CHAPTER 1

Theoretical framework

1.1   Introduction and examples

In ordinary language, we speak of a "game" as a (generally amusing) process of
interaction that involves a given population of individuals, is subject to some fixed
rules, and has a prespecified collection of payoffs associated to every possible
outcome. Here, the concept of a game mostly embodies the same idea. However, in
contrast to the common use of this term, the kind of interaction to be studied may
be far from amusing, as illustrated by the following example.

Consider the game usually known as the prisoner's dilemma (PD). It involves
two individuals, labeled 1 and 2, who have been arrested on the suspicion of having
committed jointly a certain crime. They are placed in separate cells and each of them
is given the option by the prosecutor of providing enough evidence to incriminate
the other. If only *one* of them chooses this option (i.e., "defects" on his partner), he
is rewarded with freedom while the other individual is condemned to a stiff sentence
of twelve years in prison. On the other hand, if both defect on (i.e., incriminate)
each other, the available evidence leads to a rather long sentence for both of, say,
ten years in prison. Finally, let us assume that if neither of them collaborates with
the prosecutor (i.e., they both "cooperate" with each other), there is just basis for a
relatively light sentence of one year for each.

The payoff table corresponding to this situation (where payoffs are identified
with the *negative* of prison years) is shown in Table 1.1.

Table 1.1: *Prisoner's dilemma*

1＼2	D	C
D	$-10, -10$	$0, -12$
C	$-12, 0$	$-1, -1$

What would be your prediction on the most likely outcome of this situation? It
seems clear that the prediction must be (D, D) because D is a *dominant strategy*,
i.e., it is better than the alternative C, no matter what the other individual might
choose to do; and this is so despite the fact that (C, C) would indisputably be a better
"agreement" for both. However, unless the agents are somehow able to enforce such
an agreement (e.g., through a credible threat of future revenge), they will not be
able to achieve that preferred outcome. If both individuals are rational (in the sense

1

2 **Theoretical framework**

of aiming to maximize their *individual* payoffs), choosing *D* is the only course of action that makes sense under the circumstances described.

It is important to emphasize that the former line of argument continues to apply even if the individuals are *not* isolated in separate cells and may instead communicate with each other. As long as their decisions have to be taken *independently* (e.g., in the prosecutor's office, one by one), the same reasoning applies. No matter what they might have agreed beforehand, when the time comes to implement a decision, the fact that *D* is a dominant choice should lead both of them to adopt it.

The game just outlined is paradigmatic of many situations of interest. For example, the same qualitative dilemma arises when two firms are sharing a certain market and each one must decide whether to undertake an aggressive or conciliatory price policy (see Chapter 3). Now, we turn to another example with a very different flavor: the so-called battle of the sexes. It involves a certain young couple who have just decided to go out on a date but still have to choose where to meet and what to do on that occasion. They already anticipate the possibilities: they may either attend a basketball game or go shopping. If they decide on the first option, they should meet by the stadium at the time when the game starts. If they decide on the second possibility, they should meet at the entrance of a particular shopping mall at that same time.

Let us assume they have no phone (or e-mail), so a decision must be made at this time. The preferences displayed by each one of them over the different alternatives are as follows. The girl prefers attending the basketball game rather than going shopping, whereas the boy prefers the opposite. In any case, they always prefer doing something together rather than canceling the date. To fix ideas, suppose payoffs are quantified as in Table 1.2.

Table 1.2: *Battle of the sexes*

Girl \ Boy	B	S
B	3, 2	1, 1
S	0, 0	2, 3

where *B* and *S* are mnemonic for "basketball" and "shopping," respectively, and the pairs of numbers specified quantify the utilities obtained by each individual (first the girl's, second the boy's) for each choice combination. In principle, the couple could "agree" on implementing any pair of choices on the day in question. However, only (*B*, *B*) and (*S*, *S*) represent robust (or stable) agreements in the sense that if they settle on any of them and each believes that the other side is going to abide by it, both have incentives to follow suit. Each of these agreements will be labeled a *Nash equilibrium* and either of them may be viewed as a sensible prediction for the game. The problem, of course, is that there is an unavoidable multiplicity in the task of singling out *ex ante* which one of the two possible equilibria could (or should) be played. In contrast with the previous PD game, there is no natural basis to favor any one of those outcomes as more likely or robust than the alternative one.
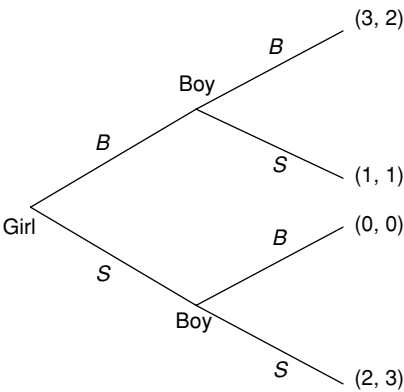
Figure 1.1: Battle of the sexes, sequential version.

