x_2 = 0.2308$, $x_3 = 0.2308$.

However, when performing the same split on the *next* timestep, we get a different result. Specifically, we maximize:

$$w(x_2, x_3) = u(x_2) + \delta u(x_3)$$

This leads to:

$$x_2 = \frac{2 - x_1 - \delta}{1 + \delta} \qquad x_3 = \frac{-1 + 2\delta - \delta x_1}{1 + \delta}$$

and for $\delta = 0.8$, $x_1 = 0.5385$ we get $x_2 = 0.3675$ and $x_3 = 0.0940$, which is *inconsistent* with the previous choice. This is very strange: a *rational* player can anticipate all of this, and make the correct final decision directly at the start! For instance, if we consider the different choices at each timestep as the choices of different players over sequential turns, we can apply backward induction to find that $x_1 = 1$. In other words, the *past self* anticipates they will not split correctly in the following turns, and so they avoid that by taking all the resources.

2.5 Multistage games

Multistage games are a particular set of dynamic games in which the same number of players acts during repeated turns. This opens the possibility of considering **intermediate payoffs**.
These can be modelled as a finite sequence of T normal form **stage games**. They all involve the *same players*, and the total payoffs can be computed by aggregating their sequence of outcomes with some function (commonly, a sum).

The usual kind of multistage games is a sequence of 2 stage games with same players, but different action sets. Each game leads to partial payoffs $u_i^{(j)}$ for each player i, which are independent of the outcome of the previous games. The total payoffs are the *discounted* sums of partial payoffs for each player, with a discount factor δ *same* for all players (and this is common knowledge):

$$u_i^{\text{tot}} = \sum_{j=1}^{T} \delta^j u_i^{(j)}$$

For example, consider Al and Bob playing the Prisoner's Dilemma as a "first turn". Then, they both go out of jail, and they can either join a gang (G) or remain *alone* (L). This second stage ("Revenge") is a static game, with the following payoffs:

- If both stay alone (L), they never meet again: payoff 0 for both.
- If they both join a gang, they fight each other, leading to a negative payoff for both.
- If only one joins a gang, he/she will receive a *small* loss, and the other a *high* loss (nobody to defend him/her).

Explicitly:

$$\mathbf{P}_1 = \begin{array}{cc} & \text{Bob} \\ & \begin{array}{cc} m & f \end{array} \\ \text{Al}\ \begin{array}{c} M \\ F \end{array} & \left\| \begin{array}{cc} 4,4 & -1,5 \\ 5,-1 & 1,1 \end{array} \right\| \end{array} \qquad \mathbf{P}_2 = \begin{array}{cc} & \text{Bob} \\ & \begin{array}{cc} l & g \end{array} \\ \text{Al}\ \begin{array}{c} L \\ G \end{array} & \left\| \begin{array}{cc} 0,0 & -4,-1 \\ -1,-4 & -3,-3 \end{array} \right\| \end{array}$$

This can be mapped into the following extensive form:

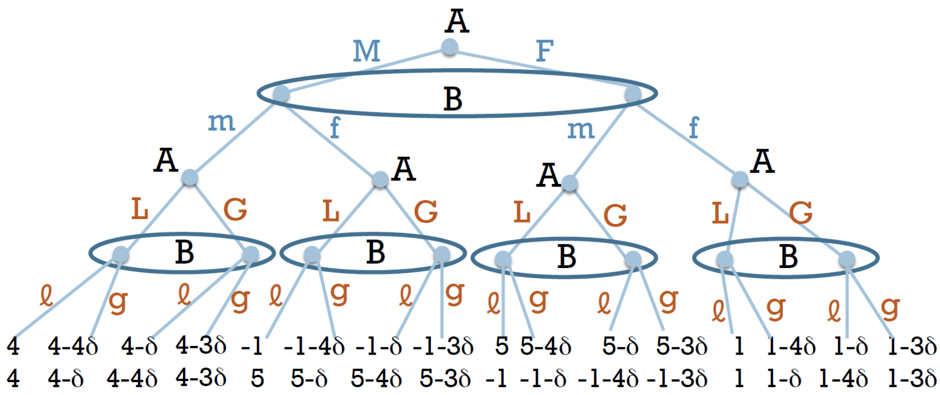

Figure (2.16)

However, from this it is not clear that the final payoffs are the combination of the partial payoffs in the two stages. Note how the number of available strategies increases *exponentially*. A strategy for Al needs to specify an action at each node she plays. Since actions are binary, we have $2^5 = 32$ possibility. Same holds for Bob, since his information sets are also 5.

We already know how to *solve* dynamic games: this amounts to find the *subgame perfect equilibria* (SPE). In multi-stage games this procedure is simplified by the fact that all stage games are **independent**. Thus, just taking a NE at each stage leads to SPE. This is formalized by the following theorem:

Theorem 2.5.1. *If s_j^* is a NE strategy profile for the j-th stage game, then there exists an SPE whose equilibrium path is $s_1^*, s_2^*, \ldots, s_T^*$.*

(But this does not mean that *all* the SPEs are of this form, as we will see!)
In the above example, we see that (F, f) is a NE of the first stage, and both (L, l) and (G, g) are NEs for the second stage. Then, we have the following SPEs:

- A plays (F, L, L, L, L), B plays (f, l, l, l, l)
- A plays (F, G, G, G, G), B plays (f, g, g, g, g)

2.5.1 Strategic connection

In the above discussion, we considered each stage by itself. This allows solving the game, but in a sense *removes* all the "strategy" from the game. For example, a player may consider *not playing* a NE at a previous stage, just to avoid incurring in some revenge at the later stages. In other words, early actions *do have consequences.* One does not want to insult the player he is expected to cooperate later on!

We will now see that if all stages have a unique NE each, then the *only* SPE is the one encompassing all these NEs, which is of the type explored above. However, if *multiple NEs* exists, there is space for more complex strategies!

First, we note that any Nash Equilibrium for the whole game must involve playing a NE in the last stage. This is because in that final part there are no more plans to be made: all the past is fixed, and now players just consider best responses, since there will be no "possibility of revenge" in the future, as the game will end immediately.

Theorem 2.5.2. *Any NE s^* (even if it is no SPE) of a multistage game (G_1, G_2, ..., G_T) must dictate a NE is played in stage game G_T.*

As a corollary, the following holds:

Theorem 2.5.3. *If G_1, G_2, ..., G_T all have a* ***unique*** *NE, then $(G_1, G_2, \dots, G_T)$ has a unique SPE.*

The idea is that all players know that, at the end on turn T, they will play a NE. But if this NE is unique, there is nothing to decide, the strategy at T is already defined. This means that the "effective final round" is now $T-1$, and from the above we know that players will choose a NE also there. Since it is unique, we can reiterate the argument up to the first turn.

However, if there are multiple NEs at the last round, then there can be SPEs involving non-NE at the previous stages! Note that this is not a contradiction: each stage is **not** a subgame, so an SPE *does not require* playing a NE at each stage! It just happens that doing so results in a particular kind of SPE.

The idea is to *steer* the game towards the "most convenient" NE at the end. For example, in the above game, there are two NEs at the end: (L, l) and (G, g). Of the two, (L, l) is clearly better for both players. The two, then, can use the *threat* of playing (G, g) as a way to enforce playing (M, m) at the first round, which is Pareto efficient, but not a NE.

Let's see how. First, consider the strategy $s_1 = (M, L, G, G, G)$ for A, and $s_2 = (m, l, g, g, g)$ for B. In other words, they choose (M, m) in the first round, and then (L, l) in the second. However, they would respond with (G, g) to any other outcome of the first round.

This is an SPE. To see that, we need to check that it is a NE in every subgame. Note that 4 subgames are in the second stage, in which both players consider only NEs. So, the only remaining check is to see if it is a NE for the whole game, i.e. if playing s_1 is a best response to s_2 and viceversa. But we already know that they are NEs for the whole second stage (i.e. nobody wants to unilaterally deviate during the second round), and so we need to effectively check only stage 1. Here, player 1 consider just two choices: M or F. Let's compare their expected utilities, knowing 2 will play s_2:

$$u_1(M, s_2) = 4 + 0\delta \qquad u_1(F, s_2) = 5 - 3\delta$$

If 1 plays M, 2 plays M, and both get $(4, 4)$ in the first stage. Then, following s_1 and s_2, they will play (L, l), getting $(0, 0)$ multiplied by the discount factor δ. Similarly, if 1 plays F instead, they will get $(5, 1)$ in the first round, but then play (G, g) in the second, leading to $(-3\delta, -3\delta)$.

Depending on the value of δ, the best response changes. In particular, M is better, meaning that (s_1, s_2) is an SPE, if δ is high enough:

$$4 > 5 - 3\delta \Rightarrow \delta \geq \frac{1}{3}$$

In other words, *betraying* the other player to get an immediate gain is convenient only if one values less long-term payoffs (low δ, meaning high discount, i.e. utility lowers quickly over time). Only in this case *threats* may be credible, and can have concrete impacts.

Moreover, this worked because we started with 2 NEs at the end, one clearly "better" (a "carrot" outcome) and one not (a "stick" outcome). This means that one of them (the *stick*) can be used as a threat.

In fact, the same procedure can work to create an SPE where the first move is *any*, for example (F, m).

2.5.2 One-stage deviation principle

These *deviations* from NEs can happen at most once, even during games with more than 2 stages. That is, it is not possible to deviate, for example, in turn 1 *and* 3. This is the statement of the **one-stage deviation principle**.

To prove it, we start with a few definitions.
A strategy s_i is **optimal** if there is no way to improve it for every information set h_i (no s'_i and h_i for which $u_i(s'_i, h_i) > u_i(s_i, h_i)$). In particular, it is **one-stage unimprovable** if there is no way to improve it by changing an action done in a *single* given information set h_i.

An optimal strategy is one-stage unimprovable by definition: the latter is a special case of the former. Surprisingly, the converse *does hold*!

Theorem 2.5.4. *A one-stage unimprovable strategy must be optimal.*

The idea is that, if we start with a non-optimal strategy a, we can always link it to an optimal strategy b by making a sequence of deviations, i.e. single changes to the one-round strategies. This forms the "optimal path" in *strategy-space* linking a to b. But then, if we consider *any intermediate strategy* c lying in that path, the optimal path linking c to b is the ones that *remains* on the initial *longer* optimal path from a to b. In other words, every part of the optimal path is an optimal path.

So, if we cannot do any step forward (i.e. we are at a one-stage unimprovable strategy), then we *must be* at the end of the optimal path (i.e. at an optimal path). Let's formalize this idea.

Proof. We proceed by contradiction. Assume s_i is one-step unimprovable but not optimal. This means that there is a way to improve it to a s_i', but it requires changing 2 or more steps. Suppose s_i' deviates from s_i under the information set h_i. The number of such deviations is finite, consider the **last** of them. Take the subgame starting at the deviation point (or at its parent node). This is a subgame for which there is a *single deviation* improving the payoff of player i, meaning that s_i can be improved by a single deviation, which is a contradiction. Thus, no better s_i' may exist, meaning that s_i must be optimal. □

2.6 Repeated Games

(Lesson 15 of 17/11/2020) Compiled: January 1, 2021

A **repeated game** $G(T, \delta)$ is a dynamic game where the *same* static game G is played as a stage game T times. All intermediate payoffs are discounted by δ and aggregated to form the total payoff. If T is finite, the game has a *finite horizon*, otherwise it has an *infinite horizon* (in this case, δ must be < 1, i.e. there must be discounting).

Unique NE

For example, consider $T = 2$ (two turns) and $\delta = 1$ (no discount), with the following normal form for each stage (Prisoners' dilemma):

$$\mathbf{P} = \begin{array}{cc} & \text{Bob} \\ & \begin{array}{cc} m & f \end{array} \\ \text{Al} \begin{array}{c} M \\ F \end{array} & \begin{Vmatrix} 4,4 & 0,5 \\ 5,0 & 1,1 \end{Vmatrix} \end{array}$$

In the final stage, there is only one NE: (F, f). Thus, it *must* be played in any rational strategy. Effectively, this means that each player gets a guaranteed payoff of 1 from the second stage. So, we can *merge* both rounds in a single one, with all payoffs increased by 1:

$$\mathbf{P}_{\text{tot}} = \begin{array}{cc} & \text{Bob} \\ & \begin{array}{cc} m & f \end{array} \\ \text{Al} \begin{array}{c} M \\ F \end{array} & \begin{Vmatrix} 5,5 & 1,6 \\ 6,1 & 2,2 \end{Vmatrix} \end{array}$$

There is, again, only one NE: (F, f). So both player will choose (F, f) in both rounds.

In general, this happens for all *finitely* repeated games with only one NE. The mere possibility of anticipating which choice to make in the *last* round, determines how all the previous ones are played.

Formally: repeated games are a particular case of multi-stage games. Thus, the outcome of last stage is a NE, *always*. Then, as a corollary of the previous theorems:

Theorem 2.6.1. *If a stage game G only has a single NE $\boldsymbol{s}^*$, then $G(T, \delta)$ has a unique subgame-perfect outcome, i.e. play $\boldsymbol{s}^*$ in every stage.*

This is merely a consequence of backward induction.

Multiple NEs

Things get more interesting if there are *multiple* NEs. For example, consider the following:

$$\mathbf{P} = \begin{array}{cc} & \text{Bob} \\ & \begin{array}{ccc} M & F & H \end{array} \\ \text{Al} \;\begin{array}{c} M \\ F \\ H \end{array} & \begin{Vmatrix} 4,4 & 0,5 & 0,0 \\ 5,0 & 1,1 & 0,0 \\ 0,0 & 0,0 & 3,3 \end{Vmatrix} \end{array}$$

Here there are two NEs: one "good" (H, H) and one "bad" (F, F).
In this case, it is possible to play a non-NE in the first stage so that the good NE may be selected. Specifically, the players choose (M, M) in the first round, and (H, H) in the second if they both played (M, M), and (F, F) otherwise (it is important to specify also responses to moves that do not happen!).
Note that this does not involve any kind of information exchange, but just *speculation*.

Strategies with repeated NE are all SPE too. From the mathematical point of view, these are all equivalent. How to *force* players to choose one rather than another is a task-specific engineering problem.

Philosophically, we can say that repeated games *tend* to introduce *cooperation*, even if players are selfish. In fact, while the last stage is always "egoistically played", the presence of multiple *egoistic* NEs can lead to "collaborative" NEs (with the *carrot and stick* mechanism). However, all NEs are equivalent, and there is no way to tell which will be played in the end. However, *external influences* (society, religion, etc.) could make players prefer the collaborative NE. In a sense, these are the "engineering ways" to optimize the outcomes!

An additional remark is that cooperative NEs require **consistency**. In other words, players should "stick" to their strategies, and don't "renegotiate".

For example, in the previous setup one of the two players may decide do *switch* their choice to F in the first round, gaining an advantage, and then convince the opponent to still play into the (H, H) *good* equilibrium. The other may threat towards (F, F), but this would hurt *both players* so it is less *credible*.
In other words, a *greedy* player may still *not cooperate* and still *gain everything*, if they can *renegotiate*, and there are no *credible* responses.
This undermines the previous philosophical argument for cooperation.

One way around that, is to have additional NEs that *punish* players that do not cooperate. In the previous example, playing (F, F) at the final round *hurts* *both players*. So, suppose we add two other NEs:

$$
\mathbf{P} = \begin{array}{cc} & \text{Bob} \\ \text{Al} & \begin{array}{c|ccccc|} & M & F & H & P & Q \\ M & 4,4 & 0,5 & 0,0 & 0,0 & 0,0 \\ F & 5,0 & 1,1 & 0,0 & 0,0 & 0,0 \\ H & 0,0 & 0,0 & 3,3 & 0,0 & 0,0 \\ P & 0,0 & 0,0 & 0,0 & 4,.5 & 0,0 \\ Q & 0,0 & 0,0 & 0,0 & 0,0 & .5,4 \end{array} \end{array}
$$

Here there are 4 NEs: (F, F), (H, H), (P, P) and (Q, Q). All except (F, F) are Pareto efficient, and (M, M) is the *best*.

Players can anticipate that, if they play (M, M) in the first round, they will collaborate with (H, H) in the second. However, if someone deviates, the other can choose the NE that will punish him/her, without losing everything. For example:

- $(M, \neg M)$ in the first round leads to (P, P), punishing the second player. Similarly, $(\neg M, M)$ leads to (Q, Q), punishing the first player.
- If both deviate $(\neg M, \neg M)$, they won't trust each other and lead to (F, F).

So, philosophically, frequent interactions with the availability of more punishment options incentives cooperation.

2.6.1 Infinitely repeated games

An infinitely repeated game, with stage game G and discount factor δ, is denoted as $G(\infty, \delta)$. To avoid diverging expected utilities, we need $\delta < 1$. This allows comparing different payoffs that are infinitely repeated. For example, is it better to receive 1 forever, or 3 forever? Both sums are infinite, so it does not matter! If we introduce payoffs, instead, it is clear that the latter is best, because we are comparing:

$$
\frac{1}{1-\delta} = 1 + \delta + \delta^2 + \cdots < 3(1 + \delta + \delta^2 + \dots) = \frac{3}{1-\delta}
$$

In infinitely repeated games we cannot apply backward induction: there is no *last* stage to begin from! As we will see, this allows *cooperation* even *without punishments*.

Since there is no last stage, there is no requirement to play a NE of G. In particular, there may be an SPE of $G(\infty, \delta)$ in which no stage's outcome is a NE of G. The key point is that, in a infinitely repeated game, *every* subgame is the same *infinitely repeated game*.

As an example, consider again the Prisoner's Dilemma. Since stages are repeated *ad infinitum*, it would be convenient to play (M, M) forever. To *force* that, players use F as a threat. The idea is to exploit the infinite nature of the game to make this threat *final*.
Thus, we define a **grim trigger** strategy (GrT) as follows:

- Start playing M at stage 1
- Play M if *all* past outcomes were of the kind (M, M). Otherwise, play F.

Note that *any deviation* of the opponent will result in an eternal sequence of F from the other! Even if this is not credible (playing (F, F) hurts both players), the severity of a misstep *infinite*, making the *grim trigger* an important strategy to consider.

It can then be shown that:

Proposition 2.6.1. *For a δ sufficiently close to 1, the joint strategy where both users play GrT is an SPE.*

Proof. First we need to show that GrT is a best response to itself. If Bob assumes that Al will play GrT, he knows that, whenever the outcome is **not** (M, M), Al will play F forever. Thus, also for Bob it is optimal to play F *forever* if the outcome is not (M, M). This proves that also Bob will play F if even a single previous outcome is not (M, M), which is *part* of the GrT strategy.

We now need to understand what Bob should play *if* all the previous outcomes are (M, M). This is the same as asking what Bob should play in his first move. In fact, in the case of no previous history, if Bob plays M, and Al the GrT, then we will have (M, M). So, the next stage will be exactly the same – with no reason to deviate. However, if Bob plays F at first, then we will get (M, F), which is advantageous for Bob. But now the GrT activates, and all the following turns will result in (F, F). So, we need to balance a one-turn gain with an infinitely long loss. The expected value V of a sequence $(F, F, \dots)$ from Bob, when Al plays the GrT, is:

$$V = 5 + \delta \cdot 1 + \delta^2 \cdot 1 + \cdots = 5 + \frac{\delta}{1-\delta}$$

If instead Bob chooses M in the first round, he will do so always (all the conditions will be the same forever, so no incentive to change), resulting in a sequence of (M, M), with a value V':

$$V' = 4 + \delta \cdot 4 + \delta^2 \cdot 4 + \cdots = \frac{4}{1-\delta}$$

Let's compare them:

$$V = 5 + \frac{\delta}{1-\delta} < \frac{4}{1-\delta} = V'$$

This holds if $\delta > 1/4$. That is, if Bob *sufficiently cares about the future*, then V' is better, and so he will play the GrT too.

This proves that, for $\delta > 1/4$, the GrT is the best response to itself, and so it is a NE for the whole game. But there are only two kinds of subgames:

- Those where all the previous stages are (M, M), which are exactly the same as the whole G (since it is infinite). We have just proven that the GrT is a NE for G.

- Those where at least one stage deviated from (M, M). In this cases, the GrT becomes "always play (F, F)", which is a sequence of NE in all stages, and so it is a NE of the entire subgame.

Since the GrT is a NE in *all* the subgames, it is an SPE. □

Friedmann Theorem

In the above example, we have seen how *cooperation* naturally emerges from infinite repetition. This is actually a particular case of a more general theorem, which – surprisingly – has no author associated. It was first cited in a paper by Friedman as a *known result*, but nobody could trace it to previous publications. That is why it is sometimes cited as "the Folk Theorem", in the sense of "something that everybody knows".

We begin with a definition. A **feasible payoff** is *any* convex combination of pure-strategy payoffs. That is, if a game is played for an infinite amount of time, any linear combination $w_1(u_{1a}, u_{2a}) + w_2(u_{1b}, u_{2b}) + \dots$ of the payoffs with weights summing to 1 ($\sum w_i = 1$) is a possible (feasible) value for the *average* payoff of a stage[4].

Graphically, these are all the points inside the *convex hull* of the pure payoffs:

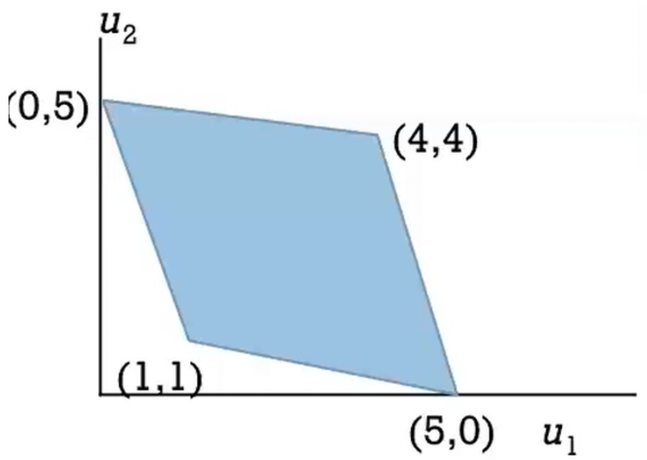

Figure (2.17) – **Feasible payoffs** are the points (u_1, u_2) inside the convex hull of the pure-strategy payoffs.

Theorem 2.6.2 (Friedman Theorem)**.** *Let G be a finite static game of complete information. Let $(e_1, e_2, \dots, e_n)$ be the payoffs from a NE of G. Let $(x_1, x_2, \dots, x_n)$ be feasible payoffs s.t. they are element-wise bigger than the above, i.e. $x_j > e_j$ $\forall j$. Then, if δ is close to 1, $G(\infty, \delta)$ has an SPE with payoffs (x_j).*

Proof. The idea is that any of these feasible payoffs can be obtained with a well-constructed GrT. □

[4]∧These are also the possible values of the expected utility of a mixed strategy. However, note that in an infinitely repeated game we can realize them by playing pure strategies alone, since we have an *infinite* number of payoffs to average!

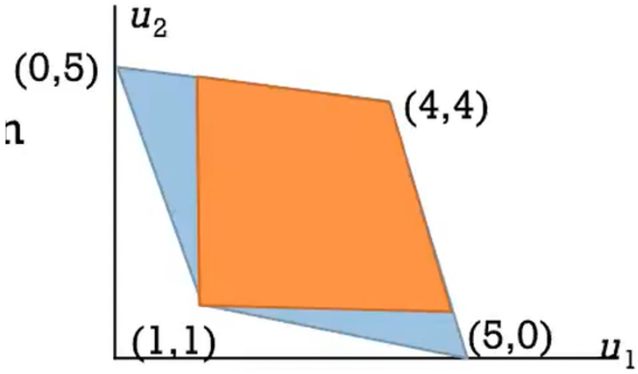

Figure (2.18) – In G, $(1, 1)$ is a NE. All the points with both utilities > 1 are inside the orange area. These are the (x_1, x_2) feasible payoffs that can form an SPE.

In the Prisoner's Dilemma, the only "punishment" available is playing (F, F). However, in general, the maximin can be used as a *worse* punishment, since it is $\leq$ the payoff at the NE. In particular, in the above we can replace $(e_1, e_2, \dots)$ with the security payoffs $(r_1, r_2, \dots)$. Still, note that any *threat* must be *credible* to be effective, i.e. *rational* from the point of view of the one that is enforcing it.

Finally, note that all of this depends on the choice of δ. A small δ makes all punishments less effective, because players focus on short-term rewards.

Finite memory

In engineering, a "grim reaper" strategy can be difficult to implement, since it requires an arbitrarily large amount of memory to hold all the past interactions. Fortunately, there are approximations (in the sense that they are NE, but not SPE) available to solve this problem.

