Roberto Lucchetti

Games suggest how to define rational behavior

Surprising aspects of interactive decision theory

January 2, 2009

Springer

0.1 Introduction

Game theory is a part of mathematics that has been developed only in recent years, its first important results going back to the beginning of last century. Its primary goal is to investigate all situations where several agents have interaction in some processes and have interests connected with the outcome of their interaction. Typically, this is what happens in games and it is also a model for many situations of our everyday life. We play when we buy something or when we take/give an exam. We are also playing a game when interacting with our coauthors or our partner: actually, we are playing games all the time. It is clear that since the beginning of the history of human thought most of the focus has been on how people interact with each other; however, this has always been done more in philosophical, religious, ethical terms than by a scientific approach. As soon as mathematics started to deal with such type of problems, it became clear that its contribution was really important. Nowadays, for instance, it is standard, for psychologists, to use simple games to understand deep, unconscious people's reaction to particular situations.

Of course, we must be very cautious in relying on results provided by the theory. We are talking about human beings, and the human being cannot be made into a mathematical formula. Hence the very first assumption of the classical theory, *the agents are fully rational*, is clearly very ideal, no matter the meaning we shall give to the term "rational". Nevertheless, the results offered by the theory provide a useful and precise term of comparison to analyze how differently people may act, and how much so. Thus, as long as it is used with ingenuity, game theory is of great help in understanding and predicting agents' behavior.

Since its inception, the theory of games has proposed interesting, though counter-intuitive, results. In some sense it seems that the human brain reacts instinctively as if the body enveloping it acts in total loneliness. At least this is my opinion, and in this article I would like to stress some situations of this type.

0.2 Eliminating dominated strategies

The first assumption about rational agents is simple, and intends to exclude, whenever possible, some of the actions available to the players. It reads like this:

A player will not choose an action x, if it is available an action z allowing him to get more, no matter which choice the other players will make.

We shall call it the rule of *elimination of dominated strategies*: x is dominated by z, and thus it will not be used. Such a rule usually does not provide the outcome of the game. All interesting situations are when the players must adapt their choices to the expected behavior of the opponents. But it can be useful in deleting some choices. And in very simple cases deleting dominated strategies actually can provide the solution. Here is a first example.

Example 1 Consider the following game[1]:

$$\begin{pmatrix} (5,2) & (3,3) \\ (4,4) & (2,5) \end{pmatrix}.$$

The first row dominates the second one, since $5 > 4$ and $3 > 2$. Moreover, the second column dominates the first one. Thus the outcome of the game is first row/second column, providing a utility of 3 to both players.

[1] A *bimatrix*, a matrix with pairs as entries like that one in the example, is a game in the following sense: player one chooses a row, player two a column. The resulting entry contains two numbers, which are the utilities assigned to the players: the first is that for the first (row) player, the second that for the second (column) player.

Example 2 In the following example instead, the above rule does not help in selecting a reasonable outcome. Something different is needed.

$$\begin{pmatrix} (5,2) & (3,3) \\ (6,6) & (2,5) \end{pmatrix}.$$

There is a category of games which are (relatively) simple to analyze. When the two players have always opposite interests, things are easier to handle since there is no possibility to switch to situations in which both are better off (or worse off). Let us consider another example.

Example 3 The game is described by the following matrix[2]:

$$\begin{pmatrix} 4 & 3 & 1 \\ 7 & 5 & 8 \\ 8 & 2 & 0 \end{pmatrix}.$$

How to analyze this game? The first simple remark is that player One is able to guarantee herself at least 5, by playing the second row (against the possibility to get 1 and 0 from the first and third, respectively). The same does player Two, with a change of sign in his mind. Thus he realizes to be able to pay not more than 5, by playing the second column. Summarizing: the first one is able to guarantee herself *at least* 5, the second one can pay *not more than* 5: the result of this game cannot be different from the first's receiving 5 from the second.

