MANUALI

UMANISTICA

– 14 –

Riccardo Bruni

Course notes on finite games and rational choice

Firenze University Press
2018

Course notes on finite games and rational choice / Riccardo Bruni. – Firenze : Firenze University Press, 2018.
(Manuali – Umanistica ; 14)

http://digital.casalini.it/9788864538310

ISBN 978-88-6453-832-7 (print)
ISBN 978-88-6453-831-0 (online)

This book is printed on acid-free paper

Università degli Studi di Firenze
Firenze University Press
via Cittadella, 7, 50144 Firenze, Italy
www.fupress.com

Printed in Italy

Table of Contents

Riccardo Bruni, *Course notes on finite games and rational choice*, ISBN 978-88-6453-832-7 (print)
ISBN 978-88-6453-831-0 (online), CC BY 4.0, 2018 Firenze University Press

Foreword

1.1. Motivations and content of the volume

As the title suggests, this book collects notes that were prepared for a university course I taught in the Spring of 2018, with a slightly different title from the one I chose for this book, *Probability and Rational Choice*, and delivered to an audience of students enrolled in the Master course on *Logic, Philosophy and History of Science* (*Laurea Magistrale in Logica, Filosofia e Storia della Scienza*), held at the *Dipartimento di Lettere e Filosofia* of the University of Florence. The goal of the course was to introduce students to some basic concepts from the area of research generally known as decision theory. Due to the vastity of it, it was necessary to sharpen the aim of the course by carefully selecting some significant topics to expose the students to. I have made my choice in this respect as is usual the case for situations like this one, namely by taking into account factors, like expertise, taste, opportunity, etc., that could make the selection optimal for the course, and for the Master programme objectives, or, I should better say, for my understanding of them.

The Master programme this course was taught for the sake of, is an interdisciplinary curriculum which aims at teaching students coming from different bachelor studies, either settled within the Humanities, or linked to scientific disciplines. Course topics vary, but are in general somehow related with actual researches centered on Logic or on the Philosophy and the History of Science. This is a unique thing among the Master programmes available at the Academic level in Italy. A substantial part of this objective is achieved by introducing students to the use of formal methods, in agreement with a tradition that has made the philosophical studies in Florence renowned in, and outside Italy. The course I have taught, and this book which, as I said, comes out from that teaching experience, was intended to serve this purpose.

Riccardo Bruni, *Course notes on finite games and rational choice*, ISBN 978-88-6453-832-7 (print)
ISBN 978-88-6453-831-0 (online), CC BY 4.0, 2018 Firenze University Press

The concept which I decided to put under the spotlight is the concept of « rational choice». I also decided that the methodological tool that could be mostly useful to deal with it was the theory of games. This is because this latter area of research has rapidly become a territory where different scholars, mathematicians, economists, philosophers, meet for discussing topics from multiple points of view. Therefore, I thought that sticking to games could best reflect the educational aim the course was supposed to have. In particular, since I was trying to minimize prerequisites needed to attend the course (see below), I decided to limit myself to the treatment of *finite* games.

Regarding the selection of topics in that direction of work, I doubt that the expert reader would consider it an original one: it is rather exactly what one may expect an introductory volume to decision theory centered around the tool of finite games to contain.

Chapter 1 is my attempt to motivate the investigation on choice – on rational choice to be more precise –, from the game-theoretic angle I will adopt and propose in the subsequent parts of the volume.

With chapter 2 I start to elaborate the main dichotomy upon which the book is built, namely as based on distinguishing the analysis of games where the order in which players play is not taken into account, which yields the theory of games *in normal form* as they are commonly referred to, from the approach to games where that feature is present, that is to games as seen in *extensive form*. Beside laying down some basic terminology, this part of the book aims at making clear the implications in 'geometrical' terms involved in taking the one route, and in taking the other instead.

Having making clear the two areas of knowledge we will be interested in, I present in chapter 3 the theory of finite games in normal form (or, I should rather say, the part of it I decided to select for the volume's sake). Once again, I must stress that nothing unexpected happens here, topic-wise: after discussing the notion of Nash equilibrium as a natural generalization of what could be seen as representing a 'solution' to a game in normal form (that is, a solution of the choice problem that can be formulated for it, namely: What action should each player choose?), I go on questioning the natural character of this concept by discussing situations, in the form of games, in which by sticking to equilibria one seems to deviate from paths that rational agents might be expected to prefer. Refinements of the original notion of equilibrium that stem from the relevant critiques to it, at least some of those that have been proposed in the literature, are presented along the way. If this, as I said, was more or less expected, what is unexpected maybe, hence counts as an original contribution that this monograph put forth in this respect, is the approach by means of which the basic properties of equilibria and their refined versions are attained at, and proved to hold. This is done by making use of formal methods which are proper to the logical investigation and are based on the setting-up of a formal language to speak of actions, utilities associated to them and their comparison. A definition of «rational strategy» is then introduced on this basis, and a study of it is pursued by means which are again standard in areas of logical research where circular concepts are relevant (like, for instance, the theory of formal truth). The whole story is told with some more bibliographical details in section 3.12.

Chapter 4, the last one of the book, contains the treatment of games in extensive form. Here, the diagrammatic form of trees by means of which games of this type are sometimes accounted for in the literature, is abandoned in favour of the mathematical model of games as sequences of natural numbers. After providing the reader with the concepts needed to familiarize with the approach, as well as a discussion about alternatives to the model that we call «canonical» and which come out from reflecting upon the common experience with games in real life, it follows a presentation of (what I chose to be) the main result on games in this form, namely the theorem of determinacy. The proof idea bears similarities with the aforementioned methodology used to deal with games in normal form. In particular, a formal language is introduced in order to speak of games in the chosen form and show that determinacy turns out as a consequence of some validity properties of formulas expressing that a strategy for winning every match of the game under scrutiny exists for one of the two players.

Presuming that by what I have said so far I succeded in providing the reader with a credible answer to the question «What do I get back if I decide to pay the price needed to go through the content of this book?», the other information required to make the decision, and eventually paying the actual price needed to buy the volume, is «What does the reading of this book presume I know already?». Being this latter question as important as the former, let us spend a few words about it as we did for the sake of answering the other one.

1.2. Notation and prerequisites

Having said that the book is supposed to be an introduction to the topics it treats, the best answer to the question on prerequisites is: nothing, no previous knowledge is presupposed here. To say that would turn out to be a lie. It would have to be regarded as a small lie probably, but still a lie in the end. Since to set up a relationship on lies is never a good idea, I will try not to lie to the reader, with whom I am suppose to engage a relationship henceforth. So, this book does indeed require some pre-existing knowledge. I have tried to avoid any assumptions about the reader knowing something about finite games and being acquainted with the notation at use in that area of study. Since the choice of sticking to games as a tool for approaching the theory of rational decision was a deliberate choice of mine, and certainly a different route could have been chosen instead, I did not want to make any presupposition regarding the reader's expertise in the area. So, the volume aims at being a genuine introduction to the theory of finite games for beginners.

As the approach that is fostered here to game-theoretic topics is partially novel, prerequisites have to be evaluated with respect to that too. Both the treatement of finite games in normal form, for what concerns the theory of equilibria and their refinements, as well as the theory of games in extensive forms as sequences, in particular the proof of the theorem of determinacy from section 4.8, are based on the use of formal methods that are proper to the logical investigation.

Although I have tried both to motivate the methodological choice and to provide the reader with the necessary information step by step, I could not avoid assuming familiarity with the way in which formal languages are dealt with notationally in the first place, and how the basic notions involved in the costruction of them, such as the concepts of terms and formulas, are commonly introduced. So, the reader who has already been exposed to the standard notation concerning predicates application, symbols chosen for the main logical operations and the like, will experience no problem in going through the sections from chapter 3 and chapter 4 where they are at use. In addition, familiarity with inductive definitions of terms and formulas, as well as with proofs by induction, might be presupposed by some of the passages the reader will be required to go through. Parts of chapter 3 (in particular, section 3.5) and chapter 4 may also require familiarity with basic mathematical notions which are only partially accounted for here. Some very basic knowledge of set theory, both notationally and conceptually speaking, and a very basic acquaintance with the mathematical notion of function could ease the reading of those parts of the volume.

That is all the reader should be warned of, I think. I have decided to publish these notes both to serve an 'internal' purpose (that is, to provide future students of other courses I may deliver on the topic with a manual they could use as a companion to actual class notes), as well as an 'outer' one, namely to share them with scholars, make them available to students and, why not, teachers of other courses that may be in search of a compact manual on these topics. I hope that anyone who will happen to stumble on this book will find it a useful tool for reading, researching, studying and teaching.

Before launching ourselves into the actual reading of the volume, I would like to save a few words for acknowledgements. As a matter of fact, there is quite a number of people I should express my gratitude to as far as the preparation of this material for publication is concerned. First of all,

to the family at large at home. Secondly, to my collegues of the research group in Logic and Philosophy of Science, and to Andrea Cantini in particular, for several discussions concerning topics treated here that finally brought me to develop the interest needed to conceive the writing of this book. To the students, who attended my course in the Spring semester of 2018 for being the first ones to test these notes and for reporting me some typos they originally contained. Finally, to the Dipartimento di Lettere e Filosofia as a whole, that granted me a generous support for making this publication possible. It is intended that I remain the only responsible for mistakes and inaccuracies that might still be found here.

Firenze, 17 September 2018

Chapter 1
Preliminary considerations

This is a a book on choice. Since choices occur in many forms and affect most, if not every aspect of our lives, the first thing that a book on choice should do is to sharpen the object of investigation. We are not dealing with choices whatsoever in the first place. Our aim is to focus on *rational choices*. So, choices we are wishing to analyze are those where what they say is the character most typical to human beings is excercised. Is this enough to clarify what kind of choices are we interested in? If so, how can the character of rational choices be specified better? Is this of any help to set up a systematic investigation of it and what kind of knowledge can we expect that this treatment will allow us to achieve? These are the kind of issues we will confront ourselves in this introductory chapter.

1.1. Types of choice

As I said already, choices are everywhere. As a result of their ubiquitary character, choices come in many different types. Choices we aim at dealing with are the rational ones. What does this mean? To make the idea more precise we could tentatively attempt a general definition like the following:

Definition 1.1 *A choice is* rational *if it can be rationally argued for, or if rational arguments can be produced against it.*

The idea of the definition is that choices are either recognized to be rational because you can rationally explain why they were made, or because you can oppose them rationally. Choices are rational if they have a rational 'content' that can be argued, for or against, by rational means. I will not even attempting at dragging the consideration of this definition

Riccardo Bruni, *Course notes on finite games and rational choice*, ISBN 978-88-6453-832-7 (print)
ISBN 978-88-6453-831-0 (online), CC BY 4.0, 2018 Firenze University Press

even longer than that, and admit straight away that the definition is a weak one. It is weak because it fails to sharpen the object of our analysis as we wanted. To speak of reasons supporting, or allowing to argue against a choice that has been made is only apparently helping us in this respect: to set up the clock at 7:00 a.m. is not really the kind of choice one is expected to pick to speak of rationality at first; however, if it turns out that I have a meeting at 8:30, and that I have to travel the distance to the meeting point in an average situation of morning traffic, this might be enough to argue for the choice I have made and make it rational in the sense of the definition. This brief observation is meant to show that it is easy to make any choice appear as if it fits the definition above. In other words, to tighten «choices» and «rationality» together as we did it makes the former notion depending upon «reasons» and «argumentations» the latter notion is also entangled with. However, the latter concepts have not a unique meaning and are subject to a variety of interpretations that suggest that they can hardly serve the purpose of making any easier the dispute about what the rational character of choices is, or is not.

Now, it is clear that the definition above could be revised and refined to make it harder to defy, but the effort would be hardly rewarding, I feel, as it is improbable that one can find a fully satisfactory version of it at this level of generality. So, I will refrain from doing this, also because it seems to me that things get really easier, only if one tries to investigate about this matter in concrete and specific situations.

Take, for instance, the following story about Alice:

> Alice is looking for a part-time job that could ensure her to earn some money while she is finishing her studies at the University. She then finds this company, which delivers support to clients for social networking and advertises available positions for perspective employees who are expected to increase the number of «likes» on websites of clients over their competitors. Alice decides to accept the offer, and, according to the contract she signs, she will be paid more, the more «likes» she is able to place.

Assuming that no legal, moral, or other sorts of infringements are at risk of occurring while Alice is performing her activity, if someone asked the question, «What is the rational choice for Alice to make?», I think that no doubt would be cast on what is the answer to such question despite the previously noted difficulty with the notion it involves. Alice is indeed expected to place as many «likes» as she is able to do during worktime, in order to increase her salary.

This suggests that questions about rational choices are easier to answer if they are raised with respect to suitably restricted settings. Yes, but is not the setting we have chosen *too* restrictive and simple? It is. However, we are trying to lay down some basic, simple remarks to start from, and there is nothing wrong to start from simple situations to do that. So, let us go on with the story, since things complicate a little bit afterwards:

> Since Alice's company is expanding, she is provided with a teammate, a new employee named Bob, and the clauses of her contract are changed accordingly: now, she will be paid more in all situations in which her choice of action agrees with Bob's, to the effect that they both like or dislike a website, while they will be paid less whenever their actions disagree. The same applies to Bob's contract.

The new situation requires the previous conclusion about Alice's way of working to be adjusted to the new conditions of Alice's, and Bob's job. From the viewpoint of Alice (the viewpoint of Bob being symmetrical), the adjustment looks easy as she is expected to place a «like» whenever Bob does, and place «dislike» again when Bob is doing the same. Simple as it may appear, it is true indeed that the situation became more complex with respect to the one we considered beforehand. By comparing them, there are some general features that is worth extracting to introduce and explain the subsequent development of these notes.

The very first issue that clearly marks the difference between the first half of the story and the second, is the number of characters involved. There is just one character at first, they are two in the subsequent part. This shift marks a first sense in which the concept of choice falls into types: some choices are done by and for ourselves, some others are done by ourselves but as part of a group of members who equally contribute to the choice we make, and have consequences for the whole group we are part of. In other words, some choices are the expression of our individuality, and are then *individual choices*, some other choices are the expression of our being part of a community, of a society, of a group of people and are named *social choices* accordingly. Examples of individual rational choices are too many to mention, since, as I was suggesting earlier, a suitable setting can make rationality appear almost everywhere. «Social choice» is similarly a concept that can be used in relation to many situations, some of the which mark important occasions that require a decision on our part collectively, like voting, investing, trading, fighting, negotiating, and so on, but may be also involved in more mundane cases which require mutual collaboration between agents.

One second feature that is worth noticing about the story of Alice and Bob comes out by reflecting some more about the kind of situation that frames it. For, it is clear that the contract conditions at which they are employed favour their collaboration: in the social schema I have just proposed, they form a coalition which, is easy to imagine, acts against other coalitions made out by employees of concurrent companies. However, it is also easy to imagine a completely different setting in which Alice and Bob are brought to avoid collaboration, since acting one against the other is more fruitful for them. Real life is alas an unlimited source of insipration for situations of this sort also. This suggests a sense in which social choices too fall into types: the type of choices made in cooperation, *cooperative social choices*, and that of choices which are *non-cooperative* instead.

Individual versus social, cooperative vs. non-cooperative are the first fundamental types of choices that can give rise to some more types. This induces enough dynamic to the setting we are laying down to suggest the potential of the systematic investigation we are aiming at pursuing over it, where situations, even simple situations like those I have started from, exemplify features that are shared by all other sorts of situations that fall into the same type, and therefore engender the kind of general reasoning about them we would like to achieve. It should be clear, however, that these are just the first building blocks of the more complex structure we wish to arrange in the end. Before going on with this, I will not leave unnoticed that there is a third character that stems from the situation I have considered here, that the reader might have indentified and that I am consciously avoiding to put the emphasis on at this stage of the analysis. I am thinking in particular of a character that can be associated with what is rational for Alice and Bob to do in the story, and that candidates itself to provide us with the very paradigm of rationality that we could use as a working hypothesis for our subsequent development. I feel it would be premature to unveal it here. Therefore, I will leave the analysis of it to my future considerations.

1.2. A decision and its components

In the previously designed setting, Alice and Bob are bound to take decisions. Once again, even though the story is a simple one, it is enough to emphasize characters that are proper to decisions in general, therefore that are universally shared by them whichever is the level of complexity of the setting they belong to (this is, at least, the conviction that we are relying to as a working hypothesis). To make up their minds for Alice and Bob means to make a choice among the actions they have at their disposal. In the chosen situation, this means that they have either to decide to 'play' *like*, or *dislike* instead. What seems to be the driving force of the choice they have to make is the consequences of actions. The story is set in such a way that to play either *like* or *dislike* yields different consequences for their lives. These consequences, in turn, affect the kind of expectation they might have: so, to play *like* in the right situation for Alice means to have wealthier life than playing *dislike* instead. The expectations Alice and Bob might have with respect to decisions they make, induce preferences over actions they choose among. In other words, decisions can be viewed as tools for turning actions into consequences they are assumed to have on the basis of an agent's preference.

These initial observations can serve the purpose of testing the goal we would like to achieve. As a matter of fact, the ambition we nurture is to be able to perform a study of choices that be rigorous and systematical. Can we? The observations we have just made, for instance, gives some hint in this respect. They may suggest indeed to represent decisions by making use of the mathematical notion of *function*. As a matter of fact, functions as they are used in mathematics are suitable to situations in which elements of a given collection, the *domain* of the function, are taken as argument to which a 'process' (in the form of an operation, or a series of operations) is applied and yields a unique result that belongs to a possibly different collection of elements, the *co-domain* of the function. Functions are then described either *intentionally*, when the process they rely upon is explicitly given, or *extensionally* as a collection of correspondences of the form argument-value. Functions in the first form have an algorithmic nature: they essentially look like a set of instructions that one needs to execute for a given argument as input to 'calculate' the result as value. Functions in the second form, instead, appear as a set (that is commonly named the *graph* of a function), a collection of ordered pairs of the form $(a, f(a))$ (i.e., a pair of elements taken in a certain order where one element counts as the first element of the pair, the other as the second one[1]), whose first element a is the argument of the function f and $f(a)$ is the value of it. To express the fact that a function with domain D takes an element a of this set as input and returns a unique element $f(a)$ of its co-domain C, it is common to use the notation $f : D \to C$.

Whatever form one chooses to present a function f, whether it is the intensional or the extensional one, this will turn out to have the following properties:

- to be defined over all elements of its domain D, in the sense of yielding a value in its co-domain C for each of them (so, for each element a belonging to domain D, there exists an element b in the co-domain C such that $f(a) = b$));
- to be right-hand unique, as it is usually said, that is to yield a *unique* value for one and the same argument (hence, for every element a and b in D, if $a = b$, then $f(a) = f(b)$ – which is the same as assuming that if values of f are different, then they must be produced out of arguments which are also different).

As far as the treatment of choices is concerned, a decision for a character like Alice, call it d_a, would then be a function from the set of her actions A to the set C of the consequences they have. The notation is

[1]Ordered pairs have a distinguished equality relation which is defined according to their ordered 'nature'. Therefore, for any given two ordered pairs (a, b) and (a', b'), it is the case that they are equal, that is that $(a, b) = (a', b')$ holds, if and only if $a = a'$ and $b = b'$. That is, ordered pairs are equal if they have the same elements in the same order.

chosen in such a way that it can be adapted to indicate decisions made by other characters, as it would be turned into d_b to indicate Bob's decision for instance. To see how satisfactory the rendition of choices as functions can be, another aspect of the situation we have been considering should be taken into account. In particular, the part of it in which decisions made by Alice combine with those made by Bob and viceversa.

The story suggests that the said interaction is a key feature. With respect to what we have noted already, this amounts to let consequences of actions being dependent upon not just the action decided by one character, but on decisions made by all characters involved in the situation. In the story by Alice and Bob, for instance, the outcomes referred to, i.e. the income they get, depends not just on Alice choosing *like* or *dislike*, but on Bob doing the same in coordination. This means that the consequence Alice aims at achieving by means of her decision depends upon the action of hers she chooses *and* the action correspondingly chosen by Bob. Does this affect our model of choices via functions? It does not seem so, as mathematical functions are tools flexible enough to accommodate for that. For instance, because functions representing choices made by single characters can be composed to simulate the combination of choices which turn out to yield a consequence which, affecting each member of a group of people, also affects the group as a whole.

To develop the mathematical details of this observation here would be like putting the cart before the horse. We are for the moment willing to flag some useful observations that could also strengthen our confidence that to take choices as the object of a rigorous study is indeed possible. So, I will just leave the remark there, also because I feel there is something more pressing that should be clarified about the schema we have been discussing.

In the representation of choices we are fostering, actions are chosen in view of their consequences and one action is preferred over the others *because* the consequence of it is preferred by the agent that performs the said choice. What made the situation simple to analyze in the case of the story I have used for the sake of argument in the previous section, was precisely that the story itself was making clear what preference an agent could excercise with respect to the consequences of her actions. This was clear since the part of the story involving Alice alone, as the choice she had to make amounted to choose between being wealthy, or being unwealthy instead. The same applies to the second half of the story, where Alice and Bob need to coordinate (and become wealthier), or avoid doing that and lowering their incomes. This means that the analysis we have performed so far presupposes that consequences are arranged in such a way that a preference is attached to each of them. So, if we are using these preliminary considerations as a test that a systematic study of choices is possible, maybe that is the idea that should be tested and argued for first.

That is, that we are indeed capable of similarly conceiving a rigorous way to express such a «system of preferences», as we shall refer to it henceforth, applied by agents over consequences of actions.

1.3. Aligning things on a scale

By a *system of preferences* I generally mean a schema according to which consequences can be compared with one another. By adopting a sympathetic view to the mathematical one we have used to show, or rather to give support to our confidence that a rigorous treatement of «choice» can be pursued, we can view such system as an relation of order over the set of consequences of a character's actions. To express preference for a consequence with respect to another is the same, according to this view, as putting the former before the latter in the list of consequences one is willing to achieve. Like the previous notion of decision as function transforming actions into consequences, this also happens to be a concept that can be accounted for mathematically speaking.

As a matter of fact, for a given set A, a *relation* R over its elements is a set of ordered pairs of them, hence a subset of what is known as the cartesian product $A \times A$ of A with itself, that contains all possible ordered pairs whose elements are elements of A (in symbols: $R \subseteq A \times A$, hence if $b \in R$, then $b = (a, a')$ for some $a, a' \in A$). For any pair (a, a') that belongs to R, it is intended that a is in relation R with a'.

This general notion of relation can be seen to turn out useful in the case we are concerned with. We only need to make it specific to the idea we are wishing to use. This idea concerns putting things in an order and we can imagine two ways of doing this. The first one is determined by the following intuition: to compare the elements of a set with one another in order to determine which one is preferred to the other, is the same as putting them on a line where every element is preferred to those which follow it. Now, if we assume the elements of A to be aligned in this way, then some obvious properties will be verified all along the line. In particular, it will hold true that if the element a of A comes prior to another element of it, say b, and b comes prior to anothe one, c, then a comes prior to c. It is also obvious that for every two elements of A, say a and b, if a comes prior to b, then b cannot come prior to a. This is obvious at least in all cases in which a and b are *different* elements of A: if the line has to be obtained by ordering the elements of A, then once an element of it has been placed it cannot be 'reused' at some later stage, or, in other words, any element can occur only once in the line. This idea is entirely satisfactory if we think of the line of elements of A as reflecting preferences over its elements. For, there is no point in expressing any preference toward an element a of A and a itself. The problem is comparing any two elements of the set, as long as they are different.

However, there is another thing about preferences that seems to clash with this model of alignment of things. As a matter of fact, if all elements of A are supposed to be aligned, and coming prior to another element in the line is taken to mean preference of the one element over the other, how can we express by means of this model the agent's absence of preference in comparing any two things? This is matter of common experience: sometimes we prefer riding the bike to walking, sometimes is the other way around, and sometimes we do not have any preference, or we prefer both things to driving but we are not willing to choose one of the two things in particular. This way of seeing preferences seems absent from the previous model of aligning the elements of A according to an 'exclusive' or *strict* way of ordering them, which is characterized by the previous property about the order being not reversible, hence such that if a is prior to b, then b is not prior to a for every a and b in A.

The other route we have just hinted at instead, which gives rise to an ordering relation that is 'inclusive' or 'large' (usually known as *partial* ordering), limits the said property to elements of A that are different, hence amounts to saying that if a is prior to b and a is not equal to b, then b is not prior to a (which is equivalent to say that if a is prior to b and b is prior to a , then a is equal to b).

The possibility of assuming either of the two modes of thinking of preferences, that is to make a system of preference coincide with ordering the elements of the set of consequences of an agent's action, is reassuring as long as our attempt of studying choice rigorously. This is because, as I said, presuming that such a preference system is at work is one of the building blocks of the analysis of the situations we have performed in the previous section. One point that may puzzle the reader about that is which of the two model should be chosen. Even though very little use of the result below will be made in the subsequent parts of the volume, since, as I shall explain in a moment, for the modest goals pursued here we will rather stick to a different representation of a consequence value (see section 4 below), I feel it is coherent with the experimental character of this introductory chapter to say a few words about this issue. As a matter of fact, the difference between a strict ordering relation and a large one, is the same that passes between the «less than» relation $<$ among, say, the elements of the set $\mathbb{N}$ of natural numbers, and the «less than, or equal to» relation $\leq$ over the same domain. It then follows by known results about the relationship between these two types of ordering that a choice between them is not really needed and can be largely considered as a matter of taste. For, if one assumes that the identity relation over elements of the domain under consideration is given[2], then a strict or-

[2]This means that there is a relation $=$ defined over elements of A, which is such that: 1. $a = a$ holds for every $a \in A$; 2. if $a = b$ is the case, then $b = a$ also holds, for every $a, b \in A$; 3. if $a = b$ and $b = c$ hold, then $a = c$ is the case, for every $a, b, c \in A$.

dering relation over a set A can be used to define a partial ordering $\leq$ of the same set[3]:

Proposition 1.1 *Let $<$ be a strict ordering relation over A, that is let it be such that: (i) $a \nless a$ for every $a \in A$, and (ii) if $a < b$ and $b < c$, then $a < c$ for every $a, b, c \in A$. Let $\leq$ be the relation over elements of A defined by putting, for every $a, b \in A$:*

$$a \leq b \Leftrightarrow_{Def} a < b \vee a = b$$

Then $\leq$ is a partial ordering of A.

Proof: first of all we prove that from (i) and (ii) it follows: (iii) $a < b$ entails $b \nless a$ for every $a, b \in A$. Suppose then that $a < b$ is the case. If $b < a$ were also the case, then $a < a$ would follow from (ii). This, however, contradicts (i). Therefore (iii) must hold.

Now, clearly $a = a$ entails $a \leq a$ by definition for every $a \in A$. Also, assume that both $a \leq b$ and $b \leq c$ holds for any $a, b, c \in A$. Since $a \leq b$ is the case, then either $a < b$, or $a = b$ (but not both, owing to (i)). If the latter is the case, then $b = a \leq c$ holds by the other assumption. If $a < b$ holds instead, then either $a < c = b$, or $a < b < c$ hold by (ii). Finally, assume that $a \leq b$ is the case, and that $a \neq b$ also holds. By definition of $\leq$, then $a < b$ must be the case. Hence, $b \nless a$ holds by (iii) above. Since $a \neq b$ entails $b \neq a$, then the definition of $\leq$ again implies that $b \nleq a$ is the case. QED

The relationship we have just proved can be reversed, as if one starts from a partial ordering of A, it is similarly possible to define a strict ordering of it:

Proposition 1.2 *Let $\leq$ be a partial ordering of A, i.e. a relation over its elements such that: (j) $a \leq a$ holds for every $a \in A$; (jj) if $a \leq b$ and $b \leq a$, then $a = b$ for every $a, b \in A$; (jjj) if $a \leq b$ and $b \leq c$, then $a \leq c$ for every $a, b, c \in A$. Let $<$ be the relation over A defined by:*

$$a < b \Leftrightarrow_{Def} a \leq b \wedge a \neq b$$

for every $a, b \in A$. Then, $<$ is a strict ordering of A.

Proof: since $a = a$ holds for every $a \in A$, then $a \nless a$ follows by definition. Moreover, suppose that both $a < b$ and $b < c$ are the cases. Then, $a \leq b$ and $b \leq c$ follow, which yield $a \leq c$. From the two hypotheses, $a \neq b$ and

[3]We assume familiarity with some standard notation from formal logic. In particular, for sentences φ, ψ, we assume that the reader knows that by $\varphi \wedge \psi$ and $\varphi \vee \psi$ are usually indicated the conjunction, respectively the disjunction of φ and ψ. We also make use of the standard notation for elementhood of sets and indicate, say $a \in A$, for: «a is an element of, or belongs to the set A».

$b \neq c$ also turn out to be the cases, which give $a \neq c$ (for, otherwise, from $a \leq b$, $b \leq c = a$ and (jj), then also $a = b$ and $a < b = a$ would follow). This means that $a < c$ is obtained by the definition above. In view of the first half of the proof of proposition 1.1, it is also the case that $a < b$ entails $b \nless a$ for every $a, b \in A$. QED

So, if one decides to align elements of a set by putting them in line according to a strict model of ordering, then it is possible to recover a larger model out of it and the same holds the other way around. Granted that, one may choose to favour either the one way or the other. But, do we really have to make this choice? As a matter of fact, the reader might have expected something different right from the very beginning by thinking of the stories we have started from, since they suggest a third, and possibly simpler way to compare the consequences of an agent's actions.

1.4. Take consequences at face value

The story of Alice we have started from, as well as the continuation of it involving Bob, raised no issue about attaching preferences to the consequences of the characters' actions. The system of preferences of the two agents, as we have referred to it above, was given to us 'for free', so to say. The matter for comparing the two actions at each agent's disposal was turning out from the story itself as with the one action was connected the increasing of the character's income, while a decrease of it was connected with the other one. One could even think of telling the story to make such connection even more explicit by speaking, for instance, of Alice as being paid one Euro more for each «like» she places, and to be given nothing if she places «dislike» instead.

The story as it was told already, once the implicit content of it in terms of consequences comparison was made clear along the lines we have just briefly mentioned, could then suggest another route for taking the matter of preferences into account.

Let us suppose that to each element of a given set A is assigned a value in the form of, say, a rational number. Let us assume that this is done by means of a function v that, with any element a in A, associates a value $v(a)$ that belongs to the set $\mathbb{Q}$ of rational numbers. Let us also suppose that such mapping of A via v is performed in such a way that no one and the same number is assigned to different elements of A. That is, let us suppose that $v(a) \neq v(b)$ whenever $a \neq b$, or, as mathematicians say, that v is an *injective* map of A into $\mathbb{Q}$. This assumption itself suggests a simple way for comparing the elements of A, as it was the case for Alice's actions in the story, namely by putting them in an order which reflects the natural ordering of numbers associated to them. In this schema, a is preferred to b just in case $v(a)$ is greater than $v(b)$.

The idea might appear both fascinating, since it is simple and easy to understand, but also simplistic at the same time. For, with the exception of situations like the one involving Alice and Bob where this possibility of 'measuring' consequences was, as I said, part of the story itself, it is unclear whether the assumption that any action has a 'measurable' consequence is feasible in all cases.

Another difficulty comes from thinking of this assignment of 'values' as representing an agent's preference: is this 'quantitative' model going to be as pleasing and satisfactory as a 'qualitative' model based on ordering relations might appear to be? Let us try to give an answer to the latter question first. Let me anticipate that the answer I will provide the reader with is largely unsatisfactory, since a proper answer to the question might require a technical detour that I consider unnecessary to take here. Since we are going to prefer the assignment of values to consequences over the relational approach to system of preferences, I do not want to linger for long about this matter.

From the superficial viewpoint I want to take, that question may reflect the following worry: how can an assignment of values in the form of numbers ever reflect the depth of reasons and justifications that might be hidden behind a system of preferences? Are not we loosing this information by sticking to values rather than to ordering relations? For instance, how can a numerical set of values account for common situations that may take place like dynamical changes of preferences? Once values have been assigned to elements of a given set A, what happens if this set gets bigger and new elements need to be equally evaluated? Is not the discreteness of (some) number systems preventing us from imagining how, having set values $v(a)$ and $v(b)$ for elements a and b of A, can we always set a value $v(c)$ for a new element c of that set to lie in between $v(a)$ and $v(b)$?

Concerns put forth in this way can easily be addressed. This is because number systems are more flexible than one can think of at first. Some number systems would indeed justify the previous preoccupation. Take $\mathbb{N}$, the set of natural numbers for instance. If we had used that as codomain of the function v assigning values to elements of A, then we could really have problems in trying to accommodate extensions of the domain set. If, for instance, we had set $v(a) = v(b) + 1$ and we wished to associate a new element c with value $v(c)$ in such a way that this is intermediate between $v(a)$ and $v(b)$, then the task would be impossible to accomplish, for there is no such value as there is no element of $\mathbb{N}$ in between $v(a)$ and $v(b)$. However, not all number systems are the same, in particular if they are like the set $\mathbb{Q}$ we have assumed to contain all values of the functions v above. $\mathbb{Q}$, as a matter of fact, is *dense* as mathematicians say, to the effect that if any two elements of it are considered, say q and q', one will always be able to find a third element of it, q'', which is an intermediate number between q and q'. So, if we are clever, we can stick to number systems flexible enough to accommodate with dynamical changes in values that do not put us at risk of requiring to always start the assignment from scratch, for every chosen change in the situation to consider.

Having thereby answered the question that came second above, at least in one of the possible interpretations of it, let us try to address the first difficulty then, which is an effect of the natural way of looking at this assignment of values that needs to be overturned. The natural view in question is the following: having assumed that a value $v(x)$ in $\mathbb{Q}$ has been assigned to any element x of the set A, and taken any two elements a and b of it, the agent prefers the former to the latter because $v(a)$ is greater than $v(b)$. Read in this way, the system of preferences that comes out from the assignment in question reflects the natural ordering of values of elements of A as numbers.

This view seems unquestionable: how else could the relationship between values and preferences be understood? Yet, this view raises the problem we have mentioned about how to assign values to consequences that do not appear to be measurable. The original remark about matters of income returning a way to measure consequences of actions for free, does not explain how the assignment of values may work with respect to actions that are evaluated with respect to different 'scales', involving principles like justice, honor, social advantage, and so on.

The way out, as I said, is to overturn the relationship between values and preferences. Yes, but how? Well, by assigning values to consequences in such a way that the value comparison reflect the system of preferences of agents. In this other scenario, having assigned values $v(x)$ in $\mathbb{Q}$ to elements of the set A, you have to assume that, for any two elements a and b, the value of $v(a)$ is greater then the value of $v(b)$ because the agent we are considering prefers a to b. To make the idea clearer with respect to the examples we have considered, the assignment of values in the case of the story about Alice should be made according to the following moral: since Alice prefers a wealthier life to a less wealthy one, then the value $v(l)$ of the consequence of choosing action «like» should be higher than the value $v(d)$ associated with the consequence of choosing the contrary action instead. And this is why an assignment of values can be as pleasing and satisfactory as a relational system of preferences in the end: just because we assume, as we shall stick to numbers as representing preferences in the course of the volume below, that values assigned to consequences reflect the agent's preference system in such a way that higher values are assigned to consequences which are preferred the more.

Having made clear this fundamental presupposition of ours, we shall now rapidly move toward considering the game-theoretic setting we want to use to frame the problem of rational choices in the rest of the book.

1.5. Decisions in a game-theoretic setting

Let us now consider the following scenario:

> Alice is put in front of two buttons, a red and a green one. She is told that each time she hits the green button she marks a score of 1. If she hits the red button instead, her score is 0. Finally, she is told that she wins if her score increases, and loses otherwise.

Let us compare this scenario with the story of Alice that was used for the sake of our initial observations in section 1.1. It is clear that there are some differences. First of all, while the first one is a story that deals with an everyday situation where the agent has to take real-life decisions, this one stems from a recreational kind of situations, a game in particular, as it seems clear by the occurrence of words like «score» and by reference to a winning, as well as to a losing condition. The actions Alice can choose among, are articulated accordingly: it is no more a matter of choosing which action is best to ensure her the greatest income, rather she is required to choose how to increase her score according to the rules. Despite the differences though, the two stories have also a lot in common: there is one agent in both who is supposed to make choices about what to do; the chosen action will have measurable consequences in the two cases (in the form of income increase on the one hand, and of score increase on the other). This suggests that, the clear differences notwithstanding, the two stories share features that are relevant to the kind of investigation we are planning to pursue. This idea is supported also by the variation of the previous scenario that corresponds to the complication of Alice's story that was considered in section 1.1:

> Alice is informed that Bob is now starting to play the same game, and the rule of it are changed accordingly. Alice will score 1 by pressing the green button if Bob also presses it, she scores the same if she presses the red button and Bob does that also, while she scores 0 in case they play differently: that is, if Alice presses green and Bob red, and if she presses red and Bob green.

The correspondence between the two situations we noticed earlier, are equally present in this case. The action by the agents is now turned into a move of the game, the consequences of which is now measured in terms of score, rather than income. As in the situation that was considered previously, the rules of the game now reflect the cooperative character of the real-life story, since Alice and Bob are favoured if they coordinate with each other, and penalised otherwise. The two situations are so close that one can think of one as being the reformulation of the other with just a different terminology. If one thinks of the features of the situations considered first, which we have flagged as crucial for the decision problem we wish to tackle, they all appear here as well, except for the fact that they go under a different name (with «players» taking the place of «agents», «moves» of «actions», and «score» of «income»).

Yet, to think of «games» rather than «stories» is not merely a matter of translation. The situations as they are described now, in the 'game form', are more clearly written with a lot of unnecessary details left off (for instance, what brought Alice to the situation where her decision-making skill is required). At the same time, those aspects we concluded to be crucial to consider when a decision come into play, are now even better emphasized and brought to the foreground, to the point that one can easily produce a schematic rendition of the situation. As a matter of fact, the game in the two-player version boils down to considering the four cases which turn out by combining the players' actions. These cases, therefore, can be identified with the pair of moves they are connected with, to indicate which we can think to a first-level formalism featuring G for «press green» and R for «press red». To keep track of which players moves what, we may use indices and write G_i, R_i, where i is either a for Alice or b for Bob. What counts then about possible outcomes is the score the players get. The formalism here comes for free again, as the score is given in the form of number of points, and to keep track of which player gets what, we can use pairs (n, m) where we conventionally assume to indicate the score of Alice first, and Bob's second. Granted that, the game conditions can be summarized by means of a simple schema:

$$
\begin{aligned}
(G_a, G_b) &= (1, 1) \\
(R_a, G_b) &= (0, 0) \\
(G_a, R_b) &= (0, 0) \\
(R_a, R_b) &= (1, 1)
\end{aligned}
$$

Since this schema contains all that we stressed as relevant for the proposed investigation and the pursuing of it, to pass from a situation in 'story form' to a situation that corresponds to it (in a fashion similar to the correspondence we found in the cases here at stake), and which is in 'game form', may be convenient for the sake of rigour. For, the example I have used for the sake of illustrating the point I want to make, suggests that the idea of game is strong enough to fill our needs in this respect. This is at least what we shall assume in the rest of this book. We will give to this assumption the form of the following *game representation hypothesis* (GRH, henceforth):

GRH: For every situation which is relevant to consider in order to investigate the nature of rational decision, one can design a corresponding game such that studying decisions made by players of this game is the same as studying decisions made by the characters involved in the original situation.

We will now pass to introduce what counts as a «game» for the sake of this hypothesis, trying to single out its features in order to give it the form of a concept that one can deal with mathematically, and then pass with the next chapter to the investigation of the properties of 'objects" falling under this concept in a systematic manner.

1.6. How to design a game

If a situation is about *agents*, a game is about *players*. Agents *act*, therefore decide which *action* to act, while players *move* (and choose among the *moves* they have available). A situation is framed by *states of affairs* that constraint or direct the agents' choices (in the story about Alice and Bob, the clauses of their contract play the role of states of affairs that are relevant to their choices), while a game is framed by its *rules* which make moves legitimate or not. In a situation where more than one agent is present, actions of agents *combine* and these *actions combinations* give rise to *consequences*. Similarly, if there is more than one player in a game, their moves combine and these action combinations give rise to *rounds* of a game match, for each of which players get a *score*. Consequences of each actions combination might be different to agents, or, to say it better, might be differently evaluated by agents which express *preferences* on them (where each agent is free to express his, or her own preference), the general rationale being that agents will takes decisions that agree with their own system of preferences (therefore, they will choose the action which they prefer the most). Correspondingly, players in a game get their score individually and, unless differently suggested by the rules of the game, it is intended that they play in such a way that their score increases as the game goes on. So, we assume that the relationship between scores and preferences be the same that was supposed to hold between values of consequences and preferences in section 1.4.

This concise summary is conceived in order to give a sample of how terminology may vary from now onwards, having decided to proceed according to GCH.

Players in a game are required to make choices that concern the action to play next to the current state of the game. More precisely, as is well-known to those who play recreational games, players are up to devising a *strategy*. In general terms, a strategy is a plan: an idea on what to play next that takes into account what has been played so far in a match and what the other players may play, and is used to determines the player's moves accordingly. If we were supposed to describe the form a strategy takes, this would likely be a series of statements of the form «If players $p_0, \ldots, p_k$ play..., then I play..., and if $p_0, \ldots, p_k$ play... instead, then I play..., and if...». That is to say, a strategy is a hypothetical plan that allows players to make choices in the actual situation. Of course, this plan is sensible to the score distribution, as the latter represents the ultimate motivation for each player's move as I said. Since strategies can be taken to represent the way decisions are made by players, what I have just said suggests that they might be represented in the same way as the latter are in ordinary situations: namely, as functions which take in this case rounds of the game as inputs (the form of which being determined by the number and the distribution of actions among the players playing in the game), and yield the score accordingly (therefore, decisions in this form would be something like a function $d : R \rightarrow S$, where R is the set of legitimate rounds of the game under consideration, while S is the set of score assignments).

Strategies are not always the same. A strategy can be more effective than another because of its 'range'. For instance, because it is so conceived that it takes more options into accounts, and therefore allows the player who conceived it to better secure her decisions with respect to what the others do. Similarly, a strategy might be stronger than others because it allows a player to rely upon that for longer (i.e., it works well for more rounds than other plans do). The main difference between strategies is between those which are «winning», and allow the player to actually achieve the best possible score, and those which are not. While, depending on the game rules, there might also be strategies that are neither winning nor loosing in all cases where ties are allowed.

Now, since strategies are crucial to players, it might be worth reasoning some more about them. As a matter of fact, there are a couple of issues that are quite naturally connected with this idea about players making choices for the sake of their performance in the match that is worth discussing preliminarly here. As games, like real-life stories and choices made there, fall into types, strategies that need to be devised accordingly fall into types too. For instance, it is clear that players make choices on the basis of their knowledge.

A player's knowledge, by the way, may in part refer to the player's individual character and skills. This aspect of the issue is rather difficult to assess as each player comes with his own talent and experience, and seems therefore to be of little help to consider that connection for the sake of an investigation like ours which aims at establishing something that could be regarded as holding in the general case.

The player's knowledge that seems important to evaluate here is the one that is related to the game characteristics. The piece of a player's knowledge that is relevant in this respect, is their knowledge of the game rules. Rules, as it was hinted at above already, determine an important restriction over the players' moves as they make these split into *legal* and *illegal moves* accordingly (the first group being made of all moves that are allowed by, or are in agreement with the game rules, the second group containing all moves which are not as such). It should be clear that this sort of information is crucial for the sake of devising a winning strategy, since only by passing from legal moves to legal moves a player will be in a position to win the game. The second part of each player's knowledge that is relevant to our goals is the knowledge of the match they are playing as the latter goes on. Knowledge of this type involves a player's awareness of the current state of the match, which in turn will make her capable of determining what moves may the opponents make, in combination with the afore mentioned knowledge of the rules of the game. Finally, there is a third part of a player's knowledge that must be mentioned, which is the players' knowledge of the score distribution and how each move will impact on their individual score in particular. To know this it is required to determine, given the current state of the match, which one of a player's legal move allows her to achieve the highest score, and could be preferred the most, and which of the opponents' move allows them to get the best score and is therefore the one they will be likely endindg up playing.

An optimal strategy, a winning strategy as I will be calling it, is the result of a careful analysis of all these aspects of a game situation. Even on the basis of this rough account, to devise a winning strategy for a game appears to be a difficult task. Most importantly, this task appears to be different according to what we assume is the knowledge level of players. Roughly speaking (again) there are two main routes that we may take here.

On the one hand, we may assume that players know everything they have to know to devise their strategies. This route applies to games which convey to players a *perfect information*, as game-theorists would say. Chess or draughts provide us with good examples of games where players always know what is needed for them to devise a plan for winning: the state of the game, which corresponds to the state of the board that is always visible to them, the rules, as well as advantages and disadvantages that legal moves represent in terms of each player's score. This makes clear that games where players play with perfect information can be as hard to play as those where some information is hidden. So, for a perfect information game, it is perfectly reasonable to ask questions about winning strategies, and whether it is always the case that a player can devise one.

There are games, however, which would be unsuitable to treat as situations where players knows everything. Think to poker, for instance. Like most games played with cards, there are always cards that remain hidden to the players' sight. So, to guess what cards are the other players holding and what cards are in the bundle is part of what each player has to try determining for the sake of her strategy. The difference between smarter and less smarter players is also measured by considering the ability of recovering this additional information. Games of this sort are generally referred to as *imperfect information* games.

The distinction between perfect and imperfect information games is naturally related to another issue about the players' choices that is worth mentioning. In a situation where all the required information is displayed, and in case the state of the match allows it, players might be in a position to make predictions about what is going to come next with absolute precision. In all these cases we shall say that players will elaborate strategies that are *pure*, as they reflect this condition of absolute certainty about the subsequent development of the match.

In the normal situation, however, this is not likely to happen as players will normally have choices to make and each player's strategy will be sensible to how these choices of players combine. Therefore, a player is rather devising strategies which take all these different combinations into account and the player's choice will be made on the basis of what combinations is favoured, i.e. regarded as more likely, and which one is not. In these cases we say that players play according to strategies that are *mixed*, which are obtained by assigning probabilities to every possible combination of moves by the players (that is, to each possible pure strategy) in a way that reflect how likely it is that they will occur in the actual game according to the players.

Pure vs. mixed strategy are pivotal concepts of two big areas in which the theory of games is divided. Also the consideration of perfect vs. imperfect information games causes the study of the subject to split. Neither of the two splittings of the general theory is reflected in the foregoing parts of this book. The observations we made here where conceived to state with exact precision what is the object of this volume and what is not, in such a way that the reader knows, to make the game corresponding to the reading of this book fair right from the beginning, what she may gain from it and what she will certainly cannot expect to gain instead. As a matter of fact, our goal is to deal with finite games, where players are always granted perfect information and devise strategies which are pure.

Chapter 2
How games are dealt with

Having sketchly summarized some of the issues related to passing from addressing the choice problem in real-life situations to dealing with it in games that corresponds to, or represent them (owing to GRH), in this chapter we are going to dig into the topic some further. Let us assume that our previous considerations from section 1.5 provide us with enough reason to think that by shifting to considering games things get simpler as long as the problem of analyzing choices made by agents is concerned. However, to say that we aim to passing from real-life situations to games is not enough. For, beside the obvious changes in perspective we have hinted at in section 1.6, the idea of «game» we have proposed is still too vague to let us properly assess advantages and disadvantages. The first sections of this chapter are conceived in a way to discuss the topic and relate the reader to standard ways in which games are dealt with by the ordinary theory. The standard vocabulary of it is also introduced along the way, in view of the thourough account that will be starting from next chapter.

2.1. Games in tree or matrix form

The story we will make use of as our running example for the overview on the theory of games we plan to pursue, is not original. On the contrary, it possibly represents the most popular and most discussed case-study in the field. Here is how the version of the story we will stick to goes:

> Alan and Bonnie have been caught red-handed, and arrested by the police for having stolen a car. However, the police has reasons for thinking that they are involved in a bigger crime and suspect they are part of a band which committed a series of robberies recently. Alas, there are only evidences to send

Riccardo Bruni, *Course notes on finite games and rational choice*, ISBN 978-88-6453-832-7 (print)
ISBN 978-88-6453-831-0 (online), CC BY 4.0, 2018 Firenze University Press

> them in jail for the car theft. However, an attempt is made to get a confession by proposing to each of them the following agreement: if you confess you took part in the robberies, make a statement that also implicate the partner and she does not confess, then you will go free and she will be put in jail for ten years; if you confess and she also does it, then you both get five years; if you do not confess and she does not too, then you will both be imprisoned for two years for the car theft. What should Alan and Bonnie do?