One example is the **Tit-for-That** strategy (TFT). At stage t, the player i chooses the move (cooperate/defect) played by the opponent $-i$ at the previous stage $t - 1$. Two players implementing TFT reach the same "optimal" equilibrium path as before (mutual cooperation). However, TFT is also *forgiving*, in the sense that a misstep from one of the parts can be *corrected* at later stages. So, if there is some error and a player makes a bad move, there is still the possibility to *converge* again to the optimal path.

However, TFT requires synchronization between the two players. Two "unsyncrhonized" TFT players will always choose (M, F) or (F, M), effectively *minimizing* their payoffs (a *death spiral*, with no cooperation).
This illustrates why TFT can't be an SPE. However, in a good implementation these outcomes *never happen*, and so the result will be the same of GrT.

There are *less forgiving* variants, e.g. "Tit for Two Tats", in which two *good choices* from the opponent are needed to forgive a past mistake.

2.6.2 Reputation

Repeated interactions can develop **trust** between the players, i.e. a higher propensity towards cooperation.

For example, consider the following game:

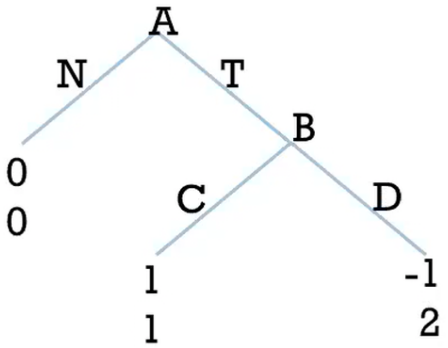

Figure (2.19) – A can either (T)rust B or not. If trusted, B can either (C)ollaborate or (D)efect.

From backward induction, we see that B defects, and so A anticipates this and does not trust him/her.

However, if the game is infinitely repeated, we can build a GrT as follows:

- In the first turn, A chooses T. Then A chooses T as long as the previous outcomes are (T, C). Otherwise, he/she will always play N.
- B chooses to always play C, which is the best response if $\delta > 1/2$.

Is there a way to give B more incentive to cooperate, other than the *grim* possibility of A playing GrT?

Guarantees

Let's introduce a third *dummy* player G, called the **guarantor**. His function is to improve B's reputation. The idea is that B *can pay* G in advance, providing an **insurance** of 2. Then, if B defects, G keeps the insurance. Otherwise, G returns the insurance to B, keeping a small fraction (0.1) for the service.
The extensive form is now:

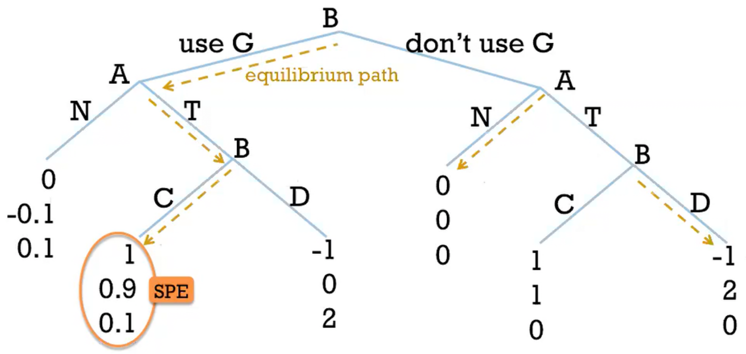

Figure (2.20) – Adding a *guarantor* G builds B's trust, allowing cooperation without the need of eternal punishments. Note that, to fully specify an SPE, we need to denote also the choices *outside* the equilibrium path, i.e. the responses to moves that do not happen.

If B pays G, he/she gains *trust*, making A want to cooperate. Note that this happens without the need of (infinite) repetition!

However, this introduces another problem. In the real case, G has an incentive to *flee with the reward*. That is, if we allow him to keep the insurance or return it to B, the strategy of keeping it *strictly dominates* the other.

Still, if the game is repeated, G may have an interest of prove himself as a trustworthy guarantor, establishing its reputation in the market. Quantitatively, if he keeps the insurance he gets a payoff of 2, and then 0 forever. If he cooperates, he gets 0.1, but remains in business. On the long run, he will get:

$$0.1 + \delta \cdot 0.1 + \delta^2 \cdot 0.1 \cdots = \frac{0.1}{1-\delta}$$

which is > 2 if $\delta > 0.95$.

2.7 Dynamic Bargain

(Lesson 16 of 20/11/2020) *Compiled: January 1, 2021*

An important application of dynamic games is that of **dynamic bargain**, where players need to split a sparse resource.

For example, consider a quantity of 1 which is shared between 2 players, so that the first gets x, and the second $1-x$.

There are two possible approaches to *decide* the optimal x.

- **Nash bargaining**, a simple axiomatic and **static** approach. Contrary to the name, this does not involve Game Theory concepts.
- **Dynamic bargaining**, in which one player proposes a split (proposer), and the other accepts/rejects it (responder), and the role are exchanged at every turn.

In this section, we will focus on the second one. The game proceeds as following:

1. In stage 1, Player 1 is the (P)roposer, and 2 is the (R)esponder. P offers shares $(x, 1-x)$, and 2 can accept, ending the game, or refuse, proceeding to the next turn.
2. In stage 2, roles are reversed: $P = 2$, $R = 1$. Then P proposes a share, and R responds to it, exactly as before.
3. In general, at stage t, $P = 1$ if t is odd, and otherwise $P = 2$.
4. Since we need a final result, the game can't go on indefinitely. So, we set a **deadline**, i.e. after T stages the game will end with both 1 and 2 getting nothing. Moreover, after each round payoffs are *discounted* by a factor δ, effectively "wasting" a fraction $1-\delta$ of resources. So, if the game ends at stage 1, they get $u_1 = x, u_2 = 1-x$, otherwise at turn t they receive $u_1 = \delta^{t-1}x, u_2 = \delta^{t-1}(1-x)$.

When $T = 1$, we get the **ultimatum game**, seen in example 2. In this case, all solutions with P proposing $(x, 1-x)$ and R accepting are NEs, and the only SPE is the one where P *keeps everything*, i.e. $x = 1$.

When T is **odd**, player 1 is the last proposer. So, at round T the setup is that of the ultimatum game: player 2 will accept everything, and 1 will propose $x = 1$, leading to $u_1^{\text{fin}} = \delta^{t-1}, u_2 = 0$.
Knowing this, player 2 will try to terminate the game *before* the final round, by making an offer that 1 can accept, i.e. $x \geq \delta$. In this way, if 1 accepts he/she gets $u_1 \geq \delta \cdot \delta^{t-2} \geq \delta^{t-1}$, which is the same or better than the payoff u_1^{fin} obtained by reaching the final round, and so 1 *will* accept such a split (there is no meaningful deviation from his/her part). The advantage is that now $u_2 \geq 0$, and it is maximized for $x = \delta$.

This reasoning can be iterated *backward*, so that the game can be concluded at the first round, leading to the following payoffs:

$$u_1 = \frac{1+\delta^T}{1+\delta}; \qquad u_2 = \frac{\delta - \delta^T}{1+\delta}$$

Similarly, if T is even, the final proposer will be player 2, and everybody knows that. So, he/she will be able to exploit this fact, and the results will be the same, but with roles reversed: $u_1 \leftrightarrow u_2$.

As a consequence of backward induction, the **unique** SPE must involve the game ending at turn 1. We can see this as the effect of two principles:

- Playing at a later round *wastes* some resource, so players want to avoid that.
- Rational players can anticipate what the final share will be, and use that to reach an agreement at previous rounds.

Note that this is **not** a repeated game, due to degradation and the termination condition (i.e. the number of rounds is not defined in advance).

Interestingly, this same reasoning can be applied to an *infinite horizon*, even though backward induction does not work there. The idea is that, even in this case, players do not want to play later rounds to avoid resource degradation. Then, for $T \to +\infty$:

$$u_1 = \frac{1}{1+\delta}; \qquad u_2 = \frac{\delta}{1+\delta}$$

and for $\delta \to 1$, we have an equal split. Otherwise, the split favors player 1, because he/she goes first, and so has a first-move advantage.

However, since we are not using backward induction, we need a different proof for the uniqueness of this SPE. This can be done by *contradiction*.

Proof. Assume there is more than one SPE. For player 1, the best payoff is v_1, and the worst is w_1. Conversely, for 2 the best payoff is $v_2 = 1 - w_1$, and the worst is $w_2 = 1 - v_1$, because he/she gets what 1 leaves out.
Now 1 *does not want* the game to go on after the first turn, and so needs 2 to accept immediately the first split. Effectively, 2 does so if going on can't improve his/her payoff. So, suppose 2 refuses the first split. Now 2 can *copy* 1's proposal, and get either $v_2 = \delta v_1$ or $w_2 = \delta w_1$ (due to the resource degradation).

If $v_2 > w_2$, it is convenient for 2 to *continue* the game, so that he/she can make this proposal. 1 anticipates this, and sets $v_2 = w_2$, so that 2 won't have any incentive to continue playing, and will accept the first split, as desired.

Putting everything together:

$$v_2 = \delta v_1 = 1 - w_1 \overset{!}{=} w_2 = \delta w_1 \Leftrightarrow v_1 = w_1 = \frac{1}{1+\delta}$$

But this means that the SPE is unique. $\square$

2.8 Dynamic duopolies

Let's consider a **Stackelberg duopoly**, which is a natural dynamic extension of Cournot duopoly.

The idea is that a *dominant* (leader, 1) firm moves first, and a *subordinate* (follower, 2) firm moves second. They both decide the quantities of goods q_1 and q_2 to produce, incurring in a cost $C(q) = cq$, with constant c. The market price is $P(Q) = a - Q$, with $Q = q_1 + q_2$, and $a > c$.

We proceed by backward induction. 1 *knows* that 2 will choose their best response. The profit of 2 is given by:

$$u_2(q_1, q_2) = q_2(a - q_1 - q_2 - c)$$

Let $R_2(q_1) = \arg\max_{q_2} u_2(q_1, q_2)$, i.e. the best response available to 2 for a given choice of q_1. By maximizing the above we see that:

$$R_2(q_1) = \frac{a - q_1 - c}{2}$$

Note that this quantity has already appeared when discussing Cournot's *mono*poly (see (1.5), pag. 35), with $a \to a - q_1$. Intuitively, 2 acts as a *monopolist* with the "left-over resources" by 1.

Now, 1 anticipates all of this, and moves accordingly by maximizing his/her utility:

$$q_1^* = \arg\max_{q_1} u_1(q_1, R_2(q_1)) = \arg\max_{q_1} q_1(a - q - 1 - R_2(q_1) - c) =$$
$$= \arg\max_{q_1} \frac{q_1(a - q_1 - c)}{2}$$

which leads to:

$$q_1^* = \frac{a-c}{2} \qquad q_2^* = \frac{a-c}{4}$$

In this setup, 1 has a clear advantage. In fact, recall that in the Cournot's duopoly we had $q_1^* = q_2^* = (a-c)/3$.

Virtual threats

Player 2 can (virtually) *threat* 1 by responding to any choice different from the *fair* one ($(a-c)/3$) with a very high q_2, which *hurts* both players. However, this is irrational, so it is an empty **non-credible** threat. 2 *does not want* to decrease u_2, and so he/she will choose $q_2 = R_2(q_1)$.

Note also that, in this setup, 2 has more information, but this does not result in any advantage for him/her. In fact, 1 can anticipate this (1 *knows* that 2 *will know*) and move accordingly (first-move advantage). In a competitive game (such as the one just considered), knowing more is actually a *disadvantage*.

More information can be bad

In fact, it would be better for 2 to *ignore* q_1, which would make the game static, and lead back to the $(a-c)/3$ fair share.

Static version

The same result can be obtained even from *fictitious play*, i.e. by allowing 1 (and then 2) to *change* their moves. For example, suppose 2 assumes 1 to play $q_1 = (a-c)/2$, and so sets $q_2 = (a-c)/4$. Knowing this, 1 plays a better $q_1 = 3(a-c)/8$. But this changes 2's best answer again. Iterating, the sequence of moves will converge to $q_1^* = q_2^* = (a-c)/3$.

Fictitious play

As previously observed, the aggregated production at the static NE is *higher* than that of a monopoly:

(Lack of) Trust

$$\frac{2}{3}(a-c) > \frac{a-c}{2} \equiv q_m$$

This means that both firms are producing more, and earning less, and they are doing so because they *do not trust* each other! Specifically, the best response to $q_1 = (a-c)/4$ is **not** $q_2 = (a-c)/4$.

However, we expect that by *infinitely repeating* the same static game there should be a way to build trust, according to Friedman's theorem. The idea is to use a Grim Trigger (GrT) strategy:

Gaining trust through infinite repetitions

- At $t = 1$, produce $q_m/2$ (half of the monopoly quantity)
- At $t > 1$, produce $q_m/2$ if in every previous stage $u < t$ the production was $q_m/2$ for both firms. Otherwise, produce $q_c = (a-c)/3$ forever after.

And this works for a *sufficiently high* discount factor δ. To find it, we can repeat the same steps used in proving that GrT is an SPE for the repeated Prisoner's Dilemma.

So, suppose 1 plays the GrT, and so:

$$q_1 = \frac{q_m}{2} = \frac{a-c}{4}$$

What is the best response of 2 during the first stage?

- **Defect**. 2 chooses the best response, which is:

  $$q_2 = \arg\max_{q_2} q_2(a - q_2 - q_m/2 - c) = \frac{3}{8}(a-c)$$

  This leads to a profit:

  $$u_D = \frac{9}{64}(a-c)^2 \tag{2.13}$$

- **Collaborate**. 2 chooses $q_2 = u_m/2$, gaining:

  $$u_2 = \frac{u_m}{2} = \frac{(a-c)^2}{8}$$

As expected, defecting at the start is advantageous for 2. However, we need to account also **future** rewards.
If 2 defects, they will get u_c forevermore, leading to a total payoff of:

$$u_D + \delta \frac{u_c}{1-\delta}$$

However, if they collaborate the final payoff will be:

$$\frac{u_m}{2(1-\delta)}$$

Inserting $u_m/2 = (a-c)^2/8$, $u_c = (a-c)^2/9$ and (2.13), and then comparing the two expressions, we see that collaborating is better if $\delta \geq 9/17$.

If $\delta < 9/17$, the GrT is no longer an SPE. However, there is still a way to improve over playing q_c. The idea is to choose a "less ambitious" GrT' with objective $q^* \in [q_c, q_m/2]$, consisting of "play q^* at the start, and after any deviation stay at q_c forever".
When both firms play q^*, they get:

$$u^* = q^*(a - 2q^* - c)$$

In this case, the ("myopic") best response at first turn is:

$$q_D = \frac{a - q^* - c}{2}$$

which leads to:

$$u_D = \frac{(a - q^* - c)^2}{4} > u^*$$

Accounting for the future rewards, the GrT' is better if:

$$\frac{u^*}{1-\delta} \geq u_D + \delta \frac{u_c}{1-\delta}$$

Inserting all the values leads to:

$$\frac{q^*(a - 2q^* - c)}{1-\delta} \geq \frac{(a - q^* - c)^2}{4} + \delta \frac{(a-c)^2/9}{1-\delta}$$

The minimum q^*, corresponding to $\max u^*$, attainable with a given δ, is the one that *saturates* the inequality:

$$q^* = (a-c)\frac{9 - 5\delta}{3(9-\delta)}$$

And for $\delta \in [0, 9/17]$, $q^* \in [q_c, q_m/2]$.

As a final remark, note that the GrT can be improved by adding a *higher threat*, i.e. not playing the NE, but a higher (and still credible) $q_{\text{threat}} > q_c$. Formally, we denote with (R)eward the action of choosing $q_m/2$, and with (P)unishment that of producing $q_{\text{threat}} \geq q_c$ (but not $\gg q_c$). Then, consider the following:

1. Start with R (Reward).

2. At stage t, choose R if both firms played R at stage $t-1$, *or even if they both defected, by playing* P (forgives a mutual betrayal). Otherwise, play P.

It can be shown that this works for $\delta = 1/2$ and $q_{\text{threat}} = 2(a-c)/5 > q_c$.

All this discussion shows that while a *static* Cournot duopoly can't lead to a "shared monopoly" (i.e. a *cartel*), infinite repetitions can lead to that! This is through GrT-like strategies, which require *no interaction between the players.* In other words, trust can develop *over time* without the need of explicit agreements. In practice, this makes countering real-life cartels a very difficult problem, since they could arise just from the *rationality* of the players.

2.9 Exercises

(Lesson 17 of 24/11/2020) Compiled: January 1, 2021

Exercise 2.9.1:

An investment fund is jointly opened by Ava (A) and Brett (B), who simultaneously invest 5000€ each at time 0. The investment fund is supposed to be left untouched for 4 years. If this happens, both investors will receive x € each at year 4. However, at the end of each intermediate year, that is, at years 1, 2, 3, A and B can decide to keep the money in the fund (K) or withdraw it (W). They make this decision independently and without consulting with one another. Early withdrawal of the money implies a penalty, so they can withdraw only $6000 + 500t$ € overall at year $t = \{1, 2, 3\}$. If either of them withdraws the money before the end date, he/she can get this entire sum, the other gets nothing. If they both apply for early withdrawal in the same year, each of them gets half of the amount allowed for withdrawal. Any kind of early withdrawal (by one or both players) closes the investment and ends the game. The player have a payoff equal to the money the eventually get, without any discount factor.

1. Represent this game in extensive form.

2. Define the players' strategies and write them down.

3. Discuss the subgame-perfect equilibria of this game, in two cases: (a) $x = 7200$; (b) $x = 8000$.

Solution.

1. The game consists of 3 turns, each involving a simultaneous decision of both players, with only one outcome (K, K) continuing the game:

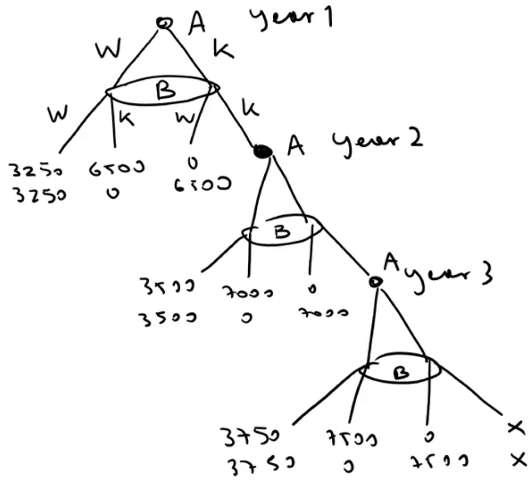

Figure (2.21)

This looks like a *simultaneous* trust game.

2. Every player takes a move in 3 possible information sets (B has 6 nodes, but can only distinguish pairs of them, due to the simultaneity of moves). The strategy of each player corresponds to designing a binary choice for each of them, for a total of $2^3 = 8$ strategies: WWW, WWK, WKW, WKK, KWW, KWK, KKW, KKK. We need to *list* all of them to identify SPEs. However, note that playing W at any round *ends* immediately the game, so in practice WWW and WWK behave the same. Effectively, to track the game evolution we only need to distinguish W, KW, KKW and KKK. More precisely, W is *equivalent* to $\{WWW, WWK, WKK\}$, and so we may write it as Wyy, with a generic $y \in \{W, K\}$.

3. Let's start with $x = 7200$. In this case, the last round can be represented in normal form as follows:

$$\mathbf{P} = \begin{array}{cc} & \text{B} \\ & \begin{array}{cc} W & K \end{array} \\ \text{A} \begin{array}{c} W \\ K \end{array} & \left\| \begin{array}{cc} 3750, 3750 & 7500, 0 \\ 0, 7500 & 7200, 7200 \end{array} \right\| \end{array} \tag{2.14}$$

Note that K is a strictly dominated strategy for both players, and so it is never played. Thus, there is only one NE in the last round, which is to play (W, W). Both players can anticipate this, and so we can *scale* the payoffs of all previous rounds, adding a guaranteed 3750 to all of them. Then we can repeat the same reasoning, finding that also in the previous round K is a strictly dominated strategy, and so it is in the first round. So, the SPE is unique, and both players choose WWW as a strategy.

If $x = 8000$, the last round is different:

$$\mathbf{P} = \begin{array}{cc} & \text{B} \\ & \begin{array}{cc} W & K \end{array} \\ \text{A} \begin{array}{c} W \\ K \end{array} & \left\| \begin{array}{cc} 3750, 3750 & 7500, 0 \\ 0, 7500 & 8000, 8000 \end{array} \right\| \end{array} \tag{2.15}$$

There are two NEs: (W, W) and (K, K) (there is also a mixed NE, but it does not matter here). This means that the SPE is not unique. In particular, both (WWW, WWW) and (KKK, KKK) are SPEs. In fact, any strategy involving playing a NE at each subgame is, by definition, an SPE. So also (WWK, WWK) and (WKK, WKK) are SPEs.

In the last round, there is actually also a mixed NE, similar to the one found in the Battle of Sexes. This can be found by applying the principle of indifference. Let p be the probability of A playing W. Then, we want B to be indifferent in choosing W or K:

$$u_B(p, W) = 3750p + 7500(1-p) \stackrel{!}{=} 0p + 8000(1-p) = u_B(p, K)$$

which results into $p = 2/13$. The expected payoffs are:

$$u_B(p, W) = u_B(p, K) = 8000\frac{11}{13} = 6769$$

However, this NE is *strictly dominated* in round 2.

Nonetheless, if we denote this strategy with m, we see that (WWm, WWm) is an SPE, because m is a NE in the last round. Still, such kind of strategy is really the same as (WWW, WWW) or (WWK, WWK), as in both the game ends immediately.

Exercise 2.9.2:

Two players A and B take turns, starting from A. They begin the game with two peas each. A legal move in the game consists in transferring to the opponent any integer number of peas. It is not allowed to transfer a quantity of peas that was transferred by someone before. Zero peas cannot be given either. Anyone who cannot make the next move according to the rules is considered a loser. Who will win this game?

Solution. This is a sequential game, and its optimal *solution* can be found through backward induction. First, we need to construct the game's extensive form. We label each node with a game's state (x, y), meaning that A has x peas, and B has y peas. We arrive to:

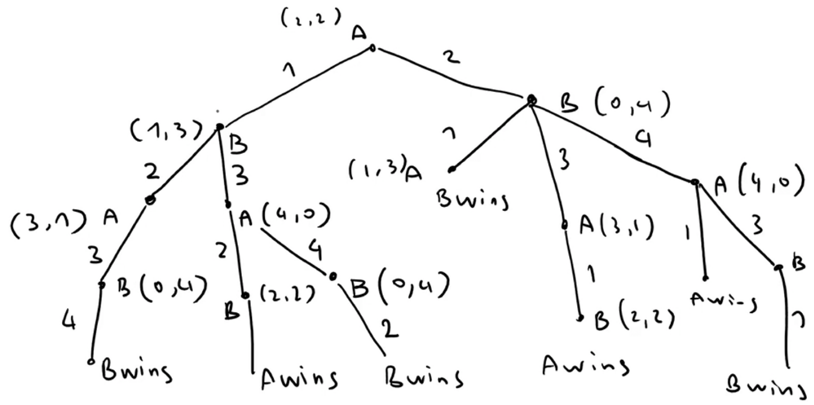

Figure (2.22)

Then we proceed with backward induction. Let's start from the left side. $(2, 2)$ and $(0, 4)$ are both final nodes, and A (at $(4, 0)$) can choose to either win, or lose. Clearly he/she will decide the former, and so the branch with $(0, 4)$ is never played. Now B (at $(1, 3)$) chooses between $(3, 1)$ and $(4, 0)$. The first is a winning move, the second leads to a loss, and so they will play the former. Thus, A at the first round can't choose 1 as an action, as this will result in B winning.

We can proceed similarly in the right half, finding that also here B will win in the end. So, independent of A action, B will *always* win.

Exercise 2.9.3:

Wife (W) and Husband (H) share the chores of the house. Due to her working shift from 10:30 am to 6:30 pm, W has two opportunities to do something, at 7:00 am and 7:00 pm. H instead works two shifts, early morning and afternoon till late evening, so he has only one opportunity to contribute, at 2:00 pm. There are two particular chores in the maily: fix some (F)ood or (C)lean the house. If either family member performs a chore, he/she pays an individual cost, but *both* members receive an identical benefit: F gives benefit 30 but costs 10, C gives benefit 50 but costs 20. Moreover, any of these actions performed by W in its first opportunity costs 10% more, but also gives 10% more benefit to both players (e.g. F costs 11, but gives benefit equal to 33). Alternatively, any player at his/her turn can "do nothing" (N) instead of choosing either chore. This can be done at any round, and gives individual benefit 10 only to the member choosing it (the social benefit is of course 0). Also, a chore cannot be repeated in the day (if food is already prepared ina previous round, the current action taker can only clean the house or do nothing). Action N can be repeated through all the rounds.