What the row player is able to guarantee herself is called her *maxmin* value. For the second player, we use the term *minmax* value[3]. They are called the *conservative values* of the players. It is clear that if each player does select one strategy which offers the player's conservative value, they reach the satisfactory outcome of the game. More precisely, if we denote by a_{ij} the generic entry of the matrix, the outcome is provided by strategies $\bar{\imath}, \bar{\jmath}$ such that

$$a_{i\bar{\jmath}} \leq a_{\bar{\imath}\bar{\jmath}} \leq a_{\bar{\imath}j}.$$

The above formula highlights the fact that there is no incentive for each player to change strategy, if the player takes from granted that the opponent will do the same. Furthermore, taking for granted that the opponent will do the same is reasonable, since there is no incentive for the opponent as well to deviate from that.
Consider now the following example:

Example 4

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

This is a very simple game, and immediately we realize that the conservative values of the players do *not* agree. Actually each player cannot guarantee himself to avoid to loose (1 indicates the victory, -1 the loss). And this clearly indicates the fact that such a game cannot have a predictable outcome. Actually, it represents for instance the case when two players declare a number at the same time, and one wins if the sum of the two numbers is odd, while the other one if the sum is even. Observe also that the conservative value of the first player is strictly less than the conservative value of the second. This is not a feature of this example, but a general fact. Thus the non-predictability of the outcome of such a game is due to the fact that a positive quantity

[2] When the game is strictly competitive, we can assume that the utilities of the players sum up to zero in every outcome. Thus, instead of having a pair of digits in each box, we just have one, representing the utility of the row player; as a consequence, the same number represents the opposite of the utility of the column player.

[3] In formulas, if the generic entry of the matrix is denoted by a_{ij}, the value of the first player is $\max_i \min_j a_{ij}$, that of the second player is $\min_j \max_i a_{ij}$.

(the difference between the minmax and the maxmin) remains so to say on the table, while each player wants it for oneself. This causes uncertainty on the outcome of the game itself. So the next question is: can we suggest the players something, or a rational analysis, and consequently rational behavior, is impossible in these cases? Von Neumann proposal is to consider *mixed strategies*, i.e. probability distributions on the strategies. The players, instead of declaring to play an odd/even number with absolute certainty, should declare to play odd with some probability. This makes the players update their utility functions by calculating their expected values[4]. It is quite easy to see that in such a way the conservative value of the first player does not decrease, while the minmax value of the second one does not increase. The wonderful result obtained by von Neumann states that actually there is again equality between the two conservative values! In other words, in this (extended) world, every (finite) zero sum game has equilibrium.

I guess this is not very easy to understand. In the odd-even game, the only possible result is that the two players tie[5]. Strangely enough, in a game where tie is not allowed...; the way to understand this is to think of the two players playing very many times: on average, each will win (approximatively) the same number of games. From the point of view of the payments, this is the same thing as tying each time...

Next, let us move away from the zero sum case. In this more general framework, we find two different approaches: to consider either the *cooperative* model, or the *non cooperative* one. Here, I will illustrate only some simple ideas of the second one. To do this, I introduce the Nash model, and Nash's idea of equilibrium.

A *two player non cooperative game* in strategic form is a quadruplet: $(X, Y, f : X \times Y \to \mathbb{R}, g : X \times Y \to \mathbb{R})$. A *(Nash) equilibrium* for such a game is a pair $(\bar{x}, \bar{y}) \in X \times Y$ such that:

- $f(\bar{x}, \bar{y}) \geq f(x, \bar{y})$ for all $x \in X$;
- $g(\bar{x}, \bar{y}) \geq f(\bar{x}, y)$ for all $y \in Y$.

X and Y are the strategy spaces of the two players, and the functions f and g their payoff functions. The idea underlying the concept of equilibrium is that no player has interest to change his own strategy, taking for granted that the opponent will do the same. Clearly, an extension of von Neumann's idea in the zero sum game, with the big difference that in general the Nash equilibrium cannot be obtained by calculating the conservative values.
Now, let us pause a second to see some examples.