The first step into trying to answer the question requires that we make decisions on how to model the situation. The whole idea is to represent it in game-form, in agreement with GRH, by interpreting the consequences of the characters' actions at face value as we said, or, to make use of the standard way of referring to them in the ordinary theory of games, by looking at the *payoff* or the *utility* that each character gets in the situation under scrutiny. Since the former are given in terms of 'quantities' already (the number of years of prison Alan and Bonnie get for each decision they make), we just make use of a direct 'translation' into payoffs by sticking to the rule that turns n (the number of years of prison in the agreement that is proposed to the characters) into utility $-n$. Therefore, we obtain the following correspondence:

Consequence	Payoff
Go free	0
Two years	-2
Five years	-5
Ten years	-10

Then, utilities are arranged into an order by making use of the usual ordering $<$ over integer numbers, which causes 0 to be the top element and -10 to be the bottom one. The story also gives us a clear indication about what actions our characters can choose, which is either to *confess* or *not confess* instead. If we apply here the notation that was suggested already in section 1.5, this gives us two options C_a, N_a for Alan, and two more, C_b, N_b, for Bonnie. These two steps together would then lead us to the schematic representation of payoffs attached to the characters' combination of actions that reads as follows:

$$(C_a, C_b) = (-5,-5)$$
$$(C_a, N_b) = (0,-10)$$
$$(N_a, C_b) = (-10,0)$$
$$(N_a, N_b) = (-2,-2)$$

Yet, the schematic form we have reduced the situation to, while certainly helpful to concisely summarize the crucial information we need in order to come up with an answer to the question about what action should each character choose, it is still not enough for achieving it. Some further aspects of the situation need to be taken into account. In particular, it seems that we are required to identify some sort of reasoning that may explain how the two characters in the story are supposed to make up their minds. In the next two sections we will scrutinize two possible options in this respect. In turn, these will naturally brought us to arrange the above data in a 'structure', so to say, which will allow us to reach the answer we are seeking for.

2.2. Dynamic choice: games as trees

Let us suppose that the following addition to the story of Alan and Bonnie is made:

> The chief inspector keeps Alan and Bonnie in the same room and makes her proposal to them at once. Then, she lets Alan choose first, and Bonnie choose second.

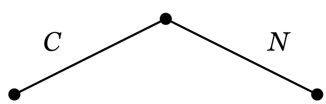
The new feature comprised in here is some sort of dynamics: it is now clear that the two characters will have to make a choice alternatively, with Alan choosing first and Bonnie having to choose once he is done. Then, one can look at the whole story from a new viewpoint, as we can focus on it by stepping into each character's shoes once at a time. When it comes to Alan to choose, he has two options to assess: one is to confess, and the other is to refuse doing it. This is the same for Bonnie, since she is given these options as well, but since she chooses second her choice combine with Alan's and give rise to four scenarios to consider: the one in which she confesses having Alan confessed as well; the one in which she does not confess, while Alan has decided to confess; the one in which she confesses contrary to what Alan does; and the one in which Bonnie does not confess and Alan has made the same decision. Payoffs are then distributed among these four alternatives.

Now, how does the addition to the story we are considering in this section change the situation as it was first given? As it was said, it gives it a structure which emerges once we try to put the information into a diagrammatic form for the sake of analysis. As a matter of fact, it is now quite natural to think to the situation involving Alan and Bonnie as taking place in stages.

Stage one occurs when the options has been declared by the chief inspector and, according to the addition to the original story, Alan is supposed to make his choice. Stage two takes place once Alan's choice has been made, and it can take one of two possible forms, depending on what Alan has decided to make. This is the stage at which Bonnie chooses. Her choice determines the third and final stage, which is when the two characters are finally granted the payoff they deserve owing to the payoffs distribution over the possible outcomes of the story.

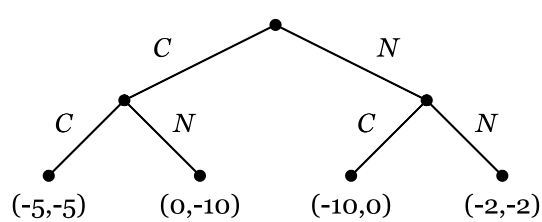
How does it happen that this structure turns into a diagram as it was suggested? Let us suppose to represent stages in the story as *nodes* in a path. Then, there should be a node representing the first stage in particular, which is the initial situation in which Alan chooses. As we said, Alan's decision leads us to the subsequent stage, which is therefore connected with the previous one. The connection between two subsequent stages is represented by *edges* liking them together. The natural interpretation of edges is to view them as determined by the reason causing the transition from the one stage to the subsequent one. The reason in the case under scrutiny is Alan's choice between the two actions at his disposal. Depending on what choice is made, a different second stage in the story will occur. Since we aim at giving a comprehensive overview of the situation for the sake of the analysis of it, it is required that both outcomes be represented: before Alan has chosen, we have no clue whatsoever to say which form of stage two will prevail.

This suggests that we should adopt the following diagram to represent stage one, stage two, and the passage from the former to the latter:



Edges carry a label that refer to the reason why they appear in the diagram, in this case this reason being Alan's choice of action. So, we have two branches: one corresponding to the situation in which from the initial stage we get to stage two and Alan has chosen to confess, the other being the situation that represent Alan's choice of not confessing instead.

The terminal nodes of the diagram represent the situation in which Bonnie's choice takes place. Her decision determines the passage to stage three, which, once again, takes a different shape owing to what decision Bonnie makes. By sticking to what we have just done for stage 1 and stage 2, this leads us naturally to the subsequent continuation of the previous diagram:



Terminal nodes of the diagram now represent the final situation to model, that is the stage at which, having both made their decisions, characters get their payoffs. This explains why nodes representing the possible forms of stage three are labelled by the payoffs assignments to agents, the latter being written as pairs of numbers as before, whose order reflect the order of agents' actions (hence, the first element of each pair represents Alan's payoff, while the second element is Bonnie's payoff instead).

As a result of the dynamic feature introduced in our story by the clause we have added here, it turns out that all the information required to assess the choice at the characters' disposal are nicely displayed in the form of a tree owing to the reading we have proposed, which we stress here again for the reader's sake: having noticed that the story dynamics causes it to split into stages, stages in our story are nodes in the tree, actually arrays of nodes, each one of those belonging to one and the same row representing a 'version' of the stage that takes place; actions leading from one stage to the other are represented by edges, the number of edges at each node corresponding to the number of options the character that is supposed to choose at that stage has; the initial node, the *root* of the tree, which is unique, then corresponds to the initial stage of the story, while terminal nodes of the tree, the *leaves*, represent the ultimate outcomes of the situation, i.e. the stage at which payoffs are given to the characters.

Games represented as trees are referred to in the literature as *games in extensive form*. This way of analyzing them depends of course upon the addition we have made to the original story, and upon the consequences of it. In particular, it depends upon the fact that characters are forced to act by sticking to an order, and the action of the one character is following the one of the other. It is clear, however, that this is not the only way things could have been going in the case of Alan and Bonnie.

2.3. Static choice: games as matrices

Let us suppose that the story we have been considering in section 2.1 goes on differently:

> The chief inspector brings Alan and Bonnie in two different rooms in such a way that they cannot communicate with one another. She makes them the offer and they are informed that the same proposal has been made to both of them. Then, they are required to make their choice without possibly knowing what choice the other is making.

It should be clear that in the new situation the previous analysis is no more of help. As a matter of fact, that was carried out under the assumption that Alan and Bonnie were choosing one after the other. The new storyline rather suggests a different approach where no preferential order between the outcomes should be assumed to hold. Hence, no list, or alignment of them seems to be justified anymore.

To represent the situation in a way to remain as close to the characters' viewpoint as possible, outcomes should be arranged in a table in some sort of neutral way. We then stick to a matrix arrangement where things are ordered lexicographically on the basis of what actions combination they correspond to. Let us briefly explain how this arrangement is achieved.

Let us assume to keep considering situations with just two agents involved for the sake of simplicity. Let us also suppose to indicate by $a_0, a_1, \ldots, a_n$ agent 1's actions, and by $b_0, b_1, \ldots, b_m$ agent 2's actions. Actions combinations are then ordered lexicographically (something that we will express symbolically by $ab \prec_l a'b'$ to mean that actions combination ab preceeds in the lexicographic order actions combination $a'b'$), owing to the clause: $a_i b_j \prec_l a_h b_k$ if and only if $i < h$, or $i = h$ and $j < k$.

The effect of the definition just given is to generate the following arrangement of actions combinations:

$$
\begin{array}{l}
a_0 b_0, a_0 b_1, \ldots, a_0 b_m \\
a_1 b_0, a_1 b_1, \ldots, a_1 b_m \\
\ldots \\
a_n b_0, a_n b_1, \ldots, a_n b_m
\end{array}
$$

Each combination in the array above represents a distinctive outcome of a situation in which agents 1 and 2 are suppose to choose among actions $(a_i)_{0 \leq i \leq n}$ and $(b_j)_{0 \leq j \leq m}$ respectively. Each outcome in the situations we shall be interested in here, is associated with a distribution of payoffs among agents. As in the arrangement of outcomes as leaves of a tree we have obtained in section 2.2, we would like the information to be 'displayed', for the sake of the evaluation of it.

The graphic representation of the situation is then achieved as follows: actions of agent 1 are displayed in a column in lexicographic order, while actions of agent 2 are similarly displayed in a row. Payoffs granted to agents are placed at crossings in the usual form of pairs where the first element is the payoff that goes to agent 1, while the second element is the payoff that goes to agent two. Notice that we use here symbol p_i^j to represent the pair of payoffs granted to players for every $0 \leq i \leq n$ and $0 \leq j \leq m$. Hence, each p_i^j is of the form $(1_i^j, 2_i^j)$ where 1_i^j is the payoff that agent 1 gets if she plays action a_i when agent 2 plays action b_j, and 2_i^j is the payoff that agent 2 gets in the same situation.

In the general situation we are here considering, this would allow us to reach the following disposition of outcomes and payoffs:

	b_0	b_1	...	b_m
a_0	p_0^0	p_0^1	...	p_0^m
a_1	p_1^0	p_1^1	...	p_1^m
...	...	...	...	...
a_n	p_n^0	p_n^1	...	p_n^m

It should be clear that starting from the leftmost, upmost corner of it, the diagram reproduces the lexicographic order of action combinations above, and it further contains the information we wanted to have displayed about what payoff each agent gets in every such situation. Hence, each row of the diagram contains all responses of agent 2 to one and the same action by agent 1, while columns provide us with the same information in the other order of agents.

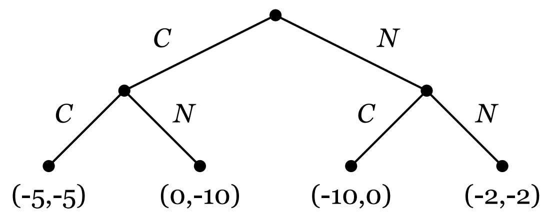
Coming then to our simpler example, the general method for representing the situation we have decided to adopt would lead us to the diagram below:

	C	N
C	(-5,-5)	(0,-10)
N	(-10,0)	(-2,-2)

It should be clear by what we have said that this way of arranging things, which leads to represent games as matrices, or to represent them *in normal form* as it is said in the literature, is as general as the previous representation of them in terms of trees. However, one may well expect that the two modes of arranging things bear differences due to the fact that they are justified by two alternative endings to our story. This is what we aim at clarifying in the rest of the chapter, by stressing two things in particular: (i) how the diagrammatic representations help us answering the original question about what is rational for Alan and Bonnie to do in the situation we are considering, and (ii) whether the two alternative ways of arranging outcomes also lead to different answers, as long as the issue of the two characters' choice is concerned.

2.4. Reasoning on trees

The first task to accomplish is to try to see how arranging all of the important information in a story such as the one about Alan and Bonnie, can help us in our attempt of determining what one should expect in terms of the characters' reaction, or what action should the characters choose. Let us take the tree case first, and let us consider then the diagram we finally displayed in section 2.2 which we reproduce here for the reader's sake:



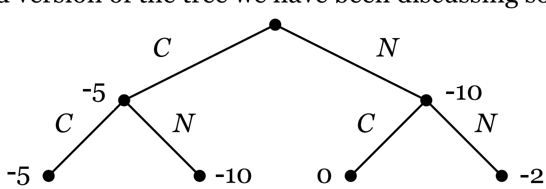
Now, it should be clear that to try determining what action between confessing and not confessing agent 1 should favour, is the hardest thing to attempt. For, agent 1's choice is not yet connected to the consequences of it, being payoffs distributed to a combined action by the two agents, and being the outcoming combination depending therefore upon agent 2's choice. Reasoning about agent 2's choice is much easier, since her actions do lead to consequences we can 'measure' by means of payoffs. Of course, this will not be an exact form of reasoning since it will be depending upon the hypothesis of how agent 1 has moved previously, but it is exact enough to put us at least in a position of discriminating between the available outcomes and select those which are most likely to happen.

So, let us first suppose that agent 1 has decided to play C, i.e., to confess that she took part in the robberies the police wanted to accuse her and her accomplice of. Then, the leftmost path in the tree is the actual one, and by looking at it we have that agent 2 is facing two alternatives: on is to confess as well, and the other is not to do it. If she does confess, the payoff distribution entails that she is going to be convicted for five years in prison (owing to the correspondence we set in section 2.1, and owing to the second element of the pair of numbers beneath the leftmost leave of the tree). If she does not confess instead, she is going to be convicted for 10 years as a result of agent 1's statement. Now, there is no doubt about what agent 2 will be doing in the end, having agent 1 decided to confess, and she will confess as well in the attempt of minimizing the effect of the conviction she has to face.

Let us assume that agent 1 decides instead not to confess in the first place. Then, the two alternatives agent 2 is facing, those coming out from the righthand path of the tree above, will have different consequences payoff-wise. As a matter of fact, if she did confess, then she will be free to go (as payoff «0» corresponds to freedom in our schema), while she will be convicted for two years if she refused to confess. Once again, it seems obvious what choice agent 2 will make and she will decide to confess.

There are two observations to flag as a result of this analysis. The first one is that agent 2 will decide to confess anyway. So, despite the reasoning is hypothetical, this means that we can determine what action agent 2 will make categorically (that is, independently of agent 1's choice). The second effect of the above analysis is that, having been able to determine what choice agent 2 will favour when it is her turn to play, we are now in a position to also state what payoff agent 1 is likely to get by making her choice of action. As a matter of fact, if agent 1 confess, then it is clear that she will get a payoff equal to -5 (which corresponds to be convicted in prison for five years), as a result of agent 2 confessing in that case. If she does not confess, since agent 2 will confess anyway, then she is going to score -10, which is the same as being convicted for ten years in prison. If this is the option, then also in the case of agent 1's decision there seems to be little doubts about what she is going to decide, since confessing offers her a clear advantage over not confessing.

Two more remarks are on order. The first one concerns the fact that, having analyzed agent 2's choices first, has finally put us in a position from which it is possible to assess agent 1's choices as well. The situation we have achieved in this way can be regarded as the result of distributing the agents' payoffs over non-terminal nodes in the tree like in the following, modified version of the tree we have been discussing so far:



As a matter of fact, the information required to perform the first step in the above analysis (determine agent 2's choice under the hypothesis of what agent 1 might be doing), is the payoff distribution for agent 2's actions at terminal nodes. No use of the information about agent 1's payoffs was made there. The latter information instead, was crucial for performing the second step of the previous reasoning (the one devoted to determine agent 1's choice at the initial node, knowing what agent 2 will do afterwards). In other words, as a result of the previous reasoning we get a new diagram where payoffs granted to characters, rather than been coupled and used to label leaves in the tree, are taken individually and attached to nodes that are reached by a choice of action made by the character that gets them.

The second observation is that, having reasoned in this way, we are able to predict how the two agents will choose, or to say how they should choose assuming (which is likely) that each one of them is willing to make the most convenient choice (this one coinciding with the one that allows them to get the minor conviction). Now, suppose that both Alan and Bonnie, our agent 1 and 2, are assumed to be capable of going through a simple reasoning such as the one that led us to the said conclusion. Notice that the side conditions of our story make the assumption realistic since all of the required information is on display, and nothing we have relied upon for the sake of the argument is hidden to the two characters. Then, under this supposition Bonnie would reason hypothetically as we did at first, just to conclude that confessing is her best chance to go for. Alan, on his part, would be able to replicate Bonnie's reasoning and would realize what she concluded. Therefore, owing to the payoff distribution he would be in a position to say that confessing is also his own best choice. In this way they both would come to a conclusion, and the chief inspector would get a confession from both of them.

2.5. Reasoning on matrices

Having reached the conclusion that Alan and Bonnie can finally make decisions on what to do in our story under the assumption that they are able to carry on the argument above, one may wonder whether this was made possible by the extra assumption on them acting in a sequence. The easiest way to determine whether the agents' order is crucial or not, is to consider the alternative ending of our story where both characters were required to make their choices simultaneously and without knowing what the other has chosen in advance. For reasons we have discussed in section 2.3, the situation in this case is best represented by the following matrix, which we reproduce here for the reader's sake:

	C	N
C	(-5,-5)	(0,-10)
N	(-10,0)	(-2,-2)

Since our agents are supposed to make their choices separately, there is no difference here in where to start for addressing the issue. Let us consider the situation from the viewpoint of agent 1 anyway. It will turn out clearly that we could have started from agent 2's choice, and nothing different would have happened. Remember that while devising the diagram, we had decided to put agent 1's actions on the leftmost column and we have assumed that her payoff is indicated as the first element of the pair that can be found at the crossing with agent 2's choice of action. Actions at agent 2's disposal are placed along the topmost row of the diagram instead, and her payoff appear in the diagram as the second element of each pairs of numbers[1]. Agent 1 may then reason as follows. She considers the possibility of confessing at first and notices that, if she did that, agent 2 is likely to confess as well since she scores -5 instead of -10, which is the score she gets if she does not confess. This follows by looking at the second element of the two pairs of numbers in the first row where payoffs pairs appear, as it results from the diagram by focusing on the portion of it that is referred to by the agent's reasoning above:

	C	*N*
C	(-5,-5)	(0,-10)

Agent 1 then considers the possibility of not confessing, in which case she notices that agent 2 is again very likely to confess since confessing would allow her to score 0, while not confessing would let her score only -2:

	C	*N*
N	(-10,0)	(-2,-2)

Then, agent 1 notices that she can conclude that agent 2 is going to confess anyway. In this case, however, it follows by looking at first elements of the two pairs of the leftmost column of the diagram where payoffs pairs appear, that to confess is best for her if agent 2 is likely to confess as she concluded, since agent 1 would indeed score -5, while not confessing would let her score -10:

	C
C	(-5,-5)
N	(-10,0)

[1]Notice that our convention about actions disposition in the matrix is unimportant here, due to agent 1 and agent 2 having the same two actions at their disposal. The same convention, however, will be maintained for the matrices to be displayed below in the volume.

Agent 2 would apply a similar line of reasoning. Actually, she applies *the same* reasoning, being the payoffs distributed symmetrically among agents. Then, she assumes that she may confess, in which case she notices that agent 1 would be willing to confess, which brings her a score of -5, rather than not confessing which would cause her to score -10 instead:

	C
C	(-5,-5)
N	(-10,0)

In case she decided not to confess, agent 2 notices that agent 1 would be equally brought to confess, which would allow her to score 0, rather than avoid doing it, since not confessing would let her score -2:

	N
C	(0,-10)
N	(-2,-2)

Having thereby established that agent 1 is likely to confess, agent 2 notices that she should confess as well to achieve a score of -5, rather than risking to score -10:

	C	*N*
C	(-5,-5)	(0,-10)

The result of the analysis we have just carried out with respect to the main issue we were up to, namely to try to determine which action combinations by the agents is likely to happen in the situation we are considering, is the following: both Alan and Bonnie are likely to confess and, under the above assumption that they are carrying out the reasoning themselves, they are also able to detect what the other is likely to choose and to make their choices accordingly.

2.6. The rationale of rational choice

The purpose of the chapter so far was the attempt to test the previous idea that, by turning situations that may resemble real-life stories into games, one would be given the opportunity of focusing on the only important features of it to consider for the sake of the investigation on rational choice we aim at pursuing. The story we have stick to in this respect has given us encouraging information. For, as far as the problem of assessing the possibility of detecting a 'solution' to the problem of choice in a given situation, that is, to determine what choice of actions agents involved in a story (which is the same as players of a game, owing to GRH), it turned out that a positive answer was attainable in both the tree-like approach to our running example, as well as in the approach to it based on matrices. Out of all the possible outcomes deriving from the combination of actions by Alan and Bonnie, we now have a plausible candidate to propose as solution to the quest of a rational choice on their part. The choice is rational because of the argument itself that has lead us to single it out. We will come back to the peculiarities of the reasoning in question later on, as there are aspects of it that are worth deepening. For the moment, we are rather happy with the possibility of approaching the choice issue via games that we have attempted here, and we would like to further test the methodology. For this very reason, we are going to flag the advantages of it, as these turn out from what we have done so far.

Both the tree and the matrix approach seem flexible enough to cope with a variety of situations we may wish to handle. They certainly rely upon a number of features that the situation under scrutiny need to present for the two approaches to be applied. First of all, they require that a distribution of payoffs be present in such a way that reasonings such as those we have exemplified in section 2.4 and in section 2.5 can be performed. As we shall see when the whole issue will be approached more closely in the next chapter, there is something more about the payoff distribution that is required for the analyses that have been carried out here to be pursued. Anyway, for the time being we shall content ourselves with the observation we have just made about it.

Another feature that we have taken for granted, as we announced it at the end of chapter 1, is that the game under scrutiny be finite. What «finite» means is now clear. For both the tree, as well as for the matrix mode of arrangement of parameters counting in a game, it is needed that both the number of players and the number of actions at each player's disposal is finite. Finiteness of the number actions, which entails finiteness of the number of actions combinations, is crucial for the sake of the kind of reasoning that we have performed in section 2.4 and in section 2.5 with respect to our exemplary situation. This is likely to be a feature that cannot be given up, and one may think that this is a minor constraint to respect: we are planning to devise a methodology that be adequate to treat, via GRH, real situations in which the choice issue occurs and no such situation would ever involve infringement of the finiteness requirement we have just mendtioned. Though the observation is plausible and unproblematic at the current stage in the analysis, we shall see in chapter 4 that by approaching the whole issue from a more abstract angle, the point of considering the extension of it to infinite cases stems quite naturally.

One further thing to notice is that the two approaches we have here envisioned, the one based on trees on the one hand, and the one based on matrices on the other, are complementary in the sense that they cope with different kinds of situations. In particular, trees are tailor-made to model situations in which the order of actions by the players counts and each player's choice is made in a certain sequence with respect to the others' choices. Matrices, on the contrary, get used in all those situations in which players are required to make their choices simultaneously, or at least whenever the order in which they are made is presumed to have no visible effect. This means that, taken together, trees and matrices should enable us to treat a number of different case studies. In view of this, it seems important to try to generalize what we have said so far about the possibility of actually getting a solution to the choice issue with respect to the toy situation we have made reference to so far. In particular, we would like to know more about combinations of actions by the players which, as it was for the combination «confess» and «confess» by Alan and Bonnie, turn out to offer concrete advantages to players with respect to the alternative ones, hence candidate themselves as solutions to the problem of detecting actions that are preferrable to others. What is their distinctive character? Can we somehow be certain that at least one of them can always be found in any possible situation we should be faced with? Are solutions of this sort really as good as they appear at a first sight? In the next two chapters we would like to explore these and other related issues. We plan to start from simultaneous games at first, and then move on to consider situations where players are required to make their moves one after the other.

Chapter 3
On games in normal form

Owing to what we have said in the previous chapter, by the name *game in normal form* goes any situation that can be schematically reproduced by means of a diagram that takes the form of a matrix. Some of the general features required for that have already been discussed in section 2.6. We just list them here again for the reader's sake:

- the number of players partaking in the situation must be finite in the first place;
- the number of actions at each player's disposal must be equally finite in number;
- each combination of the players' choices that counts as, so to say, a 'round' in game-theoretic terms, i.e. a state of the match in which 'points' or other forms of prizes are distributed among the players, must be associated with a distribution of payoffs to the players (this feature being indeed crucial for the kind of reasoning that was used in section 2.3 to reach our conviction about which of all possible action combinations was candidate to be the best one, hence the one to be chosen);
- finally, games to be put in the form of matrices must be such that no order applies to the choices made by the players, who are required to play simultaneously and must be therefore able to make up their minds while the others have not played yet.

3.1. Solving the (matrix) riddle

Let us go back to the case study we have been analyzing in the previous chapter. Let us assume the reader's confidence in the story behind it, and let us start again from the point in which the diagram summarizing it has just been laid down:

Riccardo Bruni, *Course notes on finite games and rational choice*, ISBN 978-88-6453-832-7 (print) ISBN 978-88-6453-831-0 (online), CC BY 4.0, 2018 Firenze University Press

	C	*N*
C	(-5,-5)	(0,-10)
N	(-10,0)	(-2,-2)

The actions combination made out by the pair *CC* (which is a shorthand for the proper notation $C_a C_b$, where actions carry an index referring to the player that makes it), was put forth as candidate to be the best one. What makes it special with respect to the other outcomes is the argument we considered in section 2.3. That reasoning came out as an attempt to simulate the players' own assessment of the matrix. Such an *a priori* analysis of the situation corresponds to an *a posteriori* evaluation of the outcome of it. As a matter of fact, let us assume that Alan and Bonnie, that is agent 1 and agent 2 respectively in our account from the previous chapter, are determined to confess. Then, it is clear from the distribution of payoffs that there is no convenience for neither of them to change their minds unilaterally: if, for instance, Alan did that, he would move from the leftmost, topmost cell of the diagram to the one to below it, hence he would pass from scoring -5 to scoring -10; Bonnie, on the other hand would suffer from a similar loss, because by changing her choice from «confess» to «not confess» in presence of Alan's choice of confessing, would cause the final outcome to be the one placed second in the pair of numbers to the right of the current one (hence, it would be equal to -10, instead of -5).

If we were searching for features that may characterize the outcome which appears to be the best possible on the basis of the previous line of reasoning, then the view of it which follows from the observation just made can be of help. For, it suggests that the selected outcome has a certain 'stability' property, which follows from the fact that *no player gains any advantage from changing her choice if the opponents are instead determined to stick to it*.

Now, notice that if we now re-examined the game matrix as if we never did it, and try to differentiate between good and bad outcomes on the basis of the stability property we have just singled out, then we would come to the same result we have obtained in section 2.5. For, not only the outcome *CC* has this stability property as we have just noticed, but is also the case that no other outcome has it. If we consider *NC*, for example, which corresponds to the situation where Alan does not confess while Bonnie does it, then it is convenient to Alan to change his mind if Bonnie sticks to confessing, since by confessing he scores -5 which is greater than -10 that he scores by not confessing. The same is true for Bonnie if outcome *CN* is taken into account instead. As far as outcome

NN is concerned finally, both Alan and Bonnie would get an advantage by changing their choices since their score is higher in outcomes *CN* and *NC* respectively[1].

Viceversa, if we have singled out an outcome which is stable in this sense, then it is not possible that we can find an outcome which is more convenient to stick to for one of the players in the sense of the reasoning we have been using in section 2.5. This should be clear by considering the case of a game in which we have two players, like the one we have been using so far.

The analysis from section 2.5 was made with the goal of determining which, among all possible choices of actions by a player was best in terms of the score it granted under the assumption that the opponent had made a certain choice (all possible choices of her having been considered to make the argument complete in this respect). From the viewpoint of the one we may refer to as player 1, since it is now clear that we are speaking of games, and whom we assume to be the one whose actions are displayed in the leftmost column of the diagram as Alan's actions are in our running example, this corresponds to determining which action is scorily the best reply in each row of the diagram. From the viewpoint of player 2 instead, this corresponds to do the same for each column of the matrix. Now, let us assume that an outcome which is stable in the sense we are considering has been isolated. Then, it is simply not possible that this is not coinciding with the combination of actions that results from the previous reasoning: for, assuming the outcome in question to be of the form AB, where A represents player 1's choice of action and B is the action chosen by player 2, then to assume this outcome to be not scorily the best for either of the two players would be the same as saying that there exists an action A', or an action B' (where the case in which both alternatives occur cannot be excluded as well), which allows player 1, and player 2 respectively, to score a higher payoff (i.e., such that the payoff that player 1 is granted in $A'B$ is greater than the one she gets in AB, or such that the payoff that player 2 is granted in AB' is greater than the one she obtains in AB); but then the outcome in question would not be stable in the sense we are here considering, since there would be an advantage for one, or both players to make a unilateral change in the choice of action to play in the end.

So, to analyze a game in matrix form in a fashion similar to the one we have used in section 2.5, which considers the players' viewpoint one by one, or to analyze it as we have done here, that is by analyzing each actions combination to locate the stable ones, gives the same result. Since the former line of reasoning allows one to identify for each player the actions

[1]In particular, Alan scores -2 in *NN*, but he would score 0, which is higher, if he confessed instead and Bonnie would not change her choice, that is if *CN* were the real outcome. The situation is symmetrical for Bonnie.

which are best for any given actions combination played by the others, this is also true for those outcomes which are isolated in virtue of the latter kind of scrutiny.

Now, having finally found a distinctive feature of candidate solutions to games in matrix form, there are (at least) two natural issues to address. As a matter of fact, one could ask in the first place whether solutions of this sort always exist. Secondly, whether solutions of this sort always exist as unique outcomes, or if it is possible that more than one of them may be present in one and the same game matrix instead.

3.2. Equilibria

Outcomes such as the combination of actions by Alan and Bonnie that corresponds to confessing for both in the previous situation, goes by the name of *equilibria*. The feature of solutions like the one we have singled out in our example, can be used to informally frame the concept by definition in the first stance:

Definition 3.1 *In a finite game (i.e., in a game with a finite number of players, each of which is given a finite number of possible actions), where players have to make their choice simultaneously, a combination of the players' actions is an* equilibrium *if and only if no player has benefit from changing her strategy unilaterally (i.e., without the other players changing theirs).*

The definition comes equipped with its own, easy method for determining which of all possible actions combinations in a given game in matrix form is an equilibrium in this sense of the expression: it consists in checking, for any combination of actions by the players, whether any of them gets advantages by changing her strategy on her own in terms of payoff. The method is exhaustive, or can be made as such, by checking all actions combinations in the order of their appearence in the matrix. Also, owing to the game being finite, the whole process is bound to reach an end in a finite number of steps. Neither of these two features of it, exhaustiveness and finiteness, help us under any respect with the two issues we are planning to address, namely existence and uniqueness of equilibria in all possible situations. To tackle these topics, we propose to follow a different route. We are going to specify a little bit the language we are using to speak of finite games in normal form and then show how the concept of equilibrium we have just defined can be seen to correspond to a different notion, that of fixpoint of a monotone operator, with respect to which we will be in a position to answer the said questions. Let us first start from giving a more exact shape to the notions we have been using so far.

Owing to what we have said, for the sake of the analysis of games by means of matrices it is required that: (i) the games be finite in terms of number of players and number of actions at their disposal; (ii) a distribution of payoffs be given among actions combinations that allows us to assess how convenient each of them might be for each player of the game; (iii) no assumption about what is the order of choices is made. This approach to games can be made general if we think of a game G in normal form to be given as a triple $\langle P_G, (\Sigma^i_G)_{p_i \in P_G}, u_G \rangle$ where[2]:

- $P_G = \{p_1, \ldots, p_n\}$ is a finite set of *players* of G;
- each $\Sigma^i_G = \{s^i_1, \ldots, s^i_{m_i}\}$ is the finite set of *actions* of player p_i in G;
- u_G is the function which distributes payoffs to players for any given combination of actions of theirs: let Σ_G the finite set of *strategy profiles* of G, namely the set whose elements are lists s of actions by the players of G such that s is *exhaustive*, i.e., for every player p_j of G, s features at least one action s^j_i by p_j, and *non-redundant*, i.e., for no player p_j of G, s features more than on of the actions of hers; then, u_G is a mathematical function of the form $u_G : P_G \times \Sigma_G \to \mathbb{Q}$, i.e., a right-hand unique relation between the set whose elements are pairs (p_i, s), where $p_i \in P_G$ and $s \in \Sigma_G$, and a rational number $u_G(p_i, s)$ representing the payoff of player p_i in case the strategy profile s be played by the players of G.

Notice that, owing to what we said above, elements of Σ_G can be thought of as having the following form:

$$s = s^1_{i_1} s^2_{i_2} \ldots s^n_{i_n}$$

That is, strategy profiles are lists of actions which are numbered by a superscript, indicating the player they are action of, and by a subscript indicating which action of those at the player's disposal it is.

Also observe that, at the level of generality we are aiming at achieving in this volume, the choice of the codomain of the utility function will play little role, hence different choices can be made in this respect. For

[2] For the untrained reader, the notation used here, in particular in the second item in the list, can be explained as follows: each player p_i of G has a finite set of actions among which she chooses what to play, like the set $\{s^i_1, \ldots, s^i_{m_i}\}$ where the common superscript indicates that these are all action of one and the same player p_i of G; as to the other index, this is related to the fact that it is possible that not all players in the game have one and the same number of actions to choose among. This happened to be the case in the game we have used so far, but it is not one of the features we have relied upon for the sake of the analysis of it. To make the general situation possible, the subscript m_i of action $s^i_{m_i}$ in this set, which is supposed to indicate the last action in the set of actions of p_i and carries an index that is equal to the total number of those actions, depends upon i itself since every such set Σ^i_G, which also depends upon i, will have 'its own' ending element $s^i_{m_i}$, i.e. its own number of elements this m_i corresponds to.

instance, the example we have been considering so far would have legitimate to stick to the set of integer numbers $\mathbb{Z}$. Other situations, may rather suggest that the utility function be chosen as getting values in the set of real numbers $\mathbb{R}$. The choice of $\mathbb{Q}$ is somehow intermediate between the two, due to the observations that we made back in section 1.4.

To make things easy, in the rest of the chapter we will confine ourselves to two-player games. This means that the object of our study will coincide with simpler triples than those discussed above for the sake of generality. This allows us to make some extra assumptions on the notation to use which avoid complications due to indices proliferation. For, since the set of players in this case can be assumed to be of the form $P = \{p_1, p_2\}$, we can also assume that a varies over actions by player p_1, while b does the same for actions by player p_2. Therefore, the two sets of actions of players of such a game in this case would look as follows:

$$\Sigma^1 = \{a_1, \ldots, a_n\}$$
$$\Sigma^2 = \{b_1, \ldots, b_m\}$$

In turn, the set Σ of strategy profiles for a simplified game as such will contain elements of the form $a_i b_j$ with $1 \leq i \leq n$ and $1 \leq b \leq m$. Since we plan to apply the convention of dropping the subscript «G» we have been using above to mark elements of one and the same triple representing the game G (as we have already started to do here), and let the context speak for that (as long as there is no risk of ambiguity), we can also think of simplifying the notation for the utility function and let its values be indicated as $u^1(a_i b_j)$, as long as we mean by that the utility for player 1 in case the combination of actions $a_i b_j$ be played, and as $u^2(a_i b_j)$ for the corresponding value of the utility function for player 2 instead.

Representing games in the form of the said triples has the advantage of making possible for us to speak precisely of ideas we have been referring to intuitively along the way in the analysis we have put forth in section 3.1. For instance, one can explain in clearer terms what is meant that a certain action is «more convenient» to a player than another. Let then G be a two-player game of the form $\langle P, (\Sigma^i)_{i \in \{1,2\}}, u \rangle$ where we assume that the previous conventions on the notation are at use. To analyze the situation from the viewpoint of player 1 in a way similar to what we did in section 2.5, means trying to determine which action player 1 is likely to choose for each possible choice of action made by the opponent. In turn, this depends upon the payoff distribution among strategy profiles in the following way: let b_j be an arbitrary but fixed element of Σ^2 (that is, let b_j indicate one of the actions at player's 2 disposal, and let it remain fixed throughout the argument); then, we say of any two actions of player 1, $a_i, a_h \in \Sigma^1$, that, say, a_i is more convenient than a_h (relative to b_j) if and only if $u^1(a_i b_j) > u^1(a_h b_j)$ (where $\cdot > \cdot$ is the usual «greater than» relation over elements of $\mathbb{Q}$ to which $u^1(a_i b_j)$ and $u^1(a_h b_j)$ belong). The same applies to player 2 and how choices are evaluated from her point of view. So, more convenient actions allows a player to achieve a greater payoff as replies to one and the same action chosen by the opponent, and players are supposed to asses choices on the basis of how payoffs are distributed, higher payoffs being preferred to lower ones (see also the discussion we made of this topic back in section 1.4). The whole idea is as simple as that.

Players are obviously expected to perform the said comparison between possible choices of actions for the purpose of determining the action of theirs which is *most* convenient in each case. Having specified how the comparison goes pairwise, the task in question clearly corresponds to identifyng the action that ensures a player the highest payoff, or, to say it in the previous terms of the comparison, the action which is more convenient than all the other alternatives to it with respect to one and the same choice of action by the opponent: a_i is most convenient to player 1 with respect to the choice of action b_j by player 2 if and only if $u^1(a_i b_j) > u^1(a_h b_j)$ is the case for every $a_h \in \Sigma^1$.

As it was said, each player is supposed to consider all possible cases and to perform the assessment of her own choices of action for any given choice of action by the opponent. This means that the assessment as a whole reads like the following scrutiny (which we exemplify by assuming to follow player 1's thread of thoughts, the one made by player 2 being obtained by means of the obvious modification of it):

«If b_1 is the choice of action by player 2, then ... is my choice of action»
«If b_2 is the choice of action by player 2, then ... is my choice of action»
...
«If b_m is the choice of action by player 2, then ... is my choice of action»

(where «...» is supposed to be substituted in each case by the action of player 1's which is most convenient in the previous sense of the expression). Actually, since players of games we consider are supposed to be rational agents who are performing choices in virtue of this rational character of theirs, they are assumed to single out choices which are rational. Therefore, a more appropriate rendition of the previous line of reasoning should rather read like:

«If b_1 is rational to player 2, then ... is rational to me»
«If b_2 is rational to player 2, then ... is rational to me»
...
«If b_m is rational to player 2, then ... is rational to me»

The connection between the players' obvious choice of action and the maximum payoff it ensures is so natural, that one can even imagine that any of the players in a game uses the previous reasoning also to disclose whether there is a clear choice by the opponent that is about to be made. That was actually the way in which the said argument was used first in the analysis of the game matrix we took as exemplar in section 2.5. By applying the reasoning in this way, the kind of utterances it would be made out of should rather be the dual ones to those we have previously considered. For instance, in the case of player 1 the argument in question should be something like the following:

«If a_1 is rational to me, then ... is rational to player 2»
«If a_2 is rational to me, then ... is rational to player 2»
$\dots$
«If a_n is rational to me, then ... is rational to player 2»

(where again «...» stay for the action that is most convenient to the opponent). The moral of all this is that this kind of reasoning, be that directed to disclosing which action of a player is best to choose in view of what is rational to do to the other player, or be that performed in view of determining what the other might be doing as a reply to the player's own choice, is supposed to help identifying the action which is most convenient with respect to the choice of action that is rational to the opponent. Let us try to make sense of the conclusion we have just reached by means of some modest use of formalism. For the sake of keeping this part readable, we are going to pursue the task in a way that be friendly to the reader untrained in the construction of formal languages, yet, at the same time, in a way that may suggest to the acquainted reader how things should be modified to reach the standard required to a proper mathematical logic work .

Let us try to devise a language, that we shall call $\mathscr{L}_{GM}$ henceforth (GM referring to «Game Matrix»), which could be used for the sake of literally expressing the very same reasoning we have been dealing with in a formal way. Actually, we need not to be generous with resources here, as the reasoning we aim at expressing is a parsimonious one. That reasoning speaks of players' actions in the first place, therefore the language expressing it should be equipped with means for doing that. Owing to the assumptions we have made about what kind of games we shall be concerned with, we may assume our language to contain two sorts of symbols for actions of player 1 and actions of player 2 respectively, which we may assume to coincide with those we have been using so far in the attempt to keep things simple, namely:

$$a_1, a_2, \dots, a_n, \dots$$
$$b_1, b_2, \dots, b_n, \dots$$

As this way of writing suggests, we may well assume our linguistic resources to be (denumerably) infinite in this respect in such a way that we have means for speaking of actions in any possible situation, although in each of them only a finite amount of symbols will be actually used. The reader acquainted with the use of formal methods should notice that these symbols will play the role of constants, i.e. proper names for actions in any two-player game G. We assume symbols of this sort to be juxtaposed pairwise to form names for strategy profiles, i.e. terms like $a_i b_j$ under the extra assumption that the term for action of player 1 should come prior in the juxtaposition to the name for the action of player 2.

We would like to stress for the untrained reader's sake that by calling expressions like a_i, b_j, and $a_i b_j$ as *terms*, it is because this is what they are now, that is, symbols of a linguistic apparatus that count as names for 'objects' which are in this case actions and actions combinations of players in a finite game. In particular, these symbols are part of what we, or the players themselves express in the form of the previous sentences.

Now, those very same sentences speak of *rational* actions to be more precise. Therefore, if we are willing to set up the language $\mathscr{L}_{GM}$ in such a way that those utterances become sentences of it, we have to equip it with the suitable resources. This means in particular that we have to add a symbol, say $R(\cdot)$, for the property in question, i.e. such that it applies to actions (to action terms to be precise) to produce constructions like $R(a_i)$, $R(b_j)$, which will count as sentences expressing the fact that: «Action a_i is rational» and «Action b_j is rational» respectively.

To be precise, in the propositions our sample reasoning is broken into, rationality seems to be a feature that depends upon players, which may suggest to introduce two different predicates of rationality (or, to make the one we assume to have relative to the players names). This however would be odd under a different respect, since it would suggest that players of our games might have different characters of rationality. It seems safer to assume that we have just one idea of rationality for actions, and that this idea is made relative to players only as an effect of the players speaking of actions made by someone being rational.

Another observation about reference to rationality of actions in the above statements is on order. For, while the first occurrence of it (the one that belongs to the antecedent of the «If..., then...» sentences) is totally independent of the payoffs distributions in the game, being it the effect of a blind assumption by the player who utters the statement in question, the second (which appears in the consequent of those conditional sentences) is not, since it comes out of the player making an evaluation of actions to determine which, in view of the hypothesis she has made, is most convenient, i.e. ensures the maximum payoff. This brings us to another conclusion about resources that should be added to those already available in language $\mathscr{L}_{GM}$. For, its sentences too will speak of actions being most convenient, since we are designing them to be the formal counterpart of those we are using as models, hence of payoffs and payoffs comparison. We follow here the previous prescription about keeping things simple, and we incorporate in our language $\mathscr{L}_{GM}$ the very same resources we have been using in this respect as they appeared informally: we therefore assume $\mathscr{L}_{GM}$ to have terms constructors for payoffs $u^1(\cdot), u^2(\cdot)$, which take as input terms for strategy profiles and indicate the payoff granted to player 1 and to player 2 respectively in the case of the actions combination these terms refer to. So, for instance, we take $u^1(a_i b_j)$ to be the term of our language indicating the payoff for player 1 if actions a_i and b_j are played, while $u^2(a_i b_j)$ does the same for player 2's payoff in the same situation. In addition, we require that $\mathscr{L}_{GM}$ has a binary relation symbol $\cdot > \cdot$ for payoffs comparison, which, for any given terms a_i, a_h and b_j, b_k for the players' actions, gives rise to formulas of $\mathscr{L}_{GM}$ of the form $u^n(a_i b_j) > u^m(a_h b_k)$ (where n, m can be equal to either 1 or 2, and the case in which they are equal is not excluded). The intended meaning of such formulas is to express the fact that payoff granted to player n in the strategy profile $a_i b_j$ is greater than the payoff granted to player m in $a_h b_k$.

It is clear that in the impossibility of determining which of the choice of actions at hers, or at the opponent's disposal is the rational one, the reasoning we are trying to replicate at the level of the language we are building is bound to remain hypothetical in character. This was not so in the example we have considered in section 3.1, for one precise reason: because by performing the argument in that case, it turned out that one and only one action stood out as the one to choose (i.e., confessing for both players). This was clearly an effect of the payoffs distribution, which caused, under both the hypothesis that the other may confess as well as in the alternative scenario that she may not, confessing to be more convenient than not confessing for either of the two players. That is, payoffs are distributed in the game that we have discussed in sections 2.5 and 3.1 above in such a way that the following is verified: *for every choice of the opponent, there is always a unique choice for the other player that ensures her the maximum payoffs*. Actually, in the situation we have been considering even more is true. That is, it turns out that *there exists a unique choice by any player* (to confess), *which ensures her the maximum payoff against every choice of action by the opponent*. The previous feature is just a side effect of this stronger property that the game possesses. In view of the latter, the action to substitute ellipses in the hypothetical sentences above is one and the same action. That is, if we assumed games to always replicate this property there would be an action for, say, player 1 (let it be indicated as $a^\star$) such that the following would turn out to be the case:

«If b_1 is rational to player 2, then $a^\star$ is rational to me»
«If b_2 is rational to player 2, then $a^\star$ is rational to me»
...
«If b_m is rational to player 2, then $a^\star$ is rational to me»

The same would also be true for player 2, and one action of hers $b^\star$.

In case only the other property were verified instead, then there would still always be one action to substitute ellipses in the sentences, but that would not be necessarily one and the same in all of the cases. In other words, to exemplify the situation by means of the viewpoint of player 1 again, for every action b_j of player 2 there would be a unique action $a(j) \in \Sigma_1$ (where the notation is used to indicate explicitly the dependency of the action by player 1 in question on the chosen action by player 2) ensuring the maximum payoff to player 1. Then, the previous reasoning would go in the following way:

«If b_1 is rational to player 2, then $a(1)$ is rational to me»
«If b_2 is rational to player 2, then $a(2)$ is rational to me»
...
«If b_m is rational to player 2, then $a(m)$ is rational to me»

(where the fact that all actions $a(1), a(2), \ldots, a(m)$ might be different from one another is the novel aspect with respect to the situation we were considering previously).

As a first attempt to generalize the considerations we have made on the game we took as an example, we may well require that this latter property be featured, but, to avoid our assumption to be too much controversial, we may also confine ourselves to that and let the former, the stronger one which holds in the game we took as our reference example, be not featured instead. This does not rule out that the more peculiar situation that the stronger property of payoffs distribution leads to be possible, although only as a special case of what we assume should be always verified. This will cause our analysis to be more general, in the sense of being applicable also to cases in which it fails and only the weaker one holds instead.

Now, the easiest way to explain in which terms payoffs must be distributed for the assumption to be verified in a game, that is, to make true that for every choice of action by the opponent there exists a most convenient reply by any of the player, is to say that ties in payoffs are not allowed for actions counting as reply of any player to one and the same action chosen by the opponent. That is, no two actions of any player in a game ensure one and the same payoff to her as replies to a choice of action made by the opponent. Let us formulate this property in more precise terms and, due to the fact that it corresponds to payoffs being strictly ordered along rows and columns of the game matrix (see also section 1.2 about strict ordering), let us introduce the class of *strict games* thereby:

Definition 3.2 *Let* $G = \langle \{p_1, p_2\}, (\Sigma^i_G)_{i\in\{1,2\}}, u_G \rangle$ *be a two-player game in normal form. We say that* G *is* strict *if and only if* $u^1(a_i b_j) \neq u^1(a_h b_j)$, *for every* $a_i, a_h \in \Sigma^1$ *and for every* $b_j \in \Sigma^2$, *and* $u^2(a_i b_j) \neq u^2(a_i b_k)$, *for every* $a_i \in \Sigma^1$ *and for every* $b_j, b_k \in \Sigma^2$.