1. Write down the extensive form of the game.
2. How many subgame-perfect equilibria are in this game? Also, name the process on how do you find them.

3. Actually find them.

Solution.

1. This is a sequential game with no simultaneous decisions. The extensive form is as follows:

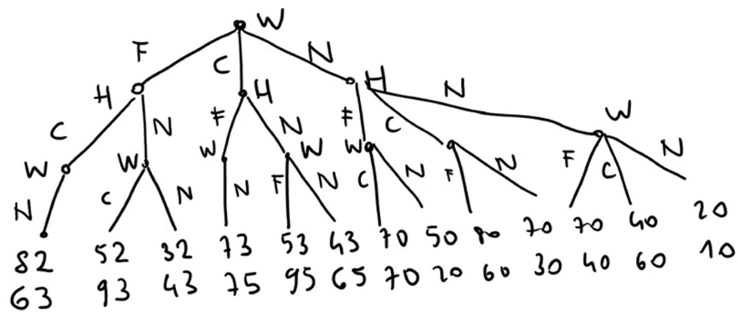

Figure (2.23)

2. SPEs can be found through **backward induction**. We *assume* W and H to be generous, in the sense that when choosing between equal payoffs for them, they opt for the outcome that favors the other. Then, since there are no repeated payoffs, the SPE must be unique, and is to be found in pure strategies.

   After removing the branches in the last round, we are left with:

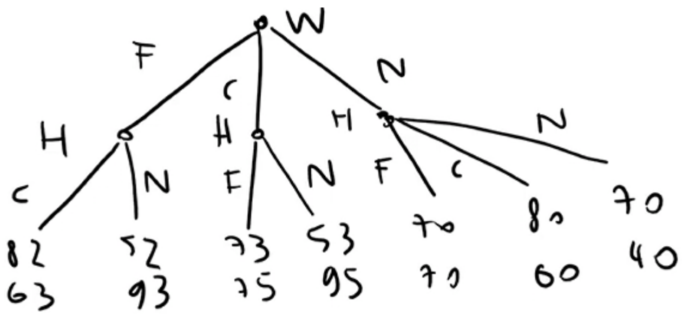

Figure (2.24)

   So, the final result will be W playing N in round 1, followed by H playing F in round 2 and finally W plays C in round 3.

   However, to *fully* describe the SPE we should actually write a more complicated strategy, specifying the action of *each* player at *each* information set. For W this amounts to $NNCNFCFF$, and for H to NNF.

   Note that the Pareto dominating outcome $(73, 75)$ is not reached, but the SPE is still a *good* $(70, 70)$. This happens because both players are egoistic, or, in this case, simply *lazy*.

Exercise 2.9.4:

Consider a repetition for 2 times with discount factor $\delta = 1$ of the stage game that is reported below in normal form:

$$\mathbf{P} = \begin{array}{c|ccc} & \multicolumn{3}{c}{\text{Player B}} \\ & M & F & H \\ \hline \text{Player A}\ M & 7,7 & -2,9 & -1,-1 \\ F & 9,-2 & 2,2 & 0,0 \\ H & -1,-1 & 0,0 & 6,6 \end{array}$$

1. What are the Nash equilibria of the stage game?

2. If any of these Nash equilibria are played at both stages, you obviously have a subgame-perfect equilibrium. Can you find a subgame-perfect equilibrium $\boldsymbol{s} = (s_A, s_B)$ of the repeated game which is not a repetition of the Nash equilibrium of the stage game?

3. If you consider an *arbitrary* discount factor δ, what are the conditions for $\boldsymbol{s}$ to still be a subgame-perfect equilibrium?

Solution.

1. This is a repeated game, for which we already have the stage's normal form. Note that it is a *symmetric* game, in the sense that swapping the players changes nothing in the matrix.

   Note that strategy M for both players is strictly dominated by F, and so we can remove it, reducing the game to a 2×2 matrix. Here we see that (F, F) and (H, H) are the two *pure* NEs, similarly to the Battle of Sexes game. There is also a *mixed* NE given by both players choosing $0.25F + 0.75H$.

2. Here we are seeking an SPE made of two strategies where *sometimes* neither (F, F) or (H, H) is played. This is possible because an SPE involves playing a NE in each *subgame*, not necessarily in each *stage* game.

   However, we know from theory that the *last round* is always played according to a NE. Moreover, since the game is *symmetrical*, it makes sense to search for a *symmetric* pair of strategies: $s_A = s_B$.

   So, we search for a possible *deviation* during the first stage. One idea would be (M, M), which is a Pareto efficient strategy. To *select* it, we can use a "carrot-and-stick" approach in the last round, as follows:

   - Play M in the first round, and then play H in the second if and only if the first outcome was (M, M). In other words, select the "better" NE, with payoff $(6, 6)$ higher for both players, if the opponent *complies* during the first round.

- Otherwise, if the opponent *deviates* from M, play F in the second round.

This results in playing (M, M) at the start, and (H, H) in the second, or (F, F) if any other deviation occurs. Players can be tempted to deviate in the first round, since (F, M) gives a bigger payoff $9 > 7$. However, this is compensated by the punishment, which is getting $2 < 6$ in the second round. Since $\delta = 1$, players regard future payoffs *the same* as immediate payoffs, so they *will* comply and play (M, M).

3. A lower δ *decreases* the incentive to comply in the first round. In particular, for $\delta = 0$, players won't play (M, M) at the start. In other words, a lower δ makes the *punishment* less effective.

   Quantitatively, let's compare the payoff of the *myopic* strategy involving a deviation in the first round, which is $9 + 2\delta$, with that of the collaborative strategy, which is $7 + 6\delta$. So, cooperating is better if:

$$7 + 6\delta > 9 + 2\delta \Rightarrow \delta > 0.5$$

   This is consistent with the theory that claims that the discount factor δ must be big enough for the punishment to be credible.

   As a final remark, note that if we had an *infinite repetition*, we could avoid the need for the "carrot" NE (H, H). The comparison is between the *myopic strategy* of always playing F (against a properly defined GrT), with payoff:

$$9 + 2\delta + 2\delta^2 + \cdots = 9 + \frac{2\delta}{1 - \delta}$$

   and the collaborative strategy of always playing M:

$$7 + 7\delta + 7\delta^2 + \cdots = \frac{7}{1 - \delta}$$

   Cooperation works if:

$$\frac{7}{1 - \delta} > 9 + \frac{2\delta}{1 - \delta} \Rightarrow \delta > \frac{2}{7}$$

Chapter 3

Bayesian Games

Introduction

(Lesson 18 of 27/11/2020) Compiled: January 1, 2021

In the previous chapters, we examined games in which all players share *common knowledge*. They know the utilities of all players, they know all the possible moves.
Equilibria are situations in which all players are forming **consistent** beliefs, i.e. they are all playing a *best response* to the action they *believe* their opponents are going to make. In other words, A *believes* that B will play b, and so A plays her best response a, and at the same time B *believes* that A will play a, and he will play his best response which is b. Both beliefs are *consistent*, they "confirm" each other, forming in a sense a "self-enforcing prophecy".
However, if A wants to *select* the right a, she needs to know what B wants, i.e. his utilities. In fact, she needs to know that B's best response to a is to play b. In this way she is certain that the loop is complete, and that her beliefs are **correct**.

In a more realistic situation, players *do not know* all opponents' utilities, and so they have to form *beliefs* over those. So, A won't know for sure that B's best response to a is exactly b. However, A can *guess* which are B's priorities, and play accordingly. This was first proposed by Harsanyi (1960).

Effectively, we are now dealing with a game of **incomplete information**. Beliefs over the characteristics of other players are captured by their **types**, i.e. how they *behave* depending on the circumstances.

Even in this situation it is possible to form *consistent* and *correct* beliefs, i.e. a form of **equilibrium**.

3.1 A first example

A player's type determines their utilities, but all players know only their type, and not that of the others. This can be modelled by drawing a type vector $(t_1, \ldots, t_n)$ at the start of the game, where t_i is drawn among all the possible types available to player i. Then each player j is shown their type t_j, and the game proceeds regularly[1].

The act of drawing a random vector can be regarded as a Nature's move, i.e. an *external source* of uncertainty, since players do not know the full details about it.

The game is both **dynamic** (Nature moves first), and of **imperfect information** (Nature's move is not known entirely), and it is denoted as a **Bayesian game**.

Types can be generalized to include also *states of knowledge* about the game. For example, a player may know *more* about the other players' types, and this information can still included in his/her type[2]. *Generalized types*

As a first example, we will use the following **Entry Game**: *Entry Game*

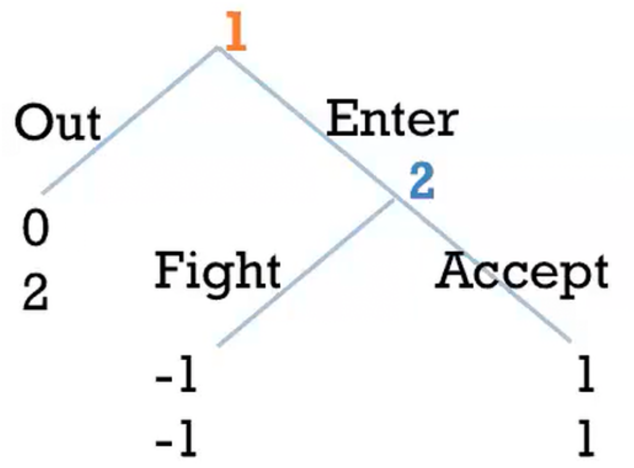

Figure (3.1)

Player 1 is a **newcomer**, who may either (E)nter or stay (O)ut. Player 2 is **incumbent**: if 1 enters, 2 may (A)ccept or (F)fight.
The SPE outcome is (E, A). Note that there is another NE, which is (O, F), but it is not an SPE, since 2's threat is non-credible. These are all the characteristics we need to show features of Bayesian games.

Let's make this game Bayesian. Suppose player 2 can be of two types:

- **Rational**: behaves as already discussed
- **Crazy**: enjoys fighting, and his/her payoff for (E, F) is 2 instead of -1.

Effectively 2 has always a *unique definite type*. The point is that 1 *does not know* which one is it!

[1]∧The same setup is used *concretely* in the Mafia/Werewolf game. Here *types* are the players' roles (e.g. *civilian*, *werewolf*, etc.), and each player knows initially only their own role.

[2]∧For example, *werewolfs* in the Werewolf game know which are the other werewolf players. This information is part of their *role* (type).

Suppose Nature decides that 2 is Crazy with probability p. Now 1 has a non-singleton information set, because he/she does not know the full outcome of Nature's decision:

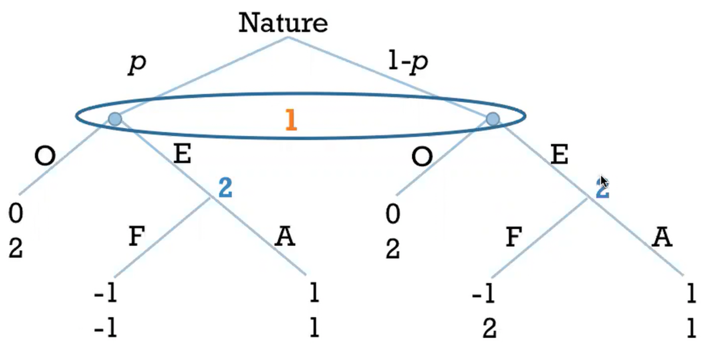

Figure (3.2)

If 2 is Crazy, the threat of playing (E, F) is now credible.

In the following, we suppose that the probability distribution for each player's types is **common knowledge** (**common prior assumption**). In other words, 1 does not know for sure if 2 is Crazy or not, however he/she knows that 2 is Crazy with probability p, and 2 knows that 1 knows, and so on.

Common prior assumption

Note that in this new extensive form, 2 has **four** pure strategies, of the form xy, where:

- x describes what a *Rational* player 2 does
- y describes what a *Crazy* player 2 does

The full list of 2's strategies is AA, AF, FA, FF. In other words, 2 "plans an action for each of their personalities". This is not important directly for player 2, who *knows* their type and so which node he acts in, but for understanding how player 1, who is uncertain, will move. In other words, 2 can "picture himself" as 1 sees him/her to *anticipate* 1's moves.

Philosophically, we can distinguish between the *actual* player 2, which has a definite type and knows it, and the *abstract type*-player 2, which is needed to model the opponents' beliefs.

Note that in incomplete information games, the number of strategies *explodes even more rapidly*, since we need to state what each *type* of player does.

After all strategies are listed, we can rewrite the game in normal form, computing the *expected utilities* for each move. For example, consider the joint strategy (E, AF). In this case, the expected payoffs are:

$$u_1(E, AF) = p \cdot 1 + (1-p) \cdot (-1) = 2p - 1$$
$$u_2(E, AF) = p \cdot 1 + (1-p) \cdot 2 = 2 - p$$

Computing all the combinations leads to:

$$\mathbf{P} = \begin{array}{cc} & \text{Player 2} \\ & \begin{array}{cccc} AA & AF & FA & FF \end{array} \\ \text{Player 1} \begin{array}{c} O \\ E \end{array} & \left\| \begin{array}{cccc} 0,2 & 0,2 & 0,2 & 0,2 \\ 1,1 & 2p-1, 2-p & 1-2p, 1-2p & -1, 2-3p \end{array} \right\| \end{array} \quad (3.1)$$

We can now find the NE using the standard techniques. However, note that, in general, the result will depend on the value of p. In other words, depending on the common prior, the NEs may be different.

3.2 Formalization

A Bayesian game consists of the following:

- A set of players $\mathcal{N} = 1, \ldots, n$.
- A strategy space S_i for each player $i = 1, \ldots, n$.
- A type space T_i for each player $i = 1, \ldots, n$.
- Type-dependent utilities of players $u_i \colon (S_1, S_2, \ldots, S_n; T_1, \ldots, T_n) \to \mathbb{R}$

A **static** Bayesian game is one in which all *real* players (i.e. all except Nature) move simultaneously. Effectively, since Nature still moves first, this is still a dynamic game.
However, in this simple case, each player's strategy is a *single action* $a_i \in A_i$.

1. The type $t_i \in T_i$ of each player i is chosen by Nature for all $i = 1, \ldots, n$, according to the **joint prior** probability distribution $\Phi(t_1, \ldots, t_n)$ which is assumed **common knowledge**. This means that the game is of **perfect** information, because all beliefs are correct, but **incomplete**, because there is uncertainty.

2. We assume that the utility u_i of player i depends **only** on player i's type, i.e. $u_i = u_i(a_1, a_2, \ldots, a_n; t_i)$ (**private values assumption**). In a more general case (**common values**), u_i is a function of *all the types*: $u_i = u_i(a_1, \ldots, a_n; t_1, \ldots, t_n)$.

   Different types have different utilities: the utility of player i of type j is $u_{i,j}(a_i, \boldsymbol{a_{-i}}) \equiv u_i(a_i, \boldsymbol{a_{-i}}; t_j)$

   Types can also *limit* the available actions. This can be done by setting the payoff for an unwanted move to a very low value ($-\infty$).

3. Since each player *know* his type, they can *infer* something about the type of others from the joint distribution $\Phi(t_1, \ldots, t_n)$, which is common knowledge. Specifically, they can compute the conditional probability:

$$\Phi(\boldsymbol{t_{-i}} | t_i) = \frac{\Phi(t_1, \ldots, t_n)}{\Phi(t_i)}$$

This means that each player can know *something more*[3], assuming that $\mathbf{\Phi}$ is not separable: if all types are completely independent, then nobody knows anything more. In fact, in this case $\mathbf{\Phi}(\boldsymbol{t}) = \prod_i \varphi_i(t_i)$, and so:

$$\begin{aligned}\Phi(\boldsymbol{t_{-i}}|t_i) &= \frac{\varphi_1(t_1)\cdots\varphi_n(t_n)}{\varphi(t_i)} = \\ &= \varphi_1(t_1)\cdots\varphi_{i-1}(t_{i-1})\varphi_{i+1}(t_{i+1})\cdots\varphi_n(t_n) \equiv \mathbf{\Phi}(\boldsymbol{t_{-i}})\end{aligned}$$

This knowledge is denoted as the **belief** of each player regarding other players' types.

Putting everything together, we denote a static Bayesian game as:

$$G = \{\mathcal{N}; A_1, \ldots, A_n; T_1, \ldots, T_n; \mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_n; u_1, \ldots, u_n\}$$

where $u_i = u_i(a_1, \ldots, a_n; t_i)$.
A **pure strategy** for i is a map $s_i \colon T_i \to A_i$, specifying the actions to play for each *type* of player. This comes directly from the definition of strategies for a dynamic game: a mapping from *each* information set to an *action*. In case of a *static* Bayesian game, there is one information set for each of i's types.
So, to fully specify a strategy for i, we need to know an action for each of i's possible types, even if i *knows* his/her type! This is not important directly for i, but for understanding the beliefs that *other players* $-i$ can form about i, and plan their reactions. Such an "over-specification" is similar to that we made in generic dynamic games, were strategies included also responses to actions that are never taken (i.e. nodes *outside* the equilibrium path). Basically, it is needed to allow beliefs of different players to be *consistent*, which is a necessary condition for finding Nash Equilibria.

Then, a **mixed strategy** for i is a probability distribution over their pure strategies.

Finally, all of this can be generalized to the **dynamic games**, in which pure strategies are "plans of actions" provided *for each* type.

3.3 Bayesian Nash Equilibria

(Lesson 19 of 01/12/2020) Compiled: January 1, 2021

A **Bayesian Nash Equilibrium** is just a Nash Equilibrium for Bayesian strategies. That is, given a (**static**) Bayesian Game:

$$G = \{\mathcal{N}; A_1, \ldots, A_n; T_1, \ldots, T_n; \mathbf{\Phi}_1, \ldots, \mathbf{\Phi}_n; u_1, \ldots, u_n\}$$

a joint strategy $\boldsymbol{s}^* = (s_1^*, \ldots, s_n^*)$ is said to be a Bayesian Nash Equilibrium if, for each player i and each type $t_i \in T_i$, $s_i(t_i)$ maximizes the expected payoff:

$$\max_{s_i \in S_i} \sum_{\boldsymbol{t_{-i}}} u_i(s_1^*(t_1), \ldots, s_{i-1}^*(t_{i-1}), s_i, s_{i+1}^*(t_{i+1}), \ldots, s_n^*(t_n), t_i)\mathbf{\Phi}_i(\boldsymbol{t_{-i}}|t_i)$$

[3] ∧Consider the Werewolf game. The number of werewolves in the game is common knowledge (this is, in fact, part of the prior distribution of types). So, if you know that there is only a single werewolf, and you pick that card, you immediately know that all others are *not* werewolf. That is, knowing your type allows you to know *more* about all the other types, due to the correlations in the prior distribution.

Note that i's *beliefs* $\boldsymbol{\Phi}_i(\boldsymbol{t_{-i}}|t_i)$ about the opponents' types $\boldsymbol{t_{-i}}$ form the *weights* for the utilities (while i's type is clearly *known*).

In other words, at a Nash Equilibrium all types of all players do not have regrets, i.e. no incentive for a unilateral deviation, because there is nothing they can do by themselves to *improve* their expected utility:

$$\mathbb{E}[u_i(s_i^*(t_i), \boldsymbol{s^*_{-i}}(\boldsymbol{t_{-i}}), t_i)|t_i] \geq \mathbb{E}[u_i(s_i, \boldsymbol{s^*_{-i}}(\boldsymbol{t_{-i}}), t_i)|t_i] \qquad \forall s_i \in S_i$$

For *dynamic* Bayesian games we need this still holds, but in that case we are more interested to an extension of Subgame Perfect Equilibria (SPEs).

3.3.1 Examples

Chicken Game

Consider the **Chicken Game**: two players drive towards each other along a narrow road. Each of them can (C)hicken (i.e. *steer*) or (D)rive toward the other. The payoffs are as follows:

- Chickens always get nothing ($u = 0$)
- Drivers gain *respect*. If only one drives, and the other chickens, then the driver will get $u = 8$. Otherwise, both players share the *respect bounty*, receiving 4 each. However, in this case an incident happens, and destroy their cars. Thus, they incur in a punishment P depending on their parents:
  - If parents are (H)ard, $P = 16$
  - If they are (L)enient, $P = 4$.

  Each player *knows* the type of their parents, and that the opponent's parents can be either H or L with equal probabilities ($p = 0.5$), and they are all *independent*. This is **common knowledge**.

Note that, in this case, the *type* of players does not refer directly to the players themselves, but on their parents. In fact, in general a type is any information that *determines* the payoffs.

The game's extensive form is as follows:

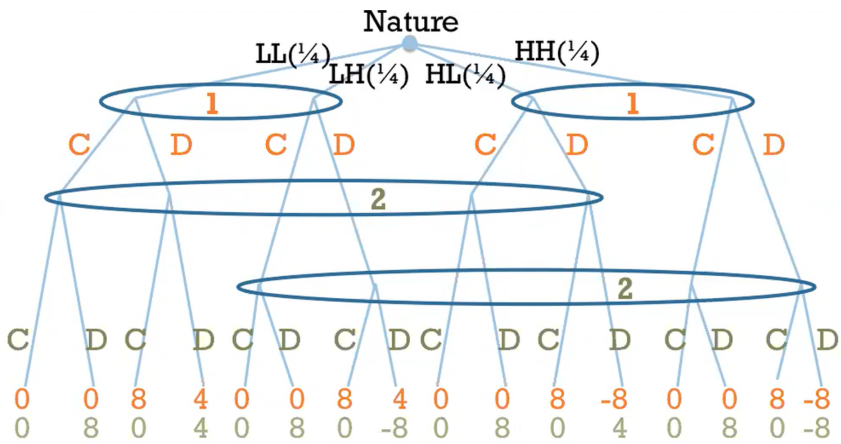

Figure (3.3) – Extensive form for the Chicken Game. The information sets reflect the fact that each player *knows* their type. LL and LH both mean that 1's parents are Lenient, and 1 is aware of that, and will behave the same in both nodes, since 1 does not know the type of 2. Similarly, LH and HH both mean that 2's parents are Hard, and so they form a unique information set too.

Each player has 2 types, and so strategies are *pairs* of actions. Since actions are binary, each player has a total of 4 pure strategies. We can compute all combinations of expected utilities and organize them in a table, putting the game in normal form:

$$\mathbf{P} = \begin{array}{c|cccc} & \multicolumn{4}{c}{\text{Player 2}} \\ \text{Player 1} & CC & CD & DC & DD \\ \hline CC & 0,0 & 0,4 & 0,4 & 0,8 \\ CD & 4,0 & -1,-1 & -1,2 & -6,1 \\ DC & 4,0 & 2,-1 & 2,2 & 1,1 \\ DD & 8,0 & 1,-6 & 1,1 & -6,-6 \end{array}$$

(Note that the game is *symmetrical*, which reduces the needed computations).

For example, consider the entry (CD, DC). We have 4 possibilities to consider:

- LL, resulting in the outcome (C, D), with payoff $(0, 8)$.
- LH, leading to (C, C), with payoff $(0, 0)$.
- HL, leading to (D, D), with payoff $(-12, 0)$
- HH, leading to (D, C), with payoff $(8, 0)$

Each possibility happens with $p = 1/4$. We can then compute the expected utilities:

$$u_1(CD, DC) = \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot (-12) + \frac{1}{4} \cdot 8 = -1$$
$$u_2(CD, DC) = \frac{1}{4} \cdot 8 + \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 0 = 2$$

Inspecting the best responses, we see that (DC, DC) is the Bayesian Nash Equilibrium (BNE).

Committee Voting

Consider a jury with just two jurors deciding whether to (A)cquit (declare *innocent*) or (C)onvict (declare *guilty*) a defendant. Every juror casts *independently* a sealed vote, and the defendant is convicted if both jurors vote C. However, it is uncertain whether the defendant is (G)uilty or (I)nnocent: the prior probability of G is $q > 1/2$, and this is common knowledge. Jurors want to make the right decision: their payoff is 1 when they make the right decision (C if G, or A if I), otherwise they get nothing.