Example 5 Let us consider again the Example 2:

$$\begin{pmatrix} (5,2) & (3,3) \\ (6,6) & (2,5) \end{pmatrix}.$$

Observe that there are two Nash equilibria: $(6, 6)$ and $(3, 3)$. Observe that 3 is also the conservative value of both players.

And now another example.

[4] Once again it is important to remember the assumption of (full) rationality of the players, which implies in this case that the utility function must be calculated as expected value. It is quite easy to think of situations in which even very intelligent players would not act in this way. I believe to act in a clever way if I choose to have 10.000.000 euros with probability one, rather than having 20.000.000 with probability a little more than 0.5. A very rich person could make a different choice, without being silly...

[5] This can be easily understood by observing that the players have symmetric options and utilities.

Example 6 The game is described by the following bimatrix:

$$\begin{pmatrix} (5,3) & (4,2) \\ (6,1) & (3,4) \end{pmatrix}.$$

Let us see that there are no Nash equilibria. The outcomes $(5,3)$ and $(3,4)$ are rejected by the first player, while the second one refuses the other two. So we are in trouble but, once again, the idea of mixed strategy is the right way to get an answer. How can we find the equilibria in mixed strategies? In the case when the two players have only two available strategies, this is particularly simple, if we observe the following. Given the equilibrium strategy played by the second player, the first one *must be indifferent* between her own strategies. Otherwise, by optimality, she will select a pure strategy! And conversely, of course. This leads to the following calculation, for finding the equilibrium strategy of the second player: denoting by q the probability of playing the first column, it must be:

$$5q + 4(1-q) = 6q + 3(1-q),$$

providing $q = \frac{1}{2}$. In the same way, it can be seen that the first player will play the first row with probability $p = \frac{3}{4}$.

0.3 Surprises and disappointing things

We have seen how to define rationality from a mathematical point of view. We have considered a first basic rule, called elimination of dominated strategies, and we then arrived to the concept of Nash equilibrium. The zero sum case suggested also to consider mixed strategies. All of this seems to be very natural and unquestionable. However, in the introduction I claimed that defining rationality in an interactive context immediately proposes counterintuitive results. In this section I want to produce some evidence about my claim. We start by looking at two of the most striking consequences of the apparently innocent assumption of eliminating dominated strategies.

The first point to consider is the following. It is very clear to everybody that for an agent it is more convenient, among two possible utility functions, to choose the one which always gives a better outcome. Let us consider an example. I want to open a new shoe manufacturing, and I have two possibilities: to do it either in the country I or in the country C. The expert I consult tells me that in the country C I will gain more than in country I, *no matter which policy I will implement.* In mathematical terms, this means that the utility function u_C relative to the country C is greater than the utility function u_I relative to the country I: for all x, $u_C(x) \geq u_I(x)$. Clearly, I do not need to decide which policy to implement, to be sure that I will built my shoe manufacturing in C. Does the same apply when there is an interactive situation? Suppose we have two games, and the outcomes of the first one are better off, *in any outcome, for both players.* You can be sure that if you ask two people which game they would like to play, they will choose the first one. Are they always right? Look at the following example.

Example 7 The two bimatrices:

$$\begin{pmatrix} (100,100) & (5,200) \\ (200,5) & (10,10) \end{pmatrix}$$

and

$$\begin{pmatrix} (50,50) & (4,10) \\ (10,4) & (1,1) \end{pmatrix}.$$

The following two facts are very clear:

- the players are better off for any outcome (i.e. pair (row,column)) in the first game than in the second game;
- elimination of dominated strategies can be applied to both games, and provide outcomes $(10, 10)$ in the first one, $(50, 50)$ in the second one.

Thus, it is *not* true that when dealing with games two players have always interest in playing the “better looking game”.