Let us now explore a slight variation of the previous story that is not subject to the aforementioned multiplicity problem. On the day set for the date, rather than both individuals being out of reach, it turns out that the boy (only he) is at his home, where he can be contacted by phone. Suppose that the girl knows this and that, initially (i.e., when the plans were drawn), the boy managed to impose the "agreement" that they both would go shopping. The girl, angry at this state of affairs, may still resort to the following course of action: she can arrive at the stadium on the specified day and, shortly before the boy is ready to leave for the shopping mall, use the phone to let him know unambiguously where she is. Assume that it is no longer possible for the girl to reach the shopping mall on time. In this case, she has placed the boy in a difficult position. For, taking as given the fact that the girl is (and will continue to be) at the stadium waiting for him, the boy has no other reasonable option (if he is rational) than to "give in," i.e., go to the stadium and meet the girl there. What has changed in this second scenario that, in contrast to the former one, has led to a single prediction? Simply, the time structure has been modified, turning from one where the decisions were independent and "simultaneous" to one where the decisions are sequential: first the girl, then the boy.

A useful way of representing such a sequential decision process diagrammatically is through what could be called a "multiagent decision tree," as illustrated in Figure 1.1. In this tree, play unfolds from left to right, every intermediate (i.e., nonfinal) node standing for a decision point by one of the agents (the boy or the girl) and a particular history of previous decisions, e.g., what was the girl's choice at the point when it is the boy's turn to choose. On the other hand, every *final* node embodies a complete description of play (i.e., corresponds to one of the four possible outcomes of the game), and therefore has some payoff vector associated to it.

In the present sequential version of the game, it should be clear that the only intuitive outcome is (B, B). It is true that, at the time when the plans for the date are discussed, the boy may threaten to go shopping (i.e., choose S) even if the girl phones him from the stadium on the specified day (i.e., even if she chooses B). However, as explained above, this is *not* a credible threat. Or, in the terminology

4   Theoretical framework

to be introduced in Chapter 4, such a threat does not belong to a (subgame) "perfect" equilibrium – only (B, B) defines a perfect equilibrium in the present case.

The representation of a game by means of a multiagent decision tree permits an *explicit* description of the order of movement of the different players as well as their information and possible actions at each point in the game. It is called its *extensive-form representation* and provides the most fundamental and complete way of defining any game. The next section formalizes this theoretical construct in a general and rigorous manner.

1.2   Representation of a game in extensive form

1.2.1   *Formalization*

The extensive form of a game requires the description of the following items.

1. *The set of players.* It will be denoted by $N = \{0, 1, 2, \ldots, n\}$, where player 0 represents "Nature." Nature performs every action that is exogenous to the game (whether it rains, some player wins a lottery, etc.). When it has no specific role to play, this fictitious player will be simply eliminated from the description of the game.

2. *The order of events.* It is given by a certain binary relation, R, defined on a set of nodes, K. More precisely, the set K is identified with the *collection of events* that can materialize along the game, whereas the relation R embodies a suitable criterion of *precedence* (not necessarily temporal, possibly only logical) applied to those events.[1] Here, the notion of event is the usual one, i.e., a description of "what is possible" at any given juncture in the game. Thus, in particular, an "elementary event"[2] is to be conceived simply as a sufficient description of a complete path of play, whereas the "sure event" refers to the situation that prevails at the beginning of the game (where still any path of play is attainable). As the players make their choices, the game advances along a decreasing (or nested) sequence of events, with a progressively narrower set of possibilities (i.e., paths of play) becoming attainable. Formally, this is captured through the relation R, which, for any pair of nodes $x, y \in K$, declares that $x R y$ whenever every path of play that is (possible) in y is (possible) as well in x. Thus, for example, if y stands for the event "both agents attend the basketball game" in the sequential battle of the sexes represented in Figure 1.1, the event x given by "the girl attends the basketball game" precedes y. Thus, by writing $x R y$ in this case,

---

[1] A binary relation R on K is defined as some subset of the Cartesian product $K \times K$. If $(x, y) \in R$, then we say that x is related to y and typically write $x R y$.

[2] In the language of traditional decision theory [see, e.g., Savage (1954)], an elementary event is the primitive specification of matters that would correspond to the notion of a "state," i.e., a description of all relevant aspects of the situation at hand. For a formal elaboration of this approach, the reader is referred to the recent (and somewhat technical) book by Ritzberger (2002).

we mean that x logically precedes y in the set of occurrences that underlie the latter event – therefore, if y occurs, so does x as well.