The strictness property as it is formulated here amounts to the assumption that payoffs are distributed among players over strategy profiles in such a way that ties are avoided as we said, i.e., for no player of the game, and for no actions combination there are two or more options that are equally valuable to her. Owing to payoffs being represented by rational numbers, it is just a matter of routine to verify that, having presumed that a game is strict in the sense of definition 3.2 above, then each player of the game can count, for each possible actions combination, on a (unique) action ensuring her the maximum payoff:

Lemma 3.1 *Let $G = \langle\{p_1, p_2\}, (\Sigma^i_G)_{i\in\{1,2\}}, u_G\rangle$ be a strict game in normal form. Then: (i) for every $a_i \in \Sigma^1$ there exists a corresponding action by player 2, say $b(i) \in \Sigma^2$, such that $u^2(a_i b(i)) > u^2(a_i b_j)$ for every $b_j \in \Sigma^2$ with $b(i) \neq b_j$;*[3] *(ii) for every $b_j \in \Sigma^2$ there exists a corresponding action $a(j) \in \Sigma^1$ such that $u^1(a(j) b_j) > u^1(a_i b_j)$ for every $a_i \in \Sigma^1$ with $a(j) \neq a_i$.*

Proof: owing to the set $\mathbb{Q}$ being equipped with its own partial ordering relation $\geq$, we have that $p \leq q$, or $q > p$ is the case for every $p, q \in \mathbb{Q}$ (which is a shorthand for: either $p = q$, or $p > q$, or $q > p$ for every $p, q \in \mathbb{Q}$). This holds in particular for elements of $\mathbb{Q}$ which are values of the utility function u_G of G. Therefore, if a_i is any element of Σ^1 and so are $b_j, b_k \in \Sigma^2$ we have either $u^2(a_i b_j) = u^2(a_i b_k)$, or $u^2(a_i b_j) > u^2(a_i b_k)$, or $u^2(a_i b_k) > u^2(a_i b_j)$. However, since G is strict, only $u^2(a_i b_j) > u^2(a_i b_k)$, or $u^2(a_i b_k) > u^2(a_i b_j)$ are the possible cases. This means that the set

$$\{u^2(a_i b_j) : 1 \leq j \leq m\}$$

necessarily has a maximum element (for, otherwise strictness would fail). Let this be $b(i)$ and part (i) of the lemma follows.

Part (ii) is proved by means of an identical argument, which then gives the desired result. QED

What does this entail at the level of the language $\mathscr{L}_{GM}$ we have been building along the way? In a game where payoffs are distributed according to the strictness prescription, for every choice of action made by the opponent, there is always a unique action ensuring the maximum payoff to each of the two players, i.e., one action that stands out as the «most

[3] Since this action $b(i)$ belongs to the set Σ^2 of actions at player 2's disposal in game G, and since actions of players in games in normal form are assumed to be given as lists, $b(i)$ must take a certain place in the order that produces the list Σ^2 of actions, i.e., it corresponds to an action of the form b_k of that list. To assume that the payoff for player 2 associated with the strategy profile $a_i b(i)$ is greater than the one associated to $a_i b_j$ for every $b(i) \neq b_j$, is the same as assuming this that this holds for every $b_j \in \Sigma^2$ such that $j \neq k$. A similar remarks applies to action $a(j)$ of player 1 mentioned in the second half of the statement of the lemma.

convenient» one to each player. Owing to the previous analysis, the action that is the most convenient reply to the rational action by the opponent is the one that players are trying to disclose by the reasoning we have considered. This is now something we are capable to express by means of a formula of the language $\mathscr{L}_{GM}$, if we just assume that a minimum amount of logical resources are also available within this language.

Let us suppose that $\mathscr{L}_{GM}$ is equipped with the standard means for expressing conjunctions and disjunctions, that is, let us assume that $\mathscr{L}_{GM}$ has symbols $\wedge$, $\vee$ for logical connectives, i.e., operations on formulas of $\mathscr{L}_{GM}$ that can be used to give rise to expressions $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$ out of any two formulas φ, ψ of $\mathscr{L}_{GM}$, with parentheses being used as customary to mark the scope of the connectives, counting as new formulas with the intended meaning: «φ and ψ» for $(\varphi \wedge \psi)$, and «φ or ψ» for $(\varphi \wedge \psi)$.

On the basis of this further assumption, the following formula of $\mathscr{L}_{GM}$

$$(R(b_1) \wedge u^1(a_1 b_1) > u^1(a_2 b_1))$$

is intended to mean that b_1 is rational, hence it is rational to player 2 owing to what we have said above about all players in our game being subject to one and the same concept of rationality, and action a_1 ensures a greater payoff to player 1 if it is the chosen reply of her to action b_1 of player 2, than action a_2. Furthermore, the formula

$$(R(b_1) \wedge u^1(a_1 b_1) > u^1(a_2 b_1) \wedge \ldots \wedge u^1(a_1 b_1) > u^1(a_n b_1))$$

(where ellipses stay for the conjunction of all formulas of the form $u^1(a_1 b_1) > u^1(a_i b_1)$ for every $3 \leq i \leq (n-1)$, where n is supposed to indicate the number of actions at player 1's disposal in G), is intended to express the fact that b_1 is rational to player 2 and a_1 ensures to player 1 the maximum payoff (i.e., a_1 is the reply to b_1 that is the most convenient to player 1).

Let us introduce a convention on the notation that might be suitable for concisely refer to an expression such as the one above in the general case. Let then

$$\bigwedge_{2 \leq i \leq n} u^1(a_1 b_1) > u^1(a_i b_1)$$

be an abbreviation for the long conjunction

$$(u^1(a_1 b_1) > u^1(a_2 b_1) \wedge \ldots \wedge u^1(a_1 b_1) > u^1(a_n b_1))$$

that appeared in the formula of $\mathscr{L}_{GM}$ above. Now, owing to the assumption on what are the logical resources of $\mathscr{L}_{GM}$ and to what we have just agreed upon, take the following expression of $\mathscr{L}_{GM}$:

$$(R(b_1) \bigwedge_{2 \leq i \leq n} u^1(a_1 b_1) > u^1(a_i b_1)) \vee \ldots \vee (R(b_m) \bigwedge_{2 \leq i \leq n} u^1(a_1 b_m) > u^1(a_i b_m))$$

where, similarly to the previous situation, ellipses substitute the disjunction of all formulas of $\mathscr{L}_{GM}$ of the form

$$(R(b_j) \bigwedge_{2 \le i \le n} u^1(a_1 b_j) > u^1(a_i b_j))$$

for every $2 \le j \le (m-1)$, m being the number of possible actions by player 2 in G. That is a formula of $\mathscr{L}_{GM}$ which expresses the fact that b_1 is rational to player 2 and a_1 is the most convenient reply of player 1 to it, or b_2 is rational to player 2 and a_1 is similarly the most convenient reply to it, or etc., up to action b_m. That is, it is a mean for expressing within the language $\mathscr{L}_{GM}$, that a_1 is the most convenient reply of player 1 to *all* choices of action made by player 2. Owing to the idea according to which «rational» is a property that applies to actions that are most convenient, this formula expresses the fact that a_1 is rational to player 1.

Let us similarly abbreviate the long disjunction featured in the expression just displayed, by means of:

$$\bigvee_{1 \le j \le m} (R(b_j) \bigwedge_{2 \le i \le n} u^2(a_1 b_j) > u^2(a_i b_j))$$

This is a formula of $\mathscr{L}_{GM}$, call it φ_1, that, as we have just concluded, expresses the fact that a_1 is rational to player 1 in a given game G in normal form.

Let us further suppose that we have produced formulas φ_2, ..., φ_n by making the obvious changes to φ_1 above, each formula φ_i of which expresses the fact that action a_i is rational to player 1.

Let us also suppose that we have similarly devised a list of formulas ψ_j of $\mathscr{L}_{GM}$, each of which expresses the fact that the action b_j of player 2 is rational to her. Granted the above conventions on the notation, this can be done by assuming each of these ψ_j's to be of the form:

$$\bigvee_{1 \le i \le n} (R(a_i) \bigwedge_{2 \le h \le m} u^1(a_i b_j) > u^1(a_i b_h))$$

Before proceeding in our attempt to answering the questions about solutions to finite games in matrix form we raised, let us try to make clear what this development of linguistic resources has allowed us to gain in this respect.

Having formulated formulas φ_i's and ψ_j's puts us in a position to have made precise under what circumstances an action turns out to be rational to play to players of a game matrix G in the previous sense of the expression. As a matter of fact, if any of the situations concerning the distribution of payoffs described by formulas $\varphi_1, \ldots, \varphi_n$ occurred, then the corresponding action would be rational to player 1 in the sense of being the most convenient to her. On the other hand, if the distribution of payoffs over strategy profiles in a game matrix G resembled the situation described by one of the formulas $\psi_1, \dots, \psi_m$, then the corresponding action would be rational to player 2.

Notice that having supposed to deal with games that are finite and also strict, means that we know for sure that one and only one of the formulas φ_i, as well as one and only one of the formulas ψ_j correctly expresses the way in which payoffs are distributed over strategy profiles of the game under scrutiny. As a matter of fact, owing to a game being strict, we can always say that for every action of a player, say b_j of player 2, there always exists one and only one action of player 1, say a_i, that is best as long as the payoff it grants is concerned (as lemma 3.1 above states). Hence, φ_i would correctly express this fact. Symmetrically, the same would hold true for a certain formula ψ_k for every choice of action made among those at player's one disposal. Since ties are avoided, due to the game being strict, and since the ordering $>$ of $\mathbb{Q}$ is non-reversible (as this was made precise for strict kinds of orderings in section 1.2), this makes impossible that 'competing' formulas of φ_i and ψ_k, say φ_h and ψ_l, be a correct description of the situation taking place in G as far as the payoffs distribution over the relevant outcomes is concerned. However, do notice that this is still different from saying that an action is the most convenient in absolute terms, for, validity of any of the formulas φ_i's and ψ_j's only entails the existence of an action of player 1 and player 2 respectively that is most convenient relatively to one action that is rational to the opponent (this action being possibly different for different assumptions concerning what is the latter, rational action of the other player), rather than the existence for both players of one and the same action which is best *independently* of which action is chosen by the opponent. As it was said before going throught the construction of language $\mathscr{L}_{GM}$, strictness embodies only the former (weaker) condition, and still leaves open whether also the latter (stronger) condition holds.

Having clarified that, let us resume the previous discussion from the conclusion we reached about rational actions in a (two-player) game in normal form G being those whose corresponding formulas among

$$\varphi_1, \dots, \varphi_n, \psi_1, \dots, \psi_m$$

are «valid», in the sense of being a correct description of how payoffs of players are distributed among strategy profiles of G. Since game matrices we are dealing with are finite in the sense of always featuring a finite number of actions at each player's disposal, we can say of any action what condition it should satisfy in order to be the rational one in a game in normal form G. In particular, it should be the one whose corresponding formula is valid in the above sense. Now, since these formulas are finitely many we can easily express that through a case-distinction type of scrutiny: for, it is either the first action of player 1 in case φ_1 holds, or it is the second action by player 1 because φ_2 holds insted, etc. This analysis can also be reflected by means of the linguistic resources of $\mathscr{L}_{GM}$, provided those we have equipped it with are modestly enhanced. In particular, it is required that we have means to refer to actions in general. This is something that is usually coped with at the level of formal languages by *variables* that indicate an unknown, hence arbitrary element out of a given domain. In the case of our language $\mathscr{L}_{GM}$, we therefore suppose that we have variables symbols $x, y, z, \ldots$ for (arbitrary) actions. In addition to that, we should be able to say that a certain action, be that arbitrary or not, 'coincides' with another (for reasons that will be clear in a minute). We then equip the language $\mathscr{L}_{GM}$ of a suitable symbol for expressing that, in the form of a binary relation for which we use the identity symbol $\cdot = \cdot$. This will provide us with the means for stating formulas such that $x = a_i$, or $x = b_j$, as well as formulas $a_i = b_j$ for every variable x and terms for actions a_i, b_j.

Finally, the previous, informal assessment of what it means for a (generic) action to be rational in a game in normal form G, can be given the form of a *definition* by making use of the previous conventions on the notation as follows:

$$R(x) \Leftrightarrow_{Def} \bigvee_{1 \leq i \leq n} (x = a_i \wedge \varphi_i) \bigvee_{1 \leq j \leq m} (x = b_j \wedge \psi_j)$$

The above expression should be read as follows: The property of «being rational» applies to any action x in a game in normal form G, just in case x coincides with a_1 and φ_1 is valid, i.e., truly describes the situation concerning payoffs distribution in G, or x coincides with a_2 and φ_2 is valid instead, or ...” (dots being replaced by a similar reading of the remaining disjuncts up to the last one where x is set to coincide with b_m).

In turn, the definition we have just given reflects a quite natural way to analyze the outcomes of a game played by two players that we have proposed as being tied up with the players' own attempt to detect the rational action out of the finitely many they have to choose among. Yet, it is unclear how this is related to the idea of equilibria as we defined them at the beginning of the section, and whether it is of any help with the two issues that were raised about them, namely whether it is always the case that such equilibria appear in any finite game in normal form, and whether it is always the case that a unique equilibrium shows up in all situations. The goal of the next section is precisely to provide the missing connection and start addressing those issues.

3.3. Getting rid of the loop on rationality of actions

In the previous section, we have tried to generalize the approach to the example of game in normal form that we pursued in section 2.5, and that sounded as a reasonable way to assess the variety of actions combinations in order to determine the most convenient one. The process started with the 'solution' of the matrix that we proposed in section 3.1, and brought to the definition of equilibria as 'stable' outcomes in a game in normal form. The reasoning that led us to that was further analyzed, and used to devise a formal language as a new tool for pursuing the assessment of finite games in normal form. By making use of this language, we managed to set up a definition of what is rational to players to choose by following the natural idea that players are likely to stick to actions that are most convenient to them. Actions that are rational in the sense of our definition above, as a matter of fact, are the most convenient reply, in terms of the payoff they ensure, to the action that is rational to the opponent.

Now, regarding the two ideas we have been working on so far, the idea of equilibria as stable outcomes, and the idea of rational actions as most convenient choices, it would be desirable that they coincided in the sense that stable outcomes turned out to be made out of actions that are rational in the latter respect, and, viceversa, that pairs of actions which are rational to the players could be proved to give rise to equilibria. The expectation is encouraged by an informal assessment of the former direction of it: if an outcome is 'stable' in the sense in which equilibria are assumed to be owing to definition 3.1, then they are made out by actions that are most convenient as rational actions should be due to the fact that, by keeping the action of the opponent fixed, the other one in the equilibrium ensures to the player who plays it the highest possible payoff (hence, makes it unconvenient to her a unilateral change in the chosen strategy, as definition 3.1 requires).

This way of putting things, however, can only serve as an indication, as an encouragement as we said, since for an action to be the most convenient one to a player is only one side of the feature it has to have in order to be rational in the sense of the definition we have come up with above. For, the definition requires that at least one among the formulas φ_i's and the ψ_j's be 'valid', in the informal sense of validity we have explained above. In turn, this requires that an action is the most convenient reply to the action that is rational to the opponent, in the same sense of the expression. That is the point where things become difficult, for it seems that to state which action is rational to a player in the sense of the defining condition we managed to extract from the analysis of games in matrix form, one must have first determined which action is rational to the opponent. To do this, by the way, it is similarly required that we know which action is rational to the other player, which (owing to the fact that we are dealing with two-player games) brings us back to the point where we started.

It seems we are facing a problem here, and there is no easy way out. Despite how natural appeared the path leading to it, the definition of «rational action» in two-player, strict games in normal form turns out to be unusable owing to the cycle it traps us into and which seems to be due to the fact that the property we are trying to assign to actions in a game, presupposes that actions that possess it are known already. Luckily there is a way out, which requires that some further reflections upon the language $\mathscr{L}_{GM}$ we have devised in the previous section be made.

The first thing that will turn out useful is to provide the reader with some more details about the language $\mathscr{L}_{GM}$ and the property of «validity» for formulas of it, that we have been sloppy about so far.

From the point of view of its syntax, the language $\mathscr{L}_{GM}$ we have been devising comprises the following types of expressions:

$$x \mid a_i \mid b_j \mid a_i b_j \mid u^{1/2}(a_i b_j) \mid R(\cdot) \mid \cdot = \cdot \mid \cdot > \cdot \mid \varphi \wedge \psi \mid \varphi \vee \psi$$

that is: terms of $\mathscr{L}_{GM}$ are either variables, terms for actions of either player 1 or player 2 of a given two-player game in normal form G, terms for strategy profiles of G, and terms for payoffs assigned to either player 1 or player 2; formulas of $\mathscr{L}_{GM}$ are either atomic formulas, which take three possible forms: (i) $R(a_i)$, $R(b_j)$, (ii) $t = a_i$, $t = b_j$, where t is either a variable or a term for action, (iii) $s > t$, where both t and s are terms for payoffs; or, they are obtained from formulas of $\mathscr{L}_{GM}$ by applying conjunction $\wedge$, or disjunction $\vee$.

The purpose of setting up this language was to speak of a game in normal form in the way we have imagined that the very players of it might be doing while analyzing all outcomes in search of the best thing to do. In particular, statements of this language are conceived in such a way that they refer to state of affairs involving features of the game matrix such as payoffs distribution over outcomes and their evaluation. The connection between formulas and states of affairs they are supposed to refer to, is usually made clear in formal languages like ours by means of a suitably specified validity relation. In turn, this validity relation comes in the form of a list of clauses, each one of which states precisely at which condition every formula of the language is valid, depending on what is its logical form (therefore, there should be a clause for each possible form that formulas of the language under consideration can take).

The previously noted difficulty about an expression of it of the form $R(t)$, where t is a term for either an action of player 1, or an action of player 2, seems to represent a stumbling block in the definition of such validity relation in the case of $\mathscr{L}_{GM}$. For, as we informally discussed the issue in section 3.2, the intended meaning of a formula $R(t)$ of the said sort is the assertion that «action named t is rational (to the player it is an action of)». The validity relation should then certify this by means of a dedicated clause stating that the formula $R(t)$ is valid, or «holds true», just in case the fact it is intended to express occur, that is in case the action named t is rational. The problem is, as we have briefly noted already, that what does the fact in question amount to is hard to say. For, according to the definition we have come up with in section 2, and which we reproduce here for the reader's sake

$$R(x) \Leftrightarrow_{Def} \bigvee_{1\leq i\leq n} (x = a_i \wedge \varphi_i) \bigvee_{1\leq j\leq m} (x = b_j \wedge \psi_j)$$

a formula $R(t)$ is supposed to hold if the term t of $\mathscr{L}_{GM}$ produces a true instance of the defining condition when substituted to the variable x, i.e. in case

$$\bigvee_{1\leq i\leq n} (t = a_i \wedge \varphi_i) \bigvee_{1\leq j\leq m} (t = b_j \wedge \psi_j)$$

which is also a formula of $\mathscr{L}_{GM}$, is valid in the above sense. If we suppose to having developed a definition of validity for formulas of $\mathscr{L}_{GM}$ that follows the standard treatment of logical operations as we would like to, then for such formula to be valid means that t is in the first place a term for action of either the form a_i (that is, a term for an action of player 1), or the form b_j (a term for an action of player 2), since only terms of this form would produce valid instances of formulas of $\mathscr{L}_{GM}$ of the form $t = a_i$ or $t = b_j$, which is part of what must happen to make $R(t)$ valid as we want. Moreover, it should also happen that at least one of the formulas φ_i, ψ_j be equally valid. However, as the reader may remember from section 3.2, both kinds of formulas feature occurrences of formulas of the form $R(s)$, where s is a term for an action of the opponent, which are then supposed to be valid to make either the expression φ_i or ψ_j they belong to, valid itself. In sum, this is the cycle we were referring to above: for a formula $R(t)$ to be valid, it is required that another formula $R(s)$ of the same form is known as valid already.

Now, the reader should remember that we arrived to the definition above as an attempt to capture the reasoning performed by the players in the attempt of evaluating outcomes of a game G. This reasoning was hypothetical in form, as it was meant to comprise utterances which, as far as player 1 was concerned to give an example of it, read like the following ones:

«If b_1 is rational to player 2, then $a(1)$ is rational to me»
«If b_2 is rational to player 2, then $a(2)$ is rational to me»
...
«If b_m is rational to player 2, then $a(m)$ is rational to me»

This is the reason for the previously mentioned occurrence of a formula of the form $R(s)$ in the defining condition for $R(t)$. By the way, it should be also noticed that, if such hypotheses were actually made, the cycle preventing us from evaluating $R(t)$ would disappear. Let us illustrate this last observation in a simpler situation.

Let us consider a game in normal form like the one we have been using as starting point and that was inspired to the story about Alan and Bonnie. Let then G be a two-player game in normal form where each player is given two actions to choose among. Only a fragment of the language $\mathscr{L}_{GM}$ we have defined is needed here. In particular, the part of it that features terms for actions a_1, a_2 and $b_1 b_2$, which give rise to terms for strategy profiles $a_1 b_1, a_1 b_2, a_2 b_1, a_2 b_2$. In turn, terms of this latter form determine the variety of terms for payoffs $u^n(t)$, where n is either 1 or 2, and t is a term for a strategy profile of G.

The definition of «rational action» in this simpler case has the same logical form of the general schema of it above, although it turns out to be less complex in the number of its components. In the sample situation we have chosen, they are so few that they can even be written down explicitly:

$$R(x) \Leftrightarrow_{Def} (x = a_1 \land \varphi_1) \lor (x = a_2 \land \varphi_2) \lor (x = b_1 \land \psi_1) \lor (x = b_2 \land \psi_2)$$

In turn, formulas φ_i and ψ_j also reflect the simple character of the situation. For instance, φ_1, which expresses the fact that action a_1 is rational to player 1, reads as follows:

$$(R(b_1) \land u^1(a_1 b_1) > u^1(a_2 b_1)) \lor (R(b_2) \land u^1(a_1 b_2) > u^1(a_2 b_2))$$

Now, in easier situations like that, imagine that a hypothesis has been made as to what is rational to do to the two players. The hypothesis needs not to be grounded on the game matrix evaluation. It may just reflect a feeling, or an intuition. Suppose then that someone hypothesized that a_1 is rational to player 1 and b_1 is rational to player 2.

Suppose also that payoffs have been distributed over outcomes, i.e. strategy profiles of G, in such a way that both $u^1(a_2 b_1) > u^1(a_1 b_1)$ and $u^2(a_1 b_1) > u^2(a_1 b_2)$ are actually the cases, corresponding to a_2 being the most convenient reply to action b_1 by player 2 to player 1, and b_1 being the most convenient choice of action by player 2 if player 1 plays a_1. This payoff distribution also causes the formulas $u^1(a_2 b_1) > u^1(a_1 b_1)$ and $u^2(a_1 b_1) > u^2(a_1 b_2)$, which express those very same facts, to be valid in the informal sense of the expression we are still referring to (that is, in the sense of being true descriptions of what really takes place in the game).

The two features we are supposing to be verified, the hypothesis of what is rational to do to the players and the distribution of payoffs, would then lead to

$$(R(b_1) \land u^1(a_2 b_1) > u^1(a_1 b_1))$$

and

$$(R(a_1) \land u^2(a_1 b_1) > u^2(a_1 b_2))$$

being formulas of $\mathscr{L}_{GM}$ corresponding to valid utterances. In turn, owing to what we said in section 3.2 about the intended meaning of logically complex formulas of $\mathscr{L}_{GM}$, this fact makes also valid formulas φ_2 and ψ_1, as we have decided to name them, since the two formulas displayed above happen to be disjuncts of them, and we informally assume of a disjunction that it is valid whenever at least one of its disjunct is. This means, in the end, that the supposition about what is rational to do for the players involved into game G, together with the actual distribution of payoffs, allows us to make valid two other formulas of $\mathscr{L}_{GM}$, namely

$$(a_2 = a_2 \wedge \varphi_2)$$

and

$$(b_1 = b_1 \wedge \psi_1)$$

Even though we are still reasoning at the informal level, the reason why this is so should turn out to be evident. It has to do of course with the intended interpretation of the conjunction that we are willing to pursue, according to which a formula of that logical form is valid whenever both of its conjuncts are valid. It has to do with φ_2 and ψ_1 turning out to be valid under the hypotheses made. Finally, it has to do with an obvious constraint that we may want to put over formulas of the form $s = t$, which are supposed to convey the idea that s and t coincide, that is, that the action named by s and the action named by t are the same. Whatever is this notion of coincidence for actions one may have in mind, one thing about it is right for sure: it certainly applies when s and t are terms naming one and the same action, which makes every instance of formulas of this type of the form $s = s$ obviously valid.

Now, this leads us to one further observation: that, by what we have just noticed, we are indeed capable of making valid two instances of the definition of rationality we came up with. The instances in question are those which are obtained by substituting everywhere in the defining condition a_2 for x on the one hand, and b_1 for x on the other. Which means that formulas $R(a_2)$ and $R(b_1)$ turn out to be valid in the end, which means that actions a_2 and b_1 can be pronounced to be rational to player 1 and to player 2 respectively, under the hypothesis that a_1 and b_1 are rational to them.

To make pronouncements of this sort on the basis of the definition we devised in section 3.2 was what appeared to us to be impossible in view of the cycle we found ourselves stuck into. It turned out in the end that this cycle is illusory, as well as the said impossibility, provided we are given hypotheses as to what is rational first. In the next section, we shall make use of this new conviction of ours, and give it the form of a method. A method by using which we are going to finally answer the two issues that were raised, and which, to be honest, is about time to tackle.

3.4. Hypotheses and their revision

Let us first flag a few features of the process which uses hypotheses about what is rational to do for players in order to disclose what turns out to be so on the basis of the definition of «rational action» we have found.

First of all, it should be noticed that the hypothesis required by that takes the form of a strategy profile. This is a general feature, but is obviously more evident in the two-player case. In the model of reasoning we have analyzed in section 3.2, players start from what might be rational to the opponent to do and try determining what turns out to be rational to them on the basis of the payoffs distribution. The defining condition of the predicate marking rational actions features occurrences of formulas in which the rational predicate R itself is applied to terms for actions of all of the players involved in the given game. Therefore, the hypothesis one needs in order to determine which other action by a player turns out to be the one that produces a valid instance of the defining condition of the rationality predicate, should comprise a complete list of actions that are supposed to be rational to all of the player's opponents.

A second observation to make concerns the result of the previous solution of the difficulty related to the definition of rational action. As a matter of fact, having begun with the hypothesis concerning what is rational to do to players of a game, one ends with a list of actions that are rational in the sense of producing a valid instance of the condition defining the rationality predicate R which also can take the form of a strategy profile, as it has then the same character of a complete list of actions of the hypothesis we have started from.

A third remark concerns the fact that in the course of making use of hypotheses to get rid of the original cycle we had spotted in the condition defining the property of being «rational» for an action, we have heavily referred to the validity property for formulas of $\mathscr{L}_{GM}$. This is for a very good reason since the two things, escaping the cycle and finding a way to state when the property we are concerned with applies, go on together. Actually, the very same trick that allows us to accomplish the former task, also allows us to accomplish the second: if we are given hypotheses as to what is rational to do for players, we also are in a position to say when it is the case that certain instances of the condition defining the predicate $R(x)$ is valid in the informal sense of «correctly reporting a state of affairs involving the game G we suppose to be given, and payoffs distribution over its outcomes». We know that because this coincidence is what we have been making use of in section 3.3. Now, we want to make this way of proceeding precise and what we need here is to make precise in the first place the definition of what it means for a formula of $\mathscr{L}_{GM}$ to be valid.

This conclusion and the form it will take should sound terribly obvious to the reader acquantied with formal languages and the way they are normally dealt with. However, since this book is also written for the untrained scholar, I feel compelled to spend a few words about what is about to happen at the cost of appearing naive to those who are knowledgeable (and who may well wish to skip this passage).

As I have tried to explain already, formulas of our language $\mathscr{L}_{GM}$ are conceived to make pronouncements about the game G under scrutiny. They are in this sense not different from statements of our natural languages which are conceived to let us speak of something going on in the physical world or in some other 'domain', to the state of affairs of which we may wish to refer. How this correspondence actually takes place is easily explained with respect to formulas of $\mathscr{L}_{GM}$, which are made out of terms, the intended meaning of which is to name either actions of the players, strategy profiles, or payoffs, and may take either the form of «simple formulas» (*atomic formulas* is the way they are named in the usual account of formal languages), which correspond in our case to formulas that intend to express properties of 'objects' named by terms or relations holding between them, or «complex formulas», which are built out of the simple ones by the use of logical operations and which, so to say, impose a 'relationship', of the logical kind, between the state of affairs referred to by their components. Now, to make this relation of meaning between formulas and state of affairs clear, as well as the validity property that follows it, we provide the reader with a couple of definitions.

The first one is supposed to state precisely what do terms of $\mathscr{L}_{GM}$ refer to (read: mean) in the context of a given game G in normal form. For the sake of it we are using the symbol $\equiv$ as a shorthand to indicate what the meaning of a given term of the language is identical to:

Definition 3.3 *Let* $G = \langle \{p_1, p_2\}, (\Sigma^i_G)_{i \in \{1,2\}}, u_G \rangle$ *be a strict, two-player game in normal form. Suppose that*

$$\Sigma^1_G = \{a_1, \ldots, a_n\}$$

and

$$\Sigma^2_G = \{b_1, \ldots, b_m\}$$

For every term t *of* $\mathscr{L}_{GM}$*, we define the* interpretation t^G of t in G *inductively as follows:*

- *if* t *is a variable* x*, then* x^G *is any action of* G*, i.e.* $x^G \in \Sigma^1_G \cup \Sigma^2_G$*;*
- *if* t *is an action term of the form* a_i *and* $i \leq n$*, then* $a^G_i \equiv a_i$*;*
- *if* t *is an action term of the form* b_j *and* $j \leq m$*, then* $b^G_j \equiv b_j$*;*
- *if* t *is a term for strategy profile* $s_1 s_2$*, where* s_1 *is a term referring to an action of player 1 (i.e.,* $s^G_1 \in \Sigma^1_G$*) and* s_2 *is similarly a term referring to an action of player 2 (i.e.,* $s^G_2 \in \Sigma^2_G$*), then* $(s_1 s_2)^G \equiv s^G_1 s^G_2$*;*

- *if t is a payoff term $u^n(s)$, where n is either 1 or 2 and s is term for a strategy profile (that is, s is of the form $s_1 s_2$ with s_1 and s_2 being like in the previous clause), then* $(u^n(s))^G \equiv u^n_G(s_1^G s_2^G)$.

The tautological character of (most of) the above clauses is obviously due to the convention we have established on the notation, which brought us to use the same symbol for terms of $\mathscr{L}_{GM}$ and their intended reference in a given game G. The interpretation of a term is set here to make precise the correspondence we intended to capture by laying down the syntax of $\mathscr{L}_{GM}$ in that way. Therefore, when we write for instance $a_i^G \equiv a_i$ in the second clause we are aiming at stating that *action* a_i on the right of $\equiv$ is what the *term* of $\mathscr{L}_{GM}$ a_i mentioned on the left of it means (or, to be consistent with the definition: is interpreted by).

Notice that language $\mathscr{L}_{GM}$ is set up to speak of *any* two-player game in normal form G, since the latter are finite objects. Therefore, no matter what size a given game G of this form has in terms of number of actions allowed to players, $\mathscr{L}_{GM}$ has terms to refer to each of them (hence, only a portion of those that would be available are actually used to speak of actions of G and needs to be given an interpretation), as well as to the strategy profiles that come out by combining them, and to payoffs that are distributed to players over the latter combinations of actions.

Next in line comes, as promised, the definition of validity for formulas, which similarly depends upon a strict game G in normal form that we suppose to be given, and it also features the mention of hypotheses, which are needed to evaluate formulas of $\mathscr{L}_{GM}$ of the form $R(s)$, or formulas which comprise occurrences of them, due to the previous discussion of cycles and cycle-breaking. Remember in this respect that hypotheses needed to break the said cycles should feature actions that are supposed to be rational to both players. More precisely, let us put the following:

Definition 3.4 *If $G = \langle \{p_1, p_2\}, (\Sigma^i_G)_{i\in\{1,2\}}, u_G\rangle$ is any strict, two-player game in normal form, a* hypothesis in G *is a set h which contains exactly one term of $\mathscr{L}_{GM}$ for an action of player 1 in G (i.e., there exists a unique $s \in h$ such that $s^G \in \Sigma^1_G$), and exactly one term of $\mathscr{L}_{GM}$ for an action of player 2 in G (i.e., there exists a unique $t \in h$ such that $t^G \in \Sigma^2_G$).*

Hypotheses defined as (unordered) pairs of action terms do correspond to strategy profile of a given game G since, if $h = \{s, t\}$, where s and t are terms of $\mathscr{L}_{GM}$ that refer to actions s^G and t^G respectively of G, then $s^G t^G$ is a strategy profile of G. To make the notation easier, we will conventionally write hypotheses in a simplified form as combinations of action terms corresponding to strategy profiles, i.e., we shall write $h = s\,t$ instead of $h = \{s, t\}$ as we should owing to definition 3.4.

Hence, we have:

Definition 3.5 *Let $G = \langle\{p_1, p_2\}, (\Sigma^i_G)_{i\in\{1,2\}}, u_G\rangle$ be as before a strict, two-player game in normal form. Let also h be a hypothesis in G. For every formula θ of $\mathscr{L}_{GM}$, then that θ* is valid relatively to G and h *(in symbols: $\models^G_h \theta$) is defined by induction on θ as follows:*

- *if θ is of the form $s = t$, hence if s is either a variable or an action term and t is an action term as well (i.e., if $s^G, t^G \in \Sigma^1_G \cup \Sigma^2_G$), then $\models^G_h \theta$ holds if and only if $s^G = t^G$ is the case*[4]*;*
- *if θ is of the form $s > t$, hence if both s and t are terms for payoffs (which entails, owing to how the utility function of G is defined – see section 2 – that $s^G, t^G \in \mathbb{Q}$), then $\models^G_h \theta$ holds if and only if $s^G > t^G$ is the case*[5]*;*
- *if θ is of the form $R(s)$, hence s is an action term (i.e., $s^G \in \Sigma^1_G \cup \Sigma^2_G$), then $\models^G_h \theta$ is the case if and only if $s \in h$;*
- *if θ is of the form $\theta_1 \wedge \theta_2$, then $\models^G_h \theta$ holds if and only if both $\models^G_h \theta_1$ and $\models^G_h \theta_2$ are the cases;*
- *if θ is of the form $\theta_1 \vee \theta_2$, then $\models^G_h \theta$ holds if and only if either $\models^G_h \theta_1$, or $\models^G_h \theta_2$ is the case (where the case that both cases hold true is not excluded).*

We invite the reader to take notice of the fact that, consistently with the discussion from sections 3.2 and 3.3 about formulas of $\mathscr{L}_{GM}$ of the form $R(s)$ representing the only problematic case with respect to exercising the intuition we had about validity of formulas, actual use of hypotheses h is made just for the sake of the clause about instances of formulas of that type. In other words, changing a given hypothesis h into a different one h' can cause changes in the attempt of assessing whether θ is valid or not in the sense of the definition only if θ is either of the form $R(s)$, or if it features occurrences of formulas of that form.

The validity relation for formulas of $\mathscr{L}_{GM}$ just defined, can be used to make the passage in section 3.3 that is crucial to overcome the obstacle connected with the definition of rational action we came up with. The passage in question is the one that, for a given strict game G in normal form, allows to refine the hypothesis about what is rational to do to players by determining which actions among those that are available in G produce valid instances of the defining clause of the predicate $R(x)$ of $\mathscr{L}_{GM}$.

Let then $G\langle\{p_1, p_2\}, (\Sigma^i_G)_{i\in\{1,2\}}, u_G\rangle$ be an arbitrary, but fixed strict two-player game in normal form. Let us indicate with $\theta_G(x, R)$ the defining

[4]The clause here presumes that we have a suitably defined notion of identity over the elements of G that interpret terms s and t, to which the identity relation between terms of $\mathscr{L}_{GM}$ corresponds to. In this respect, see what was said back in footnote 3.

[5]Like in the previous clause, this one is conceived in such a way to make valid all of the instances of formulas of $\mathscr{L}_{GM}$ of the form $s > t$ that correspond to true inequalities over $\mathbb{Q}$ between elements of this domain which interpret the terms s and t.

clause for the predicate $R(x)$, «x is a rational action in G», which is obtained from the general formulation of it in section 3.2, by making it specific to G (that is, by considering the actual number of actions that are at the players' disposal in the game under scrutiny). This means that $\theta_G(x,R)$, which keeps track in the notation of the fact that the original formula it is a shorthand for features the occurrence of a free individual variable x, as well as occurrences of the predicate symbol R, is set to indicate the following instance of the general definition we are using:

$$R(x) \Leftrightarrow_{Def} \underbrace{\bigvee_{1\leq i\leq n_G} (x = a_i \wedge \varphi_i) \bigvee_{1\leq j\leq m_G} (x = b_j \wedge \psi_j)}_{\theta_G(x,R)}$$

where

$$n_G = \max\{i \in \mathbb{N} : a_i \in \Sigma^1_G\}$$
$$m_G = \max\{j \in \mathbb{N} : b_j \in \Sigma^2_G\}$$

Now, as we were saying, the previous approach to the definition allows one to pass from a given hypothesis h, which comprises actions (in fact, action terms) that are supposed to be rational to player 1 and 2 of G, to a *revised hypothesis* h^+, that contains action terms which produce valid instances of $\theta_G(x,R)$ when substituted to x by making use of elements of h to say which instances of formulas of the form $R(a_i)$, $R(b_j)$ occurring in $\theta_G(x,R)$ are valid. This can be now expressed precisely by making use of the property of «validity» relative to a game G and a hypothesis h from definition 3.5 above. In particular, if we assume to indicate with $ATERM_{GM}$ the collection of action terms of $\mathscr{L}_{GM}$ (hence, $ATERM_{GM}$ is the collection of expressions s of $\mathscr{L}_{GM}$ such that $s^G \in \Sigma^1_G \cup \Sigma^2_G$), then for a given hypothesis h, we have that:

$$h^+ = \{s \in ATERM_{GM} : \models^G_h \theta_G(s,R)\}$$

where $\theta_G(s,R)$ indicates the expression which is obtained by substituting s for x in $\theta_G(x,R)$.

It should be noticed first that, provided h is a hypothesis in G according to definition 3.4, then also h^+ is as such. This is due to the fact that, the game being strict, for every chosen action combination there is one unique action for each player that represents the best reply, payoff-wise, to the action played by the opponent, hence also to the one that is supposed to be rational here. Therefore, for every starting hypothesis h, if $t \in h$ is such that $t^G \in \Sigma^2_G$, then there is one and only one formula φ_i that is valid relatively to G owing to t, hence just a unique $s_1 \in h^+$ with $s_1^G \in \Sigma^1_G$; viceversa, if $s \in h$ is such that $s^G \in \Sigma^1_G$, then there is a unique formula ψ_j that is valid relatively to G owing to it, hence a unique $t_1 \in h^+$ with $t_1^G \in \Sigma^2_G$. It follows that $h^+ = s_1 t_1$ in this case, and fits definition 3.4 as a consequence.

In addition, and owing to the same reason, we also have that this passage from h to h^+ is 'right-hand unique', so to say, in the sense that no two different hypotheses h' and h'' can be obtained as a result of 'revising' in the above sense one and the same hypothesis h. It is then possible to use this observation to define a *revision operator* acting on hypotheses as arguments and yielding revised hypotheses as values. The latter kind of hypotheses would then literally appear as 'function of' the hypotheses that are supposed to be given in the first place. This can be done by taking into account that hypotheses as we defined them in definition 3.4 above are pairs of action terms, one of which is a term for an action of player 1 and the other one is a term for an action of player 2. Therefore, the function that associates a hypothesis with its revised one must be defined in such a way that this aspect is taken into account.

Let us introduce the following notation to make things readable below: for every pair of sets A, B, let $P(A,B)$ be the collection of unordered pairs of elements of A and B, that is[6]:

$$P(A,B) = \{\{a, b\} : a \in A, b \in B\}$$

Let now a two-player game in normal form G be given. Put:

$$S^1_L = \{s \in ATERM_{GM} : s^G \in \Sigma^1_G\}$$
$$S^2_L = \{s \in ATERM_{GM} : s^G \in \Sigma^2_G\}$$

That is: S^1_L gathers all action terms of $\mathscr{L}_{GM}$ that are interpreted by means of actions of player 1 in G, while S^2_L does the same with terms of $\mathscr{L}_{GM}$ that are interpreted by actions of player 2 in G. Notice that in case these sets S^1_L and S^2_L are substituted for A and B respectively, then $P(S^1_L, S^2_L)$ is nothing but the set of hypotheses in G according to definition 3.4 above. Let, for the sake of brevity, this set $P(S^1_L, S^2_L)$ be indicated as H_G henceforth. Then we put:

Definition 3.6 *Let $G = \langle \{p_1, p_2\}, (\Sigma^i_G)_{i \in \{1,2\}}, u_G \rangle$ be a strict, two-player game in normal form. The* operator for revising hypotheses *in G, or the* revision operator *in short, is a function $\delta_G : H_G \to H_G$ defined by $\delta_G(h) = h^+$ for every $h \in H_G$.*

Having made precise the idea of passing from one given hypothesis to the revised one in the previous way, allows us to further specify, in the form of a rigorous and systematic method, the idea of getting in this way rid of the loop in the definition of «rational action» in (a certain class of) finite games. For instance, the reader might have noticed already that by no means the passage from an hypothesis h to h^+, or $\delta_G(h)$ as definition 3.6 tells us, should be confined to just one single application of the

[6]Notice that this is different from gathering together the set of (ordered) pairs as this is usually done by means of the Cartesian product between two given sets (see section 4.4 on that).

revision process of hypotheses. That is, the revised hypothesis h^+ which comes out of a given h, is suitable to be refined further by the same means. This will allow us to obtain a third hypothesis, say h^{++}, which can be further revised, and so on and so on. The idea of an operator δ_G that takes care of all these applications of the revision process, allows us to make the idea precise in terms of *iterations* of it as follows:

Definition 3.7 *The finite iteration* $(\delta^n_G(h))_{n\in\mathbb{N}}$ *of the revision operator* δ_G *over a hypothesis* $h \in H_G$*, for every strict, two-player game* G *in normal form, is inductively defined for every* $n \in \mathbb{N}$ *by the following clauses:*

$$\delta^0_G(h) = h$$
$$\delta^{n+1}_G(h) = \delta_G(\delta^n_G(h))$$

The definition is just a way to frame the idea according to which, given a starting hypothesis h, this is first revised according to the process we are well aware of, and the hypothesis h^+ is produced as a result. This corresponds by definition 3.6 to an application of the revision operator δ_G to h. Then, one may wish to revise h^+ itself. The result would again correspond to applying δ_G again, this time to h^+ as argument. This would lead us to value $\delta_G(h^+)$ which can also be written as $\delta_G(\delta_G(h))$ owing to h^+ being itself the result of applying the revision operator as said, and would amount to the double iteration $\delta^2_G(h)$ of the revision operator according to what we have just stated in definition 3.7. By proceeding in this way one is brought to find a third iteration of it, then a fourth one, and so on, each of them coinciding with the elements of a *revision sequence* that starts with h and features all further applications of the method for revising it:

$$h, h^+, h^{++}, \ldots, h^{++\ldots+}, \ldots$$

Now, the question stems naturally as to whether this thread of hypotheses and their revisions has an end or not. This is also relevant for the issues we are trying to use this method to give a solution to. For, were there not an end to the process of revising the hypothesis about what is rational for players of a finite game in normal form to do, then it would not be clear at all that to seek for «rational actions» could be useful for the sake of isolating 'stable solutions' to the game itself. Since to revise an hypothesis is intuitively connected to finding the most convenient replies to actions that are, owing to the hypothesis itself, rational to the players, if the process of revising hypothesis were unending, that could be taken as an indication that no stable solution can be found, even though no exact correspondence between the revision process and equilibria has been established yet. Luckily, the said process can be proved to reach an end in the case of strict games at least, this end also providing us with what we are aiming to find in terms of solutions that are stable.

3.5. Equilibria as fixpoints of the revision operator

Let us first reason a little bit about the iteration of the revision process we have been considering in the last part of the previous section, and let us think what an 'end' to it may look like.

The problem, of course, is that there is nothing in that procedure as we have described it so far that ensures us that not always new information will be acquired by passing from hypotheses to their revised forms in the way we have explained. That is, for what we know of this revision process so far, we have no clue for guessing that, by having set a starting hypothesis h, by revising it, and by keep applying the revision operator as definition 3.7 says, we will not always encounter hypotheses that have not been produced at earlier stages. Now, something about the game we suppose to start from being finite may suggest that this scenario cannot occur: the number of strategy profiles in any game $G = \langle\{p_1, p_2\}, (\Sigma^i_G)_{i\in\{1,2\}}, u_G\rangle$ is finite anyway, and since hypotheses are equal in number to strategy profiles[7], we also have a finite amount of them; therefore, any revision sequence in which every element is (setwise) different from the previous one, must come to an end in the sense of featuring at some point the very element from which it was all started, or some subsequent element of it. For, otherwise we would be building an infinite revision sequence of hypotheses in a given game G, which is impossible owing to the finite character of it.

However, this is not really the kind of end we would like to be thinking of, as it turns out clearly by reasoning a little bit about the situation that would take place in that case. This requires that we think of having set a initial hypothesis h in a game G that we suppose to be given, and that the revision operator be applied to it once to get $\delta_G(h) = h^+$ as per definition 3.6. By the supposition we are making use of, h and $\delta_G(h)$ would be different, which means that one or both elements that h contains are different from those in $\delta_G(h)$. Then, we would be iterating the process, thereby obtaining $\delta^2_G(h)$, and $\delta^3_G(h)$ after that, and so on. As we said, if we kept supposing that the sets we obtain in this way were always different, then it would be like assuming that we can build a infinite sequence of different hypotheses in G, which contradicts the fact that G is finite and no such amount of hypotheses in it is available. So, we have to suppose that the converse will take place, that is, that at some stage in our finite iteration of the revision operator over h, say at stage $\delta^n_G(h)$, we realize that what we have obtained is identical as a set to something that had been obtained before, i.e., that for some $m < n$ we have $\delta^m_G(h) = \delta^n_G(h)$. Now, having noticed that the revision operator is right-hand unique, we

[7]It should be clear, without no need of stating this in more precise terms, that strategy profiles con be thought of as ordered hypotheses and, viceversa, hypotheses are just made out of the pairs of action a strategy profile is made out of.

would be driven to conclude then that, being $\delta_G^m(h)$ and $\delta_G^n(h)$ identical, applying the operator to them would lead us to results that are identical too, hence that

$$\delta_G(\delta_G^m(h)) = \delta_G^{m+1}(h) = \delta_G^{n+1}(h) = \delta_G(\delta_G^n(h))$$

is true. In turn, the identity would be inherited, and for the same reason as before, to values obtained by applying the revision operator again, and so on. In sum, elements of the whole block of hypotheses going from $\delta_G^m(h)$ to $\delta_G^{n-1}(h)$, would be identical to the corresponding elements of the block going from $\delta_G^n(h)$ to $\delta_G^{2n-1-m}(h)$ in the end[8]. In other words, we would realize at stage $\delta_G^n(h)$ that we are stuck in a cycle, since what we obtain here correspond to what we had already achieved at stage $\delta_G^m(h)$, and what follows cannot but repeating what we had produced at stages later than that. So, we would have tried to escape a loop just to find ourselves trapped into another one, and this, as anticipated, would turn out to be no good news.

Yet, this reflection about this kind of end for the iterated revision process is not entirely useless. For, as the reader might have notice already, the existence of the said cycle between hypotheses available in a given game G depends upon the distance between the hypothesis $\delta_G^m(h)$ in our example and the next occurrence of it as hypothesis $\delta_G^n(h)$. Let us suppose that this distance be reduced to the minimum, i.e. that $n = m + 1$. As an effect of the right-hand uniqueness property of the revision operator, one would get as a consequence that the subsequent stage $\delta_G^{n+1}(h) = \delta_G(\delta_G^n(h)) = \delta_G(\delta_G^m(h)) = \delta_G^n(h)$ is also identical to $\delta_G^m(h)$, and the same holds true for every stage produced afterwards. Then, the situation can be accounted for in the following way: having started from hypothesis h, one has applied the revision process iteratively obtaining always 'new' results, until a certain hypothesis h^* has been reached (our $\delta_G^m(h)$ in the example), after which the process becomes no more creative, since no additional information in the form of new hypotheses can be obtained by iterating the revision operator any further. Now, this special case of the previous situation does sound more like the kind of 'solution' to the iteration of the revision process we were seeking for. As a matter of fact, if that were the situation, then one could be legitimated to claim that the whole procedure has reached an 'end'. If the original problem was that by keep

[8] The reader may wonder how this index was calculated. It simply follows from the reasoning we have been pursuing: if we, starting from $\delta_G^m(h)$, presume that we keep obtaining different hypotheses until $\delta_G^n(h)$ is reached, then the length of the block comprising different hypotheses measures $l = (n-1) - m$ which counts the number of times the operator was applied after $\delta_G^m(h)$ and before $\delta_G^n(h)$. Then, the end of the block comprising hypotheses that are identical to those in between $\delta_G^m(h)$and $\delta_G^{n-1}(h)$ after $\delta_G^n(h)$ is obtained by applying δ_G to it l-many times. This means that the end of it is reached at $\delta_G^l(\delta_G^n(h))$ which, by definition 3.7, corresponds to stage $\delta_G^{(n-1)-m+n}(h)$. Then, some basic arithmetical calculation yields the result.

revising hypotheses one could never end with a last one ('last' in the sense of making useless any further iteration of process), then an outcome such as h^* in the example can be put forth as solving it. The questions to answer now clearly concern whether we can expect to always find solutions such as h^*, and whether to find them is in any way connected with the search for solutions to finite games in the form of equilibria.