Let us consider another usual situation. When a decision maker is acting alone, he is always better off if he has the possibility to enlarge his decision space. This corresponds to the trivial fact that maximizing a given function f on a set $I \supset J$ provides a result which is at least as good as maximizing it on J. Observe that this is not always true in real life, where it can happen that too many opportunities get people confused. However our decision maker cannot be conditioned by emotional feelings. So, a natural question arises: is this the same in interactive situations? Here is an example:

Example 8 At first, consider the following game:

$$\begin{pmatrix} (50,50) & (5,10) \\ (10,5) & (1,1) \end{pmatrix}.$$

Its outcome is $(50, 50)$: we already saw it. Now add one more strategy to both players[6]:

$$\begin{pmatrix} (0,0) & (100,-2) & (10,-1) \\ (-2,100) & (50,50) & (5,10) \\ (-1,10) & (10,5) & (1,1) \end{pmatrix}.$$

By applying the deletion of dominated strategies, we see that the solution becomes $(0, 0)$!

It follows that adding opportunities for the players does not make them necessarily happier! The decision maker, when alone, clearly can decide to ignore strategies that do not interest him, while, in an interactive context, the players possibly *cannot* decide to eliminate outcomes which are not favorable to both of them.

Let us now see some problems arising when considering, more generally[7], Nash equilibria. A first point to be addressed, is *uniqueness*. Let us come back for a moment to the Example 5:

$$\begin{pmatrix} (5,2) & (3,3) \\ (6,6) & (2,5) \end{pmatrix}.$$

There are two equilibria: $(6, 6)$ and $(3, 3)$. The fact of having two equilibria does not bother us too much. It is likely that the two players could agree on $(6, 6)$. Quite different is the situation when the game is described by the following famous bimatrix:

$$\begin{pmatrix} (10,5) & (0,0) \\ (-5,-5) & (5,10) \end{pmatrix}.$$

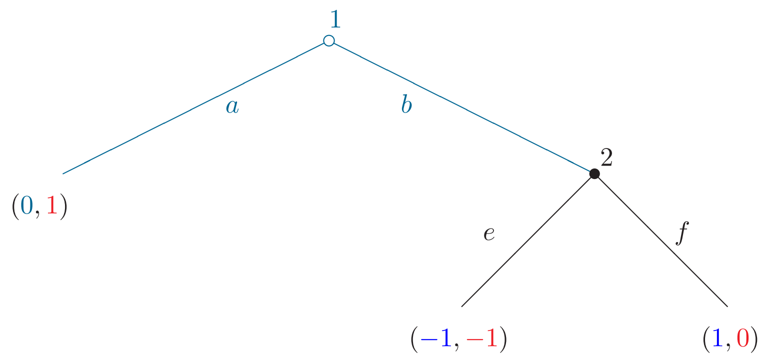
This matrix represents the so called *battle of sexes*: Gabriella and Michele want to stay together, but G would like to go to teatro alla Scala, while M likes better to go to stadio San Siro to watch Milan playing. The two Nash equilibria look very different to them. This is a problem which does *not* arise when the decision maker is alone: having two or more optimal policies does not affect him, since the final result, in term of utilities, does not change[8]. Here it does! We could suggest them to look at

[6] First row for the first player, first column for the second one.

[7] Clearly, an outcome arising from the process of deleting dominated strategies is a Nash equilibrium.

[8] Interestingly, the same is true in zero sum games. In any outcome, the result for the players is always the same: the common conservative value.

the mixed equilibrium. In this case the situation is fairer, since it is symmetric ($\frac{5}{2}$ to both), but far from being satisfactory, since they get a very low level of satisfaction. We shall come back to this point later.
Now, let us consider another game. It will be described by means of its *extensive* form, i.e. by the so called *(game-)tree*, which is a graph with some special features.



The game in plain words is as follows: Player One has two options, either choose the branch a, ending in this way the game, or passing the ball to player Two, by choosing the branch b. Once player Two has the possibility to decide, she can choose either the branch e or the branch f. The payments are attached to the final situations.