The game's normal form is:

$$\mathbf{P} = \begin{array}{cc} & \text{juror 2} \\ & \begin{array}{cc} A & C \end{array} \\ \text{juror 1} \begin{array}{c} A \\ C \end{array} & \left\| \begin{array}{cc} 1-q, 1-q & 1-q, 1-q \\ 1-q, 1-q & q, q \end{array} \right\| \end{array} \tag{3.2}$$

Since $q > 1/2$, each juror prefers to play C, and the NE is indeed (C, C).

Now, let's make this a Bayesian game by introducing some **signaling** between players. Assume each player observes the evidence and independently gets a private *signal*, representing his/her idea about the case, denoted by $t_i \in \{t_G, t_I\}$. It is more likely to receive a signal t_x if the defendant status is x, i.e. the players' types and that of the defendant are *correlated*. Specifically:

$$\mathbb{P}[t_G|G] = \mathbb{P}[t_I|I] = p > \frac{1}{2} \qquad i = 1, 2$$

Note that there is a non-zero probability of "getting the wrong signal":

$$\mathbb{P}[t_G|I] = \mathbb{P}[t_I|G] = 1 - p < \frac{1}{2}$$

Note that the types are here representing something about the *knowledge* each player has, which affects his/her payoffs.

Each player has 2 types and 2 actions, so a total of 4 pure strategies: AA, AC, CA and CC. Note that both players have the same objective, and so this is a *coordination* game.

Let's consider a single juror for now. Without any signal, they know that the defendant is more likely to be guilty, and so they would choose C.

However, the signal is an additional source of information. Since $\mathbb{P}[G] = q$, we have:

$$\mathbb{P}[G|t_G] = \frac{\mathbb{P}[G \wedge t_G]}{\mathbb{P}[t_G]} = \frac{\overbrace{\mathbb{P}[t_G|G]}^{p}\,\overbrace{\mathbb{P}[G]}^{q}}{\mathbb{P}[t_G|G]\mathbb{P}[G] + \mathbb{P}[t_G|\bar{G}]\mathbb{P}[\bar{G}]} = \frac{pq}{pq + (1-p)(1-q)}$$

Since $p > 1/2$, $1 - p < p$, and so the denominator is $qp + (1-q)(1-p) < qp + (1-q)p$, meaning that:

$$\frac{qp}{qp + (1-q)(1-p)} > \frac{q\cancel{p}}{q\cancel{p} + (1-q)\cancel{p}} = q$$

So, if the player receives t_G, $\mathbb{P}[G|t_G] > q$, i.e. he/she is even surer that the defendant is guilty.

In case t_I is received, the opposite happens:

$$\mathbb{P}[G|t_i] = \frac{\mathbb{P}[G \wedge t_I]}{\mathbb{P}[t_I]} = \frac{q(1-p)}{q(1-p)+(1-q)p} < q$$

and in this case the player is more doubtful. Exactly *how much* things are changed depends on the actual values of q and p, which are supposed to be known. If $p = 1$, then $\mathbb{P}[G|t_i] = 1$, i.e. we know for sure that the defendant is guilty. If $p = 1/2$, no new information is obtained. If $p > q$, $\mathbb{P}[G|t_I] < 0.5$, i.e. the new information can completely *change* the result, because it is more reliable than the prior information.

Knowing this, we can return to the full game. If $p > q$, we expect to find a BNE given by (CA, CA), i.e. both players "following what the signal says". To see if this is indeed correct, we first need to compute the probability of each type pair:

$$\begin{array}{cc|c}
 & & \text{juror 1} \\
 & & \begin{array}{cc} \quad t_G \quad & \quad t_I \quad \end{array} \\
\text{juror 1} & \begin{array}{c} t_G \\ t_I \end{array} & \left\|\begin{array}{cc} qp^2+(1-q)(1-p)^2 & p(1-p) \\ p(1-p) & q(1-p)^2+(1-q)p^2 \end{array}\right\|
\end{array} \tag{3.3}$$

Now, let's see if CA is a best response to itself. According to the rules of the jury, a player is decisive (i.e. *pivotal*) only if the other juror chooses C. Players *want* to be pivotal, because this would let them to know the result of their actions in advance.

Suppose that 2 plays CA, and 1 *knows* that 2 has received t_I, meaning that 2 will play A. Now 1 is not pivotal, and so he/she can play anything without changing the outcome. In other words, all *actions* are best responses if the other player chooses A, and in particular CA is a best response.

So, we need to check what happens if 2 receives instead t_G. Recall that 1 *knows* his/her type, and can add also this information to compute the posterior probabilities:

$$\mathbb{P}[G|t_1 = t_G, t_2 = t_G] = \frac{qp^2}{qp^2+(1-q)(1-p)^2} > q$$
$$\mathbb{P}[G|t_1 = t_I, t_2 = t_G] = \frac{qp(1-p)}{p(1-p)} = q$$

In both cases, it is more convenient to vote for C. So 1's best response in this case would be CC. So, CA is **not** the best response to CA.

In fact, it can be shown that the real BNE is (CC, CC), and this is so even if $p > q$, i.e. when the signal is *informative*! In other words, the initial prior information *determines* the outcome.

Note that this happens in the case where *both* players act independently. If we considered them separately, they behave as expected, following informative signals. However, when they are considered as part of the game, their *biases* reinforce each other.

3.3.2 Beliefs: mixed strategies vs types

***Mixed strategies** as beliefs*

Previously, we first discussed **beliefs** as a compelling interpretation for mixed strategies. Briefly, a player is never *uncertain* about their move, but their opponents may be. So, mixed strategies express the *beliefs* an opponent has about a player's move.

For a concrete example, consider again the Odd/Even game. Suppose A *always* plays Odd, and B *knows* this. Then B will always win! Clearly, this is not a good way to play for A. Instead, she should *act* so that B *can't know* in advance her move, for example by playing "erratically" (i.e. flipping a coin[4]). Note that A *always knows* what she is going to play (she sees the coin). However, B will react according to his *beliefs* about A's move. So, in a sense, A's strategy is tied to these beliefs, and not the actual action she takes. Then, *basically*, A playing a mixed strategy m means that B's *beliefs* about A's actions are m. In the above example, $m = 0.5\,\mathrm{Odd} + 0.5\,\mathrm{Even}$ means that B *believes* that A's action will be either *Odd* or *Even* with equal probabilities.

Note that mixed strategies are not directly important for A: they are necessary to model how B, i.e. the opponent, will *react* to A's move.
This is a prerequisite for achieving the Nash Equilibrium, i.e. a situation where both players' beliefs are **correct** and **consistent**. For example, A believes B will play $\bar{b}$, and so she plays her best response $a(\bar{b})$. At the same time, B believes A will play $\bar{a}$, and so he plays his best response $b(\bar{a})$. They are correct and consistent if $\bar{b} = b(\bar{a})$ and $\bar{a} = a(\bar{b})$. *Correct*, because each player behaves as expected: A believes B will play $\bar{b}$, and he indeed plays $\bar{b}$. *Consistent*, because their beliefs are "self-fulfilling". When A plays her best response $a(\bar{b})$, she chooses $\bar{a}$, which is exactly what B expects her to do! In other words, the players making their *optimal move* according to their beliefs about the opponents *confirms* the opponents' beliefs. Consistency makes all beliefs *sustain* each other, as *forks balancing on a toothpick*, so that the entire structure is held together.

Now, if A wants to *solve* $\bar{a} = a(\bar{b})$, she can use $\bar{b} = b(a)$, but for that she needs to know $b(a)$, i.e. how does B react to her strategies! In other words, A must have a "model of B" in her head, which clearly needs to be accurate. Now, consider that B reacts to A according to his *beliefs* about A, which are nothing more than the *mixed strategies* played by A.
In summary, a normal player thinks in terms of his moves. A *really good* player thinks about what the opponent will *think* about his/her moves. A strategy is not just moving a piece on the board, but it is changing the *state of mind* of your opponent. Pure strategies are in a 1-1 relation with *physical states*. Mixed strategies are in a 1-1 relation with *mind states*.

***Player types** as beliefs*

Now, all this discussion should feel *similar* to that of types. In fact, we have noted before how the spectrum of a player's types is not important for that player themselves (because each player knows their type), but to model *other*

[4]∧Note that no real flipping is necessary. B just needs to *believe* that the choice will be random, but maybe A is just following some unknown algorithm in her head to decide what to do.

players' beliefs about that player. In other words, types are just another form of beliefs, i.e. mixed strategies, as first highlighted by Harsanyi when formalizing Bayesian games.

***Pure BNEs** of incomplete info are **Mixed NEs** of complete info*

So, we expect that a Bayesian Game can be interpreted as a classic static game with mixed strategies. Let's show this explicitly with an example.

Consider the static Battle of Sexes:

$$\mathbf{P} = \begin{array}{cc} & \text{Brian} \\ & \begin{array}{cc} R & S \end{array} \\ \text{Ann} \begin{array}{c} R \\ S \end{array} & \left\| \begin{array}{cc} 2,1 & 0,0 \\ 0,0 & 1,2 \end{array} \right\| \end{array} \tag{3.4}$$

A mixed strategy for either player is completely specified by the probability to play R (for example). We already know that there are 3 NEs in this game, 2 in pure strategies $(0,0)$ and $(1,1)$, and 1 in mixed strategies $(2/3, 1/3)$. We will now show that this last one can be seen as a **pure** BNE of a related game with a bit of incomplete information.

The idea is to make the game Bayesian by introducing *uncertainty* on the player's payoffs:

$$\mathbf{P} = \begin{array}{cc} & \text{Brian} \\ & \begin{array}{cc} R & S \end{array} \\ \text{Ann} \begin{array}{c} R \\ S \end{array} & \left\| \begin{array}{cc} 2+c,1 & 0,0 \\ 0,0 & 1,2+d \end{array} \right\| \end{array} \tag{3.5}$$

Each player *knows exactly* their payoffs, but not the ones of the opponent: Ann and Brian do not know each other very well. So, Ann knows c but not d, and Brian knows d but not c. This effectively defines privately-known *types* for the players. Suppose that $c, d \in [0, x]$ uniformly, where x can be thought as a "perturbation".

A simple strategy s_A for Ann is to choose R if c is sufficiently big, i.e. over a certain threshold C, and otherwise play S. Similarly, Brian could choose S if $d > D$, and otherwise play R (s_B).

This joint strategy is, in fact, a Bayesian Nash Equilibrium (BNE). To prove it, we simply compute the expected payoffs, and see that Ann's strategy is the best response to Brian's strategy (and the vice versa holds for symmetry reasons). Note that $\mathbb{P}[d \leq D] = D/x$, and so:

$$\begin{aligned} u_A(R, s_B) &= \frac{D}{x}(2+c) + \left(1 - \frac{D}{x}\right) \cdot 0 = (2+c)\frac{D}{x} \\ u_A(S, s_B) &= \frac{D}{x} \cdot 0 + \left(1 - \frac{D}{x}\right) \cdot 1 = 1 - \frac{D}{x} \end{aligned}$$

So Ann's best response to s_B is to play R if $c \geq x/D - 3$, and S otherwise, i.e. to choose strategy s_A with $C = x/D - 3$.

We can repeat the same reasoning for Brian:

$$u_B(s_A, R) = \left(1 - \frac{C}{x}\right) \cdot 1 + \frac{C}{x}(2+d) \cdot 0 = 1 - \frac{C}{x}$$
$$u_B(s_A, S) = \left(1 - \frac{C}{x}\right) \cdot 0 + \frac{C}{x}(2+d) = (2+d)\frac{C}{x}$$

And so Brian plays S if $d \geq x/C - 3$, meaning that $D = x/C - 3$.

Combining these two conditions leads to:

$$\begin{cases} \frac{x}{D} - 3 = C \\ \frac{x}{C} - 3 = D \end{cases} \Rightarrow C = D,\ C^2 + 3C - x = 0 \Rightarrow C = \frac{-3 + \sqrt{9+4x}}{2}$$

The probability p_R of Ann playing R is then:

$$p_R = 1 - \frac{C}{x} = \frac{2x + 3 - \sqrt{9+4x}}{2x} \xrightarrow[x \to 0]{} \frac{2}{3}$$

which is exactly the same as Ann playing the mixed NE from the static game. In other words, when the "noise" x is eliminated, the BNE (incomplete information) reduces to the mixed NE (complete information).

3.4 Dynamic Bayesian Games

(Lesson 20 of 04/12/2020) Compiled: January 1, 2021

Nash Equilibria in Dynamic Games often involve *non-credible* threats: they are not "rational", in the sense that rational players won't naturally choose them. To get a "good" solution we need to proceed by *backward induction*, leading to the concept of a Subgame Perfect Equilibrium (SPE), which is a *special kind* of NE that is still a NE in every *subgame*.

Bayesian Games are dynamic games in which the first turn is played by Nature, which selects all players' types. If all that's left is a second turn in which all players act simultaneously, then the game is denoted as *static*, and we can use a basic extension of "plain" Nash Equilibria to *solve* them.

However, what would happen if, after removing the initial Nature's choice, the game is still *dynamic*? Can we just apply the SPE reasoning here?

Unfortunately, this is not so simple. As we will see in the next examples, the SPE does not guarantee to find a "rational" solution, because of a "clash" between incomplete information (Nature's move) and the dynamic structure of the rest of the game.

Let's start with a Bayesian game in which the SPE behaves as expected. Consider the **Entry Game**, with the following extensive form:

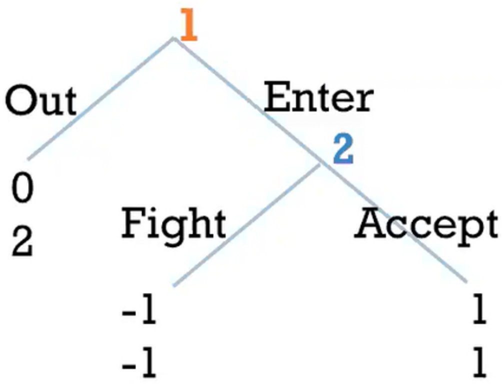

Figure (3.4) – Extensive form for the Entry Game (dynamic, non-Bayesian)

Backward induction leads to (E, A), which is the SPE. Note that (O, F) is a NE, but not a SPE.

We can make it Bayesian by considering two types for Player 2: Crazy (prefer to fight) or Normal (same as before).

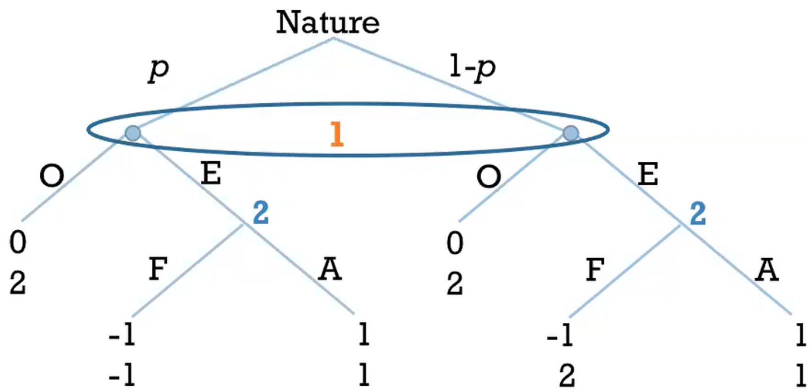

Figure (3.5) – Extensive form for the Bayesian Entry Game (v.1)

There are 4 pure strategies for player 2: AA, AF, FA and FF. So, for $p = 2/3$, the game's normal form is given by:

$$\mathbf{P} = \begin{array}{c c} & \text{Player 2} \\ & \begin{array}{cccc} AA & AF & FA & FF \end{array} \\ \text{Player 1} \begin{array}{c} O \\ E \end{array} & \left\| \begin{array}{cccc} 0, 2 & 0, 2 & 0, 2 & 0, 2 \\ 1, 1 & 1/3, 4/3 & -1/3, -1/3 & -1, 0 \end{array} \right\| \end{array}$$

There are 3 "plain" NEs: (O, FA), (O, FF) and (E, AF). However, note that it is *irrational* for 2 to play FA or FF, as this would involve 2's normal type to choose *fight* over *accept*, which is sustainable only if 1 does not enter the game. So, basically, 2 is threatening 1 to scare him/her off. However, the threat is non-credible, as it is inconvenient even for 2. A *rational* player 2 would always prefer A over F for the normal type.

So, the only *rational* strategy for 2 is to *always* play AF, as confirmed by backward induction. Then, 1 will play O or E depending on the prior probability

p of 2 being *normal*. If p is sufficiently high, as seen in the above, then it is worth entering the game (E). Otherwise, if p is low, there is a high likelihood of 2 being *crazy*, and 1 will prefer to stay out (O).

For $p = 2/3$, (E, AF) is the unique SPE, corresponding to the *rational solution* as expected. So, at least in this case, a direct application of the SPE definition works.

However, this breaks down if we consider a variant of the entry game. Suppose that player 1 has two types: (C)ompetitive, behaving as above, or (W)eak, who prefers not to enter. The probability of 1 being C is denoted by p, and in the following we take $p = 1/2$ as an example.
The game's extensive form is as follows:

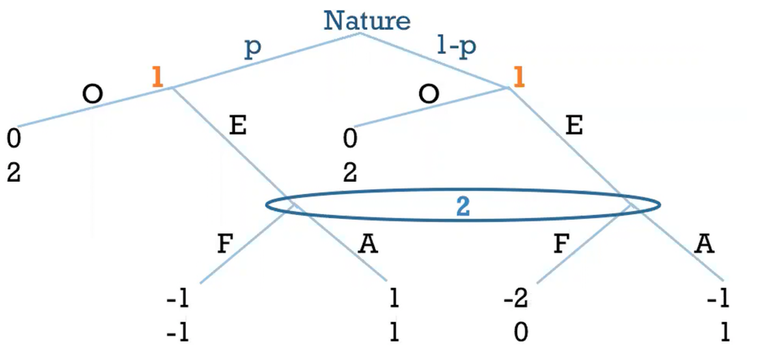

Figure (3.6) – Extensive form for the Bayesian Entry Game (v.2)

Player 1 is aware of his/her type, and can use this information to decide what to do. For example, if 1 is *weak* then it is convenient to stay out (and getting 0) rather than entering (and getting either -1 or -2).
However, 2 does not know this, and must choose a unique strategy for both types, since he/she has only one information set encompassing two indistinguishable nodes. In particular, since these are the nodes in the deepest layer, we *cannot* apply backward induction.

So, 1 has 4 pure strategies: OO, OE, EO, EE. On the other hand, 2 can either (A)ccept or (F)ight.
To write the game's normal form we need to compute the expected utilities for *all* possible joint strategies. For example, for (OE, A) we have:

$$u_1(OE, A) = p \cdot 0 + (1-p) \cdot (-1) = p - 1 = -\frac{1}{2}$$
$$u_2(OE, A) = p \cdot 2 + (1-p) \cdot 1 = 2p + 1 - p = 1 + p = \frac{3}{2}$$

Then we gather all results in a table:

$$\mathbf{P} = \begin{array}{c} \\ \text{player 1} \end{array} \begin{array}{c} \\ OO \\ OE \\ EO \\ EE \end{array} \begin{array}{c} \text{player 2} \\ \begin{array}{cc} F & A \end{array} \\ \left\| \begin{array}{cc} 0, 2 & 0, 2 \\ -1, 1 & -1/2, 3/2 \\ -1/2, 1/2 & 1/2, 3/2 \\ -1/2, -1/2 & 0, 1 \end{array} \right\| \end{array} \tag{3.6}$$

There are two NEs in pure strategies:

- (OO, F): the incumbent (2) threatens to fight
- (EO, A): only a competitive outsider (1) enters the game, and the other always accepts.

Note that OE and EE are strictly dominated strategies: as expected, a *weak* player 1 never decides to enter.

Between the two NEs, (OO, F) includes a non-credible threat, and so it does not correspond to our notion of rationality. However, it is an SPE, since the game has a single subgame, which is exactly the whole game itself.
So, the notion of SPE does not allow distinguishing between the *rational* (EO, A) and the *non-credible* (OO, F).

The main issue is that, ideally, rational players must play optimally both *on* and *off* the equilibrium path. For (OO, F), the equilibrium path covers just the "first layer" of the extensive form. However, the part *off* the equilibrium path is *merged* by the uncertainty about player 1's type, and we need a way to understand what the optimal play should be here. Intuitively, we know that 2 should always accept, since this leads to 1 always, while fighting leads to either 0 or -1. However, how can we formalize this in a unique structure?

3.5 Perfect Bayesian Equilibrium

In dynamic games, we defined the Subgame Perfect Equilibrium (SPE) as the special kind of Nash Equilibrium which is not only rational *on* the equilibrium path, but also *off* the equilibrium path.

The idea is to do the same for Bayesian games. The issue is that, as seen from the previous example, we cannot simply use *subgames* for the *off* equilibrium part.

First, given a Bayesian NE $\boldsymbol{s}^*$, an information set is said to be **on the equilibrium path** if, given the distribution of types, it is reached with probability > 0. For example, in (OO, F), the information set of node 2 is **not** in the equilibrium path.

Bayesian Equilibrium Path

Now, we need a way to somehow *split* information sets spanning multiple nodes, so that we can check in each of the resulting parts if the BNE is rational.

The idea is that, in the Bayesian setup, not all nodes in an information set are the same. Depending on the *prior distribution*, some of them may be more *probable* than others. For example, if p of 1 being *competitive* is high, then it is more likely to be on the *left* part of the graph, and not the other one.

This is formalized by introducing a **system of belief** μ, which is a probability distribution over decision nodes for every information set, that is a set of conditional probabilities of the form $\mathbb{P}[\text{being at node } x \in \mathcal{I} | \text{being at information set } \mathcal{I}]$.

System of Belief

In a game of *perfect* information, all information sets are singletons, and so the system of beliefs are all 1s: a player which is at information set $\mathcal{I}$ *knows* *for sure* that he/she is at the unique decision node inside $\mathcal{I}$. In an incomplete information game, instead, there are several possibilities, and to compute the conditional probabilities we can use Bayes' theorem:

$$\mathbb{P}[\text{node}|\text{info set}] = \frac{\mathbb{P}[\text{node}]}{\mathbb{P}[\text{info set}]}$$

This *additional information* is the key ingredient for finding the *rational solution* of a Bayesian game.

We are now ready to introduce the analogue of SPEs for Bayesian games, which is called the **Perfect Bayesian Equilibrium** (PBE), or **sequential rationality**. The idea is to *use* the systems of belief to make rational moves. Formally, a PBE is a pair $(\boldsymbol{s}^*, \mu)$ of a BNE $\boldsymbol{s}^*$ and its system of beliefs μ, meeting the following requirements: *Definition*

- A system of belief must be defined for **all** information sets of every player. In particular:
  - **On** the equilibrium path, the belief systems must follow Bayes' rule on conditional probability.
  - **Off** the equilibrium path, the belief systems are *arbitrary*. They still have to respect the normalization constraint.
- Players are **sequentially rational**: given the beliefs, they always play a **best response**.

Let's return to the previous example:

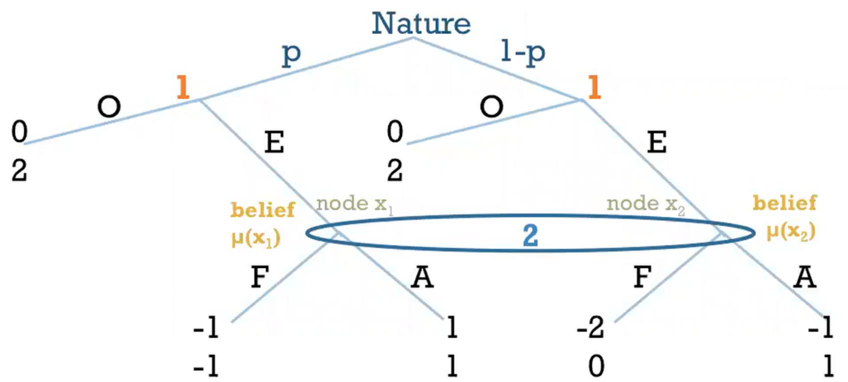

Figure (3.7)

To fully specify a BPE, we need to assign beliefs $\mu(x_1)$ and $\mu(x_2)$ to the nodes x_1 and x_2 (all the others are *trivial*, since any other information set is a singleton). Note that they must sum to 1, and so they cannot be both 0: these are *conditional probabilities*, not *absolute probabilities*! The probability of reaching x_1 or x_2 may be 0, but *if* we assume reaching it, then the probability of being in either x_1 or x_2 is 1.

If the probability of reaching $\mathcal{I} = \{x_1, x_2\}$ is non-zero, then we need to apply Bayes' rule to compute $\mu(x_1)$ and $\mu(x_2)$, starting from the **strategy** being considered.