To address the first question we have to rely on the theory of a certain class of operators to which also the revision operator we are considering here belongs to.

«Operator», as it might have been guessed already, is just another name for what one may commonly refer to by means of the word «function». The latter expression, however, is too much connected with applications to number-theoretic domains and to refer to «operators» rather than «functions» is an attempt to avoid a too narrow interpretation of the word. Sticking to the former notion rather than the latter, allows us to take a more liberal stance not just for what concerns the *nature* of elements of domain and co-domain of an operator, but also on their *type*: we are used to think of functions as being defined over *elements* of a given domain D, these elements being regarded as «individuals»; however, there is nothing wrong in thinking to a mathematical 'object' being function-like and operating over domains whose elements are not individuals but collections of individuals instead, for example, and deliver values which have the same character. By the term «operator» we are wishing to refer to operations in this wider sense of the expression. Strictly speaking, this may not appear as needed as far as the revision operator δ_G is concerned, since, as it has been defined in definition 3.6, action of δ_G reduces to a one-to-one correspondence between elements of the set H_G. However, this set does not contain individuals, since it is made of pairs, and this is why we have been calling δ_G as such rather than «revision *function*».

To familiarize the reader with this general notion, we present the main features of it. First comes the very definition of the concept:

Definition 3.8 *Given any two sets A and B, we call $\Gamma : A \rightarrow B$ an* operator *with domain A and co-domain B if and only if:*

- *for every element a in A, $\Gamma(a)$, the application of Γ to a, is an element of B;*
- *for every element a of A, $\Gamma(a)$ is the unique element of B that Γ associates with a, i.e., for every $a, a' \in A$, if $\Gamma(a) \neq \Gamma(a')$, then $a \neq a'$.*

Every operator works as a function over a given domain of elements in the sense of being everywhere defined over their totality, and being right-hand unique also. An operator is therefore exactly as a function, but, as it was said, its domain and co-domain may contain 'objects' which are not necessarily numbers, nor have an elementary nature (or, are «individuals» in the sense this expression was used in the paragraph preceeding the definition).

In what follows, we are interested into operators Γ's which are defined over subsets of a given set U and produce subsets of U as values. To avoid further complications in the notation, we will keep referring to operators of this kind as $\Gamma : U \to U$, where it is intended that the actual domain of it is the set $\mathscr{P}(U) = \{X : X \subseteq U\}$ of subsets of U, which is also its co-domain (hence, for every $X \subseteq U$, $\Gamma(X) \subseteq U$)[9]. Throughout the rest of the section we shall make no other hypothesis on U, except that this is a non-empty set to avoid trivialities.

An important related notion is the one that allows us to refer to «closure stages» in the application of an operator, like the hypothesis h^* we were previously considering in our example about the revision operator. This is made precise by means of the following definition:

Definition 3.9 *Let* $\Gamma : U \to U$ *be an operator whatsoever. We call* $X \subseteq U$ *a* fixpoint of Γ *if* $\Gamma(X) = X$.

To reconcile the definition we have just given with the example we were previously considering we need to imagine that Γ be applied iteratively, therefore to produce an output $\Gamma(Z)$ out of a subset Z of U, then once again to the value obtained thereby in order to get value $\Gamma(\Gamma(Z))$, and so on. If at some point in this iteration the value X is obtained, this is were the process stops, in the same sense as before with the revision operator δ_G and its value h^*, since any further application of Γ would reduce to $\Gamma(X)$ and, assuming X to be fixpoint of Γ as in definition 3.9, this value in turn would be X again.

One further notion to be considered is the following feature of operators:

Definition 3.10 *An operator* $\Gamma : U \to U$ *is* monotone *if and only if, for every* $X, Y \subseteq U$, $\Gamma(X) \subseteq \Gamma(Y)$ *whenever* $X \subseteq Y$ *is the case.*

Monotone operators give rise to iterative processes which are said to be *cumulative* in view of the property introduced by definition 3.10. In

[9]This apparent confusion between 'proper' functions, i.e., correspondences between elements of a set U of the form $f : U \to U$ where, for every $x \in U$, $f(x) \in U$, and 'genuine' operators, i.e., any F such that $F : \mathscr{P}(U) \to \mathscr{P}(U)$, which is motivated here to avoid adopting a more complex notation, is somehow justified at the mathematical level by a strict relationship between the two concepts. As a matter of fact, if $f : U \to U$ is a function, then there is an operator 'naturally' corresponding to it, namely $F_f : \mathscr{P}(U) \to \mathscr{P}(U)$ defined by $F_f(X) = \{f(x) : x \in X\}$, for every $X \subseteq U$. On the other hand, if $F : \mathscr{P}(U) \to \mathscr{P}(U)$ is given, one can also define a function out of it, say function $f_F : U \to U$, by setting $f_F(u) \in F(\{u\})$, for every $u \in U$. However, to get a function out of this idea one needs the axiom of choice in general (which ensures that f_F is right-hand unique in particular). So, the connection between the two concepts does not come out 'for free' in this direction, so to say.

a cumulative process the information is not dispersed, so to say, in the passage from one stage to another as long as the resources to achieve it are present: therefore, as long as Y contains the resources in X in the set-theoretic sense that the latter set is a subset of the former, then the information that X allows one to achieve, that is the value $\Gamma(X)$, is also part of the possibily bigger information that one obtains by relying on Y, and $\Gamma(X) \subseteq \Gamma(Y)$ as a consequence.

The main reason to focus our attention on monotone operator, is the following result:

Theorem 3.1 *Let* $\Gamma : U \to U$ *be a monotone operator. Then, there exists a subset of* U*,* X_Γ*, which is a fixpoint of* Γ*.*

Proof: let Γ be as in the hypothesis. We are going to prove that

$$FIX_\Gamma = \{Z \subseteq U : \Gamma(Z) = Z\}$$

that is, the collection of fixpoints of Γ, is indeed non-empty. Let

$$C_\Gamma = \{X \subseteq U : \Gamma(X) \subseteq X\}$$

It is a fact that C_Γ is not empty. As a matter of fact, take U for X: then $U \subseteq U$ obviously holds and, $\Gamma(U) \subseteq U$ also holds by definition of Γ. So, at least $U \in C_\Gamma$ is the case, which means that this latter set contains one element and is different from the empty set.

Now, let:

$$\bigcap C_\Gamma = \{z \in U : z \in X, \text{ for every } X \in C_\Gamma\}$$

To say it in ordinary set-theoretical terms, $\bigcap C_\Gamma$ is the generalized intersection over set C_Γ, that is the set of all common elements to all sets belonging to C_Γ. As a consequence, we have that $\bigcap C_\Gamma \subseteq X$ for every $X \in C_\Gamma$. Then, since Γ is monotone and $X \in C_\Gamma$, it follows that

$$\Gamma(\bigcap C_\Gamma) \subseteq \Gamma(X) \subseteq X$$

Since this holds for every element X of C_Γ, then also $\Gamma(\bigcap C_\Gamma) \subseteq \bigcap C_\Gamma$ must be the case (for, the elements of $\Gamma(\bigcap C_\Gamma)$ are common to all elements of C_Γ and $\bigcap C_\Gamma$ is meant to collect all such elements). This latter fact allows us to conclude two things. It means that $\bigcap C_\Gamma \in C_\Gamma$ holds, on the one hand. Second, by using again monotonicity of Γ, it yields that

$$\Gamma(\Gamma(\bigcap C_\Gamma)) \subseteq \Gamma(\bigcap C_\Gamma)$$

hence, that $\Gamma(\bigcap C_\Gamma) \in C_\Gamma$ is also the case. Then, it also follows that $\bigcap C_\Gamma \subseteq \Gamma(\bigcap C_\Gamma)$, because the former set is supposed to contain elements which are common to all of the elements of C_Γ, hence is a subset of all of them. Hence, we can conclude that $\bigcap C_\Gamma = \Gamma(\bigcap C_\Gamma)$, having proved that both $\Gamma(\bigcap C_\Gamma) \subseteq \bigcap C_\Gamma$, and $\bigcap C_\Gamma \subseteq \Gamma(\bigcap C_\Gamma)$ are the cases. That is, $\bigcap C_\Gamma \in FIX_\Gamma$. Hence, the theorem[10]. QED

So, monotone operators always have fixpoints. This means that if only we could prove that the revision operator δ_G is monotone, to go back to the investigation we have set up, then we would know that 'solutions' to the iteration of it, like the hypothesis h^* in the example we have been considering above, always exist. For, it should be clear with fixpoints, that once you have reach one of them there is no point into iterating the operator no matter how many times you are planning to apply it. As a matter of fact, if $Z \subseteq U$ is a fixpoint of $\Gamma : U \to U$, then it follows that is a fixpoint of $\Gamma^2 : \Gamma(U) \to U$ with $\Gamma(U) = \{\Gamma(X) : X \subseteq U\}$, meaning by that the application of Γ to subsets obtained by applying Γ to subsets of U, since

$$\Gamma^2(Z) = \Gamma(\Gamma(Z)) = \Gamma(Z) = Z$$

holds owing to double application of $\Gamma(Z) = Z$. Similarly, Z can be seen to be the fixpoint of Γ^n for every $n \in \mathbb{N}$.

Now, as long as fixpoints of the revision operator are concerned, it turns out that the answer regarding their existence can be obtained by passing through the logical approach to the notion of «rational action» in a finite game that was put forth in section 2. To be clear on that, let us re-consider what we did in that specific case from a more general persepc-tive.

Let then $\mathscr{L}$ be a first-order formal language whatsoever[11]. For every formula φ of $\mathscr{L}$, we suppose to indicate explicitly, whenever is a relevant feature of it, the fact that it contains, in its own formulation, occurrences of both an individual variable x outside the scope of the quantifiers of $\mathscr{L}$ (it contains *free occurrences* of x, as logicians are used to say – see section 4.8 below in this volume for an exact defintion of this concept in the context of another formal language), and a given predicate symbol P of $\mathscr{L}$. Let us use the notation $\varphi(x, P)$ for that[12].

[10] It should be noticed that we do not need to argue, for the sake of the theorem, that $\bigcap C_\Gamma$ is a non-empty set itself.

[11] This generically means that $\mathscr{L}$ is of the same type of the language $\mathscr{L}_{GM}$ we have been devising up to this point. For the acquainted reader, the features generally associated with first-order formal languages are: (i) the existence of an alphabet of symbols out of which expressions of $\mathscr{L}$ are made comprising denumerably-many elements; (ii) the existence of a decidable set of expressions of that sort, $TERM_L$, which are designated as the terms of $\mathscr{L}$; (iii) an equally decidable set of expressions $FORM_L$, designated as the formulas of $\mathscr{L}$; (iv) the fact that the latter collection of expressions is closed under firts-order logical operations, i.e., boolean connectives and quantifiers $\forall, \exists$ which produces formulas by bounding individual variables universally and existentially respectively.

[12] Once again the parallel with what we have done with $\mathscr{L}_{GM}$ and the definition of $R(x)$ can be of help in this case. In particular, the convention about indicating the formula of

Now, among the logical operations by which formal languages are standardly equipped with we have negation (that is usually indicated by means of the symbol $\neg$). The next definition wishes to frame a pure syntactical property of formulas having the features of $\varphi(x, P)$ we know of, that is occurrences of a free individual variable x, and occurrences of a predicate symbol P:

Definition 3.11 *Let $\mathscr{L}$ be any first-order formal language. Let $\varphi(x, P)$ be an arbitrary formula of it. Then we say that $\varphi(x, P)$ is P-*positive *if no occurrence of P in $\varphi(x, P)$ falls under the scope of an odd number of negations.*

To make the definition intelligible and ready-to-use for the goals we are aiming at, let us treat the case which is closest to the one we want to apply it to. So, let us think of P as a predicate which, like R in $\mathscr{L}_{GM}$, can be applied to a term t of $\mathscr{L}$, being this a variable or some other kind of term of $\mathscr{L}$ (like action terms in the case of R and $\mathscr{L}_{GM}$), to give rise to formulas of $\mathscr{L}$ of the form $P(t)$. Now, negation, like other logical operations, can be used to form new formulas of $\mathscr{L}$ out of them of the form $\neg P(t)$ (the intended meaning of which is «t is not P», or, more precisely, «the object denoted by t has not the property named P»). It is a general feature of logical operations that they can be applied again and again to form even more complex formulas. So is the case of negation which, in the example we are considering, gives rise to

$$\neg P(t), \neg\neg P(t), \neg\neg\neg P(t), \ldots$$

All these formulas are «negated formulas» in a sense, because they are obtained from formulas of $\mathscr{L}$ through an application of the negation connective. But, of all these formulas, only some of them express a negative state of affairs. This is due to assuming that a common principle of the logic of negation be valid, namely that two negations 'neutralize' each other to the effect that a formula like $\neg\neg P(t)$ ends up having the same meaning as $P(t)$ (that is, it means «object named t has the property P»). Clearly, this transfers to all formulas in the list above which features an even number of negations in the front: because $\neg\neg\neg\neg P(t)$ has the same meaning of $\neg\neg P(t)$ for the same reason, which happens to have the same meaning of $P(t)$; $\neg\neg\neg\neg\neg\neg P(t)$ has the same meaning of $\neg\neg\neg\neg P(t)$ and so of $P(t)$ again, and so on and so forth.

In turn, the same principle tells us that formulas in the list being prefixed by an odd number of negations express one and the same 'negative' fact, namely that «object named t has not the property P»: for, $\neg\neg\neg P(t)$ has the same meaning of $\neg P(t)$ owing to the double negation principle

$\mathscr{L}_{GM}$ constituting the defining condition of $R(x)$ by $\theta_G(x, R)$ was meant precisely to keep track of the fact that one could find occurrences of both x and R into it.

above, and $\neg\neg\neg\neg\neg P(t)$ has the same meaning of $\neg\neg\neg P(t)$ which reduces to $\neg P(t)$, and so on.

This also tells us something about definition 3.11 above. In particular, it tells us that a formula $\varphi(a,P)$ of $\mathscr{L}$ is P-positive if and only if it does not contain occurrences of negative instances of formulas involving P, or, which is just another way of saying the same thing, if and only if it contains only formulas by means of which the property P is (positively) attributed to elements of the reference domain of $\mathscr{L}$ (via their names).

Formulas of $\mathscr{L}$ like the generic $\varphi(x,P)$ we are using here for the sake of this presentation of the issue, can be used to define an operator in a similar fashion to what we have done above with the family of formulas $\theta_G(x,R)$ of $\mathscr{L}_{GM}$. This requires that an interpretation of expressions of $\mathscr{L}$ over a given domain M be given as it was done for expressions of $\mathscr{L}_{GM}$, which means in the first place that a correspondence is set between terms t of $\mathscr{L}$ and elements t^M belonging to M, whatever nature might they have, t^M playing the role of the object the 'name' t refers to.

Secondly, it requires that a relation of validity in M for formulas of $\mathscr{L}$ be defined. As long as formulas of the form $P(t)$ of $\mathscr{L}$ are concerned, as well as all other atomic formulas of this language, this definition of a validity relation would have to feature a clause stating that any given instance of such formulas «holds» or is valid in M provided the interpretation t^M of t, presuming this is the term that the predicate symbol P is applied to in the instance under consideration, belongs to a distinguished subset of the domain M_P of M which works as the interpretation of P in M, and is not valid if the contrary happens, that is if t^M does not belong to this set (which, in turn, provides us with the clause according to which negated formulas of that form, i.e. formulas $\neg P(t)$, are valid in M). So, this validity relation, as far as formulas $P(t)$ of $\mathscr{L}$ are concerned, is defined relatively to such interpretation of P being fixed in advance in the form of a set of individuals of M. This dependence of it over this latter set should be made apparent. Let us say that this is done by referring to this relation by means of the notation $\models^M_X \varphi$ to indicate that the formula φ is valid in the model M of $\mathscr{L}$ under the interpretation of occurrences of formulas $P(t)$ in φ by means of the set X.[13]

Let us assume that this has been done, hence that the relation $\models^M_X$ for formulas of $\mathscr{L}$ has been defined. Then, we have the following:

Definition 3.12 *Let $\mathscr{L}$ be a formal language. Let also $TERM_L$ the set of its terms, $\varphi(x,P)$ a formula of it, and $\models^M_X$ a validity relation for formulas of $\mathscr{L}$ over* M. *Let U_M be a subset of the collection of those elements of* M *over which interpretations of terms of $\mathscr{L}$ are defined, that is:*

$$U_M \subseteq \{a \in \mathrm{M} : a = s^M, \textit{for some } s \in TERM_L\}$$

[13]Again, the reader may clarify this by comparing this general situation to what was done with definition 3.5 above for language $\mathscr{L}_{GM}$ and formulas $\theta_G(x,R)$, by using hypotheses as interpretations of R.

Then, $\Gamma_{\varphi} : U_M \to U_M$ *is* the operator associated to $\varphi(x, P)$ *defined by*

$$\Gamma_{\varphi}(X) = \{a^M \in U_M : \models_X^M \varphi(a, P)\}$$

for every $X \subseteq U_M$.

Clearly, for one and the same formula $\varphi(x, P)$ there might be more than one operator associated with it owing to the definition we have just given. This is due to the fact that subsets U_M of M containing interpretations of terms of $\mathscr{L}$ can be selected in many ways, and operators which are defined over collections of their elements can be as many as they are.

Operators that correspond to formulas of $\mathscr{L}$ which are P-positive can be proved to have a peculiar feature:

Lemma 3.2 *Let* $\mathscr{L}$*,* $\varphi(x, P)$ *and* $\models_X^M$ *be as before. Let also* $\varphi(x, P)$ *be* P*-positive. Then, any operator* Γ_{φ} *associated with* $\varphi(x, P)$ *is monotone.*

Since we have not been precise about the definition of the language $\mathscr{L}$, as well as the definition of the validity relation $\models_X^M$ beside those that can be extracted by comparing this situation with the one we have been dealing with in the case of language $\mathscr{L}_{GM}$, it would be awkward to go through an exact proof of this result. We will confine ourselves to a sketch of it to let the reader get the idea of how the argument may go in all concrete cases.

The proof exploits the inductive definition of the set of formulas of $\mathscr{L}$ and allows one to make sure that the result holds whatever form does the formula $\varphi(x, P)$ takes through a proof by induction (see also lemma 2 from section 4.8 below for another example of this type of proofs). As it might be known to the reader already, arguments of this sort are based on proving the theorem relatively to parameters which are inductively defined, such parameter in this case being the logical form of $\varphi(x, P)$. In particular, it is first assumed that the logical complexity of $\varphi(x, P)$ be the lowest possible, that is, it is assumed that $\varphi(x, P)$ be an atomic formula, and it is shown how the proof goes under that hypothesis. In the case here at stake, one notices that if the logical form of $\varphi(x, P)$ is the simplest possible, which boils down to assuming that $\varphi(x, P)$ is $P(x)$, owing to the additional hypothesis that this formula does contain negated occurrences of P and is P-positive, then the following is verified: if X and Y are any given subsets of U_M such that $X \subseteq Y$, then $a^M \in \Gamma_{\varphi}(X)$ if and only if $\models_X^M P(a)$, that is if and only if $a^M \in X$ (since X provides us with the interpretation of P); however, if $a^M \in X \subseteq Y$, then it follows that $a^M \in Y$ is also the case, which yields $a^M \in \Gamma_{\varphi}(Y)$, hence $\Gamma_{\varphi}(X) \subseteq \Gamma_{\varphi}(Y)$ (i.e., Γ_{φ} is monotone in this case).

The next step in the proof is to take account of what goes under the name of the *induction step* of it, which has a different goal: to prove that the theorem holds by supposing that φ takes any of the possible logically complex form that the alphabet of $\mathscr{L}$ allows (depending upon which symbols for logical operations belong to it), under the hypothesis (which is called the *induction hypothesis*) that the theorem holds for formulas with lower logical complexity (so, in particular, that it holds for subformulas of φ).

As long as the lemma above is concerned, this induction hypothesis, together with the expected clauses in the definition of $\models^M_X$, are easily seen to be enough to conclude that the statement goes through this step. Having made no precise assumption on $\mathscr{L}$, this is where it is difficult to show how the proof goes in every possible case. However, we can just focus on one example and let the reader guess how the argument may work in the missing cases. So, let us assume for instance that $\varphi(x, P)$ has the form of a conjunction, namely that it is of the form $(\psi_1(x, P) \wedge \psi_2(x, P))$, where ψ_1, ψ_2 are formulas of $\mathscr{L}$ featuring occurrences of P and both being P-positive. The induction hypothesis in this case correspond to the assumption that operators $\Gamma_{\psi_1}, \Gamma_{\psi_2}$ are monotone, where these are defined as Γ_φ with ψ_1, respectively ψ_2 playing the role that φ plays in that case. One should observe that, owing to how validity relations such as $\models^M_X$ are usually defined, we have that $\models^M_X (\theta \wedge \eta)$ holds if and only if $\models^M_X \theta$ and $\models^M_X \eta$ are the cases, for every formula θ, η of $\mathscr{L}$. This has the consequence that $a^M \in \Gamma_\varphi(X)$ holds if and only if $a^M \in \Gamma_{\psi_1}(X)$ and $a^M \in \Gamma_{\psi_2}(X)$ both hold, for every $a^M \in U_M$ and $X \subseteq U_M$. Then, assuming that X, Y are any two given subsets of U_M for which $X \subseteq Y$ is the case, it follows from the assumption that $a^M \in \Gamma_\varphi(X)$, that $a^M \in \Gamma_{\psi_1}(X)$ and $a^M \in \Gamma_{\psi_2}(X)$ also hold; the induction hypothesis then leads to conclude that $a^M \in \Gamma_{\psi_1}(Y)$ and $a^M \in \Gamma_{\psi_2}(Y)$ are the cases, which yields $a^M \in \Gamma_\varphi(Y)$ as wanted.

As it was said, the argument goes through in a similar manner, both in the case of the remaining logical connectives, as well as in the case of the logical quantifiers. This eventually leads to the complete proof of lemma 3.2 above. Finally, this puts us in a position to reconcile this detour about formal languages in general to what we had been doing previously with $\mathscr{L}_{GM}$.

First of all, it might have been noticed already that:

Fact 3.1 *Any formula $\theta_G(x, R)$ of $\mathscr{L}_{GM}$ is R-positive.*

This is just a matter of observation, of course, since we have not even equipped the language $\mathscr{L}_{GM}$ with a negation symbol, hence we cannot produce negative formulas of it. This has an obvious consequence:

Corollary 3.1 *The revision operator δ_G as it was defined in definition 3.6, is an operator associated with a given formula $\theta_G(x, R)$ of $\mathscr{L}_{GM}$ that is monotone.*

Corollary 3.1 is of course a consequence of fact 3.1 and lemma 3.2 (and of definition 3.12, of which the revision operator δ_G satisfies the requirements). In turn, together with theorem 3.1, this has another important consequence:

Corollary 3.2 *For every formula $\theta_G(x,R)$ of $\mathscr{L}_{GM}$, the operator δ_G associated with it has fixpoints.*

So, the revision operator admits fixpoints. The question is: how does this help us? The shortest answer to the question is given in the form of the following result:

Theorem 3.2 *Let $G = \langle\{p_1,p_2\},(\Sigma^i_G)_{i\in\{1,2\}},u_G\rangle$ be a strict, two-player game in normal form. Let $\theta_G(x,R)$ be the formula of $\mathscr{L}_{GM}$ providing us with the defining condition of «rational action» for G. Then, any hypothesis $h = a_i b_j$ is a fixpoint of δ_G if and only if the strategy profile $a^G_i b^G_j$ is an equilibrium of G.*

Proof: actually, one can convince herself of the statement of the theorem by means of a simple, indirect argument. So, suppose that this is not the case and that, in the given two-player game G which is also strict, there exists an hypothesis $h = a_i b_j$ which is fixpoint of δ_G, but such that the corresponding strategy profile $a^G_i b^G_j$ is not an equilibrium in the sense of definition 3.1. This can only mean (see definition 3.1 again), that for at least one of the players it is beneficial to change her strategy under the assumption that the opponent will not be doing the same. Let us suppose that this holds true for player 1. Then, owing to G being strict, this can only mean that $u^1(a^G_h b^G_j) > u^1(a^G_i b^G_j)$ for some $a^G_h \in \Sigma^1_G$. However, if this were the case then $\delta_G(a_i b_j) = a_h b^*$ (where b^* is either b_j, or not), hence h cannot be a fixpoint of δ_G, contrary to the assumption. The contradiction that follows yields that no such a_h exists, and $a^G_i b^G_j$ is indeed an equilibrium of G.

A symmetrical argument can be used to show that this is also the case if player 2 is assumed to have benefits from changing her strategy unilaterally, which gives the theorem. QED

That is: fixpoints of the revision operator correspond to equilibria in a two-player game in normal form G which is also strict. This result gives us a hint for finally solving one of the two questions we raised about equilibria, namely: Does any two-player game G in normal form which is strict always admit equilibria? It turned out that the answer is positive, which, as long as equilibria are considered to provide us with a 'solution' as to what players should do in the situation depicted by G, entitles us to conclude that any game of that sort always features such solutions in the form of strategy profiles corresponding to hypotheses being fixpoint of the revision operator.

A second question was about the 'quantity' of equilibria that are given in one and the same game G. Is it always the case that we have just one of them? This is what we plan to address in the next section.

3.6. Unique equilibria?

Having shown that fixpoints and equilibria correspond to one another, one may further investigate issues regarding either of the two concepts by pursuing the path that one feels to be more inclined toward to. The next topic in our agenda is uniqueness, i.e., try solving the clue as to whether it is always the case that a game G in normal form which is strict possesses one unique equilibrium. We can anticipate here that the answer is negative. This can be easily argued for by presenting one single counterexample supporting the contrary assertion, even before discussing some general result that may help us to understand more about the collection of equilibria in finite games. As a matter of fact, to achieve the said conviction, it is enough to play a little bit with the distribution of payoffs over a game matrix and realize that there is no obstacle in detecting one which features more than one equilibria. Take for instance the following situation, that comes from modifying a little bit the distribution we considered in the game we have used in our discussion from section 2.5:

	C	N
C	(-5,-5)	(-3,-10)
N	(-10,-3)	(-2,-2)

Now, as long as the strategy profile *CC* that was indicated as an equilibrium before is concerned, this has not changed. For, neither of the two players have anything to gain from changing their strategy unilaterally. The changes we have made on the payoffs distribution has not affected that: still, player 1 would pass from a payoff equal to -5 to a lower one, -10, as well as player 2 that would suffer from an equal loss by deciding to avoid confessing.

Let us now have a look at strategy profile *NN*, the situation in which both players decide to not confessing. The payoff they both get is -2, which makes of it the most convenient situation for each of them. This was not so on the basis of the analysis of the game matrix we pursued in section 2.5. If we look at it now from the viewpoint of player 1, it is clear that not confessing is indeed the most convenient action to her in case player 2 also decided to proceed identically, since by changing her strategy she would pass from scoring -2 to scoring -3:

	N
C	(-3,-10)
N	(-2,-2)

An identical consideration can be made by looking at player 2's alternative to not confessing, which, in case player 1 decided not to confess, would bring her passing to scoring -3 instead of -2:

	C	N
N	(-10,-3)	(-2,-2)

This all entails that no player in this game would take advantage from a unilateral change in her strategy in case they have both chosen not to confess. Therefore, owing to definition 3.1, strategy profile *NN* is an equilibrium.

In full agreement with what we have concluded in the previous section by theorem 3.2, to the same conclusion one arrives by considering the method of revision of hypotheses. For, assume that *NN* is the starting hypothsis, that is, not confessing is rational to both player 1 and player 2. Then, owing to the fact that both

$$u^1(NN) > u^1(CN)$$

and

$$u^2(NN) > u^2(NC)$$

are the cases, it follows that formulas φ_2 and ψ_2, which, according to what we agreed upon in section 3.2, express that action N is the most convenient reply of player 1, respectively of player 2, in case the opponent 'plays' N, are valid relatively to the chosen game and to the hypothesis *NN*. This means then that, if $h = NN$, $\delta_G(h) = h$, that is, h is a fixpoint of the revision operator.

So, it is not always the case that one can get a finite game where payoffs are distributed among the players according to the strictness requirement of definition 3.2, with just one strategy profile that turns out to be an equilirbium of the game. As it was said, such a conclusion can be given the form of a general result. For instance, by means of the following theorem which establishes the existence, for every monotone operator, of a minimum and a maximum fixpoint. The proof of it comes from suitably modifying the argument used for the sake of proving theorem 3.1 above:

Theorem 3.3 *Let* $\Gamma : U \to U$ *be a monotone operator. Then, there are subsets* I_Γ *and* S_Γ *of* U *such that: (i)* I_Γ *is a fixpoint of* Γ *and* S_Γ *is a fixpoint of* Γ *too; (ii)* I_Γ *is the least fixpoint of* Γ *(i.e., if* X *is any other fixpoint of* Γ*, then* $I_\Gamma \subseteq X$*); (iii)* S_Γ *is the greatest fixpoint of* Γ *(i.e., if* X *is a fixpoint of* Γ *whatsoever, then* $X \subseteq S_\Gamma$ *is the case).*

Proof: we already proved, for the sake of theorem 3.1, that $\bigcap C_\Gamma$, where $C_\Gamma = \{X \subseteq U : \Gamma(X) \subseteq X\}$ would be a fixpoint of Γ. We can also prove that this is the least fixpoint. As a matter of fact, assume that $Z \subseteq U$ is also a fixpoint of Γ. That is, $\Gamma(Z) = Z$ which means that, in particular, $\Gamma(Z) \subseteq Z$ is also the case. Hence, $Z \in C_\Gamma$ and $\bigcap C_\Gamma \subseteq Z$.

Now, let

$$D_\Gamma = \{X \subseteq U : X \subseteq \Gamma(X)\}$$

and put

$$\bigcup D_\Gamma = \{z \in U : z \in X, \text{for some } X \in D_\Gamma\}$$

This means, in ordinary set-theoretical terms, that $\bigcup D_\Gamma$ is the generalized union of all the elements of D_Γ and, as such, it contains all of the elements of sets belonging to D_Γ. Therefore, $X \subseteq \bigcup D_\Gamma$ holds for every $X \in D_\Gamma$. By Γ being monotonic and X being element of D_Γ, it follows that

$$X \subseteq \Gamma(X) \subseteq \Gamma(\bigcup D_\Gamma)$$

is the case for every $X \in D_\Gamma$. So, in particular $X \subseteq \Gamma(\bigcup D_\Gamma)$ holds for every $X \in D_\Gamma$. It follows that $D_\Gamma \subseteq \Gamma(\bigcup D_\Gamma)$ and $\bigcup D_\Gamma \subseteq \Gamma(\bigcup D_\Gamma)$ is the case as well: take any $a \in \bigcup D_\Gamma$, then $a \in X$ must be the case for some $X \in D_\Gamma$, from which it follows that $a \in \Gamma(\bigcup D_\Gamma)$ since $X \subseteq \Gamma(\bigcup D_\Gamma)$ holds as stated (which proves that $\bigcup D_\Gamma \in D_\Gamma$). Also, by monotonicity of Γ we have

$$\Gamma(\bigcup D_\Gamma) \subseteq \Gamma(\Gamma(\bigcup D_\Gamma))$$

which means that $\Gamma(\bigcup D_\Gamma) \in D_\Gamma$. Owing to that, we conclude

$$\Gamma(\bigcup D_\Gamma) \subseteq \bigcup D_\Gamma$$

by the definition of the latter set. Hence, $\bigcup D_\Gamma = \Gamma(D_\Gamma)$ and $\bigcup D_\Gamma$ is another fixpoint of Γ.

Now, assume that X is another fixpoint too. This means that $X \subseteq \Gamma(X)$ holds as a consequence of $X = \Gamma(X)$. That is, $X \in D_\Gamma$ and $X \subseteq \bigcup D_\Gamma$ is the case by that.

The theorem then follows by putting $I_\Gamma = \bigcap C_\Gamma$ and $S_\Gamma = \bigcup D_\Gamma$. QED

Clearly, owing to this theorem, the following also holds:

Corollary 3.3 *Let* $G = \langle \{p_1, p_2\}, (\Sigma^i_G)_{i\in\{1,2\}}, u_G \rangle$ *be a strict, two-player game in normal form. Let* $\theta_G(x, R)$ *be the formula of* $\mathscr{L}_{GM}$ *providing us with the defining condition of «rational action» for* G*. Then, the revision operator* δ_G *has a least, and a greatest fixpoint* I_G *and* S_G *respectively.*

Proof: the result follows from corollary 3.1 and theorem 3.3. QED

The discovery that solutions to games in normal form in the form of equilibria are not always unique, raises a problem. For, while unique solutions can be more clearly connected to (the beginning of) a theory of rational choice which could be both descriptive of how rational agents act in a certain situation (provided some kind of argument about equilibria being connected to actual actions of agents), and normative about how players should act instead, the existence of two and possibly more solutions to one and the same matrix requires some further investigations on equilibria that may clarify whether or not one can differentiate one from the other for the sake of the said theory. If the definition of equilibria itself (as well as the equivalent notion of fixpoint of the revision operator), makes clear why they are preferable to strategy profiles that are not equilibria, how should we expect that rational agents could decide between different actions of theirs which belong to different strategy profiles all of which happen to be equilibria? The lack of a unique solution to finite games seems to necessarily lead to the discussion of features, and eventually to the setting of criteria which make one equilibria favourable to another.

This, however, turns out to be just one out of many problematic remarks that can be made about the concept of equilibrium as solution of a game. It seems that showing that uniqueness fails gives us the occasion to critically reconsider the preliminary assessment we made of that concept and try making it more robust.

3.7. The limits of strict games and strict equilibria

Let us go back for a minute to the modified matrix that was used in the previous section to introduce the topic of uniqueness of equilibria in finite games in normal form:

	C	N
C	(-5,-5)	(-3,-10)
N	(-10,-3)	(-2,-2)

The problem we hinted at in the very last paragraph of the previous section can be concretely formulated with respect to this given situation as follows: by searching for equilibria as solutions of the matrix we are here facing the problem that we have two of them, without any further clue on how to evaluate which one is better than the other. The problem sounds even more critical if we consider that there is an easy way to carry out the said evaluation. As a matter of fact, by deciding to both confess, player 1 scores a payoff equal to -5 and the same does player 2. By not confessing, instead, they both score more, since by sticking to action N they are both granted a payoff equal to -2. Now, since one of the criteria we have followed in our justification of equilibria as fixpoints of an operator 'extracted' from a very intuitive definition of «rational action», was based upon certain actions being mostly convenient to players owing to the fact that they ensure them the greatest possible payoff, should we not applying this criteria in this case, hence concluding that NN represents a better (i.e., more convenient) solution than CC? The proposal appears to be quite reasonable. However, there is nothing in the concept of equilibria that justifies us to apply that way of reasoning in this case (in fact, in any case). For, strategy profiles which turn out to be equilibria have been compared with strategy profiles which are alternative to them only, that is, which are obtained from the one under scrutiny by assuming just one player to change her mind. This means that strategy profile CC is compared payoff-wise with strategy profile NC and CN, these strategy profiles corresponding to the situation in which player 1 is the only player who changes her mind on the one hand, and to the one in which player 2 does that on the other hand instead. Similarly, strategy profile NN turns out to be an equilibrium of the game, owing to being most convenient to both players with respect to alternative strategy profiles CN and NC where one player at a time is supposed to change her choice of action. So, no direct comparison between CC and NN is needed for the sake of determining whether they are equilibria or not.

To stress this fact even further by relying on the geometry of the matrix in a two-player game, alternative strategy profiles which count for the sake of awarding one of them the prize in the 'equilibrium competition' are those lying on one and the same column, as long as the evaluation of the most convenient option to player 1 is concerned, and on one and the same row for what concerns player 2 instead. With respect to corollary 3.3 we have achieved in the previous section, this leads us to the following clarification of it:

Fact 3.2 *In a two-player, finite game in normal form G that is strict, no two different equilibria can lie on one and the same row, or on one and the same column.*

This is so in the example we are currently discussing, since strategy profile CC is on different row and column from those to which strategy profile NN belongs. This, however, seems just a side observation that let us progress no further with the issue we are considering, except maybe for providing us with a more solid base supporting the impression that the concept of equilibrium is rather weak since the problem of the lack of direct comparison between two different equilibria we have just spotted is somehow intrinsic to the definition of it, or to the procedure by means of which equilibria are identified.

It should be noticed, however, that fact 3.2 is directly connected with the strictness requirement of the game involved therein. As a matter of fact, the payoff distribution in a strict games G causes any two alternative strategy profiles that might be involved in a comparison for the sake of determining which one is most convenient to a player, say $a_i b_j$ and $a_h b_j$ for player 1, to be such that either $u^1_G(a_i b_j) > u^1_G(a_h b_j)$ holds, or $u^1_G(a_h b_j) > u^1_G(a_i b_j)$ is the case (where the third possibility, i.e., that $u^1_G(a_i b_j) = u^1_G(a_h b_j)$, is not possible). This means that in case the game considered is not strict, then fact 3.2 breaks down and it is also possible for two different equilibria of the game to belong to the same column or row. Let us consider an example of such situation, which will help us identifying one further critical aspect of the concept as we have been knowing it so far.

We are not going to make the example very specific, since the feature we would like to stress is of a general kind. In particular, we avoid equipping the game below with a story explaining it, and we just present its matrix, which is tailor-made to let the crucial character we would like to discuss turn out:

	b_1	b_2	b_3
a_1	(0,1)	(1,2)	(0,2)
a_2	(1,0)	(0,1)	(-1,0)

The situation depicted here is new to us, in the sense that the matrix above comes from a two-player game in normal form which is 'asymmetric' since the two players have not the same amount of actions to choose among: while player 1 has just two action at hers disposal like before, a_1 and a_2, player 2 happens to be allowed to choose between b_1, b_2 and b_3. Take notice of the fact that there is nothing wrong with that, and that our notion of finite game as we defined it in definition 3.2 allows that such a situation may occur.

The second fact to notice about the game above that further breaks the 'tradition' of examples that we have been considering so far, is the payoff distribution. In particular, the fact that, as the reader might have noticed it already, $u^2_G(a_1 b_2) = u^2_G(a_1 b_3)$. And *that* does break the notion of strict games from definition 3.2, as it prevents player 2 to have a most convenient reply to action a_1 in case the latter was chosen by player 1.

Finally, observe that strategy profiles $a_1 b_2$ and $a_1 b_3$ are both equilibria in the sense of definition 3.1. For, in the case of the former strategy profile, player 1 gets no advantage from changing her choice of action from a_1 to a_2 since she would suffer a loss in score (passing from 1 to 0), and player 2 is in a similar situation since by changing her action unilaterally, she would either pass from scoring 2 to scoring 1 (in case she decided to play b_1 instead of b_2), or she would score 2 (by deciding to play b_3 instead of b_2) which is exactly what she scores already. The analysis of strategy profile $a_1 b_3$ gives similar indications since player 1 would not like to change her choice, which would cause her to pass from scoring 0 to a lower payoff, -1, and player 2 too would again either loose or score the same as she scores already by playing b_1 and b_2 respectively. The conclusion is, as anticipated, that both $a_1 b_2$ and $a_1 b_3$ are equilibria, which makes two equilibria in one and the same row in the case considered contrary to the prescription of fact 3.2 for strict games.

Apart from failure of fact 3.2, which is due, as we said, to the game here being not strict, the situation now allows us to add a new critical remark about equilibria. For, imagine player 1 and player 2 who, having both made the analysis of the game as we have done it, are supposed to make their choice. The situation is easy for player 1 since both 'solutions' in the sense this expression was used up to now, lie on the same row, which corresponds to a_1 as her choice of action. The possibility that she may gain a positive score or gain no score at all (but avoid suffering from a loss in score at least), now depends on whether player 2 is going to choose b_2, or b_3 instead. But, how could this choice be possibily made? Player 2 is not even in the position of excercising the evaluation that was suggested with respect to the previous game (when the problem was to compare equilibrium *CC* with *NN*), since the two options she needs to weigh up, literally weigh the same.

So, by passing to games which are strict to games which are no more as such, the issue we raised about equilibria gets even worse since not even the possibility of comparing them in the way we suggested by considering the game at the beginning of this section (that is, by sticking to the equilibrium which is most convenient to the players), works anymore.

Someone may find a possibility of escaping the difficulty we are confronting, in noticing that, in the situation that was last considered, the two equilibria of the game corresponding to strategy profiles $a_1 b_2$ and $a_1 b_3$ were not really equivalent, not to both players at least. For, while player 2 would find herself stuck in the impossibility of deciding which one of them is most convenient to her, this is not so for player 1 who is still in a position to discriminate between the two strategy profiles and say which is the one she favours. As a matter of fact, player 1 scores 1 in case player

2 decided to play b_2 and the final outcome coincided with strategy profile $a_1 b_2$ then, while she scores 0 in the other case, i.e., if player 2 decided to play b_3 instead. How does this help us with player 2's indecision? In no way, if player 1 and player 2 are regarded as opposed to one another in the game. It does however help us, instead, if they are supposed to make a choice as a group. For, then one may wish to consider one outcome «more convenient» to a group of people with respect to another, if by sticking to it at least one member of the group gets a higher payoff and the rest of the group does not loose anything. By applying the said intuition to the case here at stake, one may be led to regard outcome $a_1 b_2$ to be preferable to outcome $a_1 b_3$ for player 1 and player 2 taken collectively.

However, it should be clear that this is no solution to the issue we are discussing, and for two reasons.

First of all, because this way of looking at things would work only for those situations in which the players do not act one against the other. This means that the problem we would solve there, would stay identical in all of the remaining cases.

Secondly, because once the strictness criterion has been 'betrayed', there is no reason why one should remain faithful to other aspects of it. Therefore, it would be easy to further modify the distribution of payoffs of our present example and build a new situation in which player 1 too is indifferent between the two equilibria of the game (in particular, this is easily achieved by setting player 1's payoff in outcome $a_1 b_3$ from the matrix above equal to 1).

To be honest, rather than helping us with the issue we were pondering over, this way of seeing things from the point of view of social choice adds some new material to criticize the concept of equilibrium we have been putting forth as solving games. As a matter of fact, that concept seems to be well-motivated individualistically, for, if any natural element can be found in it, that is the idea that no single person would act by preferring a lower payoff to a greater one; however, the same concept is only poorly motivated from the point of view of groups taking decisions, for, in that case, it is equally unlikely that all of its member acted without considering some sort of collective benefit. So, in the end, this viewpoint makes a solution to the problem, to the problem*s* in fact we have here discussed, even more urgent. We will consider one proposal in this respect in section 3.9 below, and another one in the section following it, and try to connect both of them with the approach we have been fostering so far. Before doing that, however, we present an extension of the methodology for analyzing games to matrices where payoffs are not distributed over outcomes in such a way that the strictness requirement is respected. This is done to let the reader familiarize gradually with modifications which we will be using for the sake of comprising those refinements to the notion of equilibrium that will be considered by further reflecting on the topic.

3.8. When strictness fails

The reader might have noticed that, in our previous discussion of a game that is not strict, we needed not to adapt the definition of equilibrium we had come up with by reasoning on a game which was strict instead. This is because the concept as it was defined by means of definition 3.1 already accommodates the event of a tie: a strategy profile $a_i b_j$ is *not* an equilibrium according to that definition if and only if there exists a strategy profile $a_h b_j$ which is more convenient to player 1 (i.e., such that $u^1_G(a_h b_j) > u^1(a_i b_j)$), or if there exists a strategy profile $a_i b_k$ which is more convenient to player 2 instead (i.e., such that $u^2_G(a_i b_k) > u^2(a_i b_j)$). In other words, a strategy profile is an equilibrium in this sense as long as the payoff that each player gets by sticking to it is greater than, or equal to the payoff that each player would get by changing her mind and decided to play differently while the opponents play according to the strategy in question.

In turn, this observation suggests an easy way to similarly adapt the definition of «rational action» that we have given in the context of strict games to make it suitable for approaching situations like the one we have been discussing lately. As a matter of fact, what we have just noticed about the concept of equilibrium legitimates to consider a strategy profile $a_i b_j$ in a two-player game G in normal form (not necessarily a strict one), more convenient to a player than another if it ensures to her a strictly higher, or an equal payoff. Therefore, for an action in a strategy profile to be rational to a player it might be regarded as a sufficient condition that it ensures to her a payoff that is no less than the payoffs ensured by all of the alternative actions she could opt for.

Let us make the discussion more concrete by considering again the example that suggested the crucial observation:

	b_1	b_2	b_3
a_1	(0,1)	(1,2)	(0,2)
a_2	(1,0)	(0,1)	(-1,0)

What we were just saying then, amounts to view a strategy profile as rational to both players as long as by sticking to it each of them obtains payoffs which are greater than, or equal to (i.e., no less than) the payoffs they would get by changing their mind with a unilateral act. This is clearly verified by both strategy profile $a_1 b_2$ (since player 1 would score 0 instead of 1 by changing her choice to a_2, that is by causing $a_2 b_2$ to be the final outcome, and player 2 would either experience a similar loss by choosing

b_1, or would achieve the same payoff by choosing b_3 instead), and $a_1 b_3$ as well (due to the same considerations as before), and only by them (as the reader can easily verify by considering all the other strategy profiles in the game).

As we did with the corresponding observation with respect to games where payoffs are distributed in such a way that the strictness requirement is verified, we would like to comprise this intuition in a formula of $\mathscr{L}_{GM}$ like the one we built and used for the sake of our discussion in previous sections. To do that we expand it by means of a new symbol for a binary relation, $\cdot \geq \cdot$, that we are going to use for the sake of payoff terms comparison, hence we use it to produce formulas of the form $u^i(s) \geq u^i(t)$, where i is an index identifying one of the players of the game, and s and t are terms for strategy profiles, with the intended meaning: «the payoff granted to player i by strategy profile s is greater than, or equal to the payoff granted to her by strategy profile t».

Let us call $\mathscr{L}^+_{GM}$ the language that is obtained from $\mathscr{L}_{GM}$ by performing the said extension. Then, one can think of using the additional resources of it to express the previous idea about an action being rational to a player by means of a formula. In the concrete case we are referring to for the sake of illustration, the formula stating that b_2 is rational in this sense to player 2, for instance, would turn out to be, owing to the convention on the notation we made beforehand and that we keep using here, the following expression:

$$\bigvee_{1 \leq i \leq 2} (R(a_i) \wedge u^2(a_i b_2) \geq u^2(a_i b_1) \wedge u^2(a_i b_2) \geq u^2(a_i b_3))$$

Coherently with the modified situation we are considering, action b_2 is mostly convenient to player 2 if it ensures no loss when it is used as reply to the action that is rational to player 1, whatever this one may be. Let us say that formulas of this sort have been devised for all the actions available to players. Let us replicate for them the notation we have been using so far for the sake of comparison, and assume then that we have expressions φ_1^+, φ_2^+ which express that action a_1, and action a_2 respectively are mostly convenient to player 1, while expressions $\psi_1^+, \psi_2^+, \psi_3^+$ (ψ_2^+ coinciding with the above formula) do the same for player 2's actions b_1, b_2 and b_3 (+'s are used here as a label for formulas of $\mathscr{L}^+_{GM}$).

In turn, these can be gathered together in a definition of «rational action» for the game above which corresponds to the following expression:

$$R(x) \Leftrightarrow_{Def} \bigvee_{1 \leq i \leq 2} (x = a_i \wedge \varphi_i^+) \bigvee_{1 \leq j \leq 3} (x = b_j \wedge \psi_j^+)$$

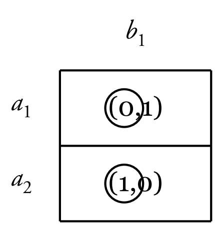
If we proceed in a fashion similar to how we did with the corresponding definition for games that are strict, and call $\theta^+(x, R)$ the defining condition of this $R(x)$, that is the expression to the right of the symbol $\Leftrightarrow_{Def}$, we can use it to define an operator for revising hypotheses concerning what is rational to do to player 1 and player 2 in the given game like we did with $\theta_G(x, R)$ and δ_G for strict games. With two provisos:

1. like in the previous situation, to carry out the procedure we are hinting at here it is required that a relation of validity corresponding to $\models^G_h$ be defined for formulas of $\mathscr{L}^+_{GM}$;
2. unlike what happened with δ_G, the revision operator defined out of the said procedure will have the character of a 'genuine' operator rather than a one-to-one mapping between hypotheses.