And now, here is the game in strategic form:

$$\begin{pmatrix} (0,1) & (0,1) \\ (-1,-1) & (1,0) \end{pmatrix}.$$

Looking at the bimatrix, we see that there are two Nash equilibria. One gives the outcome $(1,0)$, the other one gives the outcome $(0,1)$, corresponding to the choices first row/first column. Is there any difference between the two equilibria? Having the complete description of the game, we can make some more consideration. In particular, we can observe that the Nash equilibrium $(0,1)$ requires the second player to announce a strategy which is not really credible. Why? In the game, when it is her turn to make the move, she knows which are her options. And she knows that for her it is better to choose the branch f, since she get more than choosing e. So why should she declare to use the strategy f?
This argument divides the scholars in Game Theory. Some of them still believe that the Nash equilibrium $(0,1)$ can be supported by convincing arguments, some other argue that it cannot be considered credible. I do not want to enter in this discussion. Rather, let me observe that when we have a game like the above, we found an interesting procedure to find one Nash equilibrium. It is called *backward induction*, and consists in looking at what the players will choose in those situations where their move will end the game, and in this way carry on the analysis from the bottom to the top. In the above game, the equilibrium given by the backward induction is $(1,0)$. The first player knows that once the second is called to play she will choose the branch f. This will provide the first player a payoff of 1, which is better than 0, the payoff he will get if he decides to play a. For this reason, he will choose b, and player Two will choose f.

Of course, even the backward induction procedure can be source of problems. First of all, look at the following example.

Example 9 I have two precious objects, and I say to my oldest son Andrea: I offer both to you, but you must make an offer to your brother Emanuele. You can make any offer, even nothing, to him, but if he does not agree with you, I will give both objects to the youngest of you, Alberto. . .
Backward induction proposes two solutions: either Andrea will offer nothing and Emanuele will accept, or he will offer one object, and Emanuele will accept. It is however clear that this explanation will *not* help Andrea to decide! The problem is that Emanuele is indifferent to say yes or no to the offer of nothing, and this can be dangerous for Andrea. . . which however would like to keep them both.

One more time, failure of uniqueness can cause trouble. But something worse can happen. Look at the following example, this one too very famous.

Example 10 The game is depicted in Fig. 0.1. The game is played by A and S.

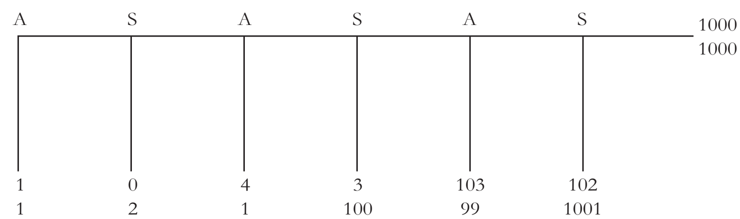


Fig. 0.1. The centipedes

Each at ones' turn decides whether to move right continuing the game, or to go down ending the game. Backward induction leads to the outcome $(1, 1)$ for both. Very disappointing: they could get 1000 each! Even more. At his last turn, A must stop the game, since he knows that S will act against him, by going down. But S, if they arrived at this stage of the game, had already shown he is willing to collaborate. . .

It is time to spend some words on the *most famous* example of Game Theory[9]. We already saw it in Example 7. The bimatrix was:

$$\begin{pmatrix} (100, 100) & (5, 200) \\ (200, 5) & (10, 10) \end{pmatrix}.$$

It is the so called *prisoner dilemma*. The story is the following one: two guys are suspected of a serious crime. The judge says to them: "I know that you both are responsible for the crime, for which the sentence is 10 years in jail. However, if one of you confesses the participation of both and the other does not, the repentant will be set free, while the other one will get a surplus of five years. If no one confesses, I will not send you in court, since I believe the jury will say that there is lack of evidence that you are guilty, and I will condemn you to one year for driving without license. . . It is an easy exercise to write down the bimatrix of the game, and to realize that both guys will spend ten years in jail.
Why this example should be the same as the above bimatrix? Simply because in that bimatrix the outcome is 10 for both players, *notwithstanding that they could get 100 each*[10]! This is well known, since a longtime. . . there are situations which could

[9] In general, it is really questionable what is "the most". But you can try and ask many people which example is the most famous one in game theory, and they all will answer as I do.