Given the interpretation of R as embodying some notion of precedence, it is natural to postulate that this binary relation is a (strict) partial ordering on K, i.e., it displays the following properties[3]:

Irreflexivity: $\forall x \in K, \neg(x R x)$.

Transitivity: $\forall x, x', x'' \in K, [x R x' \wedge x' R x''] \Rightarrow x R x''$.

Associated to R, it is useful to define a binary relation, P, of *immediate precedence* in the following manner:

$$x P x' \Leftrightarrow [(x R x') \wedge (\nexists x'' : x R x'' \wedge x'' R x')].$$

Correspondingly, we may define the set of immediate predecessors of any given $x \in K$ as follows:

$$P(x) \equiv \{x' \in K : x' P x\}$$

and the set of its immediate successors by

$$P^{-1}(x) = \{x' \in K : x P x'\}.$$

Having interpreted (K, R) as the set of partially ordered events that reflect the unfolding of play in the game, it is useful to postulate that every $y \in K$ *uniquely* defines the set of its preceding events – or, expressing it somewhat differently, that y uniquely induces the chain (or history)[4] of occurrences that give rise to it. In essence, this is equivalent to saying that (K, R) must have the structure of a *tree of events*, thus displaying the following two properties:

(a) There exists a unique root (or initial node) x_0 that has no immediate predecessor ($P(x_0) = \emptyset$) and precedes all other nodes (i.e., $\forall x \neq x_0$, $x_0 R x$). This initial node is to be viewed as the beginning of the game.

(b) For each $\hat{x} \in K$, $\hat{x} \neq x_0$, there exists a unique (finite) path of predecessors $\{x_1, x_2, \ldots, x_r\}$ joining $\hat{x}$ to the root x_0 – i.e., $x_q \in P(x_{q+1})$, for all $q = 0, 1, \ldots, r - 1$, and $x_r \in P(\hat{x})$.

As intended, (a) and (b) permit identifying each node in K with a (unique) particular history of the game – possibly partial and incomplete if it is an intermediate node, or even "empty" if it is the initial x_0. Also note that, from (a) and (b), it follows that every $x \neq x_0$ has a *unique* immediate predecessor (i.e., $P(x)$ is a singleton). Indeed, this is precisely the key feature that allows one to associate to every node the set of its preceding events (i.e., the underlying history) in a univocal fashion. A possible such

---

[3] As customary, we use the symbol $\neg(\cdot)$ to denote the negation of the statement in question, or $\wedge$, $\vee$ to join two statements by "and," "or." An alternative way of expressing negation is by superimposing / on a certain symbol, e.g., $\nexists$ stands for the negation of existence.

[4] Its temporal connotations notwithstanding, the term "history" is typically used in game theory to describe the unfolding of a path of play even when the implied irreversibility does not involve the passage of time. An illustration of this point may be obtained from some of our upcoming examples in Subsection 1.3.2.
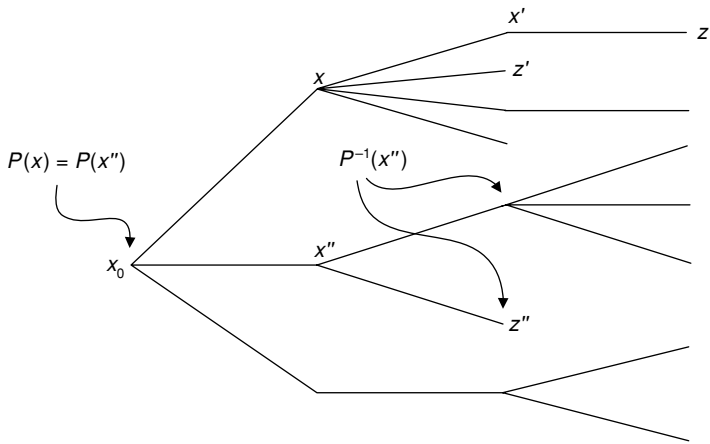
Figure 1.2: Tree of events with $x_0 \, P \, x \, P \, x' \, P \, z$; $x_0 \, P \, x \, P \, z'$; $x_0 \, P \, x'' \, P \, z''$.

tree of events is graphically illustrated in Figure 1.2, where the play of the game unfolds from left to right and any two nodes linked by a line segment are taken to be immediately adjacent according to the relation P.

For simplicity, let us posit here that every path of the game reaches a definite end.[5] Denote by $Z \equiv \{x \in K : P^{-1}(x) = \emptyset\}$ the set of final nodes, i.e., those nodes with *no* successors (for example, the nodes z, z', and z'' in Figure 1.2). As explained, the interpretation of any such node is that of a primitive event, a complete history, or simply a *game play*. It is worth emphasizing that every final node includes not only information on the "characteristics" of the final outcome of the game but also describes in full detail its underlying history. To illustrate this point, consider for example the event "wearing the two gloves" resulting from the concatenation of the intermediate events "not wearing any glove" and "wearing just one glove." Then, the two different ways in which one may end up wearing the two gloves (either the right or the left glove first) give rise to two different final nodes, even though they both display the same relevant features.