For example, consider the BNE (EO, A). Now, only the *competitive* 1 enters the game, so if we reach the information set $\{x_1, x_2\}$ we are sure to be at x_1, meaning that $\mu(x_1) = 1$, $\mu(x_2) = 0$.
Note that this reasoning can be extended to mixed strategies. For example, suppose 1 chooses E with probabilities q_C and q_W (depending on his/her type). The probability of reaching x_1 is pq_C, that of reaching x_2 is $(1-p)q_W$, and so the probability of entering $\mathcal{I}$ is the sum $pq_C + (1-p)q_W$. Then, according to Bayes' rule:

$$\mu(x_1) = \frac{pq_C}{pq_C + (1-p)q_W}$$

Returning to the BNE (EO, A), we can see that all players behave rationally. In fact, 2 is sure to be at x_1 ($\mu(x_1) = 1$), and so he/she chooses A, which is the best response. Player 1 can anticipate this, and so he/she chooses E.
Thus, (EO, A), along with the system of belief with $\mu(x_1) = 1$, is a PBE.

On the other hand, consider (OO, F). Now $\mathcal{I} = \{x_1, x_2\}$ is outside the equilibrium path, and so we cannot use Bayes' rule to compute $\mu(x_1)$ or $\mu(x_2)$, since the probability of reaching $\mathcal{I}$ is 0.
However, we still need to assign *some arbitrary* beliefs, which cannot be both 0, since $\mu(x_1) + \mu(x_2) = 1$. However, at both nodes the *best response* is to play A. So, if any of the two has a non-zero belief (as it *must* be), the rational choice would be to play A. But the BNE specifies to play F, which is different, and so this is **not** a PBE.

3.6 Further discussion

More recent literature suggests that the PBE definition could be extended even more. In particular, the notion of PBE introduced above is denoted as "weak", because beliefs *off* the equilibrium path can be any *arbitrary* probability distribution. Additional constraints, such as continuity, may be introduced. However, a full discussion of these proposals is beyond the scope of this course. In this section, we will limit ourselves to discussing an example where the PBE definition seems "odd".

Consider the following entry game:

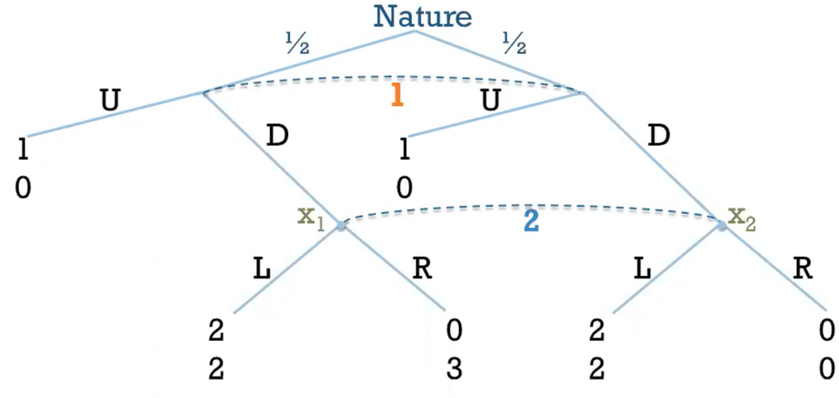

Figure (3.8)

Here Nature draws types for both players, and neither one knows anything.
If 1 plays D with some non-zero probability, then 2 will choose equal beliefs for x_1 and x_2, i.e. $\mu(x_1) = \mu(x_2) = 0.5$, since he/she has no more information than that of the prior distribution. Then, the strategies at x_1 and at x_2 must be the same, since both nodes belong to the same information set. The expected payoff for R is $3 \cdot .5 + 0 \cdot .5 = 1.5$, while that for L is $2 \cdot .5 + 2 \cdot .5 = 2$, so 2 will play L. 1 anticipates that, and plays D always, leading to the PBE (D, L).

However, suppose that 1 *never* plays D. Then x_1 and x_2 are suddenly *off* the equilibrium path, meaning that 2 can decide arbitrary beliefs for them. In particular $\mu(x_1) > 2/3$ is possible, which means that 2's best response is R, and so 1 should always play U, leading to another PBE (U, R).

Between the two, the second PBE seems a bit *forced*. The threat of choosing $\mu(x_1) > 2/3$ seems *non-credible*. This suggests that the PBE is still not "specific enough" to find the "rational solution" of Bayesian games.

3.7 Signaling Games

Bayesian games can be classified in two general categories:

- Games like the one in fig. 3.5, in which a typed player (2) moves *after* another player (1), which are called **screening games**. In this case, 1 can only guess 2's reaction based on the prior, and there is no other transfer of information. Payoffs are *hidden* from 1 by Nature's choice.

  In this case, the SPE notion suffices.

- Games as the one in fig. 3.6, where a typed player 1 moves *before* another player 2. In this case, the action taken by 1 is visible by 2, and 2 can use it to infer something about 1's type. In other words, there is a *signal* going from 1 to 2, and that is why this kind of games are called **signaling games**. Note that it is necessary that the utilities of 2 *depend* on the type of 1 (i.e. the game is *common values*), meaning that 2 is *interested* in knowing 1's type.

  In this case, SPE does not suffice to find the *rational* solution, because there is a single subgame, which is the whole game. So, we need to use PBE.

Binary signaling games are often represented as a "butterfly graph":

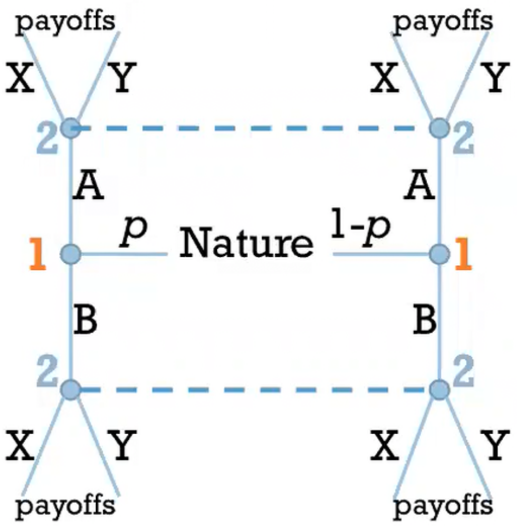

Figure (3.9)

Depending on the choice of 1, there can be different *kinds* of equilibria:

- **Separating equilibria**: all types of 1 choose a *different* action, thus perfectly revealing 1's type to 2 (simplest case, since all systems of belief are *degenerate*).
- **Pooling equilibria**: all types of 1 choose the *same* action, meaning that 2 receives no signal about 1's type.
- **Intermediate cases**, also known as "hybrid", "semi-separating" or "partially-pooling".

Let's consider an example. Brian is inveted by colleague Zöe to a coffee. Ann has two types:

- *Jealous* with probability $p = 0.8$
- *Easygoing* with probability $1 - p = 0.2$

and this is common knowledge. Ann can send a **signal** to Brian, by either staying (S)ilent about this business or (T)rashing Zöe. Brian observes the signal and can accept the (C)offee or kindly (D)ecline the offer. Payoffs are as follows:

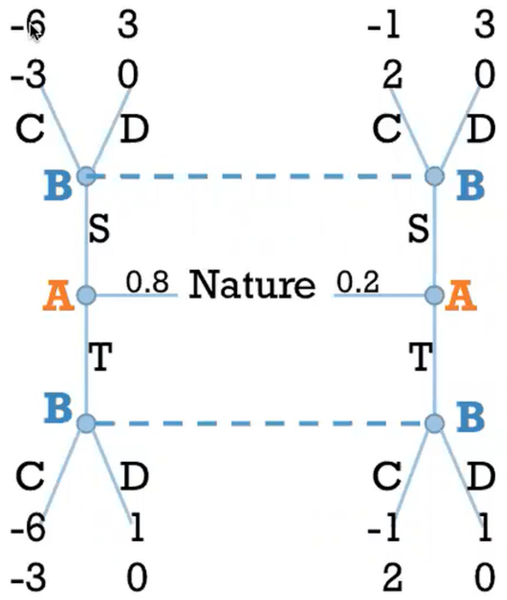

Figure (3.10) – Extensive-form for the Bayesian Signaling Game of *A coffee for Brian*.

If Brian goes out with Zöe, Ann receives a penalty, which is high (-6) if she is jealous, or lower (-1) if not. However, Ann does not want to clearly *signal* her preference by *thrashing* Zöe. If she does so, and Brian follows her signal, she will get just 1. But if Brian *understands* her preference without any need of an explicit say, she gets a higher 3.

Ann has 4 pure strategies: SS, ST, TS and TT. Brian makes binary choices on two information sets (the *upper* part, when Ann plays S and the *lower* one, when Ann plays T), and so has 4 pure strategies too: CC, CD, DC and DD.

For example, a joint strategy can be (TS, CD), meaning that Ann says *bad words* about Zöe only if she is jealous, while Brian takes the coffee only if Ann is silent (or, in other words, Brian "follows the signal").

To find the PBE, we first need to find all NEs. One simple way is to fall back to the normal form, which involves a 4×4 bi-matrix. Filling it may be challenging, so we proceed in steps.

First, consider the *simplest* expectations to compute, which are the ones where both players *stick* with only one move. For instance, if Brian plays CC, the payoffs will be the same *whatever* Ann plays, since the upper/lower part of 3.10 are *equal* for the C leaves. These will be:

$$u_A(CC, **) = 0.8 \cdot (-6) + 0.2 \cdot (-1) = -5$$
$$u_B(CC, **) = 0.8 \cdot (-3) + 0.2 \cdot 2 = -2$$

The same happens for (SS, CD), since T is never played, and so Brian will play only C.

Similarly, for (SS, DD) and (SS, DC), Brian only plays D. In these cases the payoffs in both branches of the upper part of fig. 3.10 are exactly the same, meaning that $u_A(SS, D*) = 3$ and $u_B(SS, D*) = 0$.

The same reasoning leads to $(TT, CD) = (TT, DD)$, for which the payoffs are $(1, 0)$, and $(TT, CC) = (TT, DC)$, for which we already computed $(-5, 2)$. Note that, in this case, it is the *second* letter of Brian's strategy that gets always played, since it is the response to Ann's T.

$$
\mathbf{P} = \text{Ann}\;
\begin{array}{c}
 \\ SS \\ ST \\ TS \\ TT
\end{array}
\begin{array}{cccc}
\multicolumn{4}{c}{\text{Brian}} \\
CC & CD & DC & DD
\end{array}
$$

$$
\mathbf{P} = \text{Ann}\;
\begin{array}{c|cccc|}
 & CC & CD & DC & DD \\
SS & -5,-2 & -5,-2 & 3,0 & 3,0 \\
ST & -5,-2 & & & \\
TS & -5,-2 & & & \\
TT & -5,-2 & 1,0 & -5,-2 & 1,0
\end{array}
$$

All other values need to be *carefully* computed, leading to:

$$
\mathbf{P} = \text{Ann}\;
\begin{array}{c|cccc|}
 & \multicolumn{4}{c}{\text{Brian}} \\
 & CC & CD & DC & DD \\
SS & -5,-2 & -5,-2 & 3,0 & 3,0 \\
ST & -5,-2 & -4.6,-2.4 & 2.2,0.4 & 2.6,0 \\
TS & -5,-2 & 0.6,1.6 & -4.2,-2.4 & 1.4,0 \\
TT & -5,-2 & 1,0 & -5,-2 & 1,0
\end{array}
$$

Note that DD strictly dominates CC. There are 5 NEs, 3 in pure strategies ((SS, DC), (SS, DD) and (TT, CD)) and 2 in mixed strategies[5]: $(TT, 0.5 \cdot CD + 0.5 \cdot DD)$ and $(1/6\,SS + 5/6\,TS, 2/9\,CD + 7/9\,DD)$.

To find the PBE, we need to add also the information about beliefs. For Ann, they are trivial, since her information sets are singletons. On the other hand, Brian has two information sets with two elements, and so needs to know the conditional probabilities μ_S, μ_T of Ann being *jealous* after seeing the signal (S)ilent or (T)rashing.

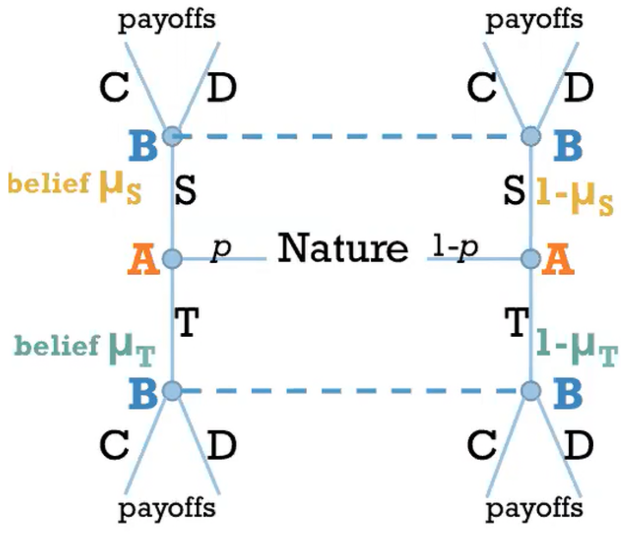

Figure (3.11) – Beliefs in the *A coffee for Brian* game.

[5] ∧Found with software

For a *separating* PBE, these would be easy to compute. For example, for ST we would have $\mu_S = 1$, $\mu_T = 0$.
Unfortunately, all the NEs found are of the *pooling* type, which are more difficult to analyze. For instance, consider SS. Since Ann *always* plays S, the signal carries no information at all, and so $\mu_S = p = 0.8$ is given by the prior knowledge. However, the *lower* part of the graph is now outside the equilibrium path, meaning that μ_T can be arbitrarily chosen, and we need to verify that *whatever the value*, the choice is always rational.

1. (SS, DD). Both players are choosing best responses along the equilibrium path, since this is a NE. The belief $\mu_S = 0.8$ is taken from the prior. *Off* the equilibrium path, i.e. when Ann plays T (which never happens), Brian has to choose between C and D, and we need to verify that D is the best response, which is true if:

$$0 \geq -3\mu_T + 2(1 - \mu_T) \Rightarrow \mu_T \geq 0.4$$

Along with these beliefs, (SS, DD) is a PBE.

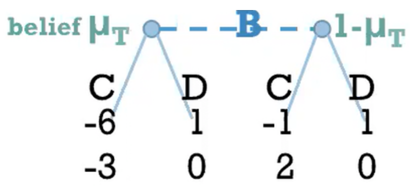

Figure (3.12)

2. (SS, DC). Brian is still playing rationally, even if he is planning to take the coffee if Ann talks bad about Zöe. In fact, this is supported by a *different* set of beliefs:

$$\mu_S = 0.8, \mu_T \leq 0.4$$

In other words, Brian is thinking that Ann plays T if she is *not* very jealous (a sort of *anti-correlated* signal). Along with these beliefs, (SS, DC) is a PBE.

3. (TT, CD). Analogous to the previous one, but with $T \leftrightarrow S$, and so $\mu_T = 0.8$ and $\mu_S \leq 0.4$. Along with these beliefs, it is a PBE.

4. $(TT, {}^1\!/_2 CD + {}^1\!/_2 DD)$. It is a pooling equilibrium in which Brian is indifferent between CD and DD. $\mu_T = 0.8$ is given by the prior. μ_S can be arbitrary, but in this case we can *fix* its values by imposing indifference, leading to $\mu_S = 0.4$. Actually, there are *infinitely* many PBEs where Brian plays $qC + (1-q)D$ with $q \geq 0.5$ and the beliefs $\mu_S = 0.4$, $\mu_T = 0.8$.

5. $({}^1\!/_6 SS + {}^5\!/_6 TS, {}^2\!/_9 CD + {}^7\!/_9 DD)$. This leads to a *semi*-separating PBE, in which most of the time Ann's signal gives full information. Intuitively, Ann is always silent when she is easygoing, but can become talkative when she is jealous, since Brian sometimes chooses C when she is silent.

Note that T is played *only* when Ann is jealous, and so $\mu_T = 1$. Depending on the value of μ_S, Brian may prefer C or D. But since he plays a mixed strategy, he must be indifferent between the two, and so $\mu_S = 0.4$, which is **not** the prior. In fact, this can be found through Bayes' rule. Let $q = \mathbb{P}[S|\text{jealous}]$ be the probability that A plays S *given* that she is jealous. Then:

$$\mu_S = \mathbb{P}[\text{jealous}|S] = \frac{\mathbb{P}[S, \text{jealous}]}{\mathbb{P}[S]} = \frac{pq}{pq + (1-p)} \underset{\substack{p=0.8\\ \mu_S=0.4}}{\Rightarrow} q = \frac{1}{6}$$

which is consistent with the mixing weights for Ann: ${}^1/_6 SS + {}^5/_6 TS$. Thus, we can use Bayes' rule to *find* semi-separating PBEs, if they exist.

Brian instead plays ${}^2/_9 CD + {}^7/_9 DD$, i.e. always responds to T with D, but takes a mixed stance after observing S, making Ann *indifferent* between her two options S and T. With T, Ann gets always 1, and with S she gets -6 if C or 3 if D. Imposing indifference:

$$1 \overset{!}{=} -6\mathbb{P}[C] + 3(1 - \mathbb{P}[C]) \Rightarrow \mathbb{P}[C] = \frac{2}{9}$$

3.8 Exercises

(Lesson 21 of 11/12/2020) Compiled: January 1, 2021

Exercise 3.8.1:

Alfred (A) and Barb (B) are new interns of a company, tasked to work together, and they barely know each other. On their first day, they are sent to the IT technician who will hand them a laptop to use in the internship. The technician is asking them what operating system they want installed, and the available choices are Mac OS (M) or Windows (W). The two interns make this choice independently of one another, also based on whether they prefer "Apple" or "Microsoft" software. Any intern has preference towards them with respective probabilities p and $1-p$. The individual preferences of A and B are independent and the value of p is common knowledge. Clearly, the interns know their individual preferences, but can only estimate the other's through the common prior p. Using the favorite OS gives a benefit quantified as $+1$, while using the lesser preferred one gives 0. However, since their internship requires working together, both A and B know that they will get an additional benefit of $+2$ if they choose the same OS, because it would be easier to exchange software. Discuss what kind of Nash equilibrium would you use for this Bayesian game and analyze the values of p for which a joint strategy where both players follow their types, that is, they choose M if their favorite OS is Mac OS, and choose W if their favorite OS is Windows, is a Bayesian NE.

Solution. There are a total of 4 possible combinations of types: mm (prob. p^2), mw ($p(1-p)$), wm ($(1-p)p$) and ww ($(1-p)^2$). The game's extensive form is as follows:

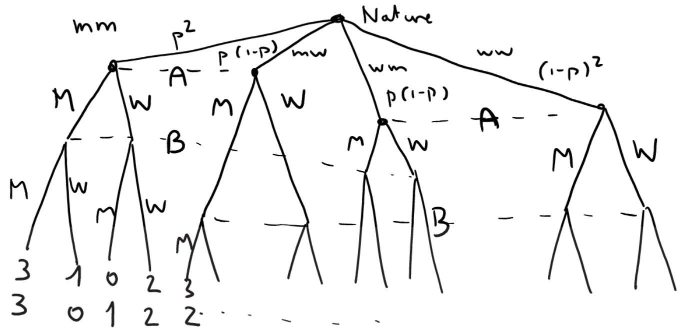

Figure (3.13)

However, this is still a **static** Bayesian game: other than the first Nature's move, all moves happen simultaneously. So, it suffices to use a **Bayesian Nash Equilibrium** (BNE).
The expected utility for (M, M) for `A` is:

$$u_A(M, M) = 3p^2 + 3p(1-p) + 2p(1-p) + 2(1-p)^2$$

Actually, *any* utility value has a form like:

$$\begin{aligned} u(\cdot,\cdot) &= ap^2 + bp(1-p) + cp(1-p) + d(1-p)^2 = \\ &= (a-b-c+d)p^2 + (b+c-2d)p + d \end{aligned}$$

For player `A`, the values of `A`, `B`, c and d are as follows:

$$\mathbf{P} = \mathtt{A}\ \begin{array}{c} \\ MM \\ MW \\ WM \\ WW \end{array} \begin{array}{c} \mathrm{B} \\ \begin{array}{cccc} MM & MW & WM & WW \end{array} \\ \left\| \begin{array}{cccc} 3,3,2,2 & 3,1,2,0 & 1,3,0,2 & 1,1,0,0 \\ 3,3,1,1 & 3,1,1,3 & 1,3,3,1 & 1,1,3,3 \\ 0,0,2,2 & 0,2,2,0 & 2,0,0,2 & 2,2,0,0 \\ 0,0,1,1 & 0,2,1,3 & 2,0,3,1 & 2,2,3,3 \end{array} \right\| \end{array} \tag{3.7}$$

In particular, we are interested in finding *when* it is a BNE to play (MW, MW), where each player follows his/her type. For symmetry reason, let's just consider `A`. We find:

$$u_A(MW, MW) = 4p^2 - 4p + 3$$

This is a NE if `A` has no meaningful unilateral deviation (and if this holds for `A`, then also `B` won't have any incentive to deviate for symmetry reasons). So, the possible deviations are MM, WM and WW, and we keep `B` fixed at MW (*unilateral*). Effectively, this means that we just need to compute the second column of (3.7). Then:

- *Deviation towards MM*: given that $u_A(MM, MW) = 3p$, *not* deviating is convenient if:

$$4p^2 - 4p + 3 \geq 3p \Rightarrow p \leq \frac{3}{4}$$

- *Deviation towards* WM. $u_A(WM, MW) = -4p^2 + 4p$, and so:

$$4p^2 - 4p + 3 \geq -4p^2 + 4p \Rightarrow \text{Always true}$$

which is expected, since WM corresponds to "always choosing the least preferred type".

- *Deviation towards* WW. $u_A(WW, MW) = -3p + 3$, and so:

$$4p^2 - 4p + 3 \geq -3p + 3 \Rightarrow p \geq \frac{1}{4}$$

As a result, (MW, MW) is a BNE if $1/4 \leq p \leq 3/4$. Intuitively, this means that *following* the type is convenient only if there is a *high* uncertainty about the other player's type, i.e. if $p \sim 1/2$. Otherwise, if $p \sim 1$ or $p \sim 0$, the opponent's type is almost certain, and so it would be better to always *pick* the most likely case, even if this is against personal preferences.

Exercise 3.8.2:

At the saloon, a young cowboy `Y` is insulted by a black hat outlaw `O`, that is feared to be the fastest gun in the state. `Y` can let it slide (S): in this case, `Y` gets utility -20 and `O` gets utility 0. Or, `Y` can challenge `O` to a duel (C), in which case `O` can either apologize (`A`) or accept the duel at high noon (D). If `O` apologizes, `Y` gets utility 10 and `O` gets utility -10. The outcome of the duel depends on whether `O` is really a sharpshooter or not. If `O` is a sharpshooter, `Y` has no chance of winning the duel. The payoffs are -100 for `Y` and 20 for `O`. If `O` is just pretending to be a fast gun, then the duel is uncertain: the probability of winning is 0.5 for both `Y` and `O`, and whoever wins gets utility 20 and whoever loses gets -100. The probability of `O` being a sharpshooter is a common prior equal to p.

1. Represent this Bayesian game in extensive form.

2. Represent this Bayesian game in normal form, with a type-agent representation of player `O`.

3. Find the Bayesian Nash equilibria of the game if $p = 0.2$ and discuss whether they are SPE and PBE.

Solution.

1. This is a **Dynamic** Bayesian game. The extensive form is as follows:

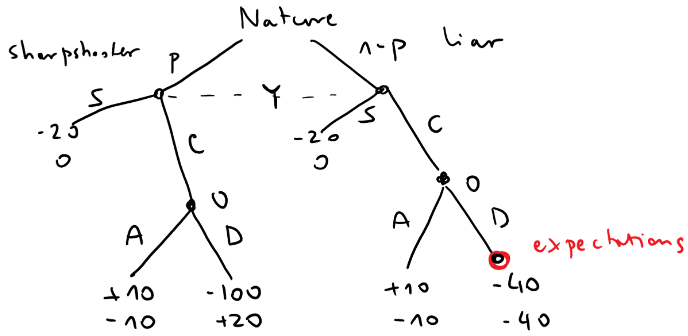

Figure (3.14)

In particular, it is a **screeening game**, since the first player to move has no type, and cannot infer any additional information about the type of the other player.