[10] It is worth mentioning here that what really matters in the examples I have shown is not the specific value I give to each digit, but the *ordering* among (some of) them. For instance, in the example above the outcome $(10, 10)$ could be substituted by any outcome (x, x), with $5 < x < 100$, without altering the nature of the game.

be better off for everybody. But in order to be able to implement such situations, a collaboration pact among agents is necessary. However, quite often these pacts are not self-enforcing, since from an individual point of view there is an advantage to adhere, but not to maintain them. And within a group of people, it is not even necessary to know that one will not maintain the pacts, to break all agreements: actually to kill the cooperation it is enough that one agent *suspects* that another agent will not maintain pacts, so that, all human life is condemned to be "solitary, poor, nasty, brutish and short"[11], unless we accept the idea of having a *dictator*, which will guarantee everybody that pacts will be maintained.

0.4 Some optimism

Beyond all surprising facts related to the definition of rationality, the conclusion of the previous section seems to suggest that game theory looks at the reality as a state of war, where cooperation is impossible. As result, the initial project to develop new tools to better understand human behavior and to improve the quality of life in a human society, is not but a dream. But is this really true? It is a constant observation, in the human setting, both from very concrete aspects to the deepest scientific theories, that there is a continuous alternation between pessimistic and optimistic attitudes. And it is a great force of the science to be always able to get from apparently negative results the strength to start again the analysis, accepting that some goals are not allowed, but that still there is always room for improvements. I think, for instance, to the theorems of Gödel and Arrow. Each asserts that some achievements were impossible, but the result, at the same time, became the starting points to develop new, important ideas.

Can game theory do the same? Can we find in this analysis of rationality some results giving a more optimistic point of view on human behavior? The answer is in the positive, always with some caution and prudence. I will produce some evidence of my claim, by displaying some simple examples.

The first point I want to stress is the fact that the players can establish, even in a non-cooperative world, some form of collaboration, which can improve their situation. I explain what I mean presenting following example.

Example 11 Two workers must contribute the same job and can commit themselves either to an high level of dedication or to a low level. The gain is equally divided between the two workers, and utility is affected by hard work. Thus a reasonable bimatrix of the game is:

$$\begin{pmatrix} (1,1) & (16,3) \\ (3,16) & (13,13) \end{pmatrix}$$

There are two Nash equilibria, providing $(3,16)$ and $(16,3)$. By using the indifference principle, we find another equilibrium, in mixed strategies. We get that the second player will play the first column with probability $\bar{q} = \frac{3}{5}$; while the first one plays the firs row with probability $\bar{p} = \frac{3}{5}$. They will get 7 each.

Is it possible to do better? It is, even in this wild world depicted by the theory. Suppose the two players agree to ask an arbitrator to propose something better. And suppose she says to the players: I attach a probability of 0.375 to the outcomes $(16,3)$ $(3,16)$ and a probability of 0.25 to the outcome $(13,13)$. Then, by selecting an outcome after a chance move agreeing with the above probabilities, I will tell both of you what to do, *privately.* Now, suppose you are the first player, and the arbitrator tells you to play the first row. In this case, you do not have any incentive to change her recommendation, since you know for sure that the outcome will be

[11] T. Hobbes, *Leviathan.*

$(16, 3)$, a Nash equilibrium. Suppose now she suggests the second row. Then you calculate the probability that the outcome will be either $(3, 16)$ or $(13, 13)$. A simple calculation shows that $(3, 16)$ has probability [12] $\frac{375}{625}$, and thus the expected gain by playing the second row is 7. If the player instead will play the first row he cannot get more (actually, he gets the same). Thus he does not have any incentive to change strategy! The same argument applies to the second player. Indeed, there is a great advantage in this situation, with respect to the Nash equilibria. Differently from the pure Nash equilibria, the players have symmetric outcome. Moreover, this outcome is strictly better than in the case of the mixed equilibrium, since in this case they both get 10.375.