Proviso no. 1 causes no difficulty as the validity relation that is needed will differ from the previous one only insofar as we are required to provide a specific clause for validity of the new type of formulas that count as expressions of the extended language, namely formulas of the form $u^i(s) \geq u^i(t)$. This, however, will be done easily, by reflecting the intended meaning we are willing to attach to them (in particular, it will be said that a formula like $u^i(s) \geq u^i(t)$ is valid in this more precise sense if and only if $u^i_G(s^G) \geq u^i_G(t^G)$ is the case).

Proviso no. 2 is a direct effect of the new kind of payoffs distribution we are facing. As a matter of fact, by supposing a given strategy profile of the game to be made out of actions that are rational to the players, it may now be the case that two or more strategy profiles should be regarded as rational owing to the previous definition. Let us illustrate this fact by means of strategy profile $a_1 b_1$ in the previous game matrix. That is, let us suppose that action a_1 is rational to player 1 and action b_1 is rational to player 2.

As to player 1, this has the effect that action a_2 turns out to be most convenient as it ensures a payoff equal to 1 against b_1, which is rational to the opponent by hypothesis, while a_1 only allows her to score 0:



From the point of view of the new definition of rational action, this causes formula φ^+_2 to be valid, and this formula only.

The difference with the previous situation comes when we evaluate the hypothesis from the viewpoint of player 2. For, having supposed that a_1 is rational to player 1, we are looking at things as if we were in player

2's shoes and glanced at scores coming second in payoffs pairs lying on the topmost row of the diagram:

	b_1	b_2	b_3
a_1	(0,1)	(1,2)	(0,2)

This would lead us to reject action b_1 for sure, which ensures the lowest payoff. As for the other two possibilities, they should both be retained as rational since they grant player 2 one and the same payoff. As a matter of fact, both formula ψ_2^+ and formula ψ_3^+ turn out to be valid in this situation, which means that, having supposed a_1 to be rational to player 1 and b_1 to be rational to player 2, we are driven to consider a_2 to be rational to player 1 in the sense of the definition, and b_2 and b_3 to be rational to player 2. In terms of strategy profiles, this causes to revise the hypothesis that $a_1 b_1$ was rational to the players by the conclusion that both $a_2 b_2$ and $a_2 b_3$ are rational in the end.

The revision operator in this case would be defined to reflect that. This means that, in particular, if it is given the hypothesis $h = a_1 b_1$ as an argument it should return the *set* $H^+ = \{a_2 b_2, a_2 b_3\}$ as value (where the use of capital letters is justified by the different kind of objects these values turn out to be).

The previous observation has an additional effect on fixpoints. For, it should be clear that no single hypothesis can be fixpoint of such an operator since, in presence of a tie in the payoffs granted to any of the players, the value of it will never be a singleton, therefore will never coincide with the starting hypothesis. The reader can easily convince herself of that by verifying, owing to an argument which is similar to the one we have just illustrated, that starting from the hypothesis that a_1 be rational to player 1 and b_2 be rational to player 2, one similarly concludes that both $a_1 b_2$ and $a_1 b_3$ are rational owing to the definition above, exactly as what happens under the assumption that $a_1 b_3$ be rational instead.

To get fixpoints in this case then, we have to think also of arguments of the revision operator as sets of hypotheses rather than single hypotheses only (which, by the way, is consistent with how fixpoints were defined in definition 3.9). Then, it turns out for instance that the *set* of hypotheses $H = \{a_1 b_2, a_1 b_3\}$ leads to a revised value $H^+ = \{a_1 b_2, a_1 b_3\} = H$.

Let us suppose that all of the previously mentioned steps have been taken care of, and that the definition of a modified revision operator, call it Δ_G to mark the difference from the previous δ_G, has been achieved[14].

[14]It might have occurred to the reader that the two provisos that make the difference

Having comprised the intuition that boosts the concept of equilibrium into the new definition of «rational action», which in turn determines the values of the revision operator exactly as in the previous treatment of strict games, it does not come as a surprise that one could easily show that a set of hypotheses H is fixpoint of the (newly defined) revision operator Δ_G, whatever two-player game in normal form G is supposed to be given, if and only if H contains only elements that are equilibria of G (i.e., if and only if, for every $h \in H$ with $h = a_i b_j$, $a_i^G b_j^G$ is an equilibrium of G). Furthermore, it is just a matter of observation to conclude that $\theta^+(x, R)$ (in fact, *any* formula $\theta_G^+(x, R)$ which is obtained by adapting it to the features of any two-player game in normal form as it was done before with formulas $\theta_G(x, R)$), is R-positive: again, it simply follows from the fact that in passing from $\mathscr{L}_{GM}$ to $\mathscr{L}_{GM}^+$ no negation sign was added. Therefore, it follows as a corollary of the these results and theorem 3.1, the following general proposition which we state here precisely (without proving, though, having hinted at all the ingredients of its proof), to let the reader appreciate the difference in scope with respect to theorem 3.2 about finite games which are strict:

Theorem 3.4 *Let $G = \langle \{p_1, p_2\}, (\Sigma_G^i)_{i \in \{1,2\}}, u_G \rangle$ be a two-player game in normal form. Let $\theta_G^+(x, R)$ be the formula of $\mathscr{L}_{GM}^+$ providing us with the defining condition of «rational action» for G. Then, there exists a set of hypothesis H which is a fixpoint of Δ_G, hence, such that $h^G = a_i^G b_j^G$ is an equilibrium of G for every $h \in H$.*

3.9. What if the opponent trembles?

So, every two-player game G in normal form has equilibria, being it strict or not. However, this does not help us with the issues that were raised in section 3.7. On the contrary, the result we have just presented can be taken as matter for arguing how pervasive those issues are, and how urgent is the need of addressing them. Reasoning on the new situation that we have started to consider here can be of help. For, it is even

between the modified route we are here describing and the original one, are not independent one of the other. For, having noticed that the revision operator should be 'genuine' as we were saying, hence that it should let a set H of hypotheses correspond to a unique set $\Delta_G(H) = H^+$, we also would like that the latter set contained hypotheses which are valid, in the sense specified according to the first proviso, granted H. This means that also the validity relation should be defined to accommodate the set-theoretic nature of H, while the previous validity relation $\models_h^G$ was introduced for single hypotheses alone. This, however, can be easily solved under the assumption that a corresponding validity relation has previously been defined, call it $(G, h) \models^+$, in such a way that $(G, h) \models^+ \theta$, where θ is a formula of $\mathscr{L}_{GM}^+$, if and only if this is valid relatively to two-player game in normal form G and to a single hypothesis h. Then, the required extension of it to sets of such hypotheses might be defined as a relation $(G, H) \models^+$ in such a way that, concerning the modified defining condition $\theta^+(x, R)$ of $R(x)$, we have, for every $s \in ATERM_{GM}$, $(G, H) \models^+ \theta^+(s, R)$ if and only if $(G, h) \models^+ \theta^+(s, R)$ for some $h \in H$.

clearer in a situation where strict distribution of payoffs breaks down, that one of the most evident shortcomings of the notion of 'solution' to a game we have been referring to so far is the fact that it only requires a player to reason individualistically: no thought is made about the other player's opportunities except for what concerns the attempt of determining where on the game matrix lies the balance of their mutual interests. Is there a reasonable way to deepen the player's concern for the opponent's choice of actions and incorporate it into her attempt of determining her best reply? In the case which we were analyzing before, which we reproduce here again for the reader's sake, such a more careful study of the situation might take the form of the following reasoning:

	b_1	b_2	b_3
a_1	(0,1)	(1,2)	(0,2)
a_2	(1,0)	(0,1)	(-1,0)

Having realized that there is no way for her to try deciding between playing b_2 or playing b_3 by some other reason, player 2 may be willing to consider unpredictable events. For instance, what happens if player 1 makes a mistake? What if, having decided to play a_1 as it is sensible to expect, player 1 trembles and ends up playing a_2 instead? Is it still indifferent whether I, player 2 would think, decided to play b_2 of b_3 instead? The consideration of this possibility causes player 2 to give another look at the matrix, by also considering what outcomes she would get by playing either of the two actions in question:

	b_2	b_3
a_1	(1,2)	(0,2)
a_2	(0,1)	(-1,0)

She would then notice that actions b_2 and b_3 differ indeed in this scenario: for, while b_2 would ensure her a payoff equal to 1, b_3 would fail to do so and she would score 0 by choosing it. So, she would conclude, action b_2 ensures the highest payoff in case player 1 plays as she should, and chooses her best action; if this did not happen, b_2 is the action that allows player 2 to still achieve the highest payoff she could possibily get. Then, it is superior to action b_3, and should be player 2's choice.

The previous reasoning suggests that definition 3.1 be changed in the following way.

Definition 3.13 *In a two-player game in normal form a strategy profile is a* trembling-hand equilibrium *if and only if no player has benefit from changing her strategy unilaterally, even under the assumption that the opponent may make mistakes.*

Let us just consider a couple of situations in addition, to familiarize further with the new idea. First of all, the following, simple one:

	b_1	b_2
a_1	(2,2)	(2,1)
a_2	(1,2)	(2,2)

This game is easily seen to have two equilibria, namely $a_1 b_1$ and $a_2 b_2$. As a matter of fact, neither player 1, nor player 2 might have any intention to 'leave' the former strategy profile, that is to play differently in case the opponent does not, since they would both decrease the expected payoff. Player 1 would indeed pass from scoring 2 to scoring 1:

	b_1
a_1	(2,2)
a_2	(1,2)

and so would be doing player 2:

	b_1	b_2
a_1	(2,2)	(2,1)

The situation is similar with strategy profile $a_2 b_2$, which is the outcome player 1 would like to stick to since, by changing her choice of action to a_1, would let her achieve the same score of 2 and get not benefit then:

	b_2
a_1	(2,1)
a_2	(2,2)

The same would hold true for player 2, and her choice of action b_2:

	b_1	b_2
a_2	(1,2)	(2,2)

If these two equilibria are compared with one another both player 1 and player 2 would be indifferent about which one to choose. To realize that they actually should be preferring strategy profile $a_1 b_1$, is something that is now possible to conclude by analizing the situation with the new intuition behind definition 3.13. As a matter of fact, player 1 notices that if she plays a_2 and player 2 trembles, then she decreases her score from 2 to 1:

	b_1	b_2
a_2	(1,2)	(2,2)

The same would hold true for player 2, and her choice of action b_2:

	b_2
a_1	(2,1)
a_2	(2,2)

If strategy profile $a_1 b_1$ is analyzed under the hypothesis of the opponent being mistaken, then it is clear that the result is different. For, player

1 would be able to maintain her score of 2 even if player 2 trembled and played b_2 instead of b_1:

	b_1	b_2
a_1	(2,2)	(2,1)

On her part, player 2 observes the same:

	b_1
a_1	(2,2)
a_2	(1,2)

This would lead them both thinking that playing a_1 and b_1 better allows them to protect their income against the opponent's trembles, and would stick to $a_1 b_1$ as anticipated.

A more complex situation that can be used to further illustrate the idea behind definition 3.13, is the following:

	b_1	b_2
a_1	(2,1)	(1,1)
a_2	(1,2)	(0,2)
a_3	(0,0)	(3,1)

As an initial observation, let us just notice that the game in question features two equilibria in the previous sense of the expression (that is, in the sense of definition 3.1). One of them is strategy profile $a_1 b_1$ since, by changing her strategy unilaterally, player 1 would pass from scoring 2 to score either 1, in case she chose action a_2, or 0 if she chose action a_3 instead:

	b_1
a_1	(2,1)
a_2	(1,2)
a_3	(0,0)

As to player 2, by changing her choice of action from b_1 to b_2, she would still score 1 (hence, no benefit) in case player 1 stuck to choosing a_1:

	b_1	b_2
a_1	(2,1)	(1,1)

If we pass to considering strategy profile a_3b_2, then we notice that similarly player 1 has no need of changing her strategy unless player 2 also does that. For, instead of scoring 3, she would get a null payoff by choosing action a_2, and just a payoff equal to 1 in case she chose a_1:

	b_2
a_1	(1,1)
a_2	(0,2)
a_3	(3,1)

Player 2, on her part, would also experience a loss, by passing from scoring 1 to scoring 0 instead:

	b_1	b_2
a_3	(0,0)	(3,1)

So, both strategy profiles, $a_1 b_1$ and $a_3 b_2$, are equilibria in the old sense. Yet, the two equilibria are indifferent from the point of view of player 2, who scores 1 in both cases. This would make it hard for her to decide whether it is better to play b_1 or b_2 instead. However, things are different if we assume that player 2 weighs the same situation by assuming the new concept of equilibrium which makes use of the trembling-hand intuition. As a matter of fact, player 2 would easily realize that only strategy profile $a_3 b_2$, hence only action b_2, manages to protect her from possible 'trembles' of player 1: for, by sticking to b_2, she would maintain the score of 1 she gets in case player 1 makes no mistake and plays a_3, which is the same she scores if player 1 mistakenly played a_1, while she even increases her payoff, scoring 2 instead of 1, if player 1 takes a_2 as her wrong action:

	b_2
a_1	(1,1)
a_2	(0,2)
a_3	(3,1)

If player 2 chose action b_1 instead, and player 1 would not play a_1 as expected by mistake, then her payoff would equally be incremental in case player 1 trembled and played a_2, as she would score 2 (the same as if she played b_2 and player 1 made the same mistake), but she would experience a loss if only player 1 played a_3 by mistake:

	b_1
a_1	(2,1)
a_2	(1,2)
a_3	(0,0)

This makes strategy profile $a_3 b_2$ the only trembling-hand equilibrium of the game, and choice of action b_2 preferable to b_1 for player 2.

Now, even before trying to address the question whether or not it is possible to incorporate the new intuition into a definition of «rational action», in a similar fashion as this was done in section 2, it would be better to anticipate here, as the reader might have already guessed that, that the new definition too has shortcomings. First of all, it is perfectly possible to imagine a situation where payoffs are distributed in such a way that equilibria which are clearly preferrable collectively, do not result to be captured even by making use of the new intuition. This is the case for instance of the following game:

	b_1	b_2
a_1	(1,1)	(2,0)
a_2	(0,2)	(2,2)

If analyzed with respect to the original idea of equilibrium, the one from definition 3.1, both strategy profile $a_1 b_1$ and strategy profile $a_2 b_2$ falls under that concept. In the evaluation of the former strategy profile, it turns out that player 1 would decrease her score by changing her mind:

	b_1
a_1	(1,1)
a_2	(0,2)

and the same would be true for player 2:

	b_1	b_2
a_1	(1,1)	(2,0)

Also, player 1 would not be gaining anything more by passing from having chosen a_2 to choose a_1 instead, in case strategy profile $a_2 b_2$ is analyzed:

	b_2
a_1	(2,0)
a_2	(2,2)

and player 2 too, would have no reason to be the one who wants to change her mind in that situation:

	b_1	b_2
a_2	(0,2)	(2,2)

Out of these two equilibria in the previous sense of the expression, only one is matching the requirement to be named as such owing to definition 3.13. However, contrary to any expectation, this is going to be $a_1 b_1$, which is also clearly not the one the players would like to choose as they get less out of it with respect to what they would get if strategy profile $a_2 b_2$ was the chosen one instead. As a matter of fact, player 1 would notice in this case that sticking to a_1 allows her to even increase her payoff in case player 2 trembled:

	b_1	b_2
a_1	(1,1)	(2,0)

and the same is true for player 2:

	b_1
a_1	(1,1)
a_2	(0,2)

This is not what happens with the players' choice if strategy profile $a_2 b_2$ is concerned, owing to the fact that player 1 would pass from a score of 2 to a null score if player 2 trembled in this case:

	b_1	b_2
a_2	(0,2)	(2,2)

and the situation for player 2 would be exactly the same:

	b_2
a_1	(2,0)
a_2	(2,2)

So, playing the strategy profile $a_2 b_2$ would mean failing for both players to protect themselves against the opponent's tremble, which means that this is not a trembling-hand equilibrium, while $a_1 b_1$ is.

Another shortcoming of the new intuition comes again by considering ties. This was the same with equilibria defined according to definition 3.1. One of the reasons we considered for refining the idea behind that concept was given to us by noticing how poorly equilibria in that sense of the word allowed us to deal with ties. The same is true with the new intuition, although, differently from what happened then, ties we shall be considering now do not involve directly the solutions to a game that are put forth on the basis of definition 3.13, but are those placed somewhere else on the game matrix. Let us just say a few words more about that to explain the problem.

The very intuition we have used to come up with the idea of strategy profiles being equilibria in the new sense, can be summarized as follows: the need for the new idea to step in is the existence of a tie in the game, that is, the existence of two or more actions which are equally good replies to one and the same action chosen by the opponent; in the worst case, such a tie involves actions occurring into strategy profiles that turn out to be equilibria in the sense we previously were referring to as solutions of the game (as it was happening in the situation we have used for the sake of our critical re-consideration of that concept in section 3.7, and that we kept using as our running example through section 3.8). Suppose we do have two strategy profiles of this sort, $a_i b_j$ and $a_h b_j$, and let us say that player 1 is in trouble since a_i and a_h are indeed the best reply of hers to action b_j of player 2's, but they both ensure her one and the same payoff. Then, player 1 would be inspecting the matrix of the game and consider all strategy profiles that turn out from it by supposing that player 2 may tremble. She would be interested in particular in those actions combinations which comprise a_i and a_h as her own reply to player 2's action, and would compare them pairwise: $a_i b_k$ with $a_h b_k$, $a_i b_l$ with $a_h b_l$, and so on. She would then be in a position to make a choice between playing a_i or a_h when she is able to locate one strategy profile in that list in which one of the two actions is best reply to player 2's action, that is, grants her a better payoff than the one granted to her by the alternative action. To make this possible and for the new intuition to be applicable, while it is true indeed that the game need not to be strict as a whole, strictness cannot fail everywhere, at least not on startegy profiles which, like those in the list before, are obtained for the sake of testing what is the achievement that any two, or more actions of a player that are best if the opponent played rational, allows her to accomplish against the opponent's trembles. We call games whose matrix meets this requirement *quasi-strict*.

Quasi-strict games do admit an exact definition, which we reproduce here for the sake of completeness. We warn the reader, however, that the combinatorics required by it can make it hard to understand. We suggest then to go through it by keeping in mind the intuitive explanation of it:

Definition 3.14 *Let $G = \langle \{p_1, p_2\}, (\Sigma^i_G)_{i \in \{1,2\}}, u_G \rangle$ be a two-player game in normal form. Let $s, s', \ldots, s^*, \ldots, s^\dagger, s^\ddagger, \ldots, s^\#, \ldots$ be used to refer to strategy profiles in G, that is for actions combinations of the form $a_i b_j$ for $a_i \in \Sigma^1_G$ and $b_j \in \Sigma^2_G$. Let S_G be a shorthand for the set of such strategy profiles in G. Also, for every strategy profile s, we indicate by s_i for $i \in \{1,2\}$, player p_i action in s (i.e., if $s = a_i b_j$ then $s_1 = a_i$ and $s_2 = b_j$). Then, we put:*

1. *For any two strategy profiles s, s' and for every $i \in \{1,2\}$, we say that $s =_i s'$ holds if s and s' contain the same strategy for player p_i but are not identical (that is, if $s_i = s'_i$ but $s_j \neq s'_j$ for p_j with $j \neq i$);*
2. *G is* quasi–strict *if, in case there exists $i \in \{1,2\}$ and $s, s' \in S_G$ with $s \neq_i s'$ and $u^i_G(s) = u^i_G(s') = \max\{u^i_G(s'') \mid s'' \in S_G, s'' \neq_i s\}$, then there are $s^*, s^\dagger \in S_G$ such that:*
   2.1 *$s^* =_i s$, $s^\dagger =_i s'$, $s^* \neq_i s^\dagger$ and $u^i_G(s^*) > u^i_G(s^\dagger)$ or $u^i_G(s^\dagger) > u^i_G(s^*)$;*
   2.2 *if $u^i_G(s^*) > u^i_G(s^\dagger)$, then for all $s^\ddagger \in S_G$, if $s^\ddagger =_i s^\dagger$, then $u^i_G(s^\#) \geq u^i_G(s^\ddagger)$ where $s^\# =_i s^*$ and $s^\# \neq_i s^\ddagger$;*
   2.3 *if $u^i_G(s^\dagger) > u^i_G(s^*)$, then for all $s^\ddagger \in S_G$, if $s^\ddagger =_i s^*$, then $u^i_G(s^\#) \geq u^i_G(s^\ddagger)$ where $s^\# =_i s^\dagger$ and $s^\# \neq_i s^\ddagger$.*[15]

Granted that quasi-strictness is a necessary requirement to the new intuition we are considering, let us focus on that. To keep things simple, let us do more. Let us consider a concrete example like the one from which this whole thread of thoughts started and let us take it up again here:

[15] Notice that, according to part 1 of this definition, in a two-player game we have that $s \neq_i s'_j$ holds if and only if $s =_j s'$ does not hold for every strategy profiles s, s'. However, this is no more true in n-player games for $n > 2$. This makes it redundant for the G considered here to require, in part 2.1 of the definition, that $s^* \neq_i s^\dagger$ holds, since $s^* =_i s$ and $s^\dagger =_i s'$ already imply that $s^*_j = s^\dagger_j$ is the case for $j \in \{1,2\}$ with $j \neq i$. Although we are not concerned with the extension of this approach to the n-player case (not until section 3.11 at least), we preferred to give the reader the 'full' definition, which goes through that general situation also. In addition, notice that in a quasi-strict game in normal form G, for no player are there actions which are fully equivalent (i.e., actions which guarantee the same payoff under all possible circumstances).

	b_1	b_2	b_3
a_1	(0,1)	(1,2)	(0,2)
a_2	(1,0)	(0,1)	(-1,0)

Roughly speaking, the idea we have put forth to deal with cases like this one goes as follows: as long as strictness is in place use that; when it fails, resort to the trembling-hand intuition to solve ties. As a matter of fact, definition 3.14 does not exclude that the game G that is considered there be strict. The quasi-strictness prescription of part 2 of it only takes place *if* strictness fails. This means that the collection of quasi-strict games does contain strict games also as elements. If we are willing to provide ourselves with a new definition of «rational action» in a game that is quasi-strict, the possibility that it must be applied to strict games cannot be dismissed. However, we have already devised an approach for quasi-strict games of this peculiar sort, have not we? Therefore, if the game we are considering is strict, an action is rational if and only if the same defining condition we were using so far is valid.

So, suppose that we wish to determine and comprise this condition in a formula of language $\mathscr{L}_{GM}$, when it is the case that a_1 is rational to player 1 in the previous game. Owing to what we just noticed, the possibility that game G be strict cannot be ruled out (even though it is clear that G is not so by looking at its matrix), because we are willing to devise a definition of rationality for actions that may work for all quasi-strict games. As we said, if G were strict, then a_1 would be rational to player 1 in case it was her best reply to the action that is rational to player 2. That is, if the following formula of $\mathscr{L}_{GM}$

$$\bigvee_{1 \leq j \leq 3} (R(b_j) \wedge u^1(a_1 b_j) > u^1(a_2 b_j))$$

is valid. Let reference to this formula be shortened by calling it $Strict^1_1$ (where indices are chosen to keep track of the fact that it 'says' that action a_1 is strictly the best choice to player 1 – the upper 1 indicating the player's number, the lower one being connected with the action term index). This just replicates for the case here at stake the intuition that was used for the sake of defining the property «rational action» for all elements in the class of strict games: it indeed allows us to pick a_1 if the latter is the most convenient reply of player 1 to player 2's rational action.

Although G might be strict, it is also possible that it is not. In that case actions that are rational to the players are selected by making use of the trembling-hand intuition. In the case of a_1 from the game matrix above, this intuition corresponds to a situation in which, given the action that is rational to player 2, say b_1, a_1 fails to be strictly the best choice that player 1 can make because it happens to grant her the same payoff that a_2 does. However, it turns out to be the action that better protects her from player 2's trembles, for instance, because it also ensures a payoff that is not less than those granted to player 1 by a_2 with respect to one of the two actions on which player 2 may tremble, say b_3, but does ensure a strictly better payoff than a_2 when conceived as a reply to the remaining action of player 2's.

This idea, literally correspond to the validity of the following formula of $\mathscr{L}_{GM}$:

$$(R(b_1) \wedge u^1(a_1 b_1) = u^1(a_2 b_1) \wedge u^1(a_1 b_2) > u^1(a_2 b_2) \wedge u^1(a_1 b_3) \geq u^1(a_2 b_3))$$

(the reader might have noticed that the identity sign = here is applied to payoff terms of $\mathscr{L}_{GM}$, which is different from how things used to be in previous sections; this however can be fixed by a routine addition of a clause in the defining condition of the set of formulas of $\mathscr{L}_{GM}$ to allow such a modified use). Alternatively, it may happen that a_1 should be chosen according to the new intuition behind the refined concept of equilibria from definition 3.13, because it does at least as good as a_2 as a reply to b_2, but it beats a_2 if chosen in response to b_3 instead. That is:

$$(R(b_1) \wedge u^1(a_1 b_1) = u^1(a_2 b_1) \wedge u^1(a_1 b_2) \geq u^1(a_2 b_2) \wedge u^1(a_1 b_3) > u^1(a_2 b_3))$$

These possibilities together can be comprised in just one single formula of $\mathscr{L}_{GM}$, by means of the disjunction sign (since it is either one case, or the other that indicates that a_1 is the best reply by player 1, in the sense of the trembling-hand intuition, to the rational action b_1 by player 2). Let again reference to this compound formula be shortened by referring to the said disjunction as $Tremb^1_{(1,1)}$, that indicates that action a_1 is the best reply that player 1 can stick to in response to rational action b_1 of player 2 (again, the upper 1 being the index that refers to the player, and the lower pair of numbers being those that keep track of the actions we are considering: the first element of the pair being the index of player 1's action, the second the index of the action of player 2).

Let us suppose that formulas $Tremb^1_{(1,j)}$ of $\mathscr{L}_{GM}$ for $j \in \{2,3\}$ have been devised as well. Owing to the convention on the notation we have made, formulas $Tremb^1_{(1,2)}$ and $Tremb^1_{(1,3)}$ express the fact that a_1 is the best reply by player 1 in the trembling-hand sense to rational actions b_2 and b_3 respectively.

Let also φ_1' be the following formula of $\mathscr{L}_{GM}$:

$$\varphi_1' = Strict^1_1 \bigvee_{1 \leq j \leq 3} Tremb^1_{(1,j)}$$

This formula expresses the fact that action a_1 represents player 1's most convenient choice because either it is strictly the best reply to the action that is rational to player 2, or because is the best way by means of which player 1 can protect herself from player 2's trembles.

Suppose that formulas φ_2', which says of a_2 what φ_1' says of a_1, as well as formulas ψ_j' with $j \in \{1, 2, 3\}$ doing the same for actions b_1, b_2 and b_3 respectively, have also been devised by obviously modifying the formulas considered above. Then, we are in a position to use these formulas to piece together a new definition of «rational action» for the game under consideration which looks like the previous one:

$$R(x) \Leftrightarrow_{Def} \bigvee_{1 \leq i \leq 2} (x = a_i \wedge \varphi_i') \bigvee_{1 \leq j \leq 3} (x = b_j \wedge \psi_j')$$

The main difference is that formulas φ_i' and ψ_j' this is made out of incorporate the intuition behind definition 3.13 of trembling-hand equilibria. As a consequence, if we suppose then that a new revision operator δ_G' has been defined by modifying what we did in section 3.4 with definition 3.6, then we would expect that fixpoints of δ_G' coincied with trembling-hand equilibria of the game. This can be easily verified in our running example by means of the following informal reasoning.

Suppose then $h = a_1 b_2$. Since b_2 is assumed to be rational, a_1 is easily seen as the best reply to it that player 1 can choose:

	b_2
a_1	(1,2)
a_2	(0,1)

It follows then that formula $u^1(a_1 b_2) > u^1(a_2 b_2)$ is valid relatively to G and the given hypothesis, which yields that $Strict_1^1$ is also valid, and φ_1' as well. Therefore, presuming that $\theta_G'(x, R)$ indicates the defining condition of the new definition of «rational action» (i.e., the big formula that is placed right-hand to the definition sign $\Leftrightarrow_{Def}$), we have that $\theta_G'(a_1, R)$ is valid relatively to this G under the hypothesis that b_2 is rational to player 2.

On the other hand, with a_1 being rational by hypothesis, it results from the game matrix that b_2 and b_3 do equally well as player 2's reply to it:

	b_1	b_2
a_1	(1,2)	(0,2)

This means that the formula $u^2(a_1 b_2) = u^2(a_1 b_3)$ of $\mathscr{L}_{GM}$ is valid in this setting. The new idea of equilibrium suggests that to solve ties of this sort, players would go and see how the actions behave with respect to possible trembles of the opponent. Here we have just one of them, and inspection of the game matrix in this case makes b_2 better than b_3:

	b_1	b_2
a_2	(0,1)	(-1,0)

This corresponds to formula $u^2(a_2 b_2) > u^2(a_2 b_3)$ being valid. This is enough to conclude that $Tremb^2_{(1,2)}$, which expresses that action b_2 is the trembling-hand best reply by player 2 to action a_1,[16] is valid and ψ'_2 is valid as a consequence of it (since it contains the previous formula as one of its disjuncts). Therefore, we conclude that $\theta'_G(b_2, R)$ is similarly valid in this case which, together with what we concluded above, yields $\delta'_G(a_1 b_2) = a_1 b_2$ as expected.

As the reader can easily verify herself, this is the only strategy profile providing us with a fixpoint of the modified revision operator. Moreover, the analysis we have performed can be taken as paradigmatic of what happens in the general case. As a matter of fact, having carried out things in all details would put us in the position of proving a general result which, like theorem 3.2, establishes the exact correspondence between fixpoint of the modified revision operator δ'_G and equilibria in the sense of definition 3.13. So, the new intuition for solving ties that allowed us to stress the defects of the previous definition of equilibrium in a finite game, is prone to being approached by the same methodology. Alas, this intuition works well only on games which are quasi-strict. This means that it is not difficult to come up with situations that we are uncapable of coping with by means of it. Hence, some new idea is required to enlarge the number of games we can solve.

[16] Recall that in the pair of numbers which serves as the lower index of formulas of this sort, the first element is the index of the action by player 1 and the second element is the one indicating player 2's action.

3.10. Proper equilibria

Let us consider now the following situation:

	b_1	b_2
a_1	(0,0)	(8,2)
a_2	(5,5)	(5,5)
a_3	(-2,7)	(3,6)

This is again a situation that requires some deal of reflection. Not really for what concerns player 1's actions, since the situation is clear from that respect. For, in case we suppose b_1 to be rational to player 2, for instance, then certainly player 1 would have no doubts in choosing a_2 as a reply to it owing to how payoffs are distributed to players in this game:

	b_1
a_1	(0,0)
a_2	(5,5)
a_3	(-2,7)

No doubts also would she have in case player 2's rational choice of action were b_2 instead, since a_1 would stands out as a natural reply to it:

	b_2
a_1	(8,2)
a_2	(5,5)
a_3	(3,6)

The situation would be different from player 2's angle. For, while she would have no doubts either, both in case action a_1 were rational to player 1:

	b_1	b_2
a_1	(0,0)	(8,2)

and in case a_3 were rational instead:

	b_1	b_2
a_3	(-2,7)	(3,6)

player 2 would have more difficulties in establishing what she is supposed to reply to player 1's choice of action, if action a_2 were the rational one to pick. As a matter of fact, player 2 in this case would have to face a tie similar to those we became acquainted with in the previous section:

	b_1	b_2
a_2	(5,5)	(5,5)

Then, she might be willing to solve it according to the «trembling-hand» intuition. She would then consider player 1's possible trembles. By looking at a_1 first, player 2 would then realize that b_2 offers her the best way to protect her score, as it is clear by looking at the two top-most rows of the game comparatively:

	b_1	b_2
a_1	(0,0)	(8,2)
a_2	(5,5)	(5,5)

But then, to make her final choice, player 2 would have to consider the other tremble of player 1's, since it might well be possible that her opponent's mistake would lead her to play a_3 instead of a_2. Then, a look at the two bottom rows of the game would lead player 2 to make an unpleasent discovery:

	b_1	b_2
a_2	(5,5)	(5,5)
a_3	(-2,7)	(3,6)

Payoffs now indicate that b_1 is instead the action to choose to prevent the opponent's mistake: for, under the said circumstance, it allows player 2 to increase her payoff up to 7, while choosing b_2 would raise it only up to 6.

Here lies the problem with all that, since we are not prepared to deal with such a split decision, when by considering one mistake of the opponent we are lead to choose one action, which is not the same one we are lead to choose by considering another possible mistake in the same game. The conclusion is not surprising since, by glancing at the game matrix as a whole first we would have realized that it does not respect the quasi-strictness requirement of definition 3.14. This means that the intuition that brought us to the modified concept of equilibria from definition 3.13, is not good to deal with it. This also means that the concept of trembling-hand equilibrium we came up with must be further refined and we need an idea for that. The idea we might are now willing to test in this respect is the following.

Finding herself in troubles, player 2 may be glancing again at the payoffs distribution just to suddenly realize that she was making a mistake: she was considering both trembles as if they were the same, while this is clearly not the case. The point is that, if player 1's trembles are actions chosen by mistake, then they should reflect how costly these are to her. For, it seems fair to assume that player 1 would not be making a mistake that costs too much to her because she would be paying a lot of attention to avoid that and double-check what she is doing to ensure she is not doing anything stupid. Rather, it is more probable that she could be more superficial while overseeing choices that are close to one another, as long as the payoff they ensure is concerned. In the situation above, this line of reasoning would lead player 2 to notice that mistaking a_3 for a_2 is indeed a costly error, since it causes player 1 to pass from a score of 5 to a score of -2 (which is a loss equal to 7 on that side of the matrix) if player 2 happened to play b_1, and from a score of 5 to a score of 3 (i.e., a total loss of 2) if player 2's choice of action were b_2 instead:

	b_1	b_2
a_1	(5,5)	(5,5)
a_2	(-2,7)	(3,6)

In particular, this is a more costly mistake than choosing a_1 instead of a_2, since player 1's loss is equal to 5 if player 2 played b_1, and is even beneficial to player 1 in case player 2 played b_2:

	b_1	b_2
a_1	(0,0)	(8,2)
a_2	(5,5)	(5,5)

So, player 2 would conclude that player 1's first mistake is unlikely, while the other is rather possible and it is the only 'proper' mistake she has to protect herself from. Since $u^2(a_1 b_2) > u^2(a_1 b_1)$, then b_2 does better against a_1 and she ends up playing b_2.

The idea we are making use of to solve ties which would remain unresolved if approached by the intution we were using up to now, is easy to comprise in a definition that refines the concept of «equilibrium» in a game in the following way:

Definition 3.15 *In a two-player game in normal form, a strategy profile is a* proper equilibrium *if and only if no player has benefit from changing her strategy unilaterally, even under the assumption that the opponent may make mistakes that are not costly to her.*

Of course, it is required to make the idea of a 'costly' mistake precise for actually making it usable. Moreover, like it was the case before with the refinement that brought us to complete the definition of «equilibrium» as it was given originally with the trembling-hand intuition, one may wonder whether it is possible to incorporate this idea into our fixpoint machinery of revision of hypotheses. The good news is that we can address both issues at once, since dealing with the latter will also make clear how one can deal with the other.

As a matter of fact, while trying to think how to use the above intuition to further modify the definition of «rational action» in finite games and make it usable for the sake of solving games that are not quasi-strict, we are required to state precisely (and in a formal way) what makes player 1's mistake of choosing a_3 instead of a_2 more costly, hence less probable than her mistake of choosing a_1 in place of a_2. This has to do, according to the analysis we have put forth, with how much she is expected to accomplish by the two mistakes in question. The idea is that the less convenient is the mistake, the less probable it is. Or, which is the same thing, the more convenient it is, the more probable it is. Concretely speaking, the whole matter is explained owing to the fact that, in the game above, we have that both

$$u^1(a_1 b_1) > u^1(a_3 b_1)$$

and

$$u^1(a_1 b_2) > u^1(a_3 b_2)$$

are the cases. The simple truth is: to play a_1 is more convenient to player 1 than a_3, whatever is the choice of action made by player 2. So, according to what we said, player 1 would not be making a mistake by choosing a_3 when she is supposed to play a_2: she would be too much careful to avoid that mistake, since she has so much to loose. Then, if any mistake of hers is possible, this will be likely the be the other one, that leads her to mess up with a_2 and a_1 instead.

Let us now recapitulate one by one all the ingredients needed to apply this new intuition to make things clear, and what is the requisite of an action in the game, say b_2 of player 2, to be picked as reply to a_2 because it embodies the intuition behind definition 3.15.

First of all, having supposed a_2 to be rational to player 1, it must be the case that b_1 and b_2 ensure one and the same payoff to player 2. If a tie was not present, no problem for player 2 would have ever occurred while going through the various alternatives. Then, this situation corresponds to what expresses the following formula of $\mathscr{L}_{GM}$:

$$R(a_2) \wedge u^2(a_2 b_1) = u^2(a_2 b_2)$$

Let us call it $1Prop^2_{(2,2)}$ to say that this is the first part of the formula of $\mathscr{L}_{GM}$ which is globally valid when action b_2 is rational to player 2 owing to the properness intuition of definition 3.15, under the hypothesis that a_2 is rational to player 1 (where, like before, the upper index indicates the player to whose action the properness intuition is applied, while the lower pair of indices contains the index of the action by player 1 which is rational by hypothesis first, and the index of the action of player 2 that is conceived as the reply to it).

Secondly, it also must be the case that the «trembling-hand intuition» must be unusable for both b_2 and b_3, which means, owing to what we did in section 9, that both formulas $Tremb^2_{(2,1)}$ and $Tremb^2_{(2,2)}$ fail to be valid. To put things in this way may cause troubles since, in order to express the failure of a formula in our formal language, we would be forced to introduce a negation sign which we have not been using so far. This, in turn, may lead us to produce a formula which is no more R-positive, with no way of making use of the results about fixpoints of monotone operators we have hinged upon up to now. Luckily this is not needed. As a matter of fact, it turns out that what we are required to refer to here is the payoff distribution that causes quasi-strictness not to apply, and this amounts to assume that the following holds:

$$\begin{array}{l}(u^2(a_1 b_1) = u^2(a_1 b_2) \wedge u^2(a_3 b_1) = u^2(a_3 b_2))\vee \\ (u^2(a_1 b_1) > u^2(a_1 b_2) \wedge u^2(a_3 b_2) > u^2(a_3 b_1))\vee \\ (u^2(a_1 b_2) > u^2(a_1 b_1) \wedge u^2(a_3 b_1) > u^2(a_3 b_2))\end{array}$$

That is: either the game is not quasi-strict because b_1 and b_2 ensure one and the same payoff to player 2 if played against both a_1 and a_3, or it fails because b_1 is strictly best in one case and b_2 is strictly best in the other, or it fails because the reversal situation is taking place. Let the formula above be indicated as $2Prop^2_{(2,2)}$.

Finally, it must be the case that b_2 does better than b_1 against the less costly mistake of player 1's, which, according to what we were noticing above, turns out to also correspond to the validity of a formula of $\mathscr{L}_{GM}$. In particular, to the one that contains the following part:

$$\begin{array}{l}(u^1(a_1 b_1) > u^1(a_3 b_1) \wedge u^1(a_1 b_2) \geq u^1(a_3 b_2) \wedge u^2(a_3 b_2) > u^2(a_3 b_1))\vee \\ (u^1(a_1 b_1) \geq u^1(a_3 b_1) \wedge u^1(a_1 b_2) > u^1(a_3 b_2) \wedge u^2(a_3 b_2) > u^2(a_3 b_1))\end{array}$$

(that is: a_3 is a less costly mistake than a_1 to player 1 and b_2 is a better reply to it than b_1), when is put together by means of a disjunction sign in between, with the following one:

$$\begin{array}{l}(u^1(a_3 b_1) > u^1(a_1 b_1) \wedge u^1(a_3 b_2) \geq u^1(a_1 b_2) \wedge u^2(a_1 b_2) > u^2(a_1 b_1))\vee \\ (u^1(a_3 b_1) \geq u^1(a_1 b_1) \wedge u^1(a_3 b_2) > u^1(a_1 b_2) \wedge u^2(a_1 b_2) > u^2(a_1 b_1))\end{array}$$

which means that a_1 is a less costly mistake than a_3 instead, and again b_2 is a better reply to it. Call $3Prop^2_{(2,2)}$ the disjunction of these two formulas, and let $Prop^2_{(2,2)}$ be the formula of $\mathscr{L}_{GM}$ that is obtained by conjoining all these three parts together, that is to correspond to the expression:

$$(1Prop \wedge 2Prop^2_{(2,2)} \wedge 3Prop^2_{(2,2)})$$

Suppose also that formulas $Prop^2_{i,2}$ for $i \in \{1, 3\}$ have also being devised to mean that b_2 is player 2's proper best reply under the hypothesis that a_1 is rational to player 1, a_3 is rational to her respectively.

Let also $Strict^2_2$ and $Tremb^2_{(j,2)}$ for $j \in \{1, 2, 3\}$ be as before. Then, let ψ''_2 be set as corresponding to the following formula of $\mathscr{L}_{GM}$:

$$Strict^2_2 \bigvee_{1 \leq j \leq 3} (Tremb^2_{(j,2)} \vee Prop^2_{(j,2)})$$

This expression can be taken to mean the following: either the game is strict and b_2 is strictly the best reply of player 2 to the rational action of player 1, or the game is quasi-strict and b_2 is the reply of player 2's to the rational action of player 1 which is the most robust one to possible mistakes of hers, or the game is not quasi-strict and b_2 is the action that player 2 should pick owing to the properness intuition.

Let ψ_1'' be the expression of $\mathscr{L}_{GM}$ that says of b_1 what ψ_2'' says of b_2. Let us also suppose that formulas φ_i'' for $i \in \{1,2,3\}$ have equally being devised. Then, the definition of «rational action» for the game under consideration turns out to be the following:

$$R(x) \Leftrightarrow_{Def} \bigvee_{1\leq i\leq 3} (x = a_i \wedge \varphi_i'') \bigvee_{1\leq j\leq 2} (x = b_j \wedge \psi_j'')$$

Notice that: (i) the logical structure of the formula has not changed; (ii) the defining condition is set to apply also to situations where quasi-strictness fails; (iii) the defining condition is R-positive, hence the revision operator δ_G'' we are now in a position to define out of that would be monotone, and subject to the application of results about the existence of fixpoints which are likely to correspond to proper equilibria.

3.11. Generalizing the analysis to all finite games

As the reader might have guessed it already, it is easy to come up with situations that would make useless even the new intuition of properness. In the end, it is just a matter of playing well with payoffs, like we have done here by modifying the score for players in the game we have considered in the previous section:

	b_1	b_2
a_1	(0,0)	(8,2)
a_2	(5,5)	(5,5)
a_3	(2,7)	(3,6)

The new state of the game is the effect of having turned what was a negative payoff delivered to player 1 in strategy profile a_3b_1, into a positive one. It should be clear that even such a small change in the game can make the previous analysis ineffective. As a matter of fact, that was depending upon one of the possible mistakes by the opponent of player 2 being costly the most to her. In turn, one action was regarded to be more costly than another in case a player has a lot more to loose by playing that, hence if the payoff she gets in all strategy profiles involving it is always strictly less than the payoff she would get by choosing some alternative action instead. This is no more so in the new version of the game: by mistakenly choosing a_1 instead of a_2, which we suppose to be the rational action to play, still player 1 gets a score of 8 in case the opponent plays b_2, that is greater than the score of 3 she would get by playing a_2 instead; however, she now gets a payoff of 2 by playing a_2 which is higher that what she gets by playing a_1, that entitles her to score 0, in case player 2 played b_1. This means that there is no more a clear indication as to what is the most costly mistake of hers. Therefore, player 2 has no way of choosing what to play under the hypothesis that a_2 be rational to player 1, and solve the tie like she was doing beforhand by means of the new idea of proper equilibrium.

So, just in case one was wondering whether the analysis by means of which we had happily resolved the problems bothering us in the previous section could be extended to all finite games, the plain answer is: no, we need something more. Rather than keep going this way, however, we feel that a change in the path we have been following could be more beneficial. This is not because we have run out of ideas on how to fix the bug. For instance, one may wish to apply a different approach by measuring the risk involved into playing certain actions rather than others, and use that as a mean for deciding a player's responding strategy. This approach could be fine-tuned in such a way to profile different kinds of players, like the risky one, who would always be inclined to run the risk involved into playing a given action, or the conservative one, who would be willing to run the risk only at certain favourable conditions.

It is not even true that we abstain from going into the deatils of these refinements because they just do not fit the methodology we have been incorporating our ideas into. Quite on the contrary, it is not difficult to express these ideas about risk and conservation into mutual relationships of payoffs in such a way that a new version of the defining condition of «rational action» in a game is reached, and a new definition of the revision operator along with it, to turn the problem of finding equilibria in the new form into the problem of calculating the fixpoints of it. However, I feel that what we have been treated here suffices for accomplishing the original aim of the chapter, which was meant to introduce some basic concepts in the treatment of games in normal form, as well as to let the reader get acquainted with the methodology for dealing with them I wanted to foster. To go on with further proposals for refining the concept of equilibrium, would not allow us to get anything more and, most importantly, would not allow us to reach a stage where all the problems we have been mentioning here get solved (on the contrary, some new, controverial issues would be piled up along the way).

Moreover, I have another reason for changing subject. This reason is connected with a new territory I feel we should explore, or, rather, with an area of the territory we have been exploring that has not been given the attention it deserves. This area is naturally connected with the idea of games in general, and with the idea of strategic thinking in general. Games offer us the opportunity to experiment with situations in which strategy is crucial, since players are required to make choices by taking into account both their goals and the goals the opponents are after. To find a solution that is optimal for the situation, it is required to think strategically. Normal form games, however, are a 'static' domain. The action of players is somehow 'freezed' and only the payoffs distribution is left as the preeminent parameter that drives the players' choices. This is so distant from how things happen in life, where the dynamic of the situation is another aspect to cope with. To get closer to the real thing, we would like to put some dynamics into games as we have considered so far. This is what we will be doing in the next chapter.

3.12. Bibliographical note

Games where introduced as objects for mathematical treatment in the pivotal book by John von Neumann and Oskar Morgenstern (von Neumann and Morgenstern 1944). Their idea was to open a new field of economic studies, which could use games as a tool for modeling agents' behaviour in a variety of situations according to a view that von Neumann in particular had envisioned earlier (von Neumann 1928). Despite the deep influential role played by this source, the actual theory of games (where by «actual» I mean closer to the type of approach to the issue we are pursuing in this volume), is the result of extensions and generalizations of von Neumann and Morgenstern original theory that took place in the decades following the publication of their book.

Among those who should be mentioned for having played a major role in this sense is the American mathematician John F. Nash, who published a few seminal papers in the early 1950's (Nash 1950a, 1950b, 1951), that eventually brought him to win the Nobel prize in Economics in 1994. Nash is mostly responsible for the theory of equilibria in finite games that we have been treating here from section 3.2 onwards. Actually, the notion of «equilibrium» as we have tried to frame it by means of definition 3.1 is commonly named after him, since it is the subject of one of Nash's most famous results about the existence of solutions of this sort in any finite game. Nash's proof was obtained by making use of other known results, primarily a theorem about the existence of fixpoints of continuous mappings named after the Dutch mathematician L.E.J. Brouwer.

The approach to equilibria we have proposed in this chapter does follow a different route, although it is one still centered on fixpoints as we have seen. There are two major sources for that. On the one hand, as far as classical results like theorem 3.1 from section 3.5 about fixpoints of monotone operators are concerned, we have been referring to results that have become standard references for work dealing with iterative constructions like (iterated) inductive definitions.

Inductive definitions play a major role in researches involving the formal methods of logic, for the simple reason that inductive definitions are, in a sense, the backbone of those methodologies themselves. To get a quick idea of how deep is this relationship, the one between logical researches and inductively defined concepts, take formal languages like the language $\mathscr{L}_{GM}$ we have been defining starting from section 3.2, in order to actually describe by means of its formulas what goes on in a game matrix. This language is based on inductive definitions, as all of its syntax is inductively defined. Moreover, so is the semantics we have associated with it via the relation $\models^{G}_{h}$.

In addition to what we have just noticed about the relation, which turns out to be a strong one, between logically driven accounts and inductive definitions of concepts, inductive definitions do also play a major mathematical role. Roughly speaking, this can be related to induction, in the form of *complete induction*, being possibly the most distinctive principle of the branch of number theory known as artithmetic. In fact, the principle of induction is a principle of proof, rather than a defining principle, but the fact that this makes proofs by induction so pervasive in areas related to the theory of natural numbers is enough to explain also why so are definitions by induction, since only collections of elements that are inductively defined are prone to be analyzed by arguments of that (inductive) sort.