2. Player `Y` has only two strategies: S and C. Player `O` is typed, and must plan a binary move for each of his two types, leading to 4 strategies: AA, AD, DA and DD. Every pair XY means "X when *sharpshooter*, Y when *liar*" (it is always best to make the notation's meaning explicit).

   Then, the game's normal form is as follows:

$$
\mathbf{P} = \mathtt{Y}\ \begin{array}{c} \\ S \\ C \end{array}
\begin{array}{cccc}
\multicolumn{4}{c}{\mathtt{O}} \\
AA & AD & DA & DD
\end{array}
$$

$$
\begin{array}{c|cccc}
 & AA & AD & DA & DD \\
\hline
S & (-20,\ 0) & (-20,\ 0) & (-20,\ 0) & (-20,\ 0) \\
C & (10,\ 10) & (50p-40,\ 30p-40) & (-110p+10,\ 30p-10) & (-60p-40,\ 60p-40)
\end{array}
$$

3. We set $p = 0.2$, leading to:

$$
\mathbf{P} = \mathtt{Y}\
\begin{array}{c|cccc}
 & \multicolumn{4}{c}{\mathtt{O}} \\
 & AA & AD & DA & DD \\
\hline
S & (-20,\ 0) & (-20,\ 0) & (-20,\ 0) & (-20,\ 0) \\
C & (10,\ 10) & (-30,\ 34) & (-12,\ -4) & (-52,\ -28)
\end{array}
$$

   There are three Bayesian NEs: (S, AD), (S, DD) and (C, DA). Since this is a screening game, the SPE notion suffices. In particular, only (C, DA) is an SPE. In fact, the other two involve a *non-credible threat*: it is not rational for the *liar* `O` to accept the **D**uel. Another way to see that is by using backward induction, which reduces the tree to:

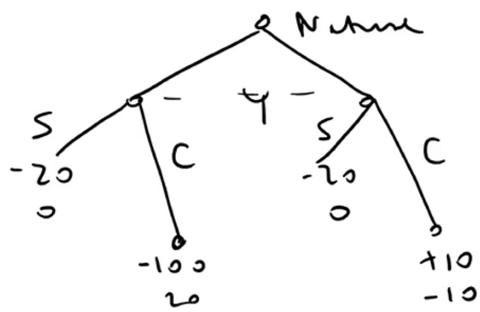

Figure (3.15)

Now, `Y` *knows* the prior probability p, and so he can compute the expected payoff when choosing S or C.

$$\mathbb{E}[u_Y(S)] = -20 \qquad \mathbb{E}[u_Y(C)] = -12$$

And so C is preferable. Thus, the full strategic choice found through backward induction is: `Y` chooses C, then `O` chooses D if *sharpshooter*, and A if *liar*.

(C, DA) is also a PBE, but showing that is not needed for this exercise. In any case, it is easy to construct. In fact, `O` has always information sets that are singletons, meaning that his beliefs are trivial. Player `Y` has a non-singleton information set, but no way to update his beliefs (he is first to move), and so he just uses the prior.

Exercise 3.8.3:

A lawyer (`L`) is representing a defendant in front of a judge (`J`). The defendant can be innocent of guilty with probability $p = 0.6$ or $1 - p$, respectively. This is a common prior. However, `L` gets to know whether the defendant is really guilty or not, and also possesses some key evidence that can be definitely exculpate or incriminate the defendant. During the trial, `L`, can decide whether to **R**eveal this key evidence or to keep it **H**idden. Choosing **R** has a cost for `L`: for example, the key evidence may be obtained from a witness that must be protected, or the plaintiff can make an objection to it. After examining the case, **J** will eventually reach a decision to either **A**cquit or **C**onvict the defendant, and wants to do so fairly, that is, `J` prefers to acquit if the defendant is innocent and to convict if guilty. Consider `J`'s utility to be either 1 or 0 depending on giving the right sentence or not, respectively. As for `L`'s utility, this is 2 if the defendant is acquitted, regardless of whether the defendant was guilty, or not, since in this case `L` gets paid and this is what matters the most. If the defendant is convicted, `L`'s utility is 0 if the defendant was anyways guilty, and -1 if he/she was innocent. Finally, subtract 1 by all of `L`'s utilities when playing **R**, since revealing the key evidence does not come for free.

1. Represent this game in extensive form.

2. Find all pooling, separating and semi-separating Perfect Bayesian equilibria.

Solution.

1. This is a *signaling* game (the typed player acts first), where one action involves revealing (or not) the private type information. Note that the lawyer's type is actually about the nature of the defendant (guilty/innocent). The extensive form is as follows:

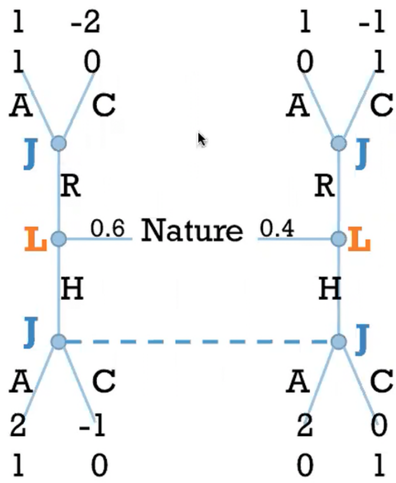

Figure (3.16) – Extensive form for the Lawyer-Judge game.

Note how the non-singleton information set is only present in the *lower* part, where `L` does not reveal (H) his type. In summary: the lawyer `L` is rewarded if the defendant is acquitted, but pays a cost if he reveals evidence. The judge `J` prefers to make the *right* decision.

2. The *upper* (R) part has only singleton information sets, and so it can be solved directly through backward induction:

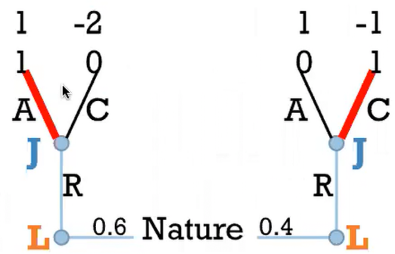

Figure (3.17)

In other words, if `L` reveals the defendant to be innocent/guilty, then `J` *knows* what to do.

After this simplification we can start looking for PBEs.

L makes a binary decision for each of his types, and so he has 4 strategies: RR, RH, HR and HH. Here XY means that X is chosen when J knows that the defendant is innocent ($p = 0.6$) and Y otherwise ($1 - p = 0.4$).

J, on the other hand, needs to plan only a binary choice (A or C) when the type is **not** revealed. If R has been chosen, J automatically knows what to do, and there are no other possibilities to consider.

For PBEs, we need to set also the beliefs. There is only one non-singleton information set, containing just two nodes, meaning that there is only a single belief to be specified: $\mu = \mathbb{P}[\text{innocent}|H]$.

There are two ways to look for PBEs. One is to write the game in normal form, find all the BNEs and examine each one. This approach is thorough, but requires many computations.

A quicker way is to *reason* about possible PBEs: we start from a "reasonable" strategy of a certain kind (e.g. separating/pooling/etc.) and see if there is a system of beliefs that supports it. Here the obvious risk is to *miss* some equilibria.

As an example, let's start with the second approach. First, consider *separating strategies*. One possibility is RH, i.e. L *reveals* the information when the defendant is innocent, and hides it when he/she is guilty. In this case, $\mu = 0$: if J sees H, he immediately knows that the defendant is guilty. The strategies are shown here:

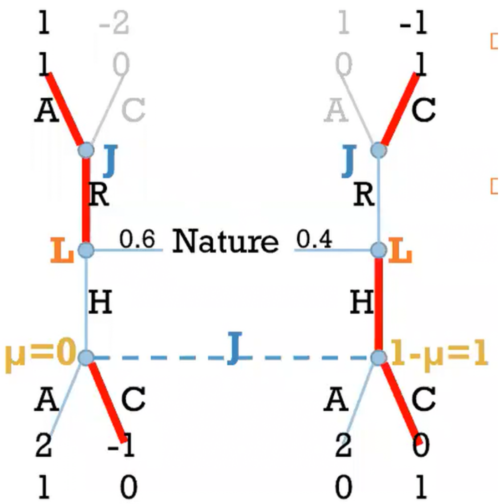

Figure (3.18)

So (RH, C) can be *sustained* by a set of beliefs. To see if it is really a PBE, we need to verify that there isn't any meaningful unilateral deviation. Clearly, J is making optimal decisions. If L changes R *when innocent* to H, then J will think that the defendant is guilty, and L will get -1 instead of 1. Similarly, changing H *when guilty* to R reduces the $\texttt{L}$'s utility from 0 to -1. So, also $\texttt{L}$ has no incentive to deviate. This confirms **sequential rationality**.

Thus, $(RH, C$ after $H)$ with $\mu = 0$ is a separating PBE.

Instead, $(HR,$ after $H)$, with $\mu = 1$, is **not** a separating PBE:

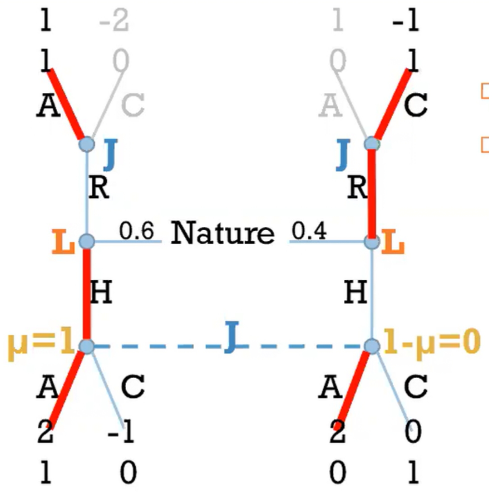

Figure (3.19)

In fact, in this case, $\texttt{J}$ plays rationally, but $\texttt{L}$ has incentive to play H *when guilty*, which would increase his payoff from -1 to 2.

We can now proceed to search *pooling* equilibria. There are two possibilities: RR or HH. The first can be discarded, since revealing *when guilty* is never good for $\texttt{L}$. So, let's consider $(HH,$ after $H)$. The belief $\mu = 0.6$ is given by the prior, since $\texttt{L}$ does not signal anything to $\texttt{J}$. The situation is the following:

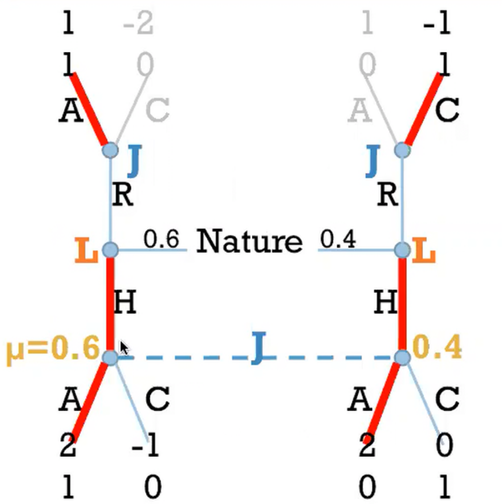

Figure (3.20)

Once again $\texttt{J}$ is playing rationally. If $\texttt{L}$ deviates *on the left*, then he would reduce utility from 2 to 1, and deviating *on the right* reduces utility from 2 to -1. So, sequential rationality holds, and this is also a PBE.

Finally, there is also a *semi-separating* PBE. The idea is that `L` always hides *when guilty*, but sometimes *reveals when innocent*, in such a way that $\mu = 0.5$, i.e. `J` is maximally confused.

The strategies are as follows: `L` plays ${}^{1}\!/\!{}_{3}RH + {}^{HH}\!/\!{}_{HH}$, while `J` plays ${}^{2}\!/\!{}_{3}A + {}^{1}\!/\!{}_{3}C$. `J` plays so that `L` is indifferent of revealing/hiding *when innocent* (but `L` always hides when guilty).

In summary, there are 3 PBEs:

- *Separating*: (RH, C), $\mu = 0$
- *Pooling*: (HH, A), $\mu = 0.6$
- *Semi-separating*: $({}^{1}\!/\!{}_{3}RH + {}^{2}\!/\!{}_{3}HH, {}^{2}\!/\!{}_{3}A + {}^{1}\!/\!{}_{3}C)$, $\mu = 0.5$

These can be found also from the normal form:

$$\mathbf{P} = \begin{array}{cc} & \text{player } \mathtt{J} \\ & \begin{array}{cc} A & C \end{array} \\ \text{player L} \begin{array}{c} RR \\ RH \\ RH \\ HH \end{array} & \left\| \begin{array}{cc} 0.2, 1 & 0.2, 1 \\ 1.4, 0.6 & 0.6, 1 \\ 0.8, 1 & -1, 0.4 \\ 2, 0.6 & -0.6, 0.4 \end{array} \right\| \end{array} \tag{3.8}$$

Note that RR and HR are strictly dominated.

Exercise 3.8.4:

Ania (`A`) and Bruno (`B`) are two students of Game theory who are engaged. They plan to meet at the movies and recreate a Battle of Sexes scenario. Quite conveniently, the movie theater has two options: "*Romantic happiness in the kingdom*" (**R**) and "*Space mutant empires*" (**S**). Still, while `B` is definitely a nerdy character, also `A` is not really a girly-girl, but more of a nerd herself too. So, while their priority is meeting at the same movie, `B` is definitely preferring movie **S** over **R**. About `A` instead, things get more uncertain: `B` believes that `A` prefers **R** with probability p: Obviously, `A` knows what she prefers: but she is also aware that `B` is not sure, and she knows the value of p, that is a common prior. In the end, if they go to different movies, both end up with payoff equal to 0. If they go to see either movie together, `B`'s payoff is 1 or 2 depending on the movie they watch being **R** or **S**, respectively. `A`'s payoff instead will be also either 2 or 1 depending on the preference. These payoff values are common knowledge too.

1. Represent this game in extensive form.
2. Represent the game in normal form with the Bayesian player as type agent.

3. Find the Bayesian Nash equilibria.

Solution.

1. This is a **static** Bayesian game, so we will need only BNEs to solve it. The extensive form is as follows:

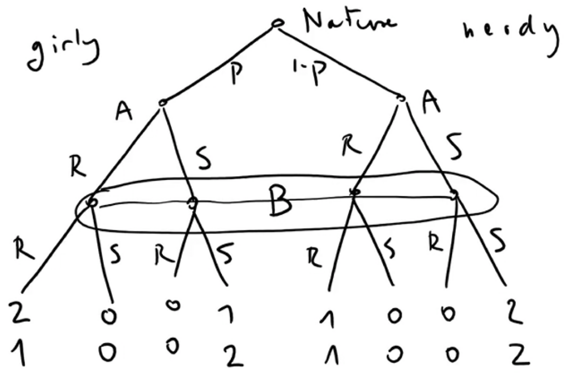

Figure (3.21)

2. The Bayesian player is the typed one, i.e. `A`. Her strategies are binary for *each* type (*type agent* representation), and so they are tuples (what to do when *girly*, what to do when *nerdy*). They are RR, RS, SR and SS. `B` is a regular player, so his strategy is just R or S.

$$\mathbf{P} = \begin{array}{c|c|} & \multicolumn{1}{c}{\text{player B}} \\ & \begin{array}{cc} A & C \end{array} \\ \text{player A} \begin{array}{c} RR \\ RS \\ SR \\ SS \end{array} & \left\|\begin{array}{cc} (1{+}p,\ 1) & (0,\ 0) \\ (2p,\ p) & (2{-}2p,\ 2{-}2p) \\ (1{-}p,\ 1{-}p) & (p,\ 2p) \\ (0,\ 0) & (2{-}p,\ 2) \end{array}\right\| \end{array} \tag{3.9}$$

SR, i.e. the strategy when `A` goes *against* her type, is strictly dominated by a mixture of RR and SS. That is, there is a coefficient ρ so that strategy $m\colon \rho\, RR + (1-\rho) SS$ is *better* than SR:

$$\begin{cases} \rho \cdot (1+p) + (1-\rho) \cdot 0 > 1-p & \texttt{B} \text{ plays } A \\ \rho \cdot 0 + (1-\rho) \cdot (2-p) > p & \texttt{B} \text{ plays } C \end{cases} \Rightarrow \begin{cases} \rho > \frac{1-p}{1+p} \\ \rho < \frac{2(p-1)}{p-2} \end{cases}$$

In general, best responses depend on the value of p. However, since $p \in [0, 1]$, in some cases one move is always a best response: for example $2 - p$ is always strictly bigger than 0. There is only one ambiguity: when `A` plays RS, `B`'s best response is either A or C, depending on p. However, neither one can be a pure NE, so this is not a problem.

3. From the normal form, we see that there are two pure NE: (RR, R) and (SS, S). Both do not depend on the value of p.

   From the original Battle of Sexes, we know that there should also be a mixed NE. First, the strictly dominated strategy can't be part of the NE. Also, the NE can't involve a mixture of RR and SS: by the indifference principle, such a mixture must satisfy $u_A(RR, \cdot) = u_A(SS, \cdot) \geq u_A(RS, \cdot)$, which is not possible. For the same reason, there is no mixed NE combining *all* 3 strategies RR, RS and SS.

   So, let's consider a mixture of RR and RS. Let's consider a strategy $q\colon q\,RS + (1-q)\,RR$ for `A` and $c\colon cR + (1-c)S$ for `B`, and apply the indifference principle.

$$u_A(RR, c) = (1+p)c \overset{!}{=} 2pc + (2-2p)(1-c) = u_A(RS, c) \Rightarrow c = \frac{2}{3}$$
$$u_B(q, A) = p \cdot q + 1 \cdot (1-q) \overset{!}{=} (2-2p) \cdot q + 0 \cdot (1-q) = u_B(q, C)$$
$$\Rightarrow q = \frac{1}{3(1-p)}$$

   Since $q \in [0, 1]$ (it must be a proper probability), we need to impose also:

$$\frac{1}{3(1-p)} \leq 1 \Rightarrow p \leq \frac{2}{3}$$

   Similarly, we can consider a mixture of RS and SS, with `A` playing $q\colon qRS + (1-q)SS$, and `B` $c\colon cR + (1-c)S$. Imposing indifference leads to:

$$c = \frac{1}{3}; \quad q = \frac{2}{3p} \quad p \geq \frac{2}{3}$$

   So, there are always at least 3 NEs: 2 pure and 1 mixed. If $p = 2/3$, there are 4: 2 pure and 2 mixed.

   Note that when p is sufficiently high, then `A` behaves as in the original Battle of Sexes, and so does `B`. If, instead, $p < 2/3$, then `A` is *different*, and her strategy changes, adding SS. However, note that in *all* these mixed NEs, `A` goes to R with probability $2/3$ *always*, independent of p. This is because `A` plans her strategy as a response to `B`, so that he will be *indifferent* about his options.

Chapter 4

Classic problems

4.1 The Master degree

(Lesson 22 of 15/12/2020) Compiled: January 1, 2021

A freshly graduated Bachelor Student (player 1) is uncertain about pursuing an MS degree. He can be a (H)ighly or (L)owly skilled student, and he is aware of which one is it. Depending on his talent, there is a different *cost* of gettin an MS: if 1 is skilled the cost is lower $c_H = 2$, otherwise it is higher $c_L = 5$. 1 can either get the (D)egree, or remain (U)ndergraduate.

Player 2 is an employer that can either give 1 a (M)anager or (B)lue collar job, with different wages: $w_M = 10$, $w_B = 6$. 1's utility is the difference between wage and cost. On the other hand, the employer's **net** profit depends on the assignment/skill match. Let's assume that the degree has *no impact* on that. Instead, a higher payoff is given if 2 assigns 1 to the *correct* position (e.g. Manager if Skilled, Blue collar if not):

$$\begin{array}{c} \\ H \\ L \end{array}\begin{array}{c} \begin{array}{cc} M & B \end{array} \\ \left\| \begin{array}{cc} 10 & 5 \\ 0 & 3 \end{array} \right\| \end{array} \tag{4.1}$$

The game's extensive form is as follows:

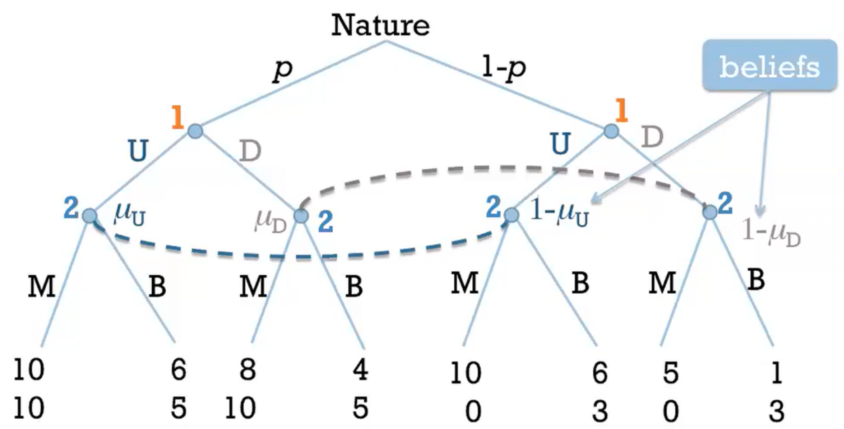

Figure (4.1)

Note that the choice taken by 1 acts as a *signal* for player 2.

In general, 2 believes that 1 is choosing a mixed strategy where:

- U is chosen with probability σ^H when type is H
- U is chosen with probability σ^L when type is L

Applying Bayes' rule we can compute 2's beliefs:

$$\mu_U = \frac{p\sigma^H}{p\sigma^H + (1-p)\sigma^L} \qquad \mu_D = \frac{p(1-\sigma^H)}{p(1-\sigma^H) + (1-p)(1-\sigma^L)}$$

Note that if D is never chosen, $\sigma^H = \sigma^L = 1$, and so only the first belief μ_U can be computed from the above formulas, while μ_D is arbitrary.

Both players have 4 possible pure strategies. In fact, 1 is typed, and has to choose a binary action for each type. 2 is not type, but has to choose a binary response to either U or D.
Let's fix $p = 1/4$ and write the game's normal form:

$$\mathbf{P} = \begin{array}{c c} & \text{Player 2} \\ & \begin{array}{cccc} MM & MB & BM & BB \end{array} \\ \text{Player 1} \begin{array}{c} UU \\ UD \\ DU \\ DD \end{array} & \left\| \begin{array}{cccc} 10, 2.5 & 10, 2.5 & 6, 3.5 & 6, 3.5 \\ 6.25, 2.5 & 3.25, 4.75 & 5.25, 1.25 & 2.25, 3.5 \\ 9.5, 2.5 & 8.5, 1.25 & 6.5, 4.75 & 4.5, 3.5 \\ 5.75, 2.5 & 1.75, 3.5 & 5.75, 2.5 & 1.75, 3.5 \end{array} \right\| \end{array}$$

There are two pure NEs, and also a mixed NE which is harder to find, and won't be discussed here.

- (DU, BM) is a *separating* NE, in which getting the degree is perfectly correlated with having high skill. All highly skilled students get the degree and are hired as managers, while the other ones remain undergraduate and get the blue collar jobs. Beliefs are $\mu_U = 0$ and $\mu_D = 1$.
- (UU, BB) is a *pooling* NE, where nobody gets a degree and everyone does menial jobs. In this case $\mu_U = 1/4$ is directly from the prior, while μ_D is arbitrary. Surely it can't be $\mu_D = 1$, since 2 is choosing B over M for graduates, which can happen only if:

  $$5\mu_D + 3(1-\mu_D) > 10\mu_D + 0 \cdot (1-\mu_D) \Rightarrow \mu_D < \frac{3}{8}$$

  This equilibrium is a PBE, i.e. it is a sustainable *rational* solution, only if 2 believes that less than 3/8 of graduates are highly skilled.

4.2 Reputation building

Uncertainty over players' types can lead to cooperative behaviors.

For example, consider a finitely repeated version of the Prisoner's Dilemma. From previous discussions, we have seen that the only SPE is that where nobody cooperates.
Cooperation can be established by considering infinite repetitions, since in that case there is a *grim trigger* strategy providing an *infinite punishment* for any deviation.
Can this be done by using *incomplete* information over a finite number of rounds?

The basic setup of a round is as follows:

$$\mathbf{P} = \begin{array}{cc} & \text{player B} \\ & \begin{array}{cc} C & D \end{array} \\ \text{player A} \begin{array}{c} C \\ D \end{array} & \begin{Vmatrix} 1,1 & -1,2 \\ 2,-1 & 0,0 \end{Vmatrix} \end{array} \tag{4.2}$$

Either player can (C)ooperate or (D)efect. Only player A is *typed*, and can be either *strategic* (with probability $1-p$), behaving normally, or *grim-trigger* (p), who starts playing C and switches *forever* to D if the opponent does not always play C. Clearly, this can happen only if payoffs for the *grim-trigger* version of B are different from the above ones.