What the arbitrator has suggested in the above example is called a *correlated equilibrium.* The set of correlated equilibria is always nonempty (since a mixed equilibrium is necessarily correlated) and it is characterized by a number of linear inequalities. On this set the players could also decide to maximize a linear function, for instance the sum of their utilities, and this is a typical linear programming problem, which can be solved by available softwares[13]. So, the correlated equilibrium is a first idea of how two people could collaborate to get better results for both.
But what about the prisoner dilemma? Are there satisfactory correlated equilibria? None at all! Unfortunately, it can be easily seen that strictly dominated strategies cannot be used with positive probability in any correlated equilibrium. So that, the only correlated equilibrium in the game is still the unsatisfactory outcome for the players.

Does anything change if I play the prisoner dilemma game several times with the same opponent? Suppose I play with her once a day for 100 days. How can I study this situation? Backward induction provides once again the right way to do it. What will I do on the last day? Of course, I will not maintain a collaboration pact, since for me defeat is dominating. And I know that we both think in the same way... Thus, when I think what to do on the day before the last one, actually, since I know what we will do on the last day, the present day becomes the last in my analysis! As a result, I will not maintain any collaboration pact, exactly for the same reason I will not do it tomorrow. And back to the first day, of course. So that, knowing that the situation will be repeated itself in the future unfortunately is of no help to improve it. Now, let us look at the following example.

Example 12 The game is described by the matrix below:

$$\begin{pmatrix} (10,10) & (0,15) & (-1,-1) \\ (15,0) & (1,1) & (-1,-1) \\ (-1,-1) & (-1,-1) & (-1,-1) \end{pmatrix}.$$

Observe that if we delete the third row and column, we have a typical prisoner dilemma game. Furthermore, the third row and the third column are strictly dominated. Thus nothing changes: by eliminating strictly dominated strategies, as it is mandatory, the result is a typical dilemma situation, and the outcome will be the same, as always, when the game is played once: the two players will get 1 each, having the possibility to share 10 each... But what about if the game is played several times, let us say N times? Of course, the temptation is to argue as before, the last day we eliminate dominated strategies, so the outcome is unfavorable, and so on... It

[12] Given the information obtained by the player, of playing the second row, updating the probability leads to say that the probability of (B, L) is $\frac{\frac{375}{1000}}{\frac{375}{1000}+\frac{250}{1000}}$.

[13] It is practically impossible solve by hands the problem of finding the set of correlated equilibria for games with more than three strategies for the players, since the number of inequalities to check grows explosively: there are examples of 4×4 games for which the polytope of the correlated equilibria has more than 100.000 vertices!

turns out that this is not the only possibile case. The very interesting and, I would say, surprising thing is that even if the players cannot guarantee themselves an average utility of 10 (the ideal situation from the collective point of view), they actually can get an utility which is very close to it[14], and this is possible thanks to the dominated strategies! It looks impossible, but let me explain the idea. The symmetric equilibrium strategy for the players is the following:
Play the first $N-k$ times the collaborative strategy (first row/column for player one/two), next play second row/column, if your opponent does the same. Otherwise, if at any stage before the $N-k$-th time the opponent plays its dominant strategy, from the following stage on play the third row/column.
By following the suggested strategy, both players will gain $(N-k)\cdot 10+k\cdot 1$. Now, let us see that for a suitable k this pair constitutes a Nash equilibrium. Suppose the second player will change his strategy. For him, the most convenient thing to do is to defeat at the last possible stage (in order to be "punished" for a shorter time), i.e. at the stage $N-k$. In this case, he will gain $(N-k-1)\cdot 10+15+k(-1)$. Thus deviating for him is *not* convenient if:

$$(N-k-1)\cdot 10+15+k(-1)<(N-k)\cdot 10+k\cdot 1.$$

This surely happens if $k>3$! Summarizing, for $k>3$ the above strategy, used by both, is a Nash equilibrium providing them on average $(1-\frac{k}{N})10+\frac{k}{N}$, close, for N large, to the ideal outcome of 10.