Both remarks we have just made explain why inductive definitions have attracted the attention of scholars dealing with issues in the foundations of mathematics. A summary of the work done by specialists up to the 1970's, is the monograph edited by W. Buchholz and others published at the beginning of the 1980's (Buchholz et al. 1981). A concise update of the work done in the area, which is more accessible to those lacking the expertise required to go through research papers on the subject, is a later article by S. Feferman (Feferman 2010) who was among the main contributors to the subject.

Anyway, iterated inductive definitions have little to do with what we have been dealing with, except for the fact that they provided scholars with a ground for the applications of results we have mentioned, as it was said, the most important of which is theorem 3.1, which states the existence of fixpoints for monotone operators. That theorem is usually named after two Polish mathematicians, Bronisław Knaster and Alfred Tarski, who are responsible for having first presented a proof of the result (Knaster and Tarski 1927). The short note is contained in the part of the *Annals of the Polish mathematical society* devoted to scientific reports coming from the Varsaw section of the society. Constructions involving the use of operators in iterative form eventually reaching fixpoints, have become very popular in the area which studies the concept of truth for formalized languages, particularly after the proposal made in the mid 1970's by philosopher Saul Kripke (Kripke 1975) for dealing with it. This also is a field of studies which features a seminal contribution by Alfred Tarski (Tarski 1935), which was first published in Polish in 1933 and then subsequently translated in English, although this is not related to the previously mentioned joint work with Knaster.

Formal theories of truth originate in turn in the study of paradoxes, which has a even longer tradition and appears nowadays as a complex area of research (Cantini and Bruni 2017, being a comprehensive source to get an idea of how back and how deep goes the topic in the history of contemporary logic, and how ramified looks today the network of its connections to other topics). Paradoxes, as well as formal truth, are connected with circularity, and this is where the approach to the concept of «rational action» in finite games we have proposed here originally comes from. As a matter of fact, the machinery of making hypothesis and revising them for the sake of escaping loops due to a concept being defined circularly, was considered as a solution to the problems affecting the concept of truth for formalized languages in a series of papers from the early 1980's by Hans Herzberger (Herzberger 1982a, 1982b), Anil Gupta (Gupta 1982) and Nuel Belnap (Belnap 1982). Gupta and Belnap later specified the methodology as a way to deal with circular definitions in general (Gupta and Belnap 1993). Then, in later contributions (Gupta 2000; Chapuis 2000, 2003), Gupta again and André Chapuis have argued that the concept of «rational choice» turns out to be circular if the intuition which is commonly assumed to underline choices by rational agents in finite games is formalized. Furthermore, it was noticed by Gupta that fixpoints of the operator that comes out by applying the revision-theoretic machinery in such setting coincide with Nash equilibria. This approach to the topic, is also the one I have decided to stick to here for the sake of presenting the reader with an introduction to this fundamental concept. It should be stressed that Gupta's contribution was confined to the class of strict games in the sense of definition 3.2 from section 3.2 above.

As it might be known to the acquainted reader already, despite the crucial role the concept has played for the history of game theory, counterexamples of various sorts have been proposed to argue that the concept of «Nash equilibrium» is not robust enough both if regarded as a description of, and if taken as a norm for how rational agents act in concrete situations. An equally diverse refinements of the concept have been proposed to make up with the criticisms that were raised. One of them, which has also been mentioned here, is the trembling-hand intuition based on the idea that optimal, hence rational solutions to agents in a game are those which best protect them against trembles of the opponents. This idea was first proposed in the literature by another Nobel laureate, Reinhardt Selten (Selten 1975).

Another idea that we have referred to here is the concept of «proper equilibrium», due to Roger H. Myerson (Myerson 1978), that we have used to show how to deal with situations where even the trembling-hand intuition is helpless. This way of presenting refinements of the concept of equilibrium originally due to Nash as ways of always enlarging the scope of the revision-theoretic machinery for isolating solutions in wider and wider classes of finite games, is inspired to the work that has recently been done by myself together with G. Sillari (Bruni and Sillari 2017). This was conceived as an attempt to show how the methodology used by Gupta to deal with the concept of «rational strategy» in a narrow set of situations can be widened, so to potentially apply to all possibile finite games.

One final note for those who, having read this introduction to classical concepts of game theory, would like to know more about this area of reasearch, or would like to view a more orthodox approach for the sake of comparison. The choice here is vast, hence we confine to a couple of suggestions: the manual by M.J. Osborne and A. Rubinstein (Osborne and Rubinstein 1994), and the more recent book by A.R. Karlin and Y. Peres (Karlin and Peres 2016). For the reader in search of a more approachable source, I suggest to take a look at the stimulating and comprehensive introduction to game theory that can be found in the entry on «Game theory» of the *Stanford Encyclopedia of Philosophy* by D. Ross (Ross 2014).

Chapter 4
Sequential play: games in extensive form

In the 1983 movie *War games*, a super-computer named Joshua, upon the operation of which homeland security of the United States depends, is led by teenager David to produce the illusion of a nuclear attack by Russia which is taken for real by everyone, and is thereby about to determine a real nuclear conflict. This is then avoided when, in a excited ending, David manages to stop the computer operations by entrapping it into an endless tic-tac-toe series of matches.

Tic-tac-toe is a game everyone knows. Two players, one is given cross marks, the other one circle marks. They are confronted with a grid where nine cells are aligned in three rows and three columns, and they fill them with their marks alternatively with the only constraint that no two marks can be put in one and the same cell. The circle player starts, the cross one will follow. The goal is to try to get three equal symbols in line, either vertically, or horizontally, or diagonally. Should this happen, the player who owns the marks wins. That is everything you have to know to play this game, in addition to the moves made by your opponent. But this is no problem, since no parts of the grid is ever hidden and you will always be able to determine the current state of the game by looking at it.

Tic-tac-toe is, as specialists would say, a game that comes with «perfect information» since nothing required to players to decide their strategies is kept secret to the players' eyes. This is not even a peculiar feature of this game, since tic-tac-toe is in quite good company with other perfect information games like draughts, chess, go and the like. The peculiar feature, quite peculiar indeed if you consider that this is a game where two players battle with each other to win, is that to win a match of tic-tac-toe you need your opponent to make a mistake. If this does not happen, then none is going to win for sure. That is the whole idea behind the use of that game for the sake of the movie's dramatic effect: if a game

Riccardo Bruni, *Course notes on finite games and rational choice*, ISBN 978-88-6453-832-7 (print)
ISBN 978-88-6453-831-0 (online), CC BY 4.0, 2018 Firenze University Press

like tic-tac-toes is played by perfect players as computers can be, none wins. So, why should anyone want to play it? This is Joshua's own conclusion: why anyone would ever be interested in 'playing' nuclear war? Why on earth playing a game whose only winning move is not to play it? We should rather play some serious game: «How about a nice game of chess?», Joshua asks in the end relieving the tension on the terrified onlookers' faces. Now chess is, as I said, a perfect information game like tic-tac-toe. We are all happy that Joshua thought chess was something more serious to turn the attention to, but are we really sure it is serious enough to play?

4.1. One player after the other

One other thing I said about tic-tac-toe in the previous paragraphs, is a feature that it shares with the other games that I have mentioned along with it. Tic-tac-toe, as I said, is a game where players have to play in a certain order. It happens to be a game where players cannot play simoultaneaously, rather a game where one player has to wait until the other has performed her move first. That is also the way a match goes in the other games I have briefly referred to, that is chess, draughts, go.

Regarding the way games were handled in the previous chapter, it should be clear that this would not fit tic-tac-toe, nor it would fit any of the other games in the short list of games I provided the reader with. In a game where players are supposed to play one after the other, you cannot get all possible combinations of moves, arrange them in a matrix, assign to each player involved in the game the payoff she would get if a certain combination was played, and try to analyze the game in this way. If you did that, you would loose an important feature of how the game is played.

All the games that have been mentioned here are perfect information games. This means that players are given the possibility to glance at everything that might be important to choose what move they should be making next. Obviously, the information they may need in this respect includes the move, or the moves the opponents have made up to the point it is their turn to play. It should be clear that, by sticking to the matrix representation of a match in a game like those, that distinctive feature of theirs would be lost. The alternation in the players' move is what we referred to as the «game dynamic» back in section 2.2, where we also considered trees as the best way to represent a game that features such dynamic. I will briefly recall how this can be done here, for the reader's sake.

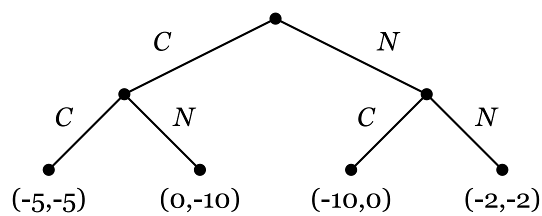
A tree of that sort will always feature a «root» from where it starts. The root represents the initial situation when player 1, the player who is supposed to play first, has not played yet. The root, as well as the subsequent stages in a match, are represented by nodes, each following node representing one among many of the possible outcomes coming out of the previous stage of the match. In addition to nodes, a tree like that will feature «branches» also, in the form of paths going through nodes, and each branch would keep track of the moves that are played as the match goes on. As a matter of fact, there would be edges coming out of the nodes of the tree connecting them to subsequent ones, and each edge would represent a move that a player has at her disposal at the stage of the game represented by the node this edge is coming out from.

For instance, if the match to be represented was a tic-tac-toe match, then the tree representing it would feature a root like all the others. Attached to that node there would be as many branches as the moves player 1 can choose among. This tree would then feature nine branches starting out from it, one branch for each of the cell player 1 can choose to put her mark in. Since the game has not started yet, all of the nine cells are empty and player 1 can freely choose the one she prefers. Each of these branches would then lead to the subsequent set of nodes of the tree. Any of these subsequent nodes would represent the stage in the match where player 2 is supposed to reply to the move made by player 1. As before, each node in this set would have attached as many edges as the moves that player 2 this time can come across. As a matter of fact, player 2 can choose to put her mark in any of the eight empty cells remaining, the one that has been chosen by player 1 beforehand being the only one excluded. Each of these branches would connect the node we are after with the subsequent one, that represents the subsequent stage of the match. By going through all the possibilities one gets a complete picture of the match, actually of the many matches that come out of a match if you collect all the possibilities that are given to the players playing it. Since all of them have been considered, the actual match will certainly coincide with one of the branches of the tree. So, the picture we have of a game by means of the tree we have built is complete, but the question is: which branch will contain the choices made by the actual players of the game?

At the bottom of the tree representation of a match you will find what we should call 'leaves'. Each leave of a tree represents the outcome of the match. Leaves are labelled with numbers, each number representing the payoff each player is getting in case the match as it is played ended up there. Players are supposed to act in a way which is not different from the one we have assumed players have played so far, namely they will try to get the best outcome they can for their part. Can we make any prediction? Is there a way, by looking at the representation of the match, to get an idea of how a match in a game where players have to wait until the opponent has played to make their move will ever end? Something we said already back in section 2.4 comes to our mind as giving information related to answering this question. To make sure our recall is correct, shall we go through it again?

4.2. Solving trees, backwards

Let us suppose for a moment that we are given a tree representing a match in a two-player game with perfect information, that is built out of the instructions from section 2.2 that we have briefly recapitulated above. Suppose we are required to evaluate its leaves only. Suppose also that the tree in question is the one that we actually built in that case to evaluate the dilemma of the two prisoners dynamically, which corresponds to the following diagram:



First of all, take notice of the fact that we are about to illustrate the point we want to make in a situation that is way too simpler than those we would have to analyze, if the trees to be considered were those coming out from any of the games we have mentioned so far in this chapter. The reason for that is merely practical, for the combinatorics which is involved into those games would make the corresponding trees impossible to handle. If you take the simpler game of them, namely tic-tac-toe, well, it turns out that the number of leaves the corresponding tree contained would be equal to

$$9! = 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2$$

(because it comes from multiplying the nine edges coming out from the root, which correspond in number to the nodes that follow it, by the number of edges coming out from the latter nodes, which is 8, and multiply the result of this by the number of edges coming out from the subsequent nodes, which is 7, and so on until all set of nodes preceeding the leaves have been considered). The result is the astonishing number of 362.880 final nodes to draw on the page! This alone would make the tree impossible to figure out, even without attempting to sketch the rest of it. The fact that we are then forced to consider a different situation, at least as long as the 'size' of it is concerned, is no harm, however, since what we are about to say is independent of the number of nodes and the size of the tree we are required to analyze.

So, let us go back to the initial question and suppose that we are required to say, of each leave in the tree above, whether it represents an outcome which is favourable to player 1, that is a final situation of the match in which player 1 wins, or if it is favourable to player 2 instead. Leaves are labelled by numbers which reflect the score each player is going to get if the outcome a leave represents turns out to be the actual outcome of the match that is being played. So, the task is an easy one. For, a leave will be regarded as a winning outcome for the player who is about to earn more from that. Suppose also that the game in question is a game where ties are not allowed. For cases like the one that originated the tree above, where tie is present, this can be achieved by assigning the match to one of the two player as is done in chess, where tie, «stalemate» as it is called, is actually regarded as a situation in which black has won. The idea behind this is compensation: the match is assigned to player 2 since player 1 has moved first, and moving first is indeed as a big advantage. So, if this rule is adopted, every leave in the tree can be evaluated as either representing an outcome of the match in which player 1 wins, or an outcome in which player 2 wins.

Suppose that we use colors for that, and we paint a node white in case it represents a favourable outcome to player 1, and we use black to do the same with leaves in the tree representing winning outcomes of player 2. Then, once we have done this with the tree depicted above we obtain the following situation:

Now, let us suppose that, having collected this information about terminal nodes of the tree, we would like to analyze the set of nodes that comes right before, i.e., right above them. In our running example, this set contains two nodes. Those are outcomes determined by the moves that player 1 has at her disposal at the preceeding stage, and represent situations in which player 2 is about to move. Elements of this latter set of nodes are connected with terminal nodes representing outcomes they lead to through edges which, in turn, represent moves that player 2 can choose. In the tree as it looks now once we have started to paint it, edges coming out from nodes preceeding the leaves would connect them with either black or white nodes. Since player 2 is moving at the nodes we are now considering, and since black outcomes represent situations in which she wins, we can now tell what are likely to be the moves she will prefer to make at that stage of the match, namely those represented in the tree by edges that lead to black leaves.

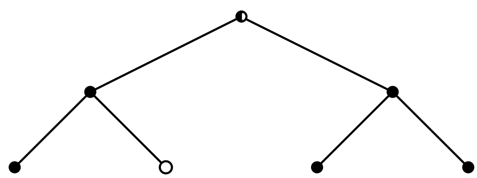
Similar considerations would apply if these were nodes representing situations in which player 1 was supposed to play. Even in this case it is hard to see why player 1 would make a move that leads to an outcome that is unfavourable to her, being it favourable to her opponent, unless she is obliged to. If players are supposed to try winning the match, they will try to reach outcomes in which they actually win. Of course, it is possible that no such edges exist and there is no way for players to avoid making decisions that will favour their opponent. In particular, it is possible that at some point in this process of analyzing set of nodes in the tree, we are facing a situation in which nodes representing a stage in the match where a player is supposed to play are only connected with nodes that are painted with the opponent's color. In order to keep this process of painting nodes in the tree going, we set up the following rule: if the nodes we are considering represent situations where player 1 is supposed play, then we paint them white just in case they are connected with at least one white node and we paint them black otherwise; if the nodes we are considering represent situations where player 2 is supposed play, then we paint them black just in case they are connected with at least one black node and we paint them white otherwise.

Let us figure out how the rule goes in the example we are using. Let us first refer to the portion of the diagram that interests us at the moment, by painting the nodes whose 'nature' we are trying to determine half white and half black until our evaluation of them by the rule we have just set is finished:

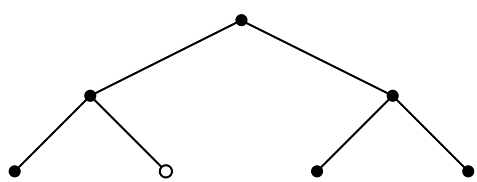
So, there are two nodes to consider. They both have two edges each coming out from them. As far as the left-hand node is concerned, the two edges connect it with the leftmost pair of leaves in the tree, one of which is painted black, while the other one is painted white. The right-hand node under scrutiny equally has two edges coming out which connect it with two black leaves. Let us consider the left-hand node first. According to the rule we have just set up, this node should be painted full black since it represents a situation in which player 2 has to play, and there is at least one edge that connects it with a black leave, that is a black subsequent node. This observation would prompt us changing the diagram as follows:

Now, let us consider the right-hand node. This also has connection with at least one black subsequent node. In fact, both nodes with which it is connected are black, hence it should be black itself, which finally brings us to the following improvement of the previous diagram:

Having painted the diagram up to this point allows us to look at the subsequent set of nodes. This is made out of just one node and, once we adopt with it the previous habit of painting uncertain nodes half white and half black until the application of the rule clarifies which color we should use for it, the situation we have to scrutinize looks as follows:



Since we are one step higher in the three, the node we are looking at corresponds to a stage in the match where player 1 is supposed to play. This node is connected with the two nodes we evaluated previously, and that were both painted black. Since there is no other node the root of the tree is connected with, the rule we have set applies and we paint it black, being it impossible for us to paint it white. We thus finally obtain the following diagram:



Now, the whole idea behind painting the diagram's nodes was to have a quick way of realizing whether, at any stage of the match, the player who is about to move has a preferred choice. This happens if one of the subsequent nodes is painted with the player's own color. If this is not the case, the node itself is painted with the opponent's color so that we know, when this happens, that the player who is in charge of making the next move will not be able to reach a node that will possibly lead her to win the match.

The process that we started with the leaves of the tree brought us to paint black the tree's own root. This is the stage where the match starts and it should be clear, in view of what we have just said, that this is painted with the color of whoever is in a position to win the game. By the way, on the basis of the premises we have been careful to set, only one player can be in that position. So, after the process of painting nodes in a tree has ended, you will know the only player who can ever win the game by looking at the color of the root.

It should be noticed also that the rule as we have set it can never be broken. For, either a node is connected by edges with nodes all painted with one and the same color, or at least one of them must be painted with the contrary color. So, whatever tree we are applying the process to, and whatever is the size of this tree, we would be able to carry it through. Then, one other thing should be clear about our painting game trees at this point: that this process could serve the purpose of allowing us to always determine the only player who is in a position to choose a path through the nodes and reach an outcome that is winning to her. No matter how complicated the game you are referring to is, if it is subject to the tree representation, then it is also subject to the painting procedure. And if it subject to the painting process, then it is possible to anticipate who is going to win.

This conclusion appears to be so strict and convincing, that it could take the form of a mathematical theorem. Yet, the conclusion cannot leave us indifferent. For, since all the games we have been discussing so far are indeed subject to the tree representation, what does not this conclusion tell us about them? Does not it allow us to conclude, in particular, that the difference between a simple game like tic-tac-toe and a complicated one like chess is only the complication itself? For, as long as the problem of trying to determine who is going to win, there seems to be now a very tiny difference (assuming we would like to consider the difference in combinatorics, which is huge, a tiny aspect). This remark is in fact so surprising that we cannot remain silent to it and forces us to do something more to test it. Maybe there is something wrong with the previous account. There has to be a mistake somewhere.

We now feel compelled to go through this matter more seriously, leave this conclusion aside for the moment along with white and black paint, and lay down some serious tools to deal with it with greater precision and be able to get back to it again.

4.3. An abstract, mathematical approach

Following the previous preliminary analysis of game with perfect information as trees, and the surprising conclusion we have reached, we shall now begin a systematic investigation of this subject. The first step of ours may indeed turn out to be unexpected to the reader as we plan to change the object of our interest, or at least to change the way it appears. As a matter of fact, we shall get rid of trees in the following and substitute the model of games where players play one after the other by a more abstract, mathematical one.

This may sound puzzling at first as I said, for trees are mathematical objects already as the acquainted reader may know, so to say that we are going to build a mathematical model for games in extensive form, as this type of games was referred to back in section 2.1, does not explain too much since a model of that kind we had already. What I meant to say is that we are going to prefer another mathematical model to the one we have been sticking to so far. This is mostly due to the fact that trees are hard to handle. They have indeed the advantage of being representable diagrammatically, which allows one to put forth live examples that usually make the explanation easier. However, this advantage has a cost since tree diagrams can easily become impossible to frame on the page space. This was already hinted at in the discussion we went through in the previous section. In addition to that, if a systematic investigation has to be pursued, then diagrams of trees would have to be abandoned anyway in favour of the mathematical definition of trees, and trees as mathematical objects are not as easy to handle as their representations lead to think. So, we are rather going to choose a model which is definitely more abstract than trees, but which is simpler, mathematically speaking, at the same time. Since to change the model means to change the basic notions, the language, and the notation too, the first thing before going through the said investigation is to introduce the features of the new mathematical setting to let the reader be able to familiarize with it.

4.4. Preliminary notions and basic notation

The main mathematical structure we will be interested in, is the one which is based upon the set $\mathbb{N}$ of natural numbers:

$$\mathbb{N} = \{0, 1, 2, \ldots\}$$

Having spoken of ordered sets already, natural numbers will be referred to as a totally ordered set. In particular, elements of the set $\mathbb{N}$ will be taken to be arranged in their 'standard' order. The mathematical notation commonly at use for that is the Greek letter ω, which denotes the least infinite ordinal number that corresponds to the ordered set $(\mathbb{N},<)$. The difference between using $\mathbb{N}$ and ω should be clear, and, since this is also a book about games, it is the same that occurs when one is referring to a group of cards and to the same cards arranged in such a way to form a straight: the first is a simple set, while the second is what one obtains when the elements of that very same sets are arranged in their order of size. That said, however, I will take a relaxed approach to the topic and consent myself to often treat ω and $\mathbb{N}$ as equivalent notations (in particular, by letting the former symbol refer to the latter set).

We will denote by ω^n the n-Cartesian product of ω, where n is a natural number itself. This means that ω^n denotes the result of repeating the application of the Cartesian product to ω n-times to produce the set

$$\underbrace{\omega \times \omega \times \ldots \times \omega}_{n-times}$$

which is made out of n-tuples, i.e. ordered tuples $(k_0, \ldots, k_{n-1})$ with n elements from ω. For every $s \in \omega^n$ we assume to indicate with $(s)_i$ for $0 \leq i \leq (n-1)$, the i-th element of s. This amounts to assume that s comes in the form:

$$((s)_0, \ldots, (s)_{n-1})$$

This also coincides with identifying finite sequences of this sort (i.e., elements of sets ω^n for $n \in \mathbb{N}$) with functions $f_s : (0, 1, \ldots, n-1) \to \omega$, where, for every i such that $0 \leq i \leq (n-1)$, $f_s(i)$ denotes the i-th element of the sequence s, i.e. $f_s(i) = (s)_i$.

Functions regarded extensionally (see section 1.2) are identified with their *graph*, which is the set of pairs coupling each argument in the function's domain with the value of it it produces. For functions of the sort we are considering here like f_s, this means that they are extensionally coincident with the set $G_{f_s} \subseteq (0, 1, \ldots, n-1) \times \omega$ of pairs of elements from $(0, 1, \ldots, n-1)$ and from ω (for example, pairs of the form $(i, f_s(i))$ where $0 \leq i \leq n-1$), satisfying the right-hand uniqueness conditions (i.e., such that $b = b'$ for every $(a, b), (a, b') \in G_{f_s}$).

A note on the empty sequence as the empty function. Owing to what we have just said, the empty sequence ( ) would be identified with a function $f_\emptyset : \emptyset \to \omega$, the graph of which is a subset $G_{f_\emptyset}$ of $\emptyset \times \omega$, which is an element of ω^0. The latter, by the way, is the set of all pairs (x, y) such that $x \in \emptyset$ and $y \in \omega$, but since there are no elements in $\emptyset$, it turns out that $\emptyset \times \omega = \emptyset$. This also means that $G_{f_s} = \emptyset$ as $\emptyset \subseteq \emptyset \times \omega$ (in fact, $\emptyset$ is the only subset of $\emptyset \times \omega$). It follows that there is a *unique* function $f_\emptyset$ whose graph $G_{f_\emptyset}$ is the empty set. So, this $f_\emptyset$ represents no correspondence at all (and, in this latter sense, is suitable to represent the empty sequence). However, it *is* a function since the statement «for all x, if $x \in \emptyset$, then there exists a unique $y \in \omega$ such that $f_\emptyset(x) = y$» is vacuosly true by logic alone because, since there is no such x, the antecedent «$x \in \emptyset$» is true of all of them, which, in turn, makes true the conditional utterance as a whole.

Notationally, we also assume to indicate with $\omega^{<k}$ the set of all finite sequences of length less than k, which in standard set-theoretical notation is indicated as:

$$\omega^{<k} = \bigcup_{n<k} \omega^n$$

As an extension of the above convention, the set of *all* finite sequences will be denoted by $\omega^{<\omega}$. That is:

$$\omega^{<\omega} = \bigcup_{n\in\mathbb{N}} \omega^n$$

which contains all n-tuples of natural numbers of every finite length (i.e., $s \in \omega^{<\omega}$ if and only if there exists a natural number k such that $s = ((s)_0, \ldots, (s)_{k-1})$ which is the same as saying that $s \in \omega^k$ holds).

For every finite sequence s, we denote by $|s|$ its *length*: therefore, $|s| = k$ if and only if $s \in \omega^k$ holds (notice that, due to the conventions we have made, such k is unique, and $|(\,)| = 0$).

The repeated application of the Cartesian product can be extended to infinite ordinal numbers, and ω^ω would represent the least set which is obtained in this way. Despite the fact that grasping this idea is not trivial, since the very same concept of an infinite set can be a bit disturbing, things should be made simpler owing to what we have said so far. The whole idea of identifying finite sequences with functions can be explained in the following way. To come up with a finite sequence $s \in \omega^n$, i.e. with an arrangement of n elements in an order, is the same as labelling them with natural numbers in such a way that: (i) each element gets a unique labels (hence no two different elements get one and the same label), and (ii) elements are arranged in the order of their label. To denote the i-th element $(s)_i$ of the sequence s by $f_s(i)$, is a way to take notice of the fact that the element in question stems from the i-th application of the labelling process (here identified with the function f_s). Condition (i) implies that a finite number of labels will allow to produce only the ordering of a finite number of elements, i.e. a finite sequence. Infinite sequences of elements is what you get by extending this idea to a situation where we assume that the number of labels can be infinite and coincide with the total set ω. This means that a sequence s is infinite if it is the result of arranging a given infinite amount of elements in an order. Or, to extend to this situation the previous notation, if s is obtained as the result of applying the labelling function f_s infinitely many times. This is for instance what happens if the labelling function is any function f whose domain is the set ω itself (which corresponds to the simplest way to extend the previous schema, in which labels came out of finite subsets – finite subsequences, one should better say –, of ω). This is precisely what one would assume by denoting the set of all *infinite sequence* by ω^ω. An element s of this set is a sequence of natural numbers of infinite length, i.e.:

$$s = ((s)_0, (s)_1, \ldots, (s)_n, \ldots)$$

(where we are exploiting the same notation convention as before, and $(s)_i$ denotes the i-th element of s), or, equivalently, s is the result of the application of the function

$$f_s : \omega \to \omega$$

which takes as input the whole set ω of natural numbers (and therefore $f_s(i) = (s)_i$ for every such i).

Now, it should not come as a surprise that one can have different infinite sequences out of the elements coming from one and the same infinite set ω of natural numbers, as the reader might be acquainted of this fact already (the sequence s of even natural numbers, and the sequence s' of odd natural numbers providing us with a simple example for illustrating that). That goes back to the feature which can be taken to be the most prominent characteristic of an infinite set, that is the property of containing as subsets different sets with the same (infinite) number of elements. As long as distinct *sequences* of natural numbers are considered, i.e. different order arrangements of infinitely-many elements of ω, this feature is reflected by the fact that the size of ω^ω is the same as the size of the set of all functions of type $f : \omega \to \omega$, and the latter set can be proved by the well-known «diagonalization argument» conceived by the mathematician Georg Cantor to be 'bigger' in size than the former set ω itself. This, however, is just a digression since here we shall be concerned only with finite sequences. So, let us get back to them.

Let now $s, s' \in \omega^{<\omega}$ be finite sequences of numbers. The main comparison relation between objects of this sort will be denoted by:

$$s \sqsubseteq s'$$

to means that s is a subsequence of s', which takes place when: (i) $|s| \leq |s'|$ (i.e., the length of s is not greater than the length of s'), and (ii) $(s)_i = (s')_i$ for every $0 \leq i < |s|$ (elements of s are equal to the corresponding elements of s', namely to the element which occupies the same place in the order). Requirements (i) and (ii) makes clear that when $s \sqsubseteq s'$ is the case, then s is the initial part of s' or, as it should be said, the *initial segment* of s'.

Two basic facts should be obvious about the subsequence relation just defined, namely:

$$(\,) \sqsubseteq s$$
$$s \sqsubseteq s$$

for every finite sequence s. That is, the empty sequence is the universal subsequence owing to the fact that it contains no element (hence, its elements, being nowhere, are everywhere, so to say), and every finite sequence is the subsequence of itself since conditions (i)-(ii) above are easily seen to apply if $s' = s$.

With respect to finite sequences such as s and s', the main operation we will be dealing with is *composition*. This will be indicated by:

$$s \bowtie s'$$

We will make use of this notation to refer to the sequence that is obtained by prefixing all elements of s to all elements of s', that is:

$$s \bowtie s' = ((s)_0, \ldots, (s)_{|s|-1}, (s')_0, \ldots, (s)_{|s'|-1})$$

(where clearly $|s \bowtie s'| = |s| + |s'|$ then, and $|s|, |s'| \leq |s \bowtie s'|$ as a consequence)[1].

Composition of finite sequences s, s' is inductively defined over $i \in \mathbb{N}$ such that $0 \leq i < |s'|$, by:

$$(s \bowtie s')_i = \begin{cases} (s)_i, \text{ if } i < |s| \\ (s')_{i-|s|}, \text{ if } i \geq |s| \end{cases}$$

Notice that $s \sqsubseteq s \bowtie s'$, but this does not hold for s' in general. In particular, $s' \sqsubseteq s \bowtie s'$ holds if either $s = (\,)$, or $s' = (\,)$, or $s = s'$ are the cases. In addition, one has that the following cases also hold:

- $(s \bowtie s) \neq s$, if $s \neq (\,)$;
- $(s \bowtie s') \neq s' \bowtie s$, if $s \neq (\,)$, $s' \neq (\,)$, $s \neq s'$;

(as a matter of fact, $s \bowtie (\,) = (\,) \bowtie s = s$). Also, notice that one can more easily express the previous observation of s being the initial segment of s' in case $s \sqsubseteq s'$ holds, owing to the fact that $s \sqsubseteq s'$ is the case if and only if $s' = s \bowtie s''$ for some $s'' \in \omega^{<|s'|}$ (the proof is an easy exercise).

Both the subsequence relation and the composition operation can be extended to infinite sequences. In particular, if $s \in \omega^{<\omega}$ and $s' \in \omega^{\omega}$, then $s \sqsubseteq s'$ holds just in case $(s)_i = (s')_i$ for every $0 \leq i \leq |s'| - 1$ (since $|s| < |s'|$ then), while, at the same conditions, $s \bowtie s'$ denotes the infinite sequence:

$$((s)_0, \ldots, (s)_{|s|-1}, (s')_0, (s')_1, \ldots)$$

(where again $s \sqsubseteq s'$ is obviously the case).

[1]As a consequence of using natural numbers as labels for elements of a finite sequence, i.e., of counting the elements of a finite sequence s starting from the first natural number, which is 0, we have that $|s| = k$ if and only if s has k-many elements, which means that, if its first element is $(s)_0$, then its last one is $(s)_{k-1}$, the k-th element of the sequence, namely $(s)_{|s|-1}$.

4.5. Two-person, zero-sum, perfect information games

As declared at the very beginning of this chapter, the goal is to deal with two-player, perfect information games.

The two players playing the game will be denoted as Player I and Player II to differenciate here the notation we have used for the sake of the analysis of games in normal form. Since the main feature of games we would like to consider is the fact that players are supposed to play in a certain order, we shall assume that Player I plays first, and Player II responds. One additional feature of the kind of games we are going to model is that tie is not possible. So, either Player I wins, in which case Player II loses, or Player II wins instead and Player I loses. In game-theoretic terms games as such are called *zero-sum*. Finally, in games we are considering Player I and Player II knows everything and both knows that. In particular, there is no asymmetrical distribution of knowledge among the players of our game.

The consequences of our assumptions in terms of what we are actually modeling, and what we are leaving out of consideration, are summarized in the following table:

$\checkmark$	$\times$
two players alternate moves perfect information	multiple players simultaneous moves imperfect, or asymmetrical information

Having made clear that, we are now going to introduce the mathematical model in a stepwise manner, and to provide the reader with all of the information required to reconcile the model with the original intention.

4.6. The model of finite games

The mathematical structure we take in order to represent finite games comes in the form of a triple. A finite game G is as follows:

$$G = \langle k, P, A \rangle$$

where:

- $k \in \mathbb{N}$ is the *number of moves* in matches of G;
- $P = \omega^{2k}$ is the set of *matches* of G;
- $A \subseteq P$ is the set of *winning conditions* for Player I in G.

In particular, G is the model of a game played as follows: Player I plays first, and Player II plays second at each turn; Player I plays by picking a number out of $\mathbb{N}$, and Player II responds by picking out a natural number too; after exactly $2k$ rounds, hence after both players have picked k natural numbers each, the match is over and Player I wins in case the sequence $s \in \omega^{2k}$ of numbers that contains all of the players' choices belongs to A, and looses (that is, Player II wins) if that is not the case. In the following it should be clear that by G we will always refer to a model of a game of this sort, hence, to a game in which players play according to the previous description.

As it was said at the end of section 3.11, the main concept we would like to incorporate in our model is the concept of *strategy*. This should be done in a way that reflects the intuition according to which a player in a game has a strategy if she has a plan that allows her to decide what to do at each stage of a match. In order to model a situation in which all moves that a player chooses are part of a unique plan (hence, that there is a particular strategy that a player follows), it was suggested that we could make use of the concept of function which acts on the information required to make a choice. In a perfect information game, all the required information to decide what to do next is displayed as the match goes along. Here we assume that the information over which strategies of players 'act' (i.e., the information that is used for the sake of deciding the next move), is the portion of the match that has actually been played.

This idea is comprised in the following definition:

Definition 4.1 *Let $G = \langle k, P, A \rangle$ be a finite game as before. Let $E^{<2k} = \{s \in \omega^{<2k} : |s| = 2m, m \in \mathbb{N}\}$ indicate the set of subsequences of a match in G with even length, and let $O^{<2k} = \{s \in \omega^{<2k} : |s| = 2m+1, m \in \mathbb{N}\}$ be the set of subsequences of matches of G with odd length instead. Then:*

1. *a strategy for Player I in G is a function $\sigma : E^{<2k} \to \omega$;*
2. *a strategy for Player II in G is a function $\tau : O^{<2k} \to \omega$.*

The definition is a consequence of the conventions we have made. Since Player I plays first, the state of the game, namely the portion of the match that has been played before she makes the next move, is always represented by a sequence with a even number of elements. This is true for the sequence we find at first: none has played yet, so the portion of the match that Player I has to consider to make her choice is the empty sequence $(\,)$, which is such that $|(\,)| = 0$ and 0 is an even number (then, $(\,) \in E^{<2k}$). The same holds to every further stage of the game, since moves by Player I are always preceded by situations where Player II has also played. So, Player I will need the help of a strategy, in the form of the next natural number to play, for every given a piece of information that takes the form of a sequence comprising all numbers that have been played before the match ends, that is a sequence whose maximum length is equal to the greatest even number strictly less than $2k$ (which is even, but marks the end of the game and no strategy is required by any player at that stage). For the same reason, Player II only plays when a odd number of natural numbers have been played, and a strategy for her then takes the above form.

Now, it should be clear that, following the intuition above, to play according to a strategy means to have a plan which motivates a player's choice at each stage in a game, but it does not mean that the plan is winning! So, for the moment, that a player plays according to a strategy does not mean that she is going to win. Before attacking this particular situation, we have to adapt the previous conventions on the notation to the new concept of strategy. We do that by abusing the previous symbolism, in order to avoid an increasing complication of it:

Definition 4.2 *1. Let $s \in \omega^k$ be any sequence, and let σ be a strategy for Player I. Then, $\sigma \bowtie s$ is used to indicate the sequence where numbers occurring at odd places are produced by means of σ, while numbers occurring at even places come from s in its own order. That is:*

$$\sigma \bowtie s = (x_0, (s)_0, x_1, (s)_1, \ldots, x_{k-1}, (s)_{k-1})$$

where, for every $0 < n \leq k-2$:

$$\begin{cases} x_0 = \sigma(()) \\ x_{n+1} = \sigma((x_0, (s)_0, \ldots x_n, (s)_n)) \end{cases}$$

2. Let $s \in \omega^k$ be any sequence, and let τ be a strategy for Player II. Then, $s \bowtie \tau$ is used to indicate the sequence where numbers occurring at odd places are produced by means of s in its own order, while numbers occurring at even places are given by applications of τ instead. That is:

$$s \bowtie \tau = ((s)_0, y_0, (s)_1, y_1 \ldots, (s)_{k-1}, y_{k-1})$$

where, for every $0 < n \leq k-2$:

$$\begin{cases} y_0 = \tau((s)_0) \\ y_{n+1} = \tau(((s)_0, y_0 \ldots y_n, (s)_{n+1})) \end{cases}$$

This basic definition can be used to define, for a given game G, the set of matches of it where players play according to some given strategy:

Definition 4.3 *Let $G = \langle k, P, A \rangle$ be any finite game, σ be a strategy for Player I in G, and τ be a strategy for Player II. Then:*

$$P_\sigma = \{\sigma \bowtie s : s \in \omega^k\}$$
$$P_\tau = \{s \bowtie \tau : s \in \omega^k\}$$

denote the set of matches of G where Player I plays according to σ, Player II plays according to τ respectively[2].

[2] It should be clear that, having set the collection of matches of a game G to be equal to ω^{2k}, as a consequence of which every sequence s with length equal to $2k$ is a match of G, then every sequence with length equal to k can be taken to represent the sequence of moves played by one of the players of G in a match of it.

Finally, we come to the crucial notion of *winning strategy*, which will be used for the sake of analyzing those situations in which not only a player plays according to a plan, but this plan is a winning one, that is allows her to *always* win (i.e., to win independently of what the opponent decides to play):

Definition 4.4 *Let $G = \langle k, P, A \rangle$ be any finite game. Then:*

1. *a strategy σ for Player I in G is* winning *if, for every $s \in \omega^k$, $\sigma \bowtie s \in A$;*
2. *a strategy τ for Player II in G is* winning *if, for every $s \in \omega^k$, $s \bowtie \tau \notin A$;*

As a result of the definitions we have given so far, we can obtain a deductive confirmation that games as we have described them are indeed zero-sum as we wanted, in the sense of the next result:

Proposition 4.1 *Let $G = \langle k, P, A \rangle$ be any finite game. Then, not both Player I and Player II can have a winning strategy in G.*

Proof: let G be a finite game as in the statement of the lemma. Let, for any strategy σ for Player I and strategy τ for Player II

$$\sigma \bowtie \tau = (x_0, y_0, \ldots, x_{k-1}, y_{k-1})$$

be the sequence inductively defined as follows:

$$\begin{cases} x_0 = \sigma((\,)) \\ y_0 = \tau(x_0) \\ x_{i+1} = \sigma(x_0, y_0, \ldots, x_i, y_i) \\ y_{i+1} = \tau(x_0, y_0, \ldots, x_i, y_i, x_{i+1}) \end{cases}$$

(that is: $\sigma \bowtie \tau$ is the sequence obtained by applying σ to the previous elements to produce new ones occurring at odd places, and by applying τ to produce elements occurring at even places of it).

Then, to assume that both Player I and Player II have a winning strategy in G means that $\sigma \bowtie \tau \in A$ and $\sigma \bowtie \tau \notin A$ may hold for some strategies σ and τ, which is impossible. Hence, the lemma. QED

4.7. Alternative models of finite games

Before going on with the investigation on finite games as sequences, we plan to stop and discuss some issues that the definitions we have set so far may be puzzling the reader. This apparent *detour* will have the purpose to help the reader understand the model better and to reconcile it with some intuitions that are naturally connected with the idea of games due to our experience with them. It also has the goal of showing the model flexibility, therefore to justify why we left trees behind to stick to sequences instead.

With respect to things we were discussing preliminarly about games (see section 4.6 and chapter 1), the first thing that may have struck the reader in the definition of finite game we have given in the previous section, is the length of matches. It is part of the common experience with games that different matches of one and the same game may have different lengths, this feature being indeed related to many factors including the players' ability. In view of this, it appears rather odd to assume that every match of the same game should have always the same length as it follows by setting P to contain as elements *only* $2k$-tuples of numbers.

A second issue turns out by looking at the same set of matches, which, being P equal to ω^{2k}, is also supposed to contain *all* such sequences. This has the effect that the definition we have given does not allow to distinguish legal matches, which are those where players play only according to the rules, from illegal ones, which contain at least one move that is not allowed by the rules of the game. However, this is again part of the common experience with playing a game we have.

Now, it is easy to show that the assumptions we have made do not affect the scope of the model, and the variety of types of games we can capture by means of it.

4.7.1. Finite games with rules

Among those features that are apparently missing in the model we have proposed, we first want to try encapsulating the distinction between moves that are *legal*, i.e. made in full agreement with the rules of the game, and moves that are *illegal*, i.e. not allowed by these rules instead. Here is how the model from section 4.6 could be modified to do so.

Let then G^R («R» staying for «rules») be the following triple:

$$G^R = \langle k^R, P^R, A^R \rangle$$

where:

- $k^R \in \mathbb{N}$ is the number of moves of matches of G^R;
- $P^R \subseteq \omega^{2k^R}$ is the set of matches of G^R;
- $A^R \subseteq P^R$ is the set of winning conditions for Player I in G^R.

The reader should notice that the difference from the previous definition of finite games, is the fact that the set of matches now is only a portion, a subset, of the set of all sequences of length equal to $2k^R$. This means that it may happen that this set P^R is a proper subset of ω^{2k^R} (which is commonly denoted by $P^R \subset \omega^{2k^R}$) with the consequence that the difference set $C = \omega^{2K^R} \setminus P^R$ between the total set and this subset is not empty as P^R contains only some, but not all the elements of the latter set. If this was the case, one could take P^R as representing the set of legal matches as wanted, and take every $s \in \omega^{2k^R}$ which is not an element of it (hence, such that $s \in \omega^{2k^R} \setminus P^R$), to represent an illegal match of G^R instead.

To make the model correspond to the intended situation, the following extra-condition is assumed over the set P^R of matches. Let, for every $n \leq 2k^R$, P_n^R indicate a subset of ω^n. Intuitively, each P_n^R represents a 'layer' of P^R. To the latter set we want to assign the role of the set of legal matches. Hence, each P_n^R could be viewed as the set of legal portions at stage n of legal matches of G^R. To this purpose, we assume that the following conditions hold:

1. for every $n < 2k^R$, $P_{n+1}^R \subseteq \{s \in \omega^{n+1} : s' \sqsubseteq s$, for some $s' \in P_n^R\}$;
2. $P^R \subseteq \omega^{2k^R}$ such that, for every $s \in P^R$ and for every $n < 2k^R$

$$((s)_0, \ldots, (s)_{n-1}) \in P_n^R$$

   is the case.

Now, the first condition is thought of in order to ensure that every legal portion of a match has a legal continuation[3] (i.e., having Player I and Player II played fairly up to stage n, they can always play a next fair move if they want to). The second condition ensures that legal matches of G^R are only those sequences which are made out of legal portions and, as a consequence of it, we have that, for every $s \in \omega^{2k^R}$, $s \in P^R$ holds if and only if for every $s' \in \omega^{2k^R} \setminus P^R$ there exists $0 \leq i < 2k^R$ such that $(s)_i \neq (s')_i$. Hence, it is enough for a sequence to contain just one illegal move to be illegal as a whole (which also yields $s \in \omega^{2k^R} \setminus P^R$ if and only if for some $0 \leq i < 2k^R$, for every $s' \in P^R$, $(s)_i \neq (s')_i$, since illegal moves can be found nowhere in a legal match).

Owing to what we have just said, if we restrict the attention to moves made by each player, i.e., to elements of the set ω^{k^R}, we have that every such element s is made out of legal moves only if and only if the only elements it has in common with an illegal match, say s', are those that are legal, i.e., also belongs to some legal match s'''. Call s a *legal set of moves* of a player in that case. Remembering that moves by Player I occurr in a match at even places and moves by Player II occurr at odd places instead, we have the following:

Definition 4.5 *Let $s \in \omega^{k^R}$ be given. Then we say:*

[3]Notice that $(\,) \in P_0^R$ since $(\,)$ is the universal subsequence, i.e., every match starts as a legal match (which is quite obvious: you can cheat only when the match has started).

- s *is a* legal set of moves by Player I *just in case* $(s)_i = (s')_{2i}$ *for some* $s' \in P^R$ *also;*
- s *is a* legal set of moves by Player II *just in case* $(s)_i = (s')_{2i+1}$ *for some* $s' \in P^R$ *also;*

Having noticed that the main difference from the new version of the model, and the 'canonical' one amounts to P being possibily a proper subset of ω^{2k^R}, also means that we can obtain the main model as a limit case, the one in which no such distinction is needed since any move is legal and $P^R = \omega^{2k^R}$ as a consequence. This means that the new model contains the canonical one. We would like to see if the contrary also holds, as far as strategic choice in this more complicated model is concerned at least, and if there is a way to encompass the distinction between legal and illegal matches by means of the previous model.

First of all, we introduce by means of the following definition the idea of *legal segment* of a match that will turn out to be useful afterwards:

Definition 4.6 *Let G^R be a triple as above where $P^R \subset \omega^{2k^R}$ holds. Let $s \in \omega^{2k^R}$ be any sequence of numbers. Then, we say:*

1. $s' \sqsubseteq s$ *is a* legal *initial segment of* s*, if there exists* $s'' \in P^R$ *such that* $s' \sqsubseteq s''$ *holds;*
2. $s' \sqsubseteq s$ *is a* maximal legal *initial segment of a match, if* s' *is a legal initial segment of* s *but* $s' \bowtie (s)_{|s'|}$ *is not.*

A legal segment of a sequence would contain only moves of it that are legal. The maximal legal segment would count, and contain, all legal moves in a sequence starting from the beginning of it up to and excluding the first illegal one. Now, clearly, if $s \in P^R \subseteq \omega^{2k^R}$, then the maximal legal segment of it is s itself. A maximal legal initial segment of an illegal match, is a subsequence of legal moves of maximal length that can be extracted from the latter starting from its first element. Notice that any subsequence of a legal initial segment of a match in G^R is also legal.

The following has an easy proof:

Lemma 4.1 *If G^R is as in the previous definition, and $C = \omega^{2k^R} \setminus P^R \neq \emptyset$, then, for every $s \in C$, if s' is a maximal legal initial segment of s, then s' is unique.*

Proof: Suppose $s \in C$ has two maximal legal initial segments s' and s''. Then, either $|s'| < |s''|$ and $s' \sqsubseteq s''$ is the case, which means that s' is not maximal, or $|s''| < |s'|$ and $s'' \sqsubseteq s'$ hold, which means that s'' is not maximal, or, finally, it happens $s' \not\sqsubseteq s''$ and $s'' \not\sqsubseteq s'$, which means that either $s' \not\sqsubseteq s$ or $s'' \not\sqsubseteq s$ is the case. Since no other case is possible and all the previous ones contradict the hypothesis, it follows that the maximal legal initial segment of a sequence must be unique as wanted. QED

As a consequence of this lemma, let, for every $s \in C$, $ml(s)$ indicate the maximal legal initial segment of it.

The maximal legal initial segment of a sequence can be used for the sake of determining who among the two players of a game G^R was the last to play legally, or, equivalently, who was the first to play an illegal move. Owing to the global assumption about how matches of G^R are played, with Player I playing first and Player II playing second, it should be clear that the following holds for every $s \in C$:

- if $|ml(s)| = 2m + 1$ for some $m \in \mathbb{N}$, then Player II was the first to play illegally;
- if $|ml(s)| = 2m$ for some $m \in \mathbb{N}$, then Player I was the first to play illegally.

Since, for a given $s \in C$, $ml(s)$ contains only legal moves of it, if its length is odd the last move by Player I it contains was a legal one and the move that Player II played in response to it was the first illegal move of the match (because if you couple each move with the one that has been played as a reply to it, you will be left with a single action, the last one of $ml(s)$, which must be a move made by Player I). Symmetrically, if the length of $ml(s)$ is even then the opposite must be true.