Effectively, one way to think of this setup is *as if* there were 3 players: a strategic B, a strategic A and the *grim-trigger* A.

One round

If the game is repeated once, then B's strategy will be D, and A's will play C if *grim-trigger* and D if strategic, i.e. CD.

Two rounds

If the game is repeated twice, there is a unique PBE in which, at the second round, the *strategic* player A and player B both choose D. This is because strategic players *must* play a NE at the last round.
However, at the first round, strategic A will still behave normally and *defect*, but B does not know A's type. So, if $p > 1/2$, i.e. if there is a high probability of A being the grim-trigger, it is more convenient for B to play C.

Let's formalize this. In the second round, rational players must play a NE. There are two possibilities: (D, CD) if B has not defected in the first round, and (D, DD) if the grim-trigger has been angered. The expected utilities for B in both cases are:

$$u_B(D, CD) = 2 \cdot p + 0 \cdot (1-p) = \boxed{2p}$$
$$u_B(D, DD) = 0 \cdot p + 0 \cdot (1-p) = \boxed{0}$$

In the first round, A still plays CD, and so B's expected utilities *over the whole game* are as follows:

$$u_B(C, CD) = 1 \cdot p + (-1) \cdot (1-p) + \boxed{2p} = 4p - 1$$
$$u_B(D, CD) = 2 \cdot p + 0 \cdot (1-p) + \boxed{0} = 2p$$

Comparing the two:

$$u_B(C, CD) = 4p - 1 > 2p = u_B(D, CD) \Rightarrow p > \frac{1}{2}$$

If the game is repeated over 3 stages, and $p > 1/2$, then there is a **unique** PBE in which everybody starts the game with *cooperation*: (C, CC).

Three or more rounds

In fact, consider 3 rounds. If B plays D at the start, all subsequent rounds will be played as (D, DD), which is not in the best interest of B. But if B plays C at the start, then the second and third round are exactly the same as the *two-round* game examined above.
Note that also the strategic A plays C, because he/she does not want to *reveal* A's type to B. In fact, only *strategic* A can play D at the first round, and if this happens all subsequent rounds will be played as (D, DD).

Let's check this computationally. Denote with μ_G B's belief that A is a grim trigger. Suppose B plays C at the first round. There are two possibilities:

- Also strategic A plays C, i.e. the join strategy is (C, CC), a *pooling* strategy. In this case, *no signal* is passed to B, who can only exploit the prior to set $\mu_G = p > 1/2$, and so will be lead to believe that A is the *grim-trigger* (more likely). Thus, strategic A will get 1 in the first round. Then, the game proceeds as the *two-rounds* version, with B playing C again, and strategic A switching to D, and getting 2. Then in the last round both B and strategic A play D, and get 0. So, by not revealing themselves at the start, strategic A gets the maximum total payoff of $1 + 2 + 0$.

- The strategic player A chooses D, revealing themselves: $\mu_G = 0$. Following rounds are all Ds, and so the overall payoff for strategic A is $2 + 0 + 0$, which is less than the other possibility.

So if B plays C at the start, the game will be played as (C, CC), (C, CD), (D, CD). The expected payoff for B is then:

$$\underbrace{p \cdot (1 + 1 + 2)}_{\text{Grim-trigger}} + (1 - p) \cdot (1 - 1 + 0) \cdot (1 - p) = 4p$$

Instead, if B plays D at the start, the *grim-trigger* is activated immediately, and all subsequent rounds are played as (D, DD), with 0 utilities for both players. There are two cases:

- Strategic A plays C, leading to the joint strategy (D, CC), where B gets 2 for sure, and strategic B gets -1.

- Strategic A plays D, leading to (D, CD). Here B gets $2p \leq 2$, and strategic B gets 0.

The best case gives 2, and comparing that with the payoff of playing C leads to $4p > 2 \Rightarrow p > 1/2$, and so B will choose to start with C.

In fact, this happens even if B believs strategic A will still play D. In this case, the 3 rounds are (C, CD), (C, CD) and (D, CD), with an expected utility for B:

$$p \cdot (1 + 1 + 2) + (1 - p) \cdot (-1 - 1 + 0) = 5p - 1$$

And if B plays D, the 3 rounds are (D, CD), (D, DD) and (D, DD), with expected payoff for B equal to $2p$. Comparing the two:

$$u_B(C) = 5p - 1 > 2p = u_B(D) \Rightarrow p > \frac{1}{3}$$

So, cooperation is *stable*!

The interesting thing is that also strategic A has incentive to play C at the first round, to "build a reputation" for the following rounds. The final betrayal is inevitable, but it is *delayed*.

This result is generalized by Kreps, who finds that for finitely repeated games with $T \gg 1$ stages, the number of rounds with defection of a strategic player is $< M$, where M depends only on p, and not T.

4.3 The tough negotiator

Consider a ultimatum game with just two offers from player 1: a (H)igh share or a (M)oderate share for him/herself. Player 2, the responder, can (A)ccept or (R)efuse.

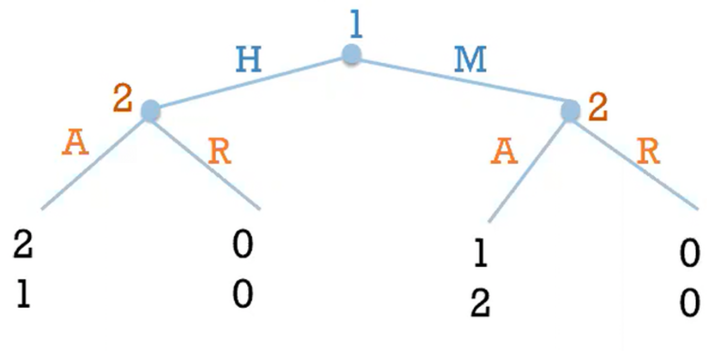

Figure (4.2)

The only SPE is (H, AA), i.e. 2 always accepts. There are two others NEs (non-SPE):

- (H, AR), which is the same as (H, AA), with a *irrational* choice *off* the equilibrium path.
- (M, RA), with a *non-credible* threat, in which 2 plans to refuse H, i.e. a bad share for him/her. Still, H is better than receiving nothing, so a rational 2 should accept it.

Let's introduce types to model some kind of *irrationality* for player 2, who can be either (N)ormal (rational) or a (J)erk (thinking H is worse than receiving nothing).
The extensive form is as follows:

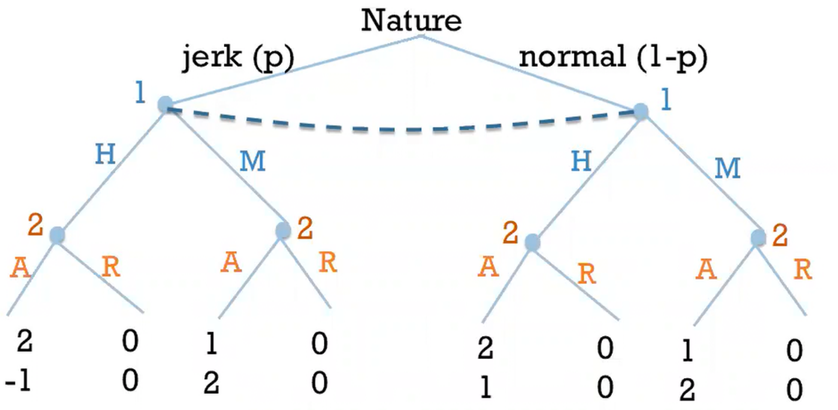

Figure (4.3) – Extensive form for the Bayesian Ultimatum game.

This is a *screening* game, since the typed player acts at the end, "hiding" the payoffs beyond a layer of uncertainty. In this case beliefs require few computations. In fact, for 1, the beliefs are just given by the prior. Then, if $p > 1/2$, 1 should offer M: if the *left* branch is the more likely one, then 2 will probably refuse H and accept M. This PBE can be found directly from backward induction, since this is a *screening* game, and so we can use effectively just the SPE definition.

Two rounds

So, let's make the game more complex by including a *second round*. If the first offer is rejected, 1 can give another offer, which 2 can accept/refuse. We also include a discount δ for payoffs, so that if 1 offers *the same amount* during the second round it will be refused: if it was acceptable, it would have been accepted at the start.

In this case, 2's action in the first round forms a *signal* from which 1 can infer something about 2's type. Let μ_H the probability of 2 being a *jerk* if he refuses a H share, and similarly μ_M for the M share.
The game's extensive form is quite complex:

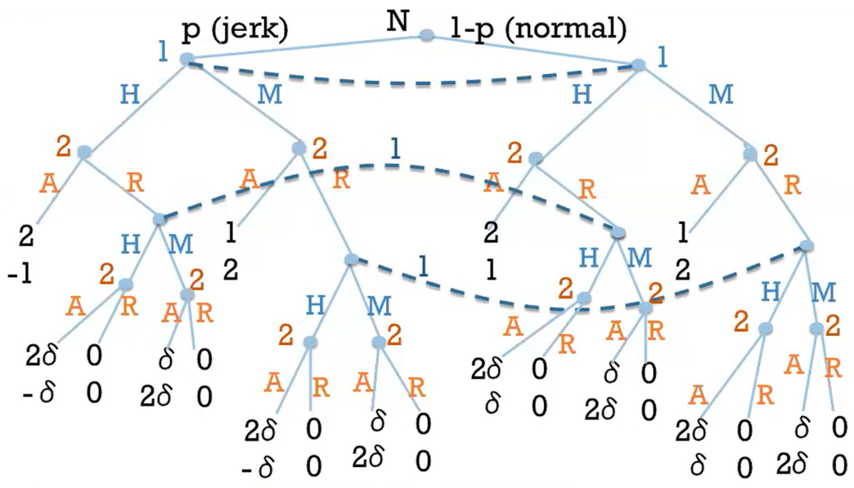

Figure (4.4)

and the normal form is non-tractable by hand. This complexity is required to make the game more interesting. In fact, suppose 2 is always *normal*. Then

at the final round 2 will accept everything. Knowing this, 1 offers always H, and 2 will accept it *at the first round* to avoid paying the discount. There is no way for 2 to refuse the first offer and get something better, since the *threat* of refusing at the end is *non-credible*.
However, if 2 is typed, the added uncertainty makes *bargaining* possible. In fact, if 2 refuses the first offer, 1 cannot know exactly 2's type. He/she may be a *jerk*, who refuses by nature, or a rational player who is *emulating* a jerk, and tries to *avoid* 1 discovering his/her type. In other words, when types are added, even the *rational* 2 may not accept everything, contrary to the non-typed game.

Even if we understand the need for such complexity, we still need a way to *simplify* the game in order to analyze it. In particular:

- If M is offered, 2 always accepts it: refusing it cannot lead to a better outcome (and this is common knowledge). More precisely, accepting M at the first round gives $u_2 = 2$, but the maximum payoff at the second round is $2\delta \leq 2$. This is a consequence of **sequential rationality**. Clearly, M is always accepted even at round 2, since both 2's types prefer it.

- Backward induction can be applied to solve 2's last move.

The simplified[1] extensive form is as follows:

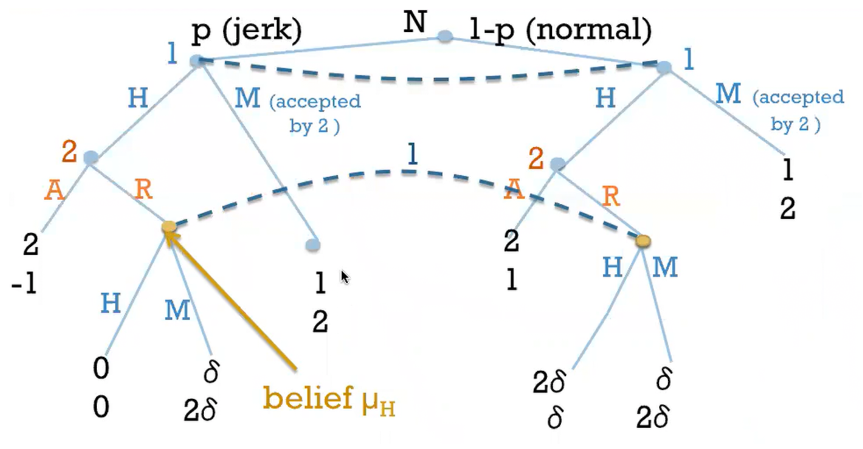

Figure (4.5)

The only remaining belief is μ_H, i.e. the conditional probability of 2 being a *jerk* if he/she refused H.
Player 1 makes two binary offers, and so has 4 pure strategies. Player 2 plans only his/her binary response to H for each of his/her two types, and so has 4 pure strategies too.

This simplified game can be put in normal form, where we let $\delta = 0.9$ and

[1] ∧Effectively, this is a *different* game, since we are entirely removing some actions.

$p = 0.1$:

$$
\mathbf{P} = \begin{array}{cc|cccc|}
 & & \multicolumn{4}{c}{\text{Player 2}} \\
 & & AA & AR & RA & RR \\
\hline
\multirow{4}{*}{\rotatebox{90}{Player 1}} & HH & 2, 0.8 & 1.82, 0.71 & 1.8, 0.9 & 1.62, 0.81 \\
 & HM & 2, 0.8 & 1.01, 1.52 & 1.89, 1.08 & 0.9, 1.8 \\
 & MH & 1, 2 & 1, 2 & 1, 2 & 1, 2 \\
 & MM & 1, 2 & 1, 2 & 1, 2 & 1, 2
\end{array}
$$

Note that MH and MM are strictly dominated strategies: offering M in the first round is a bad choice, it is always better to "probe the waters" and see if 2 can accept something lower. Interestingly, HH is one of the *viable* strategies: this is because 2 is unlikely to be a *jerk*, meaning that an early refusal is unlikely, and also the discount factor is close to 1, so it is not *too bad* to go to the second round anyways.

Then, AA and AR are *dominated* strategies (but not strictly). Intuitively, when 2 is a *jerk*, he/she will not accept H. However, after removing MH and MM, we can remove also AA and AR. The simplified normal form is:

$$
\mathbf{P} = \begin{array}{cc|cc|}
 & & \multicolumn{2}{c}{\text{Player 2}} \\
 & & RA & RR \\
\hline
\multirow{2}{*}{\rotatebox{90}{Player 1}} & MH & 1.8, 0.9 & 1.62, 0.81 \\
 & MM & 1.89, 1.08 & 0.9, 1.8
\end{array} \tag{4.3}
$$

There is no NE in pure strategies, so there must be one in mixed strategies, due to Nash Theorem, which is 1 playing $^8/_9 HH + {}^1/_9 HM$ and 2 playing $^8/_9 RA + {}^1/_9 RR$.

This is a PBE, at least for the reduced game, and in fact it makes the belief $\mu_H = 0.5$, i.e. *maximizes* 1's uncertainty about 2's type. Since we have used **sequential rationality**, this is also the *core* of the PBE for the full game: we just need to add *all* the information *off* the equilibrium path. In particular, μ_M is arbitrary, but must be in $[0, 0.5]$.

Note that H is *always* offered at the first round. *Jerk* always rejects it, and *normal* does so $^8/_9$ times. Player 1 has a probability $^8/_9$ of sticking to H also during the second round, and $^1/_9$ of changing it to M.

This is a *semi-separating* PBE: *pooling* and *separating* strategies are inconvenient per se. In fact:

- *Pooling on accept* (AA) is not convenient, since *jerk* would accept H, but wants to reject it. *Pooling on reject* (RR) means that 1's beliefs follow the prior, and so since *jerks* are rare, 1 will never offer M in the second round, and *normal* 2 will need to accept H. But this is suboptimal: if the game was to end this way, why not accept at the first round, avoiding the discount δ?

- The only considerable *separating* strategy is RA, so *normal* 2 accepts H at the first round, while *jerk* continues to the second one, and gets offered

M. However, it is slightly better for *normal* 2 to "pretend being *jerk*" so that he/she has a chance of accepting M in the second round (since the discount is not too severe).

4.4 Auctions

(Lesson 23 of 18/12/2020) Compiled: January 1, 2021

Auctions seen in movies are *dynamic*: several players respond to each other bids, raising the value until the object is sold. However, this is very difficult to analyze, and so in this section we will limit ourselves to *sealed-bid* auctions, in which *bids* are sent in closed envelopes, and revealed simultaneously.

More precisely, there is one seller and N bidders. Each bidder i has a **personal valuation** $v_i \in \mathbb{R}^+$ of the object, and takes a move $b_i \in \mathbb{R}^+$. All moves happen simultaneously, and the highest bid wins. In case of a tie, the object is *shared* between all winners.

4.4.1 First price auction

The simplest way to model a *sealed-bid* auction is to let winners *pay* their bids to the seller. Their payoffs are given by the *difference* between their personal valuation of the object and the bid they placed (shared in case there are multiple winners):

$$u_i(b_1, \dots, b_n) = \begin{cases} v_i - b_i & i \text{ is a single winner} \\ (v_i - b_i)/W & i \text{ is among } W > 1 \text{ winners} \\ 0 & i \text{ is not a winner} \end{cases}$$

Note that bidding $b_i \geq v_i$ always yields $u_i \leq 0$, so rational players will not ever do that. In fact this is a *dominated* strategy (but not *strictly* dominated). Also, $b_i = 0$ is weakly *dominated* by any other $b_i' > 0$ with arbitrarily small b_i'. So, a rational choice is to bid something $0 < b_i < v_i$, but there is no further indication on *what* choice to make.

Note that, in general, bidders will try to stay significantly *under* their valuations v_i, meaning that the object will be sold at a discounted price, which is bad for the seller.

4.4.2 Second price auction

A model that is better for the seller is the one proposed by Vickrey in 1961. In this case, the winner pays the **second highest bid**. Let the maximum of all *other* bids (of $-\boldsymbol{i}$) be $S = \max_{b_j \neq b_i} b_j$. Then, i's payoffs are as follows:

$$u_i(b_1, \dots, b_n) = \begin{cases} v_i - S & b_i > S \\ (v_i - S)/W & b_i = S \\ 0 & b_i < S \end{cases}$$

In fact, if $b_i < S$, then b_i is not the highest bid, and so i is not the winner. Otherwise, if $b_i > S$, then S is the second highest bid, and i pays it. Finally, $b_i = S$ means that i has the highest bid, but there are other players that bid the same, and so the cost is shared.

At a first glance, paying the *second highest bid* would seem to lead to *even less money* going to the seller. However, note that now bidding $b_i = v_i$ *always* yields $u_i \geq 0$, or even $u_i > 0$ if the winner is unique.
In fact, bidding v_i is a *weakly dominant* strategy for i. We can see that by first proving that bidding $b_i < v_i$ is not a good move. In general u_i depends on S, and so we compare the payoff of playing b_i with that of playing v_i in all possible cases:

- If $S < b_i < v_i$, then both b_i and v_i are unique winning bids, and for both $u_i = v_i - S$.
- If $S = b_i < v_i$, then b_i is a non-unique winning bid with payoff $u_i = (v_i - S)/W$, which is less than the payoff $v_i - S$ of playing the unique winning bid v_i.
- If $S > b_i$, b_i loses ($u_i = 0$), but v_i *may* win ($u_i \geq 0$).

Similarly, playing $b_i > v_i$ is also not a good move:

- If $S < v_i < b_i$, both b_i and v_i are unique winning bids, with the same payoff $u_i = v_i - S$.
- If $S = v_i < b_i$, v_i is a non-unique winning bid with payoff 0, while b_i is a unique winning bid with the same payoff 0.
- If $S > v_i$, v_i loses ($u_i(v_i) = 0$). Now, either b_i loses ($u_i(b_i) = 0$) or it is one of $W^* \geq 1$ winning bids, with a *bad* payoff $u_i(b_i) = (v_i - S)/W^* < 0$.

Then, a joint strategy made of dominant strategies **is** a Nash Equilibrium. So, the second price auction has a NE given by $(v_i, v_2, \ldots, v_n)$, and so this model *encourages* bidders to offer what they think it is the real object's value. In fact, if a player wins, they will still play something *less* than their valuation.

This strategy is even Pareto efficient. The utility of the bidder is $v - S$, and that of the seller is S, meaning that the total utility is v, i.e. exactly the **highest** evaluation of the object. So, total utility (social surplus) is *naturally* maximized by a second price auction.

4.4.3 Vickrey-Clarke-Groves auction

A second price auction can be generalized to sell multiple object with the Vickrey-Clarke-Groves (VCG) model. The idea is that each bidder submits to the seller a different value for each **combination** of the items. This is because sometimes a combination of items is worth *more* than the sum of the values of its parts, and this changes the players' strategies.

A player *wins* an object if his/her bid for *that* set of objects is the **highest** between all other possible bids for that set, including also "joint bids" of different players.

For example, consider two objects A and B. 1 offers 5 for A and 3 for B, while 2 offers 10 for $A+B$, while 3 offers 4 for A and 6 for B. Then, 1 wins A and 3 wins B, because the combined bid for $A+B$ obtained by joining 1's bid for A and 3's bid for B reaches 11, which is better than 2's bid for $A+B$ (10). Note that if objects are many, finding the best combination requires a lengthy combinatorial search (NP-hard).

Then, each player pays according to "how he/she hurt the others". If $\boldsymbol{i}$ gets item $\boldsymbol{x}$, the price paid for it is:

$$V(N \setminus \{\boldsymbol{i}\}, M) - V(N \setminus \{\boldsymbol{i}\}, M \setminus \{\boldsymbol{x}\})$$

where N is the set of all players, and M that of objects in the auction. Then V is the total valuation made by all the players to the current best outcome. In other words, payment is equal to the difference between the *added valuation* by $\boldsymbol{i}$ and the loss of valuation if $\boldsymbol{x}$ were not present.

For example, consider 3 players bidding for two objects:

- A bids 5 for one object.
- B bids 4 for one object.
- C bids 7 for the pair, but does not want a single object.

(All bids *correspond* to the players' valuations for the objects)
The optimal allocation is selling one object to A and one to B ($5+4>7$).

To compute what A should pay, note that *without* A, the total valuation would be 7 (C wins), and without A *and* the object, B would have won (4), so we get $7-4=3$. In other words, if A were not present, B would have lost (-4) the single object (which has *thanks to* A), but C would have gained the pair (+7).

Similarly, without B, A loses the object (-5) but C gains the pair (+7), and so B pays $7-5=2$. Finally, C has lost, and so he/she pays nothing.

However, the VCG is not commonly used in practice for the following reasons:

- Non-monotonic behavior: increasing the bid can improve the utility of some other player, since this can "kick out" some other players. For example, suppose that in the above example C offers 10 for the pair, and so he/she is the winner. However, if A changes the bid from 5 to 7, C is *kicked out* from the auction, and suddenly B is winning too, without having to change anything, and *has to pay* his/her share!
- Vulnerable to collusions: some players can group and share their valuations, find a common-ground and use that to gain an advantage in the auction, acting *as if* they had *different* valuations.
- In general players will tend to *estimate* valuations of the others, leading to low revenues for the seller.

4.5 Mechanism design

Consider n players that choose among a set of **public alternatives** X according to their *privately known* types $\boldsymbol{t} = (t_1, t_2, \dots, t_n)$, denoted as **state of the world**. The prior distribution of types is common knowledge.

In general, players might not want to reveal their type to others. A **mechanism** is a game specifically designed to *discover* the players' types. In this way, "appropriate goods" can be assigned to each type, obtaining a "globally optimal solution" (i.e. Pareto efficient).

In a sense, this is very similar to what happens in a second price auction: rational players *naturally* reveal their types (i.e. their valuation), the object is assigned to the *most appropriate* player, and the social surplus is maximized.

Let's formalize this. Given an outcome $o \in \mathcal{O}$, player i gets a utility $u_i(o, t_i)$, depending on their type t_i. In mechanism design, we want to *fix* the final outcome as a function of the types: any joint type $\boldsymbol{t} = (t_1, \dots, t_n)$ of the players is mapped to a desired outcome $f(\boldsymbol{t})$ by the so-called **social choice function** $f : T_1 \times \cdots \times T_n \to \mathcal{O}$.

A mechanism is a game constructed so that rational players of types $\boldsymbol{t}$ are lead at equilibrium to the desired outcome $f(\boldsymbol{t})$. More precisely, the types $\boldsymbol{t}$ determine the joint strategy $\boldsymbol{s}^*$ which is played at the Bayesian Nash Equilibrium. Given any strategy $\boldsymbol{s}$, its outcome is denoted as $g(\boldsymbol{s})$, where g is the appropriately named **outcome function**. Thus, a mechanism implementing f is a game with a BNE $\boldsymbol{s}^*$ such that $g(\boldsymbol{s}^*(\boldsymbol{t})) = f(\boldsymbol{t})$. In other words, players behaving rationally, who only try to maximize their own utilities, are lead to a *designed* outcome.