This is very interesting. How can we explain, without using formulas or mathematical concepts, a situation like the one described in the example above? I would say that adding the possibility of a (credible) threat to the players, forces them to maintain a "good behavior". Not at every stage, however, since there must be some room to have an aggressive behavior: on the last day it is clear that there is no possibility of collaboration: this would go against individual rationality, and explains why we need to include k final stages where the players will be aggressive.

We have learned by the previous example how to exploit, in a repetition of the game, apparently useless strategies for the players, i.e. dominated strategies. Is it possible to exploit also other facts which usually seem to be of no help? For instance, lack of information? I will only give a very qualitative idea of the last result I want to mention. Again, it deals with repetitions of a game.
Studying repeated games is very important. Clearly, the theory needs the analysis of the "one shot" games. But since a game wants to be a model, it is clearly interesting to consider the case when a player faces several times the same game with the same opponent(s). Thus, we need a more sophisticated model to deal with this type of situations, and game theory shows very well that repetition can change the attitude of the players. The result I want to mention to conclude essentially says that it is possible to construct a model of *repeated game* in such a way that if the prisoner dilemma is repeated an *unknown* number of times, and if the players are patient enough[15], then the collective-optimal outcome can be sustained by a Nash equilibrium.

The above result is in some sense for me the example of how game theory can be of so great interest, beyond mathematics, in understanding human behavior. I want to state clearly here that the result does *not* depict the best of the possible worlds, from a philosophical point of view. First of all, the quoted equilibrium is just one of the many possible. Actually, one criticism made to the concept of Nash equilibrium is that it produces too many outcomes, especially in repeated games. Secondly, the

[14] By this I mean that if I let N go to infinity, the average payment converges to the value 10.

[15] If a player is too impatient it can be for him more convenient to get more today by not collaborating and be punished all the successive steps (because the payments in the future are essentially uninteresting for him), rather than collaborating all the time.

result certainly does not assert that the players show to be irrational when they do not collaborate. On the contrary, the non-collaborative behavior fits perfectly with the rationality scheme of Game theory. This is what we continuously observe in our lives: making pacts is fundamental for improving the quality of life: providing ourselves with rules has the primary goal that we all live better lives. However, according to Hobbes, everyday somebody breaks the rules: this cannot be considered craziness. But game theory helps to understand that we do not need Hobbes's dictator: what we really need is to convince everybody that collaboration is useful, and thus it must be continuously promoted and stimulated, since it is not so natural, but at the very end produces better results for everybody.

0.5 Conclusion

The conclusion I want to draw here is that game theory is really helpful in understanding behavior of interacting agents, even if often they do not behave as the theory predicted. Moreover, it has the merit of stimulating us not to take for granted some facts that seem obvious. Often the results, at least the first results of the theory look natural, even trivial. But quite often, before seeing them, people are convinced that the *opposite* is true! I tried to give some example to explain this.
To conclude, I want to mention that Game Theory need not only apply to humans, since life is interaction also for animals. For example, bats and sticklebacks, which are expert players, according to some recent models, explaining that they play a form of prisoner dilemma (with several players) with other elements of the same species. They collaborate: the bats by exchanging fresh blood, essential to survive, the sticklebacks by organizing small groups of them which swim close to every big fish in order to detect if it is aggressive: if so, only a small amount of them will not survive, but this allows the rest of the shoal to look for food without loosing too many energies to escape a big fish any time is coming close.

Fig. 0.2. Flying bats

Fig. 0.3. A stickleback