The next item on the agenda is to equip the new model of games with a suitable notion of strategy, for which, by the way, we stick to the one we have used for the canonical model of finite games:

Definition 4.7 *Let $G^R = \langle k^R, P^R, A^R \rangle$ be a finite game as above. Let $E^{<2k^R} = \{s \in \omega^{<2k^R} : |s| = 2m, m \in \mathbb{N}\}$ indicate the set of subsequences of sequences of length equal to $2k^R$ whose length is even, and let $O^{<2k^R} = \{s \in \omega^{<2k^R} : |s| = 2m+1, m \in \mathbb{N}\}$ be the set of subsequences with odd length instead. Then:*

1. *a strategy for Player I in G^R is a function $\sigma : E^{<2k^R} \to \omega$;*
2. *a strategy for Player II in G^R is a function $\tau : O^{<2k^R} \to \omega$.*

Notice that by defining strategies in this way, we admit them to be defined everywhere, even on sequences representing portions of illegal matches (and for very good reasons: one can also cheat strategically!).

Now, before proceeding further with the notion of winning strategy for the new sort of games, let us first dig a little bit more into this idea of legal vs. illegal moves to see if we can get some helpful insights in this respect. Now, suppose that Player I and Player II actually come up with one of these illegal sequence of numbers. This must have happened because, having played legally up to some point, either I or II suddenly plays a move that is not allowed by the rules of the game. So, an illegal match turns out because one of the two players is mistaken, and starts playing illegally, or if she deliberately decides to do so. This remark suggests two possible interpretations for elements of the set C of illegal matches: in the one view these are the result of mistakes made by one of the players that could, or should be amended; in the other view, these are the effect of deliberate choices and should rather be sanctioned. There are two possible attitudes as a consequence: a *permissive* attitude which allow players to make up for mistakes, and a *strict* attitude which blames every transgression of rules to those who are responsible for them. These two attitudes are reflected by the two ways in which we define the notion of «winning strategy» for a game such as G^R above.

The first definition is the effect of regarding illegal matches permissively, and let the player amend mistakes:

Definition 4.8 *Let $G^R = \langle k^R, P^R, A^R \rangle$ be a triple as above. Then:*
1. A strategy σ for Player I in G^R is winning *if and only if,*

$$\sigma \bowtie s \in A^R$$

for every $s \in \omega^{k^R}$ that is a legal set of moves by Player II.
2. A strategy τ for Player II in G' is winning *if and only if,*

$$s \bowtie \tau \in P^R \setminus A^R$$

for every $s \in \omega^k$ that is a legal set of moves by Player I.

Notice that owing to the definition just set, illegal matches are not counted as far as strategic winning is concerned, and winning strategies are considered only insofar as they do what is expected over legal matches (while their value over illegal matches is not taken into account).

The alternative definition of winning strategy, the one comprising the strict approach toward illegal moves we have spoken of above, would make use of the concept of maximal legal initial segment of a sequence to blame a player who plays illegally:

Definition 4.9 *Let $G^R = \langle k^R, P^R, A^R \rangle$ be a triple as above. Then:*
1. A strategy σ for Player I in G'' is winning *if and only if,*

$$\sigma \bowtie s \in A^R$$

for every $s \in \omega^{k^R}$ which is a legal set of moves by Player II, or

$$\sigma \bowtie s \in C$$

and $|ml(\sigma \bowtie s)| = 2m+1$ alternatively.
2. A strategy τ for Player II in G' is winning *if and only if,*

$$s \bowtie \tau \in P^R \setminus A^R$$

for every $s \in \omega^{k^R}$ *which is a legal set of moves by Player I, or*

$$s \bowtie \tau \in C$$

and $|ml(s \bowtie \tau)| = 2m$ *alternatively.*

So, for a strategy to be winning in this second sense of the expression it is not enough to meet the previous requirement as long as legal matches are concerned, it is also needed for it to be a strategy that never causes a player to be the first who played illegally. Notice that this definition is made possible by the extra requirement that was assumed over P^R, and which amounts to always give players the possibility of playing legally. This assumption makes both part 1 and part 2 of this definition non-vacuous.

We now turn to the problem of defining a game in canonical form that may correspond to a game where legal and illegal matches are kept distinct as long as strategic winning is concerned. We do that in such a way that the new game will make possibile to prove that such a result holds independently of which attitude is taken with respect to illegal moves (i.e., whether winning strategies for players in G^R are defined according to definition 4.8, or according to definition 4.9 instead). The idea is simple and amounts to extending to the set of matches as a whole the intuition behind the strict approach to illegal moves that was fostered for the sake of definition 4.9, namely to blame illegality on the one who is responsible for it. So, we will consider a match to be lost in the new game by whom was the first to play outside of the rules of G^R. This idea can be made precise for a game in canonical form as follows.

Let in the following $G^R = \langle k^R, P^R, A^R \rangle$ be an arbitrary but fixed triple of the above sort. Let $G^\star$ be similarly given in the form of a triple:

$$G^\star = \langle k^\star, P^\star, A^\star \rangle$$

where:

- $k^\star = k^R$ is the number of moves of matches of $G^\star$;
- $P^\star = \omega^{2k^\star}$ is the set of matches of $G^\star$;
- $A^\star \subseteq P^\star$ is the set of winning conditions for Player I in $G^\star$.

The set $A^\star$ is defined according to the following condition: $s \in A^\star$ if and only if $s \in A^R$, or $s \in C = \omega^{2k^R} \setminus P^R$ and $|ml(s)| = 2m + 1$.

So, $A^\star$ is the subset of $P^\star$ whose elements are either matches of G^R that are won by Player I, or illegal matches of G^R in which Player I was the last one to play legally. Strategies and winning strategies for game $G^\star$ are defined as usual (so, no modification except the obvious ones[4]

[4]By the «obvious» modifications we mean those which follow from substituting the elements of a generic game in canonical form G, with those which are part of the triple $G^\star$ that is here at stake.

are required with respect to the general definitions of these notions for games in canonical form given according to definition 4.1, and definition 4.4 above).

The correspondence we would like to achieve sounds as follows: a player has a winning strategy profile in G^R if and only if the same player has a winning strategy profile in $G^\star$. We would like to prove that with respect to both winning strategies being defined permissively like in definition 4.8, and winning strategy being approached strictly like in definition 4.9. Due to the arguments being symmetrical in the case of Player II, we illustrate this point by proving that this holds for Player I alone.

So, we first prove that the following holds:

Proposition 4.2 *Let G^R and $G^\star$ be as before, and let the notion of winning strategy for Player I and II in G^R be defined according to definition 4.8. Then, Player I has a winning strategy in G^R if and only if Player I has a winning strategy in $G^\star$.*

Proof: we first prove the direction from left to right of the statement, and then assume that σ is a (permissive) winning strategy for Player I in G^R. Let $\sigma^\star$ be the strategy for Player I in $G^\star$ defined out of σ by the following condition for every $s \in E^{<2k^\star}$:

$$\sigma^\star(s) = \begin{cases} \sigma(s), \text{ for every } s \in E^{<2k^\star} \text{ such that } s \not\sqsubseteq s' \text{ for every } s' \in C \\ \min\{z \in \mathbb{N} : z \neq (s')_{|s|}, \text{ for every } s' \in C\}, \text{ otherwise} \end{cases}$$

Notice that for every $s \in E^{<2k^\star}$ the set

$$\{z \in \mathbb{N} : z \neq (s')_{|s|}, \text{ for every } s' \in C\}$$

contains numbers which do not correspond to illegal moves in C at stage $|s|$ in a match of G^R. Therefore, they represent a next legal move by Player I in the match s played so far (s represents in fact a portion of a match of $G^\star$). The whole idea of this definition is as follows: strategy $\sigma^\star$ coincides with σ as long as no illegal move has been played. Otherwise, it provides Player I with a next legal move. This means that, for every $s \in \omega^{k^\star}$, we have:

$$\sigma^\star \bowtie s = (x_0, (s)_0, \dots, x_{k-1}, (s)_{k-1})$$

where $x_0 = \sigma(())$, and:

- $x_{i+1} = \sigma(x_0, (s)_0, \dots, x_i, (s)_i)$, if $(s)_j \neq (s')_{j+1}$ for every $s' \in C$ and $0 \leq j \leq i$;
- $x_{i+1} = z^*$, with

$$z^* = \min\{z \in \mathbb{N} : \sigma^\star(x_0, (s)_0, \dots, x_i, (s)_i) \bowtie z \not\sqsubseteq s', \text{ for every } s' \in C\}$$

otherwise.

It follows then that $\sigma^\star \bowtie s \in A^\star$ is always the case, i.e., holds for every $s \in \omega^{k^\star}$, for:

- either s is a legal set of moves by Player II in G^R, hence $\sigma^\star \bowtie s = \sigma \bowtie s$ and $\sigma \bowtie s \in A^R \subseteq A^\star$ since σ is a winning strategy for Player I (therefore, $\sigma^\star \bowtie s \in A^\star$ as wanted);
- or, s contains illegal moves[5] and $\sigma^\star \bowtie s \in C$ as a consequence, but if one takes the least $i \in \mathbb{N}$ for which $(s)_i = (s')_i$ for some $s' \in C$ holds, then $\sigma^\star$ is instructed in such a way that $(\sigma^\star \bowtie s)_{i+1} = \sigma^\star((\sigma^\star \bowtie s)_0, \ldots, (s)_i)$ is a legal move by Player I (i.e., $(\sigma \bowtie s)_{i+1} \neq (s'')_{i+1}$ for every $s'' \in C$); this means that $|ml(\sigma^\star \bowtie s)| = 2m+1$ for some $m \in \mathbb{N}$, and $\sigma^\star \bowtie s \in A^\star$ by the definition of it.

Let, for the contrary direction of the statement, $\sigma^\star$ be a given winning strategy for Player I in $G^\star$. Then, for every $s \in \omega^{k^\star}$ which is a legal set of moves by Player II, $\sigma^\star \bowtie s \in A^\star$. However, it is also the case that $\sigma^\star \bowtie s \notin C$ (for, otherwise $|ml(\sigma^\star \bowtie s)| = 2m$ for some $m \in \mathbb{N}$, hence this would correspond to an illegal match of G^R where Player I was the first to play illegally; this is not the way $A^\star$ was defined and it would contradict $\sigma^\star \bowtie s \in A^\star$, that we know holds instead owing to $\sigma^\star$ being a winning strategy). Hence, it must be $\sigma^\star \bowtie s \in A^R$, which means that $\sigma^\star$ is winning for Player I in G^R and the statement holds. QED

The same model $G^\star$ turns out to allow us to prove the corresponding result to proposition 4.2 when the alternative, strict attitude toward illegal matches is taken and winning strategies for players are defined according to definition 4.9 instead:

Proposition 4.3 *Let G^R and $G^\star$ be as before, and let the notion of winning strategy for Player I and Player II in G^R be defined according to definition 4.9. Then, Player I has a winning strategy in G^R if and only if the same holds in $G^\star$.*

Proof: as before, we start from the left-to-right direction of the statement, and assume that σ is a (strict) winning strategy for Player I in G^R. Then, if s is an arbitrary but fixed element of $\omega^{k^R} = \omega^{k^\star}$, either $\sigma \bowtie s \in A^R$ in case s is a legal set of moves by Player II, which yields $\sigma \bowtie s \in A^\star$ since $A^R \subseteq A^\star$ holds by definition, or s is not a legal set of moves by Player II, hence (what stated in footnote 5 holds of it and) $\sigma \bowtie s \in C$ with $|ml(\sigma \bowtie s)| = 2m+1$ for some $m \in \mathbb{N}$, from which it follows that $\sigma \bowtie s \in A^\star$ owing again to the definition of it. That is: σ is also a winning strategy for Player I in $G^\star$, and (this direction of) the statement of the proposition holds.

[5] Owing to definition 4.5, this holds if $(s)_i = (s')_{2i+1}$ for some $0 \leq i < k^R$ and $(s)_i \neq (s'')_{2i+1}$ for every $s'' \in P^R$.

Conversely, let us assume that Player I has a winning strategy in $G^\star$ in the form of a function $\sigma^\star$. Let then $s \in \omega^{k^R}$ be such that s is a legal set of moves by Player II. Then, $\sigma^\star \bowtie s \in A^\star$ since $\sigma^\star$ is a winning strategy and $\omega^{k^R} = \omega^{k^\star}$ by definition. This can be either because $\sigma^\star \bowtie s \in A^R$, or because $\sigma^\star \bowtie s \in C$ and $|ml(\sigma^\star \bowtie s)| = 2m+1$ for some $m \in \mathbb{N}$. Having supposed that s is a legal set of moves by Player II, the latter cannot be the case and $\sigma^\star \bowtie s \in A^R$ must be true instead.

If s contains illegal moves instead, and $\sigma^\star \bowtie s \in C$ as a consequence, since $\sigma^\star \bowtie s \in A^\star$ also holds having assumed that $\sigma^\star$ is winning for Player I in $G^\star$, then $\sigma^\star \bowtie s \in A^R$ cannot be the case owing to $\sigma^\star \bowtie s \in C$, which means that $|ml(\sigma^\star \bowtie s)| = 2m+1$ must hold for some $m \in \mathbb{N}$. That is: $\sigma^\star$ is a winning strategy for Player I in G^R, and the left-to-right direction of the theorem holds. QED

One may wonder, having proved by proposition 4.2 and proposition 4.3 that for Player I to have a winning strategy in $G^\star$ is the same for her as to have either a permissive or a strict winning strategy in G^R, whether from the two results it follows that these two latter concepts, that were apparently looking differently, are the same instead in the end. This is not so, and the reason is that the two results we have proved have different consequences. In particular, proposition 4.3 proves more than stated: it actually proves that the set of winning strategies for Player I in $G^\star$ is *the same as* (i.e., it contains the same elements of) the set of strict winning strategies for Player I in G^R. Proposition 4.2 instead does not do the same to the set of winning strategies for Player I in $G^\star$ and the set of permissive winning strategies for Player I in G^R. For, while it proves that a winning strategy for Player I in $G^\star$ is a permissive winning strategy for Player I in G^R, it only proves that out of a permissive winning strategy of this latter sort, one winning strategy for Player I in $G^\star$ *can be defined*. This means that the two sets of winning strategies in question have a non-empty intersection, but are not necessarily the same set of strategies in the end.

4.7.2. Games with matches of different lengths

Another possible modification of the canonical model of games we have decided to stick to, would be prompted by the following observation: it is part of our common experience with games that matches have not always one and the same length; as a matter of fact, there might be different factors causing matches to last differently, including the ability of a player to find a more efficient strategy in one match, that helps her to win earlier than in another match of the same game. This simple observation makes the original requirement over matches in a game to have always one and the same length a bit odd. In order to take this issue into account, the canonical model is modified as follows. Take now a finite game G^{DL} (for «games with (matches of) different lenghts») to be a triple like the following:

$$G^{DL} = \langle k^{DL}, P^{DL}, A^{DL} \rangle$$

where:

- $k^{DL} \in \mathbb{N}$ is the bound over the number of moves of matches of the game G^{DL};
- $P^{DL} \subseteq \omega^{<2k^{DL}}$ is the set of matches of G^{DL};
- $A^{DL} \subseteq P^{DL}$ is the set of winning conditions for Player I in G^{DL}.

Game G^{DL} is played exactly in the same way as games in canonical form are played. In particular, matches always have even length, which means that Player II is always the last to move. Yet, they have different durations and one match $s \in P$ may end earlier (i.e., when a minor number of moves have been played) than another $s' \in P$ (i.e., $|s| < |s'|$ holds in this case). Notice that we assume that the bound over the number of moves allowed to each player is never reached, since for no $s \in P^{DL}$ we have $|s| = 2k^{DL}$. This motivates the view we propose below of the main result of this part of the section, proposition 4.4 below, as establishing the means for passing from such a model to one in which all matches of the given game are extended to have length equal to $2k^{DL}$.

Now, the definition of winning strategy in this type of games needs to be modified with respect to how it was given in definition 4.4, so to take into account the new features of matches of G^{DL}. This is done by means of definition 4.11 below, where, for the sake of completeness, we have defined the new concept for both players even though only the new definition of winning strategy for Player I is actually needed, as far as proposition 4.4 below is concerned.

Before going into that, however, we need as usual to define strategies in general. For technical reasons that make life easier with the proof of the said proposition, we leave it as it stands. As we shall briefly comment afterwards, this has some apparently counter-intuitive drawbacks that should be taken into account:

Definition 4.10 *Let G^{DL} in the form of the previous triple be given. Let $E^{<2k^{DL}}$ and $O^{<2k^{DL}}$ be as in definition 4.1, that is:*

$$E^{<2k^{DL}} = \{s \in \omega^{<2k^{DL}} : |s| = 2m, m \in \mathbb{N}\}$$
$$O^{<2k^{DL}} = \{s \in \omega^{<2k^{DL}} : |s| = 2m+1, m \in \mathbb{N}\}$$

Then:

$$\sigma : E^{<2k^{DL}} \to \omega$$

is a strategy for Player I in G^{DL}, and:

$$\tau : O^{<2k^{DL}} \to \omega$$

is a strategy for Player II in G^{DL}.

Strategies being defined in this way are capable of returning values also for sequences which are not supposed to be matches of G^{DIL}, due to the fact that, owing to $P^{DIL} \subseteq \omega^{<2k^{DL}}$, P^{DL} might be a proper subset of that total set. This, however, is no harm, and is just motivated to make proofs easy: we might have been more strict here and redefine strategies as functions which return values only when a proper match is concerned. Notice that as we defined them, strategies are doing what they are supposed to be doing since thay are defined over a superset of P^{DL}, hence they return values for every initial segment of an arbitrary element of the latter set.

As it was the case for the canonical model, also in the case of G^{DL} the notion of winning strategy is obtained as a specification of the previous definition of strategy:

Definition 4.11 *Let G^{DL} be a triple as before. Then:*

1. A strategy σ for Player I in G^{DL} is winning *if and only if, for every $s \in \omega^{<k^{DL}}$ such that $(s)_i = (s')_{2i+1}$ for some $s' \in P^{DL}$ and for every $0 \leq i < |s|$,*

$$\sigma \bowtie s \in A^{DL}$$

2. A strategy τ for Player II in G^{DL} is winning *if and only if, for every $s \in \omega^{<k^{DL}}$ such that $(s)_i = (s')_{2i}$ for some $s' \in P^{DL}$ and for every $i \leq |s| - 1$,*

$$s \bowtie \tau \in P^{DL} \setminus A^{DL}$$

Notice that out of all values that winning strategies produce as strategies, those that matters for their being winning are only the values returned with respect to portions of actual matches of G^{DL}. The extra conditions on $s \in \omega^{<k^{DL}}$ are intended to ensure that elements of s can count as moves by Player II and Player I respectively (hence, corresponds to elements occurring at odd places in such a match in the first case, which is where moves by Player II can be found, and corresponding to elements occurring at even places in a match, which is where moves by Player I stay instead).

As it was the case before with the modification of the original model to encapsulate the idea that some matches may contain illegal moves, we are interested in showing that for every game defined according to the idea that matches might have different lengths there exists a game in canonical form that corresponds to it in the sense that strategic winning of players is preserved. As anticipated above, we are going to partially achieve the goal by showing that, for every game G^{DL} in the previous form there exists a game G^{+} whose matches are obtained by extending the length of any match in G^{DL} up to length $2k^{DL}$, and such that it corresponds to G^{DL} in the previous sense of the expression. This does not yet mean that we have obtained a version in canonical form of the latter model as the set P^{+} of matches of G^{+} is, as it will be clear in a minute, still a subset, and possibly a proper subset of the total set $\omega^{2k^{DL}}$. We will make some comments about this at the end of the section.

For the time being, let us content ourselves with the goal just described, and let G^{+} be as follows:

$$G^{+} = \langle k^{+}, P^{+}, A^{+} \rangle$$

where:

- $k^{+} \in \mathbb{N}$ is the maximum number of moves of matches of G^{+};
- $P^{+} \subseteq \omega^{2k^{+}}$ is the set of matches of G^{+};
- $A^{+} \subseteq P^{+}$ is the set of winning conditions for Player I in G^{+}.

The conditions we put over the elements of G^{+} are:

1. $k^{+} = k^{DL}$;
2. $P^{+} = \{s \in \omega^{2k^{+}} : s' \sqsubseteq s$, for some $s' \in P^{DL}\}$;
3. $s \in A^{+}$ if and only if $s' \sqsubseteq s$ for some $s' \in A^{DL}$, and $(s)_i = 0$ for every $|s'| \leq i < |s| - 1$ and $i = 2m$ for some $m \in \mathbb{N}$, while $(s)_i = 1$ for every $|s'| \leq i \leq |s| - 1$ and $i = 2m + 1$ with $m \in \mathbb{N}$.

The first two conditions of the list make clear what we meant to emphasize above: matches of G^{+} are obtained by extending matches of G^{DL} until length $2k^{DL}$ is reached. The third condition reveals the whole idea on which winning matches in G^{+} is based. As far as Player I is concerned, she is required to first play in such a way the portion of the match that corresponds to a match in G^{DL} is won, and then trivialize her game behaviour to the point of playing the constant move «0» until the end of the match is reached. In other words, the hard part of winning a match in G^{+} for Player I is winning the part of it that is a match in G^{DL}. Player II should play no differently, as once she has made sure to have won the portion of the match in G^{+} that corresponds to a match of G^{DL}, then she just have to play numbers different from 1.

The conditions on P^{+} and A^{+} justify the following observations:

Remark 4.1 *For every $s' \in P^{DL}$, there exists $s \in P^{+}$ such that $s' \sqsubseteq s$ (i.e., every match of G^{DL} is extended by some match in G^{+}), while for every $s' \notin P^{DL}$ and for every $s \in P^{+}$, we have $s' \not\sqsubseteq s$ (i.e., no match in G^{+} is the extension of a sequence that is not in P^{DL}).*

Remark 4.2 *For every $s' \in A^{DL}$, there exists $s \in A^+$ such that $s' \sqsubseteq s$, $(s)_i = (s')_i$ for every $0 \leq i < |s'|$, and $(s)_j = 0$ for every $|s'| \leq j < |s|$ with $j = 2m+1$ for some $m \in \mathbb{N}$ (i.e., every match of G^{DL} where Player I wins has an extension in G^+ where Player I also wins), while for every $s' \notin A^{DL}$ and for every $s \in A^+$, $s' \not\sqsubseteq s$ is the case (i.e., no match of G^{DL} where Player I does not win has an extension in G^+ where Player I wins).*

Next come the definitions of strategy and winning strategy for game G^+. As far as the first is concerned, nothing changes with respect to the canonical model:

Definition 4.12 *Let G^+ be given as the above triple. Let the sets $E^{<2k^+}$ and $O^{<2k^+}$ be defined as follows:*

$$E^{<2k^+} = \{s \in \omega^{<2k^+} : |s| = 2m, m \in \mathbb{N}\}$$
$$O^{<2k^+} = \{s \in \omega^{<2k^+} : |s| = 2m+1, m \in \mathbb{N}\}$$

Then:

$$\sigma : E^{<2k^+} \to \omega$$

is a strategy for Player I in G^+, and:

$$\tau : O^{<2k^+} \to \omega$$

is a strategy for Player II in G^+.

Next comes the definition of winning strategy. If G^+ were a game in canonical form itself, we could just repeat here what was stated in definition 4.4 above. However, that definition must be modified owing to the fact that $P^+ \subseteq \omega^{2k^+}$ only, which leaves open the possibility that $C^+ = \omega^{2k^+} \setminus P^+ \neq \emptyset$ (as it was the case for games such as G^R above). Therefore, we apply a similar modification to the original definition in this case:

Definition 4.13 *Let G^+ be a triple as before. Then:*

1. A strategy σ for Player I in G^+ is winning *if and only if, for every $s \in \omega^{k^+}$ such that $(s)_i = (s')_{2i+1}$ for some $s' \in P^+$ and for every $0 \leq i < |s|-1$,*

$$\sigma \bowtie s \in A^+$$

2. A strategy τ for Player II in G^+ is winning *if and only if, for every $s \in \omega^{k^+}$ such that $(s)_i = (s')_{2i}$ for some $s' \in P^+$ and for every $0 \leq i < |s|$,*

$$s \bowtie \tau \in \omega^{2k} \setminus A^+$$

Finally, we are in a position to conclude this part of the section by proving, as before, that having a strategy condition for Player I in a game like G^{DL} above can be put in correspondence to having a winning strategy for Player I in G^+. Once again we only provide the reader with this result as a sample of how the symmetrical argument needed to conclude the same for Player II would look like.

Then, we have:

Proposition 4.4 *Let G^{DL} and G^+ be as before. Then: Player I has a winning strategy in G^{DL} if and only if Player I has a winning strategy in G^+.*

Proof: we first prove the left-to-right direction of the statement, and suppose that σ is a given winning strategy for Player I in G^{DL} then. Take $s \in \omega^{<2k^+}$ to be such that $|s| = 2m$ for some $m \in \mathbb{N}$ and $s \sqsubseteq s^+$ for some $s^+ \in P^+$ (that is: s represents a portion of an actual match of G^+). Then, owing to the definition of P^+ above, $s' \sqsubseteq s$, wfor some $s' \in P^{DL}$, and $s = s' \bowtie s''$ where either $s'' = (\,)$ or not. Let, for every $s \in \omega^{<2k^+}$ meeting the extra condition above, s_{DL} indicate the initial sequence of it that belongs to P^{DL} (hence, in our running example, $s_{DL} \in P^{DL}$ and $s_{DL} \sqsubseteq s \sqsubseteq s^+$ where $s^+ \in P^+$). Let, for every such s, $s - s_{DL}$ be the remaining part of the sequence (that is, $s = s_{DL} \bowtie (s - s_{DL})$).

This is something we use for the sake of defining an extension σ^+ of σ that may work as winning strategy for Player I in G^+. Put:

$$\sigma^+(s) = \begin{cases} \sigma(s), \text{ if } s - s_{DL} = (\,) \\ 0, \text{ otherwise} \end{cases}$$

if s is as above, and let

$$\sigma^+(s) = 1$$

if $s \in \omega^{<2k^+}$ with $|s| = 2m$ for some $m \in \mathbb{N}$, but $s \not\sqsubseteq s^+$ for every $s^+ \in P^+$. Then, σ^+ is a function defined on the whole set $E^{<2k^+}$ and hence is a strategy for Player I in G^+. Notice in particular that $\sigma^+((\,)) = \sigma((\,))$ since $(\,)_{DL} = (\,) = (\,) - (\,)_{DL}$.

Let now $s \in \omega^{k^+}$ be such that $(s)_i = (s')_{2i+1}$ for some $s' \in P^+$ and for every $0 \leq i < |s|$. Then,

$$\sigma \bowtie s = (\sigma((\,)), (s)_0 = (s'_{DL})_1, \sigma((\sigma(\,), (s'_{DL})_1)), \ldots, (s' - s'_{DL})_1, 0, 0, \ldots 0)$$

since $s' = s'_{DL} \bowtie (s' - s'_{DL})$. It turns out that $\sigma^+ \bowtie s \in A^+$ since $\sigma^+ \bowtie s = (\sigma \bowtie s'_{DL}) \bowtie s''$ where:

- $\sigma \bowtie s'_{DL} \in A^{DL}$ since σ is winning for Player I in G^{DL} and $s'_{DL} \in P^{DL}$ (see also remark 4.1 above);
- s'' is such that $(s'')_i = 0$ for every $|s'_{DL}| \leq i < 2k^+ - 1$, with $i = 2m$ for some $m \in \mathbb{N}$.

hence the proposition.

Conversely, let σ^+ be a winning strategy for Player I in G^+. Let $s \in \omega^{<k^{DL}} = \omega^{<k^+}$ be such that $(s)_i = (s')_{2i+1}$ for some $s' \in P^{DL}$ and for every $i \leq |s| - 1$ and let us suppose that

$$\sigma^+ \bowtie s \notin A^{DL}$$

is the case (notice that $\bowtie$ is defined here owing to σ^+ being a strategy for Player I in G^+, and therefore counting also as a strategy for Player I in G^{DL} owing to our main assumptions and to how definition 4.12 was devised). Take $s^\star \in \omega^{k^+}$ to be such that $(s^\star)_i = (s^\dagger)_{2i+1}$ as in part 1 of definition 4.12, and, moreover, such that $s \sqsubseteq s^\star$ also holds.

First of all, notice that such $s^\star$ exists: having assumed that $(s)_i = (s')_{2i+1}$ for some $s' \in P^{DL}$ and for every $0 \leq i < |s|$, and since we also have $s' \sqsubseteq s''$ for some $s'' \in P^+$ owing to remark 4.1 above, we have $(s)_i = (s'')_{2i+1}$ for every $0 \leq i < |s|$; hence, it is enough that $s^\star = s''$ for the previous assumption to be justified.

Then,

$$\sigma^+ \bowtie s \sqsubseteq \sigma^+ \bowtie s^\star$$

and

$$\sigma^+ \bowtie s^\star \in A^+$$

both hold since σ^+ is winning for Player I in G^+ by assumption. This means that from our hypothesis it would follow that there is an element $\sigma^+ \bowtie s^\star$ of A^+ whose initial subsequence $\sigma^+ \bowtie s$ is not an element of A^{DL}, which contradicts remark 4.2. By that we conclude $\sigma^+ \bowtie s \in A^{DL}$, which yields, being s fixed but arbitrary, that σ^+ is winning for Player I in G^{DL}. Hence the proposition. QED

As we warned already, the result we have achieved in this way is not yet a strategical correspondence between games whose set of matches is a subset of the total set of sequences with a given length and games in canonical form. This is because in passing from games such as G^{DL} above to G^+, we have passed from the set $P^{DL} \subseteq \omega^{<2k^{DL}}$ to the set $P^+ \subseteq \omega^{2k^+} = \omega^{2k^{DL}}$. Hence, all matches in the given game have been extended to matches with length equal to $2k^{DL}$. A more clever definition of P^{DL} would have allowed us to achieve a broader result. However, I felt that this putting more extra conditions would have made things less clear to read. Therefore, I have decided to leave things simple and illustrate what I thought was really worth conveying.

4.8. Determinacy of finite games

In section 4.6 we managed to clarify what is the main model of finite games we plan to study, and, in view of what we have just seen in the subsequent sections, we now know that the model is sufficiently flexible that it can cope with many of the features that games have in reality.

The concept of strategy in the model was first used to show that games in the sense of our model are zero-sum games, hence it is not the case that both Player I and Player II have a winning strategy. Now, we would like to investigate a closely related question: Is it always the case that in every finite game either Player I has a winning strategy, or Player II has a winning strategy instead?

Let us first introduce some terminology:

Definition 4.14 *A finite game $G = \langle k, P, A\rangle$ is* determined *if it is either the case Player I has a winning strategy, or Player II has a winning strategy instead.*

We are going to attack the problem of determinacy of finite games by following a 'logical route" similar to the one that brought us to study equilibria in games in normal forms as fixpoints of a certain class of operators. As one could expect then, this proof strategy requires to be more precise about the linguistic resources we have been using so far.

The main relationship we have spoken of, and the one that we have to stick to as long as the goal of this section is concerned, is the elementhood relation between *individuals*, in the form of either natural numbers, or sequences of natural numbers, and *sets*. Relations of this sort are denoted by $n \in S$, or $s \in S$, as we have conventionally decided to indicate numbers with symbols $k, m, n, \ldots, k', m', n', \ldots, k_0, m_0, n_0, \ldots$, to indicate sequences by means of symbols $s, s', s'', \ldots, s_0, s_1, \ldots$, and save capital letters for sets.

So far we have not needed to dig up and discuss the relation between each of these symbols and the 'objects' they refer to. This becomes more important now, due to to the peculiar proof strategy we plan to pursue to accomplish the said task. As long as the latter is concerned, however, we need to formally state with precision only a very little fragment of the linguistic resources we have made use of so far. In particular, we need to set up a formal language that allows us to speak of some very fundamental facts occurring in a domain such as a game G in canonical form. These fundamental facts concerns sequences that belong or do not belong to the set A of winning conditions for Player I in G.

The language we are going to build will be called $\mathscr{L}_S^-$ («S» being a shorthand for «sequences», the main objects we will be referring to by means of it, and the minus sign «$-$» being a reminder of the quite weak expressive power of it). We assume this language to have 'proper names' for all the elements of $\mathbb{N}$ in the form of constant symbols $\overline{k}$ for every $k \in \mathbb{N}$. In addition, we assume it comprises means for referring to elements of $\mathbb{N}$ generically, that is by means of variable symbols $x_0, x_1, \ldots, x_n, \ldots$. What we mean by saying that these symbols of $\mathscr{L}_S^-$ 'refer to' (or, 'range over' as it should be better said) numbers, it is something that has already been exploited for the sake of building the language $\mathscr{L}_{GM}$ to speak of game matrices in the previous chapter and, as far as the new language is concerned, will turn out clearly by definition 4.19 below.

Terms of the said sort, either proper names or variables, are used to construct «complex terms» of $\mathscr{L}_S^-$, which indicate *sequences of numbers*. This will be done by supposing that the alphabet of $\mathscr{L}_S^-$ also contains (denumerably many) n-ary function symbols $(\ )^n$ for every $n \in \mathbb{N}$. As it turns out from definition 4.15 below, these are conceived as (term-)functions that take as input a finite list of n terms of $\mathscr{L}_S^-$, for instance $t_0, \dots, t_{n-1}$, and return the term $(t_0, \dots, t_{n-1})^n$ as value (in the following, the superscript n will be omitted in the case of complex terms of this form, whenever the indices of terms occurring into it make it redundant). The intended meaning for such complex construction is to refer to the sequence whose elements are the objects of our reference domain named by terms $t_0, \dots, t_{n-1}$.

In addition to «individual terms» of this three-fold sort (constants $\overline{k}$, variables x_i, and sequence terms $(t_0, \dots, t_{n-1})$), $\mathscr{L}_S^-$ is equipped with a constant set name $\overline{A}$, that we intend to use to refer to the set of winning conditions for Player I of a given finite game. As long as sentences of this language are concerned, we assume that $\mathscr{L}_S^-$ contains a binary relation symbol $\in$, that will be used to express the usual elementhood relation (see definition 4.16 below), as well as some symbols which have a standard logical meaning like $\neg$ for negation, $\forall$ for universal quantification over individuals, and $\exists$ for existential quantification. Expressions of $\mathscr{L}_S^-$ are obtained also by making use of parentheses (,) as auxiliary symbols.

Having said that, the exact definition of this notion of «term of $\mathscr{L}_S^-$» is given below, along with the definition of the notion of «free variables» for expressions that count as terms of $\mathscr{L}_S^-$:

Definition 4.15 *1. The set $TERM_S^-$ of* terms *of $\mathscr{L}_S^-$ is the smallest collection that contains all terms $\{\overline{k} : k \in \mathbb{N}\}$, all variables $x_0, x_1, \dots, x_n, \dots$, and is closed under application of sequence symbols $(\ )^n$ to them, that is contains expressions of the form $(t_0, \dots, t_{n-1})$ for every $n \in \mathbb{N}$ where each t_i (with $0 \leq i < n$) is either a constant (i.e., $t_i \equiv \overline{m}$ for some $m \in \mathbb{N}$), or a variable (i.e., $t_i \equiv x_j$)*[6].

2. For every element t of $TERM_S^-$, the set of free variables *of t, $FV(t)$, is defined as follows: $FV(t) = \emptyset$ in case $t \equiv \overline{k}$ ($k \in \mathbb{N}$); $FV(t) = \{x_n\}$ if $t \equiv x_n$; $FV(t) = \{x_1, \dots, x_m\}$, in case $t \equiv (t_0, \dots, t_{n-1})$ and $\bigcup_{1 \leq i \leq n-1} FV(t_i) = \{x_1, \dots, x_m\}$.*

Let us just add a couple of comments:

[6] I am using the symbol $\equiv$ to indicate coincidence between two formal expressions. So, $e \equiv e'$ holds if e and e' are expressions of $\mathscr{L}_{\text{seq}}^-$ – finite lists of symbols from the alphabet of the language – that coincide with one another (i.e., cointain exactly the same symbols in the same order).

- part 1 of the previous definition aims at identifying the collection of expressions that can be formed out of the symbolic resources of $\mathscr{L}_S^-$ and count as well-formed 'names' of it. As anticipated, names we count are either proper or generic names for natural numbers, as well as names for sequences of them;
- by means of part 2 of the definition we are simply isolating, for every expression t that counts as a term of $\mathscr{L}_S^-$, the variables that occur into it (i.e., which are part of the list of symbols of $\mathscr{L}_S^-$ that t is made out of); the reason why variables as such are called «free» is something that will turn out clearly (and will be thereby stressed) after definition 4.16 below.

Terms of $\mathscr{L}_S^-$ are then used to make statements regarding sequences of numbers belonging, or not belonging to the set A of a given game G. This is done by means of a binary relation $\cdot \in \cdot$ whose intended meaning is: «$\cdot$ is an element of $\cdot$». The definition of the set of expressions that we regard as *formulas of* $\mathscr{L}_S^-$ is given below (along with the extension to expressions of $\mathscr{L}_S^-$ of this sort of the previous notion of «free variables»). For the sake of it, we make use of the logical tools of $\mathscr{L}_S^-$ for extending its own expressive capacity:

Definition 4.16 *1.The* atomic formulas *of* $\mathscr{L}_S^-$ *are all, and only the expressions of the form* $t \in \overline{A}$, *where* t *is in* $TERM_S^-$ *and*

$$t \equiv (t_0, \ldots, t_{n-1})$$

for some $n \in \mathbb{N}$, *with* $t_0, \ldots, t_{n-1} \in TERM_S^-$.

1.1 If $t \in \overline{A}$ *is an atomic formula of* $\mathscr{L}_S^-$, *then the set* $FV(t \in \overline{A})$ *of free variables of it is given by* $FV(t \in \overline{A}) = FV(t)$.

2. The set $FORM_S^-$ *of formulas of* $\mathscr{L}_S^-$ *is the smallest set of expressions of it that contains all of the atomic formulas of* $\mathscr{L}_S^-$, *and is further defined by the following clauses:*

- *if* φ *is an element of* $FORM_S^-$, *then* $\neg\varphi$ *is also an element of it, with* $FV(\neg\varphi) = FV(\varphi)$;
- *if* φ *is an element of* $FORM_S^-$ *with* $x_i \in FV(\varphi)$, *then* $\forall x_i \varphi$ *is an element of* $FORM_S^-$ *too, with* $FV(\forall x_i \varphi) = FV(\varphi) \setminus \{x_i\}$;
- *if* φ *is an element of* $FORM_S^-$ *with* $x_i \in FV(\varphi)$, *then* $\exists x_i \varphi$ *is an element of* $FORM_S^-$ *too, with* $FV(\exists x_i \varphi) = FV(\varphi) \setminus \{x_i\}$.

Part 2.2 and 2.3 of the previous definition makes clear in what sense a variable can be free in a expression of $\mathscr{L}_S^-$, by establishing what it means for it to be no more free (or, «bounded» by a quantifier as logicians say): the latter is said of a variable occurring in an expression φ that counts as a formula of $\mathscr{L}_S^-$ and which is subject to the application of a quantifier (either $\forall$, or $\exists$) prenexed to it. That is, x_i is bounded in $Qx_i\varphi$, where Q is either $\forall$ or $\exists$ and φ is a formula of $\mathscr{L}_S^i$, if it occurs free in φ (i.e., if φ is *not* an expression of the form $Qx_i\psi$ where again Q is either $\forall$ or $\exists$ and ψ is a formula of $\mathscr{L}_S^-$). In other words, «free» for a variable means: «within the range of no quantifier symbol».

Having devised the syntax of $\mathscr{L}_S^-$ in this way, it is now time to equip expressions that are well-formed with respect to the syntactical rules we have listed with a *meaning*. This is done by similarly devising a basic semantics for $\mathscr{L}_S^-$.

As it was already mentioned for language $\mathscr{L}_{GM}$ from the previous chapter, this is a consequence of our habit of looking at languages as tools for *speaking of* something, i.e. allowing us to refer of something occurring in a domain of «facts», or «state of affairs», or «situations» etc., we would like to report on. The same goes with formal languages like the language $\mathscr{L}_S^-$ we are devising. The intended domain of reference is a finite game in the canonical form they are supposed to take, namely as triples $G = \langle k, P, A\rangle$, and in particular we want to speak by means of sentences of $\mathscr{L}_S^-$ about facts concerning the set A which take the form $s \in A$ (that corresponds to the 'object' s belonging to the 'collection of objects' A) or $s \notin A$ (which corresponds instead to s not belonging to A). For our intention to be met, we need to state a correspondence between sentences and facts of the said sort that also comprises the means for saying when a sentence truly refers to the fact it corresponds to. As it is customary for formal languages, this is done by specifying a relation of validity for formulas of $\mathscr{L}_S^-$. Since this relation presupposes a reference domain to be given (in the form of a game G) and is therefore relative to that domain, and since it concerns formulas which report on facts referring to sequences of natural numbers belonging or not belonging to a collection of them (which makes it also relative to the 'nature' of the objects these formulas speak of), we call it «$\mathbb{N}$-validity relative to G».

Before going into that, we need to establish the required notation, and explain what we mean in the following by writing

$$\varphi[x_0/t_0, \ldots, x_{n-1}/t_{n-1}]$$

for a given formula φ of $\mathscr{L}_S^-$, individual variables $x_0, \ldots, x_{n-1}$ and terms $t_0, \ldots, t_{n-1}$. This is intended to indicate the expression that is obtained from φ by literally replacing the symbols $x_0, \ldots, x_{n-1}$ occurring into it by means of the symbols $t_0, \ldots, t_{n-1}$ in that order. Since free variables of a formula are inherited by free variables of the terms occurring into it (see definition 4.16 above), for this notation to be defined precisely we are required to first define with a similar degree of precision what is meant by

$$t[x_0/t_0, \ldots, x_{n-1}/t_{n-1}]$$

where again $t, t_0, \ldots, t_{n-1}$ are in the set $TERM_S^-$ of terms of $\mathscr{L}_S^-$ and

$x_0, \ldots, x_{n-1}$ are variables of $\mathscr{L}_S^-$. This we use as a notation to indicate the expression that is obtained from t by substituting every occurrence of $x_0, \ldots, x_{n-1}$ with terms $t_0, \ldots, t_{n-1}$ in that order.

This is defined over elements of $TERM_S^-$ as follows:

Definition 4.17 *1. If* $t \equiv x_i$ *for some* $0 \leq i \leq n-1$ *(hence,* $FV(t) = \{x_i\}$*), then we have that the term* $t[x_0/t_0, \ldots, x_{n-1}/t_{n-1}]$ *is* t_i*, for every variables* $x_0, \ldots, x_{n-1}$ *of* $\mathscr{L}_S^-$ *and* $t_0, \ldots, t_{n-1} \in TERM_S^-$.

2. If $t \equiv x_j$ *for some variable* x_j *of* $\mathscr{L}_S^-$ *such that* $x_j \neq x_i$ *for every* $1 \leq i \leq n-1$ *(hence,* $FV(t) = \{x_j\}$ *and* $x_i \notin FV(t)$ *for every* $1 \leq i \leq n-1$*), then* $t[x_0/t_0, \ldots, x_{n-1}/t_{n-1}] \equiv x_j$ *for every variables* $x_0, \ldots, x_{n-1}$ *of* $\mathscr{L}_S^-$ *and* $t_0, \ldots, t_{n-1} \in TERM_S^-$.

3. If $t \equiv \overline{k}$ *for some* $k \in \mathbb{N}$ *(hence,* $FV(t) = \emptyset$*), then*

$$t[x_0/t_0, \ldots, x_{n-1}/t_{n-1}] \equiv \overline{k}$$

for every variables $x_0, \ldots, x_{n-1}$ *of* $\mathscr{L}_S^-$ *and* $t_0, \ldots, t_{n-1}$ *in* $TERM_S^-$.

4. If $t \equiv (t'_0, \ldots, t'_{m-1})$ *for some* $m \in \mathbb{N}$ *and for some* $t'_0, \ldots, t'_{m-1} \in TERM_S^-$ *(hence,* $FV(t) = \bigcup_{0 \leq i \leq m-1} FV(t_i)$*), then*

$$t[x_0/t_0, \ldots, x_{n-1}/t_{n-1}] \equiv (t^*_0, \ldots, t^*_{m-1})$$

where, for every $0 \leq i \leq m-1$, $t^*_i \equiv t'_i[x_0/t_0, \ldots, x_{n-1}/t_{n-1}]$ *(for every* $x_0, \ldots, x_{n-1}$ *variables of* $\mathscr{L}_S^-$ *and* $t_0, \ldots, t_{n-1}$ *in* $TERM_S^-$*).*

It follows from the definition that if $\{x_0, \ldots, x_{n-1}\} \notin FV(t)$, then

$$t[x_0/t_0, \ldots, x_{n-1}/t_{n-1}] \equiv t$$

for every variables $x_0, \ldots, x_{n-1}$ of $\mathscr{L}_S^-$ and terms $t, t_0, \ldots, t_{n-1}$ in $TERM_S^-$ (proof is left as an exercise – hint: it is enough to carry out an inductive argument over $TERM_S^-$).

As said, we are now in a position to similarly define the corresponding notation for formulas of $\mathscr{L}_s^-$:

Definition 4.18 *1. If* $t \equiv (t'_0, \ldots, t'_{m-1})$ *for some* $m \in \mathbb{N}$ *and for some* $t'_0, \ldots, t'_{m-1} \in TERM_S^-$*, and* $\varphi \equiv t \in \overline{A}$ *(hence,* $FV(\varphi) = FV(t)$*), then*

$$\varphi[x_0/t_0, \ldots, x_{n-1}/t_{n-1}] \equiv t[x_0/t_0, \ldots, x_{n-1}/t_{n-1}] \in \overline{A}$$

for every variables $x_0, \ldots, x_{n-1}$ *and* $t_0, \ldots, t_{n-1}$ *in* $TERM_S^-$.

2. If $\varphi \equiv \neg\psi$ *for some* $\psi \in FORM_S^-$ *(hence,* $FV(\varphi) = FV(\psi)$*), then*

$$\varphi[x_0/t_0, \ldots, x_{n-1}/t_{n-1}] \equiv \neg\psi[x_1/t_1, \ldots, x_n/t_n]$$

for every variables $x_0, \ldots, x_{n-1}$ *and* $t_0, \ldots, t_{n-1}$ *in* $TERM_S^-$.

3. If $\varphi \equiv \forall x_i \psi$ for some $\psi \in FORM_S^-$ (hence, $FV(\varphi) = FV(\psi) \setminus \{x_i\}$), then

$$\varphi[x_0/t_0,\dots,x_{n-1}/t_{n-1}] \equiv \forall x_i \psi[x_0/t_0,\dots,x_{n-1}/t_{n-1}]$$

for every variables $x_i, x_0, \dots, x_{n-1}$ of $\mathscr{L}_S^-$ with $i \neq j$ for every $0 \leq j \leq n-1$ and $t_0, \dots, t_{n-1} \in TERM_S^-$.[7]

3. If $\varphi \equiv \exists x_i \psi$ for some $\psi \in FORM_S^-$ (hence, $FV(\varphi) = FV(\psi) \setminus \{x_i\}$), then

$$\varphi[x_0/t_0,\dots,x_{n-1}/t_{n-1}] \equiv \exists x_i \psi[x_0/t_0,\dots,x_{n-1}/t_{n-1}]$$

for every variables $x_i, x_0, \dots, x_{n-1}$ of $\mathscr{L}_S^-$ with $i \neq j$ for every $0 \leq j \leq n-1$ and $t_0, \dots, t_{n-1} \in TERM_S^-$ (notice again that it is not assumed that $x_i \in \{x_0, \dots, x_{n-1}\}$).

As in the previous case, it is an easy task to prove that, for every variables $x_0, \dots, x_{n-1}$ and terms $t, t_0, \dots, t_{n-1}$ in $TERM_S^-$,

$$\varphi[x_0/t_0,\dots,x_{n-1}/t_{n-1}] \equiv \varphi$$

holds if $\{x_0, \dots, x_{n-1}\} \notin FV(\varphi)$ (proof is again left as an exercise and the hint is the same: an inductive argument over $FORM_S^-$ this time is needed).