Mechanism design is often used for constructing auctions. In this case, we consider an *allocation* game. An outcome o consists of an allocation $x \in X$ (i.e. *who* gets the object) and a vector of payments $\boldsymbol{m} = (m_1, \dots, m_n)$ made by all players to the seller. Following the formalism from the above analysis of second price auctions, we assume utilities to be *quasi-linear*:

$$u_i(\underbrace{(x, \boldsymbol{m})}_{o}, t_i) = v_i(x, t_i) - m_i$$

That is, each player has a **valuation function** $v_i(x, t_i)$, determining the value of x. For an auction, $v_i(i, t_i) = v$ and $v_i(j, t_i) = 0$ for $j \neq i$, i.e. each player *wants* the object for him/herself.

With this notation, the social choice function can be written as $f : T \to Y$, where $T = T_1 \times \cdots \times T_n$, and Y is the set of outcomes given by:

$$Y = \{(x, m_1, \dots, m_n) : x \in X, \sum_i m_i \leq 0\}$$

The outcome function $g(\boldsymbol{s})$ can be decomposed in a pair $(x(\boldsymbol{s}), \boldsymbol{m}(\boldsymbol{s}))$, where $x(\boldsymbol{s})$ is the *decision rule*, i.e. how to assign the correct x given a joint strategy $\boldsymbol{s}$, and $\boldsymbol{m}(\boldsymbol{s})$ is the *transfer rule*, i.e. how to decide how much each player should *pay* given a joint strategy $\boldsymbol{s}$. *Why negative?*

The simplest way to implement a mechanism is to *just ask the types*. That is, introduce an action for each available type, and design the equilibrium to be fully separating. In this way each action corresponds to a type ($A_i = T_i$), and a joint strategy $\boldsymbol{s}$ *is* the same as a type vector $\boldsymbol{t}$. Then we set $g(\boldsymbol{t}) = f(\boldsymbol{t})$. This kind of game is denoted as a **direct revelation mechanism** Γ.
We say that a choice rule f is **truthfully implementable** in a Bayesian Nash Equilibrium if, for all $\boldsymbol{t} \in T$, the direct revelation mechanism $\Gamma(T, f)$ has a BNE $\boldsymbol{s}^*$ where $s_i^*(t_i) = t_i$ for all i.

An interesting result is the **revelation principle**, which says that a mechanism implementing a choice rule f exists **if and only if** f is truthfully implementable in a Bayesian Nash equilibrium. In other words, if Γ does not work for a certain f, there is no other "clever" mechanism that will work.

4.5.1 Coalitional games

Most games we analyzed involved *competition* between different players. However, Game Theory can be used also for collaborative games.

For example, consider a set $\mathcal{N}$ of players. We define a coalition S as any subset of $\mathcal{N}$. Each player i can choose any combination of *coalitions* to join, i.e. a set of elements from the *parts of* $\mathcal{N}$ ($\mathcal{P}(\mathcal{N})$). For example, if i chooses $\{i\}$ then he/she will be a *loner*.
Each coalition S has a value $v(S)$ given by the **value function** $v\colon \mathcal{P}(\mathcal{N}) \to \mathbb{R}$, which is split among all the participants to the coalition. So, denoting with x_i the individual payoff of $i \in S$, the following holds:

$$\sum_{i \in S} x_i = v(S)$$

This is the case of **transferrable utilities**. A different kind of coalitional games involve *indivisible* values $v(S)$, and are said to be **non-transferrable** utility games.

In **simple** coalitional games the value $v(S)$ of any coalition S is either 1 or 0: some coalitions win, the others lose. A **veto player** is a player that must belong to a coalition for it to be winning.
The set $\mathcal{N}$ is also called the **grand coalition**, and an interesting question is whether all players will cooperate, i.e. if they form a grand coalition.
Superadditive games are a famous type of coalitional games where *bigger* coalitions have a value that is *higher* than that of the sub-coalitions forming it:

$$v(S_1 \cup S_2) \geq v(S_1) + v(S_2) \qquad \forall S_1, S_2 \subseteq \mathcal{N},\ S_1 \cap S_2 = \varnothing$$

In this case cooperation is always beneficial, and the grand coalition is guaranteed to form.
An interesting question is then whether the grand coalition is **stable** over time, i.e. if the share of each player suffices to hold them inside the coalition. This can be explored through the notion of *core*.

The **core** is the set of allocations $(x_1, x_2, \dots, x_N)$ such that:

$$\sum_{i \in \mathcal{N}} x_i = V(\mathcal{N}) \qquad \sum_{i \in S} x_i \geq v(S) \quad \forall S \subseteq \mathcal{N}$$

In other words, the core consists of the ways to split the value of the grand coalition ($v(\mathcal{N})$) between all players so that nobody has incentive to form different coalitions, because this would *decrease* their share x_i. The core is non-empty if and only if the grand coalition is stable.
Often it is empty, and when it is not it holds *many* elements, and there is no way to choose which one is "better". Many allocations, in fact, can be *unfair*.

One way to find if the core is non-empty is by examining if a game is **balanced**. Suppose that players can change coalitions over time, and let $a(S)$ be the fraction of time that player $i \in S$ should spend inside S. It must hold:

$$\sum_{S \colon i \in S} a(S) = 1 \qquad \forall i \in \mathcal{N}$$

since i will spend 100% of the time *in some coalition*.

The game is **balanced** if:

$$\sum_{S \in \mathcal{P}(\mathcal{N}) \setminus \varnothing} a(S) v(S) \leq v(\mathcal{N})$$

that is if the average value obtained by sharing time over *any combination* of coalitions S is less than the value obtained by staying always in the grand coalition.
A theorem by Bondareva-Shapley shows that a game has a non-empty core (and thus a stable grand coalition) if and only if it is balanced.

Another way of ensuring a non-empty core is to examine **convexity**. For any $S_1, S_2 \subseteq \mathcal{N}$, the game is **convex** if:

$$v(S_1 \cup S_2) + v(S_1 \cap S_2) \geq v(S_1) + v(S_2)$$

A convex game has a non-empty core. This can be used to prove that a simple coalitional game with a veto player has a non-empty core, because everyone wants to be with the veto player.

In engineering, one can simply try to see if some specific allocations (e.g. fair shares) lie inside the core. Finding even a single example proves that the core is non empty, and so that the grand coalition is stable.

4.5.2 Nash bargaining

Nash bargaining, contrary to the name, is a non-game theoretic approach to solve bargaining problems. The game-theoretic approach is *dynamic bargaining*, which was seen in sec. 2.7.

Nash bargaining is a purely static approach, based on a set of axioms specifying how a share is *fair*.

As an example, consider two players: (R)ich and (P)oor. They share 1 unit, and keep their share only if both agree on the division. The idea is that for R a single unit is an insignificant fraction of his/her fortune, while for P even receiving a fraction of 1 would be a very good outcome. So R can *easily* refuse a share, getting nothing without any real penalty, while P has a high incentive of accepting even *unfair* shares, as long as he/she gets *something*.

First, we define the **utility region** S, i.e. the compact and convex set of all joint utilities for the players. For a 2 player game, this would be a plane, containing all points (u_1, u_2) of possible shares. The **disagreement** point $\boldsymbol{d}$ is the minimum outcome achievable in the game. In the above example, it is $(0, 0)$, reached by players *refusing* the share.

Now, the idea is to define a set of axioms that, given a pair $(S, \boldsymbol{d})$, leads to find the game's *solution*.

1. **Feasibility**. Clearly the solution $\boldsymbol{s}^*$ must be in S, and in particular: $(s_1^*, s_2^*) \geq (d_1, d_2)$.

2. **Pareto efficiency**: nobody can improve their share without lowering that of others.

3. **Symmetry**: if $d_1 = d_2$ and S is symmetric, then $s_1^* = s_2^*$. This is not the case for the above example, but there will be a way to *adapt* it nonetheless.

4. **Invariance to linear transformation of utilities**

5. **Independence of Irrelevant Alternatives** (IIA). If $U \subset S$ is a compact and convex set, and $\boldsymbol{s}^*(S, d) \in U$, then $\boldsymbol{s}^*(S, d) = \boldsymbol{s}^*(U, d)$, that is *adding* more "irrelevant" options does not change the solution. For example, giving more options with a better u for 1 does not change the solution.

Nash showed that the only $\boldsymbol{s}^*$ satisfying all the above is:

$$\boldsymbol{s}^* = (s_1^*, s_2^*) = \max_{(s_1, s_2) \in S} (s_1 - d_1)(s_2 - d_2)$$

In particular, if $d_1 = d_2 = 0$, this is the same as maximizing the *product* of utilities:

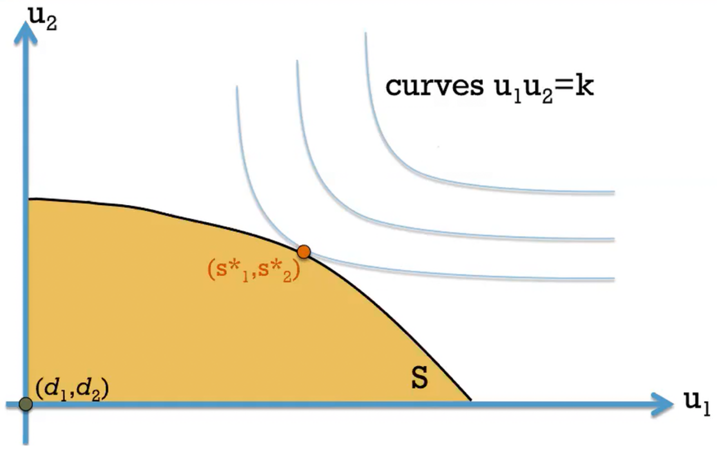

Figure (4.6) – Graphical representation of a Nash Bargaining Solution in the case of asymmetric S. Note that u_1 has a higher range of acceptable shares d_1, and so will receive more at equilibrium ($s_1^* > s_2^*$). In other words, 1 is (R)ich and has *more* bargaining power. Note that the solution is unique only if S is convex (as required).

This can be generalized (as in the figure) to cases where S is not symmetric, leading to the *generalized* Nash Bargaining Solution (NBS). The idea is to apply *linear transformations* to make S symmetric.

NBS is not the only possible approach for solving static bargaining problems. For example, Kalai and Smorodisnky (1974) suggests replacing the IIA axiom with an axiom of **individual monotonicity**. the idea is that if one player has *more options*, this should matter more at equilibrium. The unique bargaining equilibrium is then found to be:

$$(s_1^*, s_2^*) = \max_{\lambda:\ (s_1,s_2)\in S} \{(d_1, d_2) + \lambda[(i_1, i_2) - (d_1, d_2)]\}$$

where (i_1, i_2) is the ideal point of max utilities:

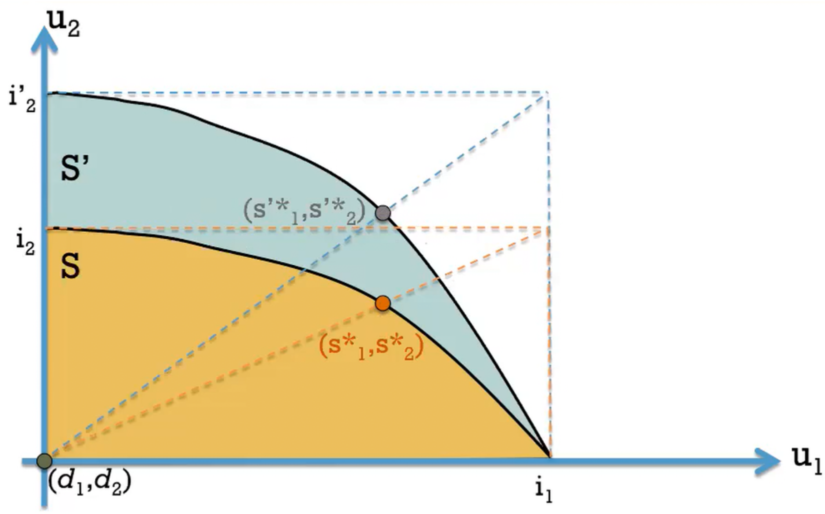

Figure (4.7) – Graphical representation of the Nash Bargaining Solution with the Kalai and Smorodinsky approach (KSBS).

In all cases the solution lies on the *black line* at the edge of S, representing the Pareto efficient solutions.

One advantage of KSBS is that it can be extended also to games with non-convex S. However, in general, the choice between using NBS or KSBS is philosophical: mathematically they are both “equivalently good” solutions.

4.5.3 Bayesian cooperative games

(Lesson 24 of 22/12/2020) Compiled: January 1, 2021

Consider two students participating in a joint project for a course exam. They can either put (E)ffort, or (S)hirk. The payoff of each student is the difference between the project's *value* and the *cost* of doing it. The cost is a fixed constant c, equal for both students, and is paid only by who is putting effort. However, the two assign different values to the project, which are modelled as the students' types. Specifically, t_i is uniform in $[0, 1]$, and this prior is common knowledge. Then, the value for player i is t_i^2 if the project is successful, which happens if *at least* one student plays E.

The payoff matrix is given by:

$$\begin{array}{cc} & \text{student B} \\ & \begin{array}{cc} E & S \end{array} \\ \text{student A} \begin{array}{c} E \\ S \end{array} & \left\| \begin{array}{cc} t_A^2 - c, t_B^2 - c & t_A^2 - c, t_B^2 \\ t_A^2, t_B^2 - c & 0, 0 \end{array} \right\| \end{array} \tag{4.4}$$

Note that this is **not** the game's normal form, which would include an *infinite* number of strategies, since type space is continuous.

Fortunately, we can simplify the description by noting the following:

- If i plays E, her payoff is $t_i^2 - c$, regardless of what $-i$ does.
- If i plays S, her payoff is $\neq 0$ only if $-i$ chooses E. The expected payoff is then $t_i^2 \mathbb{P}[s_{-i}(t_{-i}) = E]$.

So, i prefers E over S if:

$$t_i^2 - c \geq t_i^2 \mathbb{P}[s_j(t_j) = E]$$

Solving for t_i leads to:

$$t_i \geq \sqrt{\frac{c}{1 - \mathbb{P}[s_j(t_j) = E]}}$$

This is a threshold behavior: i plays E only if her type is *beyond* a threshold, i.e. if she sufficiently values the project.

For comparison, let's imagine that i is playing *alone*. In this case, there is no other player that can “save” i: if i plays S, she will get 0. Still, it is convenient to play E if the cost can be justified:

$$t_i^2 - c \geq 0 \Rightarrow t_i \geq \sqrt{c}$$

which is again a threshold behavior. Note that in the two-player case the threshold is *increased* by the presence of the other player.

So, we can focus just on the thresholds q_i, since the only viable strategies are the ones in which $t_i > q_i$ leads to playing E, and $t_i \leq q_i$ leads to S.

A **belief** held by player i is also a value q_j for the threshold that i believes that j will use. So, if i *believes* q_j, this means that $\mathbb{P}[s_j(t_j) = E] = 1 - q_j$, and i plays E if and only if $t_i > \sqrt{c/q_j}$.
To have a NE, this must also be the best response played by i, and the converse must be true for j, so that:

$$\sqrt{c/q_j} = q_i$$

And for symmetry reasons:

$$\sqrt{c/q_i} = q_j$$

which leads to $q_i = q_j = c^{1/3} > \sqrt{c}$ since $c < 1$. So, at the NE the threshold to play E is *bigger* for both players than the *alone* case.

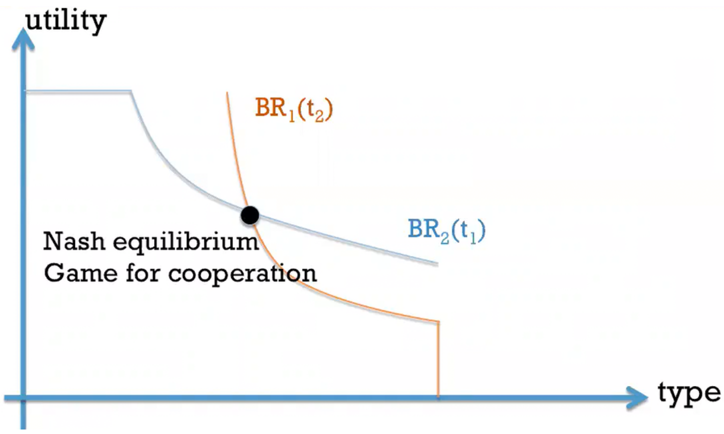

Figure (4.8)

4.5.4 Cheap talk

An *extreme signaling* game is one in which Nature chooses a state of the world and discloses it only to player 1, who can then communicate it to player 2. Finally, 2 takes action, and the game ends, with no direct payback for player 1.

For example, consider the following. 1 lives in city B, 2 is a friend of 1 and has found a job in city A. However, since commuting takes time, 2 wants to move to a different place. She has 5 options: 1, go to city A, $2-4$, move in some city in between A and B, or 5, stay at B. She does not know which city is better, and she asks 1 his/her suggestion. However, 1 is *biased*, because prefers 2 living close.

The utility of 2 depends on the *type* of the chosen city, which can be $t \in T = \{1, 3, 5\}$, and is given by:

$$u_2(a_2, t) = 5 - (t - a_2)^2$$

However, only 1 *knows* t, but he is biased towards 5:

$$u_1(a_2, t) = 5 - (t + b - a_2)^2$$

where b denotes the bias.

Let the prior distribution of types be uniform ($\mathbb{P}[1] = \mathbb{P}[3] = \mathbb{P}[5] = 1/3$), and $b = 1.1$.

We avoid discussing a full solution of this game, and only highlight the interesting results:

- There is no PBE in which 1 reports the true state of the world: 1 has incentive (bias) to *lie*.
- There is a "babbling equilibrium" corresponding to a pooling equilibrium in which 1 always sends the same message, and 2 ignores it and just uses the prior.
- There is a PBE where 1 *partially* reports the true state of the world. Specifically, he is truthtelling if $t = 1$, but pools information of $t = 3$ and $t = 5$. However, this happens only if the bias is sufficiently low.

4.5.5 Utility transformations

When setting the payoffs (utilities), most of the time we do not have a precise definition, but just an *intuitive* notion of "preference" which is often difficult to quantify. We are interested to see if NEs are *stable* with respect to utilities. That is, if we consider a game $G = (S_1, \dots, S_n; u_1, \dots, u_n)$ with certain NEs, what can be said about a *different* game $G' = (S_1, \dots, S_n; u_1', \dots, u_n')$ where utilities are assigned in a *slightly* different manner?

Suppose u_i and u_i' differ through some affine transformation with positive a_i:

$$u_i' = a_i u_i + b_i \qquad a_i \in \mathbb{R}^+,\ b_i \in \mathbb{R}$$

then G and G' have the same best responses, dominant strategies, etc. Thus they have the *same* set of NEs.

The same holds if a constant $k \in \mathbb{R}$ is added to all utility values of player i that share the *same* opponent strategy $\boldsymbol{s}_{-i}^*$:

$$\begin{aligned} u_i'(s_i, \boldsymbol{s}_{-i}^*) &= u_i(s_i, \boldsymbol{s}_{-i}^*) + k \\ u_i'(s_i, \boldsymbol{s}_{-i}) &= u_i(s_i, \boldsymbol{s}_{-i}) \qquad \boldsymbol{s}_{-i} \neq \boldsymbol{s}_{-i}^* \end{aligned}$$

Let's see this in practice. Consider an entry game with normal form:

$$\mathbf{P} = \begin{array}{cc} & \text{Pri} \\ & \begin{array}{cc} E & O \end{array} \\ \text{Sec} \begin{array}{c} E \\ O \end{array} & \left\| \begin{array}{cc} 0,0 & 2,2 \\ 1,4 & 1,4 \end{array} \right\| \end{array} \tag{4.5}$$

and extensive form:

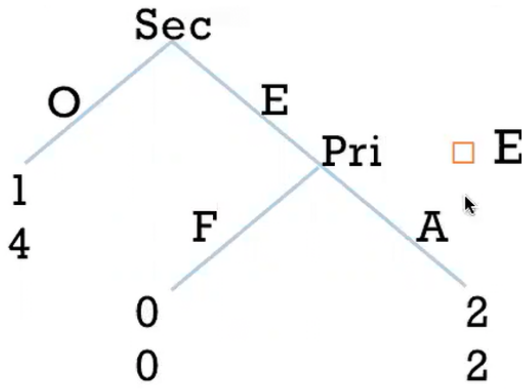

Figure (4.9)

There are two NEs:

- (E, A), which is also an SPE
- (O, F), including a *non-credible* threat.

The same results hold if we subtract 1 to the payoff of player 1 for *the same action* (A) of the second player:

$$\mathbf{P} = \begin{array}{cc} & \text{Pri} \\ & \begin{array}{cc} E & O \end{array} \\ \text{Sec} \begin{array}{c} E \\ O \end{array} & \left\| \begin{array}{cc} 0,0 & {\color{red}1},2 \\ 1,4 & {\color{red}0},4 \end{array} \right\| \end{array} \tag{4.6}$$

Likewise, we can subtract 4 from player 2's entry of row O:

$$\mathbf{P} = \begin{array}{cc} & \text{Pri} \\ & \begin{array}{cc} E & O \end{array} \\ \text{Sec} \begin{array}{c} E \\ O \end{array} & \left\| \begin{array}{cc} 0,0 & 1,2 \\ 1,{\color{red}0} & 0,{\color{red}0} \end{array} \right\| \end{array} \tag{4.7}$$

And finally we can apply an affine transformation, dividing all payoffs of Pri by 2:

$$\mathbf{P} = \begin{array}{cc} & \text{Pri} \\ & \begin{array}{cc} E & O \end{array} \\ \text{Sec} \begin{array}{c} E \\ O \end{array} & \left\| \begin{array}{cc} 0,0 & {\color{red}1},{\color{red}1} \\ {\color{red}1},0 & 0,0 \end{array} \right\| \end{array} \tag{4.8}$$

While the NEs are the same as before, now it is more difficult to get an *intuition* about the game.

However, these transformations can be used to fully *classify* **symmetric two-players** games, i.e. games $G = \{S_1, S_2; u_1, u_2\}$ with $S_1 = S_2$ and $u_1({\color{blue}a}, {\color{red}b}) =$

$u_2(b, a)$. In this case, the normal form is as follows:

$$\mathbf{P} = \begin{array}{cc|c} & & \text{Pri} \\ & & \begin{array}{cc} 1 & 2 \end{array} \\ \text{Sec} & \begin{array}{c} 1 \\ 2 \end{array} & \left\| \begin{array}{cc} a, a & b, c \\ c, b & d, d \end{array} \right\| \end{array} \tag{4.9}$$

Note that we can report only 1's payoffs, since 2's payoffs can be immediately deduced from them:

$$\mathbf{P} = \begin{array}{cc|c} & & \text{Pri} \\ & & \begin{array}{cc} 1 & 2 \end{array} \\ \text{Sec} & \begin{array}{c} 1 \\ 2 \end{array} & \left\| \begin{array}{cc} a & b \\ c & d \end{array} \right\| \end{array} \tag{4.10}$$

If $b = c$, the game is said to be **doubly symmetric**. Any symmetric game can be reduced in such form by applying the above transformations:

$$\mathbf{P} = \begin{array}{cc|c} & & \text{Pri} \\ & & \begin{array}{cc} 1 & 2 \end{array} \\ \text{Sec} & \begin{array}{c} 1 \\ 2 \end{array} & \left\| \begin{array}{cc} a - c & 0 \\ 0 & d - b \end{array} \right\| \end{array} \tag{4.11}$$

Then, the signs of $a - c$ and $d - b$ fully determine the kind of NEs. There are 4 cases:

- $a - c < 0$, but $d - b > 0$. Then 2 is strictly dominant. The game can be solved by IESDS, and the game is analogous to the Prisoner's dilemma.
- $a - c > 0$ and $d - b > 0$. This is a coordination game (such as the Battle of Sexes), with two pure NEs, and one mixed NE.
- $a - c < 0$ and $d - b < 0$: same as before, but this time it is an *anti*-coordination game (Hawk-Dove game). There are again two pure NEs, and one mixed.
- $a - c > 0$ and $d - b < 0$. This is the same as the first case, but with the *other* strategy as dominant.

A *discoordination* game is not symmetric, and so it does not appear in the above cases.