Having said that, the announced definition of $\mathbb{N}$-validity relative to a game G for formulas of $\mathscr{L}_S^-$, that is limited only to those expressions which are of interest for the sake of theorem 4.1 below, goes as follows:

Definition 4.19 *Let an arbitrary finite game $G = \langle k, P, A \rangle$ be given. Then we say:*

1. *a formula of $\mathscr{L}_S^-$ of the form $t \in \overline{A}$ where $t \equiv (\overline{n}_0, \dots, \overline{n}_{2k-1})$ for some $n_0, \dots, n_{2k-1} \in \mathbb{N}$ is $\mathbb{N}$-valid relative to G if and only if*

$$(n_0, \dots, n_{2k-1}) \in A$$

2. *a formula of $\mathscr{L}_S^-$ of the form $t \in \overline{A}$, where $t \in TERM_S^-$ $FV(t) = \{x_0, \dots, x_{m-1}\}$ ($m \leq 2k$), and $t \equiv (t_0, \dots, t_{2k-1})$ with $t_0, \dots, t_{2k-1}$ in $TERM_S^-$,*[8] *is $\mathbb{N}$-valid relative to G if and only if*

$$(t \in \overline{A})[x_0/\overline{p}_0, \dots, x_{m-1}/\overline{p}_{m-1}]$$

   is $\mathbb{N}$-valid relative to G for every $p_0, \dots, p_{m-1} \in \mathbb{N}$.
3. *a formula of the form $\neg\varphi$ for some $\varphi \in FORM_S^-$ is $\mathbb{N}$-valid relative to G if and only if φ is not $\mathbb{N}$-valid relative to G;*

[7] The extra assumption on x_i, which has the effect that this is different from all those x_j that are substituted by terms t_j in the given expression, is made to keep the definition simple. It is known that it causes no loss of generality.

[8] Since $t \in \mathrm{TERM}_S^-$ each term from the list $t_0, \dots, t_{2k-1}$ is required – see definition 4.15 – to be either of the form $\overline{n}$ for some $n \in \mathbb{N}$, or a variable x_j of $\mathscr{L}_S^-$. Hence the free variables occurring in t can be $2k$ at most.

4. *a formula of the form* $\forall x_i \varphi$ *for some* $\varphi \in FORM_S^-$ *and* x_i *variable of* $\mathscr{L}_S^-$, *is* $\mathbb{N}$*-valid relative to* G *if and only if* $\varphi[x_i/\overline{n}]$ *is* $\mathbb{N}$*-valid relative to* G *for every* $n \in \mathbb{N}$;
5. *a formula of the form* $\exists x_n \varphi$ *for some* $\varphi \in FORM_S^-$ *and* x_i *variable of* $\mathscr{L}_S^-$, *is* $\mathbb{N}$*-valid relative to* G *if and only if* $\varphi[x_i/\overline{n}]$ *is* $\mathbb{N}$*-valid relative to* G *for some* $n \in \mathbb{N}$;

Henceforth, we shall write

$$\models_{\mathbb{N}}^{G} \varphi$$

to mean that φ is $\mathbb{N}$-valid relative to G, and

$$\not\models_{\mathbb{N}}^{G} \varphi$$

to mean that φ is not $\mathbb{N}$-valid relative to G instead.
A few comments on definition 4.19 are due:

- The rationale behind clause no. 2 in the definition comes from wishing to assign to variables the value of generic names for natural numbers. Therefore, a formula φ that contains a generic name as such in the form of a free variable x_i should be regarded as $\mathbb{N}$-valid relative to G if and only $\varphi[x_i/\overline{n}]$ is for every possible value of n (i.e., for every $n \in \mathbb{N}$). This is stated explicitly for atomic formulas alone. However, it follows from an easy inductive proof on φ, that it extends to $FORM_S^-$ as a whole (that is: if $x_i \in FV(\varphi)$, then $\models_{\mathbb{N}}^{G} \varphi$ if and only if $\models_{\mathbb{N}}^{G} \varphi[x_i/\overline{n}]$ for every $n \in \mathbb{N}$).
- By looking at clauses 4 and 5, it turns out clearly what is the intended meaning of $\forall$ and $\exists$. In particular, $\forall x_i \varphi$ counts as the statement according to which «φ holds in G of every $n \in \mathbb{N}$» (i.e., $\models_{\mathbb{N}}^{G} \varphi[x_i/\overline{n}]$ for every $n \in \mathbb{N}$), while $\exists x_i \varphi$ counts as the statement according to which «φ holds in G of some $n \in \mathbb{N}$» (i.e., $\models_{\mathbb{N}}^{G} \varphi[x_i/\overline{n}]$ for some $n \in \mathbb{N}$). These are known as the *universal quantifier* (over individuals) and the *existential quantifier* (over individuals). Also, notice that the following holds owing to definitions 4.18 and 4.19: if $x_i \notin FV(\varphi)$, then $\models_{\mathbb{N}}^{G} \forall x_i \varphi$ and $\models_{\mathbb{N}}^{G} \exists x \varphi$ are the cases if and only if $\models_{\mathbb{N}}^{G} \varphi$.
- It follows from definition 4.19 that $\models_{\mathbb{N}}^{G} \forall x_i \exists x_j \varphi$ holds if and only if for every $n \in \mathbb{N}$, there exists $m \in \mathbb{N}$ such that $\models_{\mathbb{N}}^{G} \varphi[x_i/\overline{k}, x_j/\overline{h}]$.
- Observe instead, that $\models_{\mathbb{N}}^{G} \exists x_i \forall x_j \varphi$ if and only if there exists $n \in \mathbb{N}$, such that for every $m \in \mathbb{N}$, $\models_{\mathbb{N}}^{G} \varphi[x_i/\overline{n}, x_j/\overline{m}]$.

We now focus on generic formulas involved into remarks (iii) and (iv) above, since they turn out to play a crucial role in proving theorem 4.1 below. First, we state and prove the following ancillary result to that theorem, as a consequence of definition 4.19 above. For the sake of the readability of it, we allow us to use also symbols $y_0, \ldots, y_n, \ldots$ as ranging over variables of $\mathscr{L}_S^-$.

We have:

Lemma 4.2 *Let $G = \langle k, P, A \rangle$ be given. Then, for every $n \leq k$*

$$\models_{\mathbb{N}}^{G} \neg \exists x_0 \forall y_0 \ldots \exists x_{n-1} \forall y_{n-1} \varphi$$

holds if and only if

$$\models_{\mathbb{N}}^{G} \forall x_0 \exists y_0 \ldots \forall x_{n-1} \exists y_{n-1} \neg \varphi$$

is the case.

Proof: lemma is proved by induction on n. Assume first that $n = 0$ is the case. Then, it is clear that there cannot be a sequence of variables such as $x_0, y_0, \ldots, x_{n-1}, y_{n-1}$. That is,

$$\neg \exists x_0 \forall y_0 \ldots \exists x_{n-1} \forall y_{n-1} \varphi \equiv \neg \varphi \equiv \forall x_0 \exists y_0 \ldots \forall x_{n-1} \exists y_{n-1} \neg \varphi$$

So, we should prove that $\models_{\mathbb{N}}^{G} \neg \varphi$ holds if and only if $\models_{\mathbb{N}}^{G} \neg \varphi$, which is trivial.

Now, suppose that the lemma holds up to $n < k$ (this is the *induction hypothesis*), that is

$$\models_{\mathbb{N}}^{G} \neg \exists x_0 \forall y_0 \ldots \exists x_{n-1} \forall y_{n-1} \varphi$$

holds if and only if

$$\models_{\mathbb{N}}^{G} \forall x_0 \exists y_0 \ldots \forall x_{n-1} \exists y_{n-1} \neg \varphi$$

We want to prove that this holds true also for formulas

$$\neg \exists x_0 \forall y_0 \ldots \exists x_n \forall y_n \varphi, \forall x_0 \exists y_0 \ldots \forall x_n \exists y_n \neg \varphi$$

Notice that what counts for the application of the induction hypothesis is the *number* of alternations of quantifiers in the prefix of a formula that starts with a negation, which is followed by a list of quantifiers where an existential comes prior to a universal one, and ends up with the formula φ. We are assuming that the lemma we are proving holds whenever this number is n at most to prove that holds for formulas which have $n+1$ alternations of quantifiers too.

Let us suppose then that

$$\models_{\mathbb{N}}^{G} \neg \exists x_0 \forall y_0 \ldots \exists x_n \forall y_n \varphi$$

holds. Owing to definition 4.19, this is the same as saying that

$$\not\models_{\mathbb{N}}^{G} \exists x_0 \forall y_0 \ldots \exists x_n \forall y_n \varphi$$

is the case, that is (again by definition 4.19 and observation (iv) after that), it is not true that there exists $p \in \mathbb{N}$ such that for every $q \in \mathbb{N}$

$$\models^G_{\mathbb{N}} \exists x_1 \forall y_1 \ldots \exists x_n \forall y_n \varphi[x_0/\overline{p}, y_0/\overline{q}]$$

holds. However, this is equivalent to say that for every $p \in \mathbb{N}$, there exists $q \in \mathbb{N}$ such that

$$\not\models^G_{\mathbb{N}} \exists x_1 \forall y_1 \ldots \exists x_n \forall y_n \varphi[x_0/\overline{p}, y_0/\overline{q}]$$

holds, which corresponds by definition 4.19 to saying that

$$\models^G_{\mathbb{N}} \neg\exists x_1 \forall y_1 \ldots \exists x_n \forall y_n \varphi[x_0/\overline{p}, y_0/\overline{q}]$$

is the case for every $p \in \mathbb{N}$ and for some $q \in \mathbb{N}$. The formula just displayed however has n alternations of quantifiers (which are counted starting from 1 to n), where an existential quantifier always comes before a universal one. That is, has all features that are needed for applying the induction hypothesis to it, from which it follows that the previous situation holds if and only if

$$\models^G_{\mathbb{N}} \forall x_1 \exists y_1 \ldots \forall x_n \exists y_n \neg\varphi[x_0/\overline{p}, y_0/\overline{q}]$$

for every $p \in \mathbb{N}$ and for some $m \in \mathbb{N}$. Finally, by applying the relevant clauses of definition 4.19 (see the observation (iii) after that), we have that this holds true if and only if

$$\models^G_{\mathbb{N}} \forall x_0 \exists y_0 \ldots \forall x_n \exists y_n \neg\varphi$$

We conclude that the above lemma holds for every $n \leq k$. QED

As announced, this lemma leads us to the sought-for result:

Theorem 4.1 *Every finite game is determined, i.e. for every given finite game $G = \langle k, P, A\rangle$, either Player I has a winning strategy, or Player II has a winning strategy instead.*

Proof: let a finite game $G = \langle k, P, A\rangle$ be given. First, we briefly review the proof strategy, and explain it. We start by proving that (1) Player I has a winning strategy σ_G for G if and only if

$$\models^G_{\mathbb{N}} \exists x_0 \forall y_0 \ldots \exists x_{k-1} \forall y_{k-1}((x_0, y_0, \ldots, x_{k-1}, y_{k-1}) \in \overline{A})$$

is the case. From this, definition 4.19 and lemma 4.2, it follows that if that is not the case and Player I has not a winning strategy, then

$$\models^G_{N} \forall x_0 \exists y_0 \ldots \forall_{k-1} \exists y_{k-1} \neg((x_0, y_0, \ldots, x_{k-1}, y_{k-1}) \in \overline{A})$$

holds instead. By proving that (2) the latter is the case if and only if Player II has a winning strategy for G, we get the theorem.

For the sake of the proof, let us first assume the convention to use φ_A in the following as a shorthand for $(x_0, y_0, \ldots, x_{k-1}, y_{k-1}) \in \overline{A})$. It should be then clear that

$$FV(\varphi_A) = \{x_0, y_0, \ldots, x_{k-1}, y_{k-1}\}$$

Having stated that, the proof of part (1) comes from observing that Player I has a winning strategy σ_G in G if and only if $\sigma_G \bowtie s \in A$ for every $s \in \omega^k$. It turns out by exploiting this fact that we can conclude, at the level of the language $\mathscr{L}_S^-$ we have been constructing, that

$$\varphi_A[x_0/\overline{\sigma_G(())}, y_0/\overline{(s)_0}, \ldots, x_{k-1}/\overline{\sigma_G(\sigma_G(()), \ldots, (s)_{k-1})}, y_{k-1}/\overline{(s)_k}]$$

is $\mathbb{N}$-valid relative to G for every $s \in \omega^k$ as a consequence. Hence

$$\varphi_A[x_0/\overline{\sigma_G(())}, y_0/\overline{(s)_0}, \ldots, x_{k-1}/\overline{\sigma_G(\sigma_G(()), \ldots, (s)_{k-1})}, y_{k-1}/\overline{n}]$$

for every $n \in \mathbb{N}$, in particular for all those $n \in \mathbb{N}$ such that $n \neq (s)_k$, since this amounts to say that $\sigma_G \bowtie s' \in A$, where $s' \in \omega^k$ is such that $(s)_i = (s')_i$ for every $i < k-1$ but $(s')_{k-1} = n \neq (s)_{k-1}$.

Therefore, it follows from definition 4.19 that

$$\forall y_{k-1}\varphi_A[x_0/\overline{\sigma_G(())}, y_0/\overline{(s)_0}, \ldots, x_{k-1}/\overline{\sigma_G(\sigma_G(()), \ldots, (s)_{k-1})}]$$

is $\mathbb{N}$-valid relative to G (where, if we indicate by ψ the displayed formula for the sake of brevity, then $FV(\psi) = \{x_0, y_0, \ldots, x_{k-1}\}$ and $y_{k-1} \notin FV(\psi)$), which, owing to definition 4.19, is the same as saying that

$$\exists x_{k-1}\forall y_{k-1}\varphi_A[x_1/\overline{\sigma_G(())}, x_2/\overline{(s)_0}, \ldots, x_{k-2}/\overline{(s)_{k-1}}]$$

is $\mathbb{N}$-valid relative to G too. By keep using this reasoning, which allows to substitute elements of the sequence of the form $(s)_i$ by means of any natural number n, and substitute elements of the form $\sigma_G(\ldots(s)_i\ldots)$ by $\sigma_G(\ldots n\ldots)$ accordingly, it follows from definition 4.19 that if the previous condition is satisfied then

$$\models_{\mathbb{N}}^{G} \exists x_0\forall y_0 \ldots \exists x_{k-1}\forall y_{k-1}\varphi_A$$

holds as wanted.

Viceversa, assume that

$$\models_{\mathbb{N}}^{G} \exists x_0\forall y_0 \ldots \exists x_{k-1}\forall y_{k-1}\varphi_A$$

is the case. definition 4.19 entails then that

$$\models_{\mathbb{N}}^{G} \forall y_0 \ldots \exists x_{k-1}\forall y_{k-1}\varphi_A[x_0/\overline{n_0}]$$

holds for some $n_0 \in \mathbb{N}$. This means that, if we put

$$\psi_0 = \forall y_0 \exists x_1 \forall y_1, \ldots, \exists x_{k-1} \forall y_{k-1} \varphi_A$$

(and notice that $FV(\psi_0) = \{x_0\}$), then $\psi_0[x_0/\overline{n_0}]$ is $\mathbb{N}$-valid relative to G. Let then

$$n^* = \min\{z \in \mathbb{N} : \models^G_{\mathbb{N}} \psi_0[x_0/\overline{z}]\}$$

that is, n^* is an element of $\mathbb{N}$ for which the following holds:

- $\psi_0[x_0/\overline{n^*}]$ is $\mathbb{N}$-valid relative to G;
- n^* is the minimum element of $\mathbb{N}$ for which this is the case, i.e. if $n' \in \mathbb{N}$ is such that $\models^G_{\mathbb{N}} \psi_0[x_0/\overline{n'}]$ also holds, then $n^* < n'$.

Now, n^* exists because of the properties of the $<$ relation over $\mathbb{N}$. In particular, it follows from the fact that every non-empty subset of $\mathbb{N}$ has a minimum element with respect to $<$. By exploiting this very same idea for every variable x_i in the prefix of the above formula we come up with the following definition of σ_G.

First of all let, for every $0 \leq i \leq (k-1)$, ψ_i indicate the formula

$$\forall y_i \exists x_{i+1} \ldots \exists x_{k-1} \forall y_{k-1} \varphi_A$$

Notice that $FV(\psi_i) = \{x_0, y_0 \ldots, x_i\}$.

Then, we inductively define a function σ_G for every sequence s such that

$$s = ((s)_0, \ldots, (s)_{2m}) \in E^{<2k}$$

with $0 \leq m < k$, as follows:

$$\begin{cases} \sigma_G(()) = n^* \\ \sigma_G(s) = \min\{z \in \mathbb{N} : \models^G_{\mathbb{N}} \psi_{|s|-1}[x_0/\overline{\sigma_G(())}, y_0/\overline{(s)_0}, \ldots, x_i/\overline{z}]\} \end{cases}$$

The claim is that σ_G is a winning strategy for Player I in G.

First thing to check is that $\sigma_G : E^{<2k} \to \omega$, hence that σ_G fits definition 4.1. This amounts to checking, in particular, that for every $s \in E^{<2k}$, $\sigma_G(s) \in \omega$, that is σ_G is everywhere defined over $E^{<2k}$. However, granted the definition we have just given, we have:

$$\sigma_G(s) = \min\{z \in \mathbb{N} : \models^G_{\mathbb{N}} \psi_{|s|-1}[x_0/\overline{\sigma_G(())}, \ldots, x_i/\overline{z}]\} \in \omega$$

Then, we have to prove in addition that σ_G is right-hand unique. This follows again by the definition of it we have just given, and $<$ being a strict order of $\mathbb{N}$.

It remains to prove that σ_G is winning for Player I (see definition 4.4). Let then $s \in \omega^k$. It follows by the definition of σ_G above and definition 4.2 that:

$$\sigma_G \bowtie s = (n^*, (s)_0, \sigma_G(n^*, (s)_0), \ldots, \sigma_G(n^*, (s)_0, \ldots, (s)_{k-1}), (s)_k)$$

where, for every $i \leq k$

$$\sigma_G(n^*, (s)_0, \ldots, (s)_{i-1}) = \min\{z \in \mathbb{N} : \models^G_{\mathbb{N}} \psi_{i-1}[x_0/\overline{\sigma_G(())}, \ldots, x_i/\overline{z}]\}$$

This yields that

$$\models^G_{\mathbb{N}} \psi_{k-1}[x_0/\overline{\sigma_G(())}, \ldots, x_{k-1}/\overline{z}]$$

is the case, that is

$$\models^G_N \forall y_{k-1}\varphi_A[x_0/\overline{\sigma_G(())}, y_0/\overline{(s)_0}, \ldots, x_{k-1}/\overline{\sigma_G(\sigma_G(()), \ldots, (s)_{k-2})}]$$

Definition 4.19 entails that this holds true if and only if

$$\models^G_N \varphi_A[x_0/\overline{\sigma_G(())}, y_0/\overline{(s)_0}, \ldots, x_{k-1}/\overline{\sigma_G(\sigma_G(()), \ldots, (s)_{k-2})}, y_{k-1}/\overline{n}]$$

for every $n \in \mathbb{N}$, from which it follows that

$$\models^G_N \varphi_A[x_0/\overline{\sigma_G(())}, y_0/\overline{(s)_0}, \ldots, x_{k-1}/\overline{\sigma_G(\sigma_G(()), \ldots, (s)_{k-2})}, y_{k-1}/\overline{(s)_{k-1}}]$$

holds in particular. By definition 4.19 again, this is the case if and only if

$$(\sigma_G(\,), (s)_0, \ldots, \sigma_G(\sigma_G(()), \ldots, (s)_{k-2}), (s)_{k-1}) \in A$$

Hence, σ_G is winning for Player I in G.

Having proved that step (1) in the sketch of the proof we started from holds, it remains to prove step (2), as it was said, to get the theorem. In turn, this amounts to proving that Player II has a winning strategy in G if and only if

$$\models^G_N \forall x_0 \exists y_0 \ldots \forall_{k-1} \exists y_{k-1} \neg((x_0, y_0, \ldots, x_{k-1}, y_{k-1}) \in \overline{A})$$

which, owing to the notation we are using, is the same as

$$\models^G_N \forall x_0 \exists y_0 \ldots \forall_{k-1} \exists y_{k-1} \neg \varphi_A$$

The direction from left to right of this proof goes along lines that are similar to the corresponding case of (1) above. Let us suppose then that Player II has a winning strategy τ_G in G. That is,

$$s \bowtie \tau_G = ((s)_0, \tau_G((s)_0), \ldots, \tau_G((s)_0, \ldots, (s)_{k-1})) \notin A$$

for every $s \in \omega^k$. If we let s to be an arbitrary but fixed element of ω^k, to say that $s \bowtie \tau_G \notin A$ is, by definition 4.19, the same as saying that

$$\neg\varphi_A[x_0/\overline{(s)_0}, y_0/\overline{\tau_G((s)_0)}, \ldots, x_{k-1}/\overline{(s)_{k-1}}, y_{k-1}/\overline{\tau_G((s)_0, \ldots, (s)_{k-1})}]$$

is $\mathbb{N}$-valid relative to G. So, in particular, it follows by applying the relevant clause of definition 4.19 that

$$\exists y_{k-1} \neg\varphi_A[x_0/\overline{(s)_0}, y_0/\overline{\tau_G((s)_0)}, \ldots, x_{k-1}/\overline{(s)_{k-1}}]$$

is $\mathbb{N}$-valid relative to G. Moreover, since s is arbitrary and τ_G is winning for Player II in G,

$$\exists y_{k-1}\neg\varphi_A[x_0/\overline{(s)_0}, y_0/\overline{\tau_G((s)_0)}, \ldots, x_{k-1}/\overline{n}]$$

is $\mathbb{N}$-valid relative to G for every $n \in \mathbb{N}$, and

$$\forall x_{k-1}\exists y_{k-1}\neg\varphi_A[x_0/\overline{(s)_0}, y_0/\overline{\tau_G((s)_0)}, \ldots, y_{k-2}/\overline{\tau_G((s)_0, \ldots, (s)_{k-3})}]$$

is $\mathbb{N}$-valid relative to G as a consequence. By keep using this reasoning, it should be clear that if τ_G is a winning strategy for Player II in G, then

$$\models^G_N \forall x_0 \exists y_0 \ldots \forall_{k-1} \exists y_{k-1} \neg\varphi_A$$

as wanted.

Conversely, let us suppose that the latter situation occurs. Hence, assume that

$$\models^G_N \forall x_0 \exists y_0 \ldots \forall_{k-1} \exists y_{k-1} \neg\varphi_A$$

holds. First of all, notice that, for every formula θ of $\mathscr{L}^-_S$ such that $FV(\theta) = \{x_i, x_j\}$, we have that

$$\models^G_{\mathbb{N}} \forall x_i \exists x_j \theta$$

holds if and only if there exists $f : \omega \to \omega$ such that

$$\models^G_{\mathbb{N}} \theta[x_i/\overline{n}, x_j/\overline{f(n)}]$$

is the case for every $n \in \mathbb{N}$. The direction from right to left of such a statement is clear (just observe that if such f exists, one can apply the relevant clauses of definitions 4.19 to conclude that $\forall x_i \exists x_j \theta$ is $\mathbb{N}$-valid relative to G). Viceversa, if

$$\models^G_{\mathbb{N}} \forall x_i \exists x_j \theta$$

is the case, then it follows from definition 4.19 that, for every $n \in \mathbb{N}$

$$M^n_\theta = \{m \in \mathbb{N} : \models^G_{\mathbb{N}} \theta[x_i/\overline{n}, x_j/\overline{m}]\} \neq \emptyset$$

For reasons we have already made use of above, it follows that this set M^n_θ has a minimum element with respect to $<$. Let m^n_θ indicate such element. This allows to set up the definition of a function $f : \omega \to \omega$, to be such that $f(n) = m^n_\theta$ for every $n \in \mathbb{N}$. This is a legitimate definition of a function: f is defined over $\mathbb{N}$ as a whole and yields a value that belongs to the same set; in addition, f is right-hand unique as functions are supposed to be. Finally, one observes that, by the definition of f,

$$\models^G_{\mathbb{N}} \theta[x_i/\overline{n}, x_j/\overline{f(n)}]$$

holds for every $n \in \mathbb{N}$. Now, granted that, and assuming that $\theta \equiv \neg\psi$ for some formula ψ of $\mathscr{L}_S^-$, it follows from what we have just observed and definitions 4.19 that if

$$\models_N^G \forall x_0 \exists y_0 \ldots \forall x_{p-1} \exists y_{p-1} \neg\psi$$

is the case, then there exist functions $f_0, \ldots, f_{p-1}$ with $f_i : \omega \to \omega$ for every $0 \leq i < p$, such that

$$\models_N^G \neg\psi[x_0/\overline{n_0}, y_0/\overline{f(n_0)}, \ldots, x_{p-1}/\overline{n_{p-1}}, y_{p-1}/\overline{f_{p-1}(n_{p-1})}]$$

holds for every $n_0, \ldots, n_{p-1} \in \mathbb{N}$ (p being an arbitrary element of $\mathbb{N}$ throughout the statement). It should be noticed that each of these functions f_i depend upon the n_i's (since each element of $\mathbb{N}$ has 'its own' function that does what the f_i's turn out to do)[9]. Therefore, the sequence of functions $f_0, \ldots, f_{p-1}$ depends upon the sequence $s = (n_0, \ldots, n_{p-1})$. Let us make this dependence evident and let us put, for every sequence $s \in \omega^{\leq 2k}$, $f_0^s, \ldots, f_{|s|-1}^s$ be the functions which exist by the previous observation in correspondence to the elements $(s)_0, \ldots, (s)_{|s|-1}$ of s.

Let us now go back to the initial hypothesis. Then, owing to what was just observed, we have that

$$\models_N^G \forall x_0 \exists y_0 \ldots \forall x_{k-1} \exists y_{k-1} \neg\varphi_A$$

is the case if and only if

$$\models_N^G \neg\varphi_A[x_0/\overline{(s)_0}, y_0/\overline{f_0^s((s)_0)}, \ldots, x_{k-1}/\overline{(s)_{k-1}}, y_{k-1}/\overline{f_{k-1}^s((s)_{k-1})}]$$

holds for every $s \in \omega^k$. Let, for every $s \in O^{<2k}$, τ_G be the function defined by $\tau_G(s) = f_{|s|-1}^s((s)_{|s|-1})$. Again, this is a legitimate definition of function owing to what we have said. Moreover, for every $s \in \omega^k$,

$$\models_{\mathbb{N}}^G \neg\varphi_A[x_0/\overline{(s)_0}, y_0/\overline{\tau_G((s)_0)}, \ldots, x_{k-1}/\overline{(s)_{k-1}}, y_{k-1}/\overline{\tau_G((s)_0, \ldots, (s)_{k-1})}]$$

is the case, by definition of τ_G itself. By definition 4.19, it follows that:

$$s \bowtie \tau_G = ((s)_0, \tau_G((s)_0), \ldots, (s)_{k-1}, \tau_G((s)_0, \ldots, (s)_{k-1})) \notin A$$

as needed. QED

Theorem 4.1 and proposition 4.1 together yields the following immediate corollary:

[9]Actually, these functions also depend upon the formula ψ we are considering. For the sake of readability, we leave this dependence implicit since, despite the generality of the observation, we are interested in a specific instance of it, namely the one that applies to ψ being the formula $(x_0, y_0, \ldots, x_{k-1}, y_{k-1}) \in \overline{A}$ above. However, it should be clear that this dependence should be emphasized also.

Corollary 4.1 *For every finite game $G = \langle k, P, A \rangle$, either Player I has a winning strategy, or Player II has a winning strategy, but not both Player I and Player II have a winning strategy.*

4.9. GRH at work

As the reader may recall, we have started the whole investigation on a mathematical model of finite games driven by the hypothesis that for every real-life situations one could find a finite game of that sort such that to solve the choice problem in the former would correspond to solve it in the latter (GRH we decided to call the hypothesis in question, back in section 1.5). Now, the previous result about every finite game being determined can be viewed as stating that the problem of (strategic) choice can always be solved in such a setting.

As long as this attempted interpretation of the result we have given is concerned, two things should always be taken into account: (i) the strength of the result we have proved on the one hand, which amounts to yield the existence of a strategy for either of the two players in case of every game, that is the existence of a uniform method for winning every match (which is not the same as, given any match of a game, finding the way to win it, and then find another way to win another one, and so on for all matches of it); (ii) that a substantial part of what would be needed to turn the theoretical possibility it suggests into actuality is missing, for, the result gives no clue on whether it is possible for either of the two players to really devise the strategy we proved it exists (i.e., to *know* the strategy that exists by the theorem).

Anyway, the result as it is proved does tell us something about real-life games. In particular, it can be used to finally give our preliminary analysis of games with perfect information via trees in section 4.2 the form of a mathematical theorem. To make the connection with those remarks apparent, let us speak of the same game we were dealing with there. The theorem we have in mind then, reads as follows:

Theorem 4.2 *In chess, either White has a winning strategy, or Black has a winning strategy, or both White and Black have a strategy for drawing.*

Now, this theorem is obtained as a simple corollary of theorem 4.1.

First of all notice that with chess we have a new alternative situation to the one we have retained as the canonical one, for it comprises situations where draw is possible. If we wanted to model this in a similar fashion to the modifications of the canonical model we considered in section 4.7, we would have to consider for instance a game G^D as a quadruple $\langle k, P, A_1, A_2 \rangle$, where $k \in \mathbb{N}$, $P = \omega^{2k}$, $A_1, A_2 \subseteq P$ with $A_1 \cap A_2 = \emptyset$ (the latter conditions ensuring that G^D is zero-sum since there is no match in P

that is both won by Player I and by Player II). Then, A_1 would work as the set of winning conditions for Player I (therefore, a strategy σ for Player I would be winning if, for every $s \in \omega^k$, $\sigma \bowtie s \in A_1$), A_2 would be the set of winning conditions for Player II (therefore, a strategy τ for Player II would be winning if, for every $s \in \omega^k$, $s \bowtie \tau \in A_2$), while $D = P \setminus (A_1 \cup A_2)$, which contains those $s \in P$ for which both $s \notin A_1$ and $s \notin A_2$ are the cases, i.e., matches where a draw takes place.

The reason why this modification was not considered earlier, is that there is a very natural way for handling with that: define two games in canonical form out of G^D, namely G_1^D and G_2^D respectively, where $G_1^D = \langle k, P, A \rangle$ where $A = A_1 \cup D$, while $G_2^D = \langle k, P, A' \rangle$ where $A' = A_1$. If the usual notions of strategy and winning strategy for both players are retained, then G_1^D is the game where Player I has a winning strategy if and only if she has a strategy for either winning or drawing any match of G^D, and Player II has a winning strategy if and only if she has a strategy for winning every match in G^D. Game G_2^D, instead, is the game where Player I has a winning strategy if and only if she has a winning strategy in G^D, and Player II has a winning strategy if and only if she has a strategy for either winning or drawing any match of G^D. Having noticed that every match of G_1^D is a match of G_2^D (and also a match of G^D) and viceversa, the consequence of theorem 4.1 for game G^D are easily evaluated. Since both G_1^D and G_2^D are finite games in canonical form, theorem 4.1 applies and we get the following results out of it:

Corollary 4.2 *Either Player I has a winning strategy in G_1^D, or Player II has a winning in G_1^D.*

Corollary 4.3 *Either Player I has a winning strategy in G_2^D, or Player II has a winning in G_2^D.*

By considering the combinations of these two outcomes we get four alternatives:

1. Player I has a winning strategy in both G_1^D and G_2^D;
2. Player II has a winning strategy in both G_1^D and G_2^D;
3. Player I has a winning strategy in G_1^D and Player II has a winning strategy in G_2^D;
4. Player I has a winning strategy in G_2^D and Player II has a winning strategy in G_1^D.

In turn, these combinations tell us what happens to G^D in view of Corollaries 4.2 and 4.3 of theorem 4.1. For, suppose that alternative no. 1 takes place. Hence, Player I has a winning strategy in G^D (since she has a strategy for either winning or drawing in G_1^D, but since drawing a match in G_1^D counts as loosing the corresponding match in G_2^D and she is supposed to never loose a match in this game too, then she never draws a match in G_1^D which means that she always wins the corresponding match in G^D by following the said strategy). By a symmetrical argument, alternative no. 2 yields that Player II has a winning strategy in G^D.

Alternative no. 3, instead, has the consequence that both Player I and Player II have a strategy for drawing every match in G^D (for, assuming that Player I has a winning strategy in G_1^D then it is clear that this cannot be because she has a winning strategy in G^D, for otherwise Player II would not have a winning strategy in G_2^D as she has instead).

Finally, alternative no. 4 is easily seen as impossible: for, owing to the assumption that Player I has a winning strategy σ in G_2^D, it follows that she has a winning strategy in G^D, which contradicts the assumption that Player II has a winning strategy τ in G_1^D, since this implies that she is supposed to have a winning strategy in G^D as well (from which it follows that $\sigma \bowtie \tau \in A_1 \cap A_2 = \emptyset$; contradiction).

So, from the two corollaries above, we get theorem 4.2 with Player I being White and Player II being Black instead. As a consequence of it, we finally have a confirmation that there was no mistake in our analysis of finite games with perfect information that we performed in section 4.2. In particular, we now know that there is no difference between a simple game such as tic-tac-toe and a complex one like chess. They are both determined, hence there is a way for one of the two player to devise a method for winning every match of it. When this conclusion appeared on our horizon then, we panicked. We simply could not figure out how this was possible. However, we now know more about this result to make sure that we should not panic owing to it in the end. In particular, we know more about what it means for a winning strategy to exist, and we know very well the difference between knowing that a winning strategy 'is there' and actually have it 'grasped'. Some of the results we have gone through, as a matter of fact, required us to actually devise a winning strategy and that was no easy task at all. Should then the determinacy property change our view of finite games? Not really, as a matter of fact. When we start playing a match of a perfect information, zero-sum game we know a few things. Among the things we know, there is the simple fact that we loose the match we have started if our opponent wins it. This could be because she was clever enough to devise a better approach to that match. It could also be that our opponent was so clever that she devised the best possible approach, and grasped the winning strategy of the game. How can we make sure? In the end, there is only one way to say which of the two is the real case: playing another time and, if we loose again, keep playing.

How about a nice game of chess?

4.10. Bibliographical note

By leaving trees behind, and by dealing with games where players alternate moves with one another by means of the model based on sequences of numbers, we wanted to connect ourselves to the way games are dealt with in ordinary mathematical terms. This was not at all intended as being motivated by the attempt of leaving the study of games of this sort by means of trees behind. Rather, our decision was driven by wishing to provide the reader with tools that might encourage her to keep deepening her acquaintance with the ordinary mathematical theory of games in its usual form. By the way, this was done by starting from a classical way for dealing with trees representing games. The argument presented in section 2 to suggest what we have proved only in section 8, namely that all finite games with no tie is determined, is based on a well-known analysis of games as trees known as *backward induction*, the earliest notable application of it is in the theorem about determinacy of (real) chess we reported on here as theorem 4.2 and which is commonly attributed to the mathematician Ernst Zermelo. Zermelo's theorem is contained in a little note from the early 1910's (Zermelo 1913), which has became a classical reference on the topic although it is unclear what one should acknowledge Zermelo's proof to achieve. The story of this note and of Zermelo's contribution is re-considered by U. Schwalbe and P. Walker in the attempt of assessing the German mathematician's role in the early history of game theory (Schwalbe and Walker 2001).

What should be stressed with clarity here is that the portion of topics covered by this volume does not even count as a small scratch on the surface of the enormous area of study that the theory of games (in both normal and extensive form) appears to be nowadays. To get a quick idea of this, it is enough to have a look at the 4-volume *Handbook of Game Theory* (Aumann and Hart 1992, 1994, 2002; and Young and Zamir 2014), which covers both the basic concepts of the theory (mostly dealt with in vol. 1), as well as a wide range of applications of it. This is obviously a valuable source to start from, for anyone interested in a thorough study of the topic.

As far as the modest aim that we were pursuing here was concerned, we based ourselves as reference for most of the topic we wished to cover on the initial part of the notes on infinite games by Yurii Khomskii (Khomskii 2010) made available by Yurii Khomskii (which, on our latest search on the web – September 2018 – were still up for free download). Most of what we have been doing here in sections 4.6, 4.7 and 4.8 can be viewed as filling the missing details in the very first part of those notes which, as their title suggests, are more focused on the extension of the ground theory to games which are infinite. Owing to the compatibility in style and notation, that could be a reasonable source to look at for expanding knowledge attained at by reading this volume.

Having made clear how little of the current theory of games we have covered by means of the material presented here, it really would be hard to fill in what is missing through bibliographical references. A couple of remarks in this direction we found that could be useful anyway, the first one of which concerns the study of games in extensive form from a philosophical perspective.

Like with finite games in general, the theory of games as trees has attracted a lot of attention by scholars working in fields directly connected with topics stemming from agents interactions. This is no surprise since the very idea of rational agents making decisions which influence each others has a lot in common with classical topics from a variety of philosophically related areas of research. This connection turns out quite clearly from sources attempting at surveying the most important work which has been done in the field. To give the reader some concrete indication in this respect, Johann van Benthem's numerous books on the topic contain a huge amount of information, both from the point of view of the issues involved and, most importantly, of the literature they survey (see, for instance, van Benthem 2011, 2014) – even though also Ross 2014 we already referred to, might also be a good starting point).

Another direction of work that departs from what we have been accounting for here has already been mentioned, or better hinted at from time to time (for instance, in section 4.4), when we spoke of extension of the above theory to games which are infinite. Infinite games in this sense take the form of infinite (ω-long) sequence of natural numbers. Games in this form have been studied extensively, since the investigation on this topic proved to be tied up with some remarkable directions of research with high mathematical importance. This holds true in particular for the main property we have been dealing with here with respect to finite games, namely determinacy. As a matter of fact, the property of an infinite game to be determined has early been seen to be connected with relevant topological properties of sets. Several results have been provided in this respect, starting from a theorem in by D. Gale and F.M. Stewart (Gale and Stewart 1953) about infinite games with perfect information being determined provided some extra features of the set of winning condition is met. This result was then extended once by Philip Wolfe (Wolfe 1955), and then by Donald A. Martin (Martin 1975), whose main theorem applies to a wider class of infinite games (though, not the whole class of it).

The connection of these studies with topological concepts that are relevant for the branch of set theory known as descriptive set theory is not the only one, as it turned out that the study of infinite games is also interesting for its connections with the investigation on set theory itself. As a matter of fact, it turns out that Martin's result is the best that someone can expect to prove by means of the ordinary set-theoretical axioms. The statement according to which «All infinite games are determined», with no extra condition on the class of games it refers to, is indeed disprovable by the axioms of set theory if these are taken to include the axiom of choice. This has not prevented scholars from investigating it. Quite on the contrary, the fact that the statement in question, which is known in the literature as the «axiom of determinacy», is instead consistent with the axioms of the theory minus the axiom of choice, has made it attract a lot of attention due to the strong deductive power that is associated with it (more information about this direction of work can be found in calssical textbooks on sets like Jech 2003).

It would be pointless, however, to keep the reader busy over considerations of this sort owing to the distance, in terms of the mathematical expertise they require, between the level at which these kind of investigations locate themselves and the level at which we have (hopefully) helped her to rise by means of the reading of this volume. These remarks were put here with the only goal of trying to tease the reader why some more information that may further stimulate her interest in the topic, and hopefully give her some reasons to keep digging the matter she has started to study with this book.

Bibliography

Aumann R.J., Hart S. (editors) (1992) *Handbook of Game Theory with Economic Applications, volumes 1-3*. Amsterdam: North Holland.

Belnap N. (1982) Gupta's rule of revision theory of truth. *Journal of Philosophical Logic* 11, 103-116.

Bruni R., Sillari G. (2018) A rational way of playing: revision theory for strategic interaction. *Journal of Philosophical Logic* 47 (3), 419-448.

Buchholz W. *et al.* (1981) *Iterated inductive definitions and subsystems of analysis: recent proof-theoretical studies*. Berlin: Springer.

Cantini A., Bruni R. (2017) Paradoxes and contemporary logic. *The Stanford encyclopedia of philosophy* [online]. <https://plato.stanford.edu/entries/paradoxes-contemporary-logic/> [Last seen 12/2018].

Chapuis A. (2000) Rationality and circularity. In: Chapuis A., Gupta A. (editors) *Circularity, Definition, and Truth*. New Delhi: Indian Council of Philosohical Research, 49-77.

— (2003) An application of circular definitions: rational decision. In: Räsch T. *et al.* (editors) *Foundations of the formal sciences II: applications of mathematical logic in philosophy and linguistics*. Dordercht: Kluwer, 47-54.

Feferman S. (2010) The proof theory of classical and constructive inductive definitions. A 40 year saga, 1968-2008. In: Schindler R. (editor) *Ways of proof theory*. Hausenstamm: Ontos Verlag, 7-30.

Gale D., Stewart F.M. (1953) Infinite games with perfect information. In: Kuhn H.W., Tucker A.W. (editors) *Contributions to the Theory of Games II*. Princeton: Princeton University Press, 245-266.

Gupta G. (1982) Truth and paradox. *Journal of Philosophical Logic* 11, 1-60.

— (2011) On circular concepts. In: Id., *Truth, Meaning and Experience*. Oxford: Oxford University Press, 95-134.

Gupta A., Belnap N. (1993) *The Revision Theory of Truth*. Cambridge:

Riccardo Bruni, *Course notes on finite games and rational choice*, ISBN 978-88-6453-832-7 (print) ISBN 978-88-6453-831-0 (online), CC BY 4.0, 2018 Firenze University Press

MIT Press.
Herzberger H.G. (1982a) Naive semantics and the liar paradox. *Journal of Philosophy* 79, 479-497.
— (1982b) Notes on naive semantics. *Journal of Philosophical Logic* 11, 61-102.
Jech T. (2003) *Set Theory. The Third Millennium Edition Revised and Expanded*. Berlin: Springer.
Karlin A.R., Peres Y. (2017) *Game theory, alive*. Providence: American Mathematical Society.
Khomskii Y. (2010) *Infinite games* [online]. <https://www.math.uniha mburg.de/home/khomskii/infinitegames2010/InfiniteJuly> [Last seen 09/2018]
Knaster B., Tarski A. (1927) Un théorème sur les fonctions d'ensembles. *Annales de la société polonaise de mathématique* 6, 133-134.
Kripke S. (1975) Outline of a theory of truth. *Journal of Philosophy* 72, 690-716.
Martin D.A. (1975) Borel determinacy. *Annals of Mathematics* 102 (2), 363-371.
Myerson R.B. (1978) Refinements of the Nash equilibrium concept. *International Journal of Game Theory* 7 (2), 73-80.
Nash J.F. (1950a) The bargaining problem. *Econometrica* 18 (2), 155-162.
— (1950b) Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences* 36 (1), 48-49.
— (1951) Non-cooperative games. *Annals of Mathematics* 54 (2), 286-295.
Osborne M.J., Rubinstein A (1994) *A Course in Game Theory*. Cambridge: MIT Press.
Ross D. (2018) Game theory. *The Stanford encyclopedia of philosophy* [online]. <https://plato.stanford.edu/entries/game-theory/> [Last seen 12/2018].
Schwalbe U., Walker P. (2001) Zermelo and the early history of game theory. *Games and Economic Behavior* 34 (1), 123-137.
Selten R. (1975) Re-examination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory* 4 (1), 25-55.
Tarski A. (1983) The concept of truth in formalized languages. In: Corcoran J. (editor) *Logic, Semantics, Metamathematics*. Cambridge: Hackett.
van Benthem J. (2011) *Logical Dynamics of Information and Interaction*. Cambridge: Cambridge University Press.
— (2014) *Logic and Games*. Cambridge: MIT Press.
von Neumann J. (1928) Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen* 100 (1), 295-320.

von Neumann J., Morgenstern O. (1944) *Theory of Games and Economic Behaviour*. Princeton: Princeton University Press.
Wolfe P. (1955) The strict determinateness of certain infinite game. *Pacific Journal of Mathematics* 5 (5), 841-847.
Young P., Zamir S. (editors) (2014) *Handbook of Game Theory, volume 4*. Amsterdam: North Holland.
Zermelo E. (1913) Über eine Anwendung der Mengenlehre auf die Theorie des Schachpiels. In: Hobson E.W., Love A.E.H. (editors) *Proceedings of the Fifth International Congress Mathematics (Cambridge 1912), volume 2*. Cambridge: Cambridge University Press, 501-504.

MANUALI

Biomedica

Branchi R., *Le impronte nel paziente totalmente edentulo*

Branchi R., *Riabilitazione protesica del paziente oncologico testa-collo*

Martinelli E., *Rieducazione del piede. Problematiche biomeccaniche e posturali statiche e dinamiche*

Martinelli E., *Sviluppo del dolore rachideo in gravidanza. Mutamenti della biomeccanica rachidea, problematiche posturali, prevenzione e attività fisica adattata pre e post parto*

Martinelli E., *Rieducazione posturale. Fondamenti per la progettazione della postura*

Martinelli E., *Prevenzione del mal di schiena di origine meccanica con attività motoria e comportamentale. Approfondimenti di patomeccanica e biomeccanica rachidea*

Rossetti R., *Manuale di batteriologia clinica. Dalla teoria alla pratica in laboratorio*

Rucci L., *Testo Atlante di embriologia clinica della Laringe. La chirurgia conservativa compartimentale della regione glottica*

Rucci L., *Clinical Embryology of Human Larynx for Conservative Compartmental Surgery. A Text and Atlas*

Scienze

Bart J.C.J., *Polymer Additive Analytics. Industrial Practice and Case Studies*

Caramelli D., *Antropologia molecolare. Manuale di base*

Lo Nostro P., Peruzzi N., *Spontaneamente. Esercizi Risolti di Chimica Fisica*

Ricci R., *Lezioni di Sistemi Dinamici*, a cura di Lorenzo Barletti, Angiolo Farina, Lorenzo Fusi, Federico Talamucci

Scialpi A., Mengoni A. (a cura di), *La PCR e le sue varianti. Quaderno di laboratorio*

Simonetta M.A., *Short history of Biology from the Origins to the 20th Century*

Spinicci R., *Elementi di chimica* (nuova edizione)

Scienze Sociali

Ciampi F., *Fondamenti di economia e gestione delle imprese*

Giovannini P. (a cura di), *Teorie sociologiche alla prova*

Maggino F., *L'analisi dei dati nell'indagine statistica. Volume 1. La realizzazione dell'indagine e l'analisi preliminare dei dati*

Maggino F., *L'analisi dei dati nell'indagine statistica. Volume 2. L'esplorazione dei dati e la validazione dei risultati*

Marchettoni L., *Breve storia della democrazia.*

Da Atene al populismo

Magliulo A., *Elementi di economia del turismo*

Petretto A., *Economics and Institutions and Law*

Visentini L., Bertoldi, M., *Conoscere le organizzazioni. Una guida alle prospettive analitiche e alle pratiche gestionali*

Scienze Tecnologiche

Borri C., Pastò S., *Lezioni di ingegneria del vento*

Borri C., Betti M., Marino E., *Lectures on Solid Mechanics*

Cidronali A., Paolo Colantonio P. e Lucci L., *Antenne Integrate Attive*

Gulli R., *Struttura e costruzione / Structure and Construction*

Maddio S., *Introduzione ai Sistemi di Localizzazione Indoor*

Policicchio F., *Lineamenti di infrastrutture ferroviarie*

Umanistica

Bertini F., *Risorse, conflitti, continenti e nazioni. Dalla rivoluzione industriale alle guerre irachene, dal Risorgimento alla conferma della Costituzione repubblicana*

Bombardieri L., Graziadio G., Jasink A.M., *Preistoria e Protostoria egea e cipriota*

Bombi A.S., Pinto G., Cannoni E., *Pictorial Assessment of Interpersonal Relationships (PAIR). An analytic system for understanding children's drawings*

Baldini S., Marini D., Vorrei. *Corso di lingua italiana di livello elementare 1. Libro di testo e libro degli esercizi*

Borello E., Mannori S., *Teoria e tecnica delle comunicazioni di massa*

Brandi L., Salvadori B., *Dal suono alla parola. Percezione e produzione del linguaggio tra neurolinguistica e psicolinguistica*

Bruni R., *Course notes on finite games and rational choice*

Coniglione F., Lenoci M., Mari G., Polizzi G. (a cura di), *Manuale di base di Storia della filosofia*

Marcialis N., *Introduzione alla lingua paleoslava*

Michelazzo F., *Nuovi itinerari alla scoperta del greco antico. Le strutture fondamentali della lingua greca: fonetica, morfologia, sintassi, semantica, pragmatica*

Peruzzi A., *Il significato inesistente. Lezioni sulla semantica*

Peruzzi A., *Modelli della spiegazione scientifica*

Sandrini M.G., *Filosofia dei metodi induttivi e logica della ricerca*

Trisciuzzi L., Zappaterra T., Bichi L., *Tenersi per mano. Disabilità e formazione del sé nell'autobiografia*