3. *Order of moves.* The set $K \backslash Z$ of intermediate nodes is partitioned into $n + 1$ subsets $K_0, K_1, \ldots, K_n$. If $x \in K_i$, this simply means that when the event reflected by x materializes, it is player i's turn to take an action. For convenience, it is typically assumed that, if Nature moves in the game, it does so first, thus resolving once and for all any bit of exogenous uncertainty that may affect the course of play. In terms of our previous formalization, this amounts to making $K_0 \subseteq \{x_0\}$ – of course, K_0 is empty if Nature does not have any move in the game.

4. *Available actions.* Let $x \in K_i$ be any node at which some player $i \in N$ moves. The set of actions available to player i at that node is denoted by

[5] Some of the game-theoretic models proposed at later points in this book (cf., for example, Subsections 5.2.1 and 8.2) admit the possibility that the game never ends, a case that requires a natural extension of the present formulation. Then, every infinite history must be interpreted as a different "end node," which again embodies a full description of the whole turn of events that underlie it.

$A(x)$. Naturally, the cardinality of $A(x)$ must be identical to that of $P^{-1}(x)$, the set of immediate successors of x. This simply reflects the fact that it is player i who decides how the game proceeds after x along one of the possible ensuing directions. Formally, what is required is that the sets $A(x)$ and $P^{-1}(x)$ be isomorphic, i.e., each immediate successor of x must have a *unique* and *different* action a in the set $A(x)$ associated to it, and *vice versa.*

5. *Information sets.* For every player i, we postulate that her corresponding set of decision nodes K_i can be partitioned into a set H_i of disjoint sets, i.e., $K_i = \bigcup_{h \in H_i} h$ with $h \cap h' = \emptyset$ for all $h, h' \in H_i$ ($h \neq h'$). Each of these sets $h \in H_i$ is called an *information set* and has the following interpretation: player i is unable to discriminate among the nodes in h when choosing an action at any one of them. Intuitively, if player i cannot distinguish between two different nodes $x, x' \in h$, it must be that player i did not observe (or has forgotten – see Section 1.4) the preceding occurrences (choices) on which x and x' differ. Obviously, this interpretation requires that $A(x) = A(x')$ – that is, there must exist the same set of available actions at both x and x'. Otherwise, the inconsistency would arise that player i could in fact distinguish between x and x' on the basis of the different set of actions available at each node (an information that of course player i should have because she is the *decision maker* at both of those nodes).

6. *Payoffs.* Associated with every possible game play (i.e., final node or complete history of the game) there is a certain payoff for each of the different players. Thus, for every one of the final nodes $z \in Z$, we assign an n-dimensional real vector $\pi(z) = (\pi_i(z))_{i=1}^n$, each $\pi_i(z)$ identified as the payoff achieved by player $i = 1, 2, \ldots, n$ if the final node z is reached. These real numbers embody how players evaluate any possible outcome of play and thus reflect every consideration they might deem relevant – pecuniary or not, selfish or altruistic. Payoffs for Nature are not specified since its behavior is postulated exogenously. (Fictitiously, one could simply posit constant payoffs for Nature over all final nodes.)

Payoff magnitudes are interpreted as von Neumann–Morgenstern utilities and, therefore, we may invoke the well-known theorem of expected utility[6] when evaluating random outcomes. That is, the payoff or utility of a certain "lottery" over possible plays (or final nodes) is identified with its expected payoff, the weights associated with each one of those plays given by their respective *ex ante* probability. This implies that payoffs have a cardinal interpretation (i.e., payoff *differences* have meaning) and embody players' attitude to risk. Formally, it amounts to saying that the specification of the payoffs in the game is unique *only* up to monotone *affine* transformations.[7]

Finally, note that even though "payoff accounting" is formally performed at the end of the game (i.e., payoffs are associated with final nodes alone),

---

[6] See, e.g., Kreps (1990) for a classical textbook treatment of this topic.

[7] A monotone affine transformation of a utility function $U(\cdot)$ is any function $\tilde{U}(\cdot)$ over the same domain, which may be written as follows: $\tilde{U}(\cdot) = \alpha + \beta U(\cdot)$ for any real numbers α, β, with $\beta > 0$.

8   **Theoretical framework**

this does not rule out that partial payoffs may materialize at intermediate
stages. In those cases, the payoff associated with any final node is to be
interpreted as the overall evaluation of the whole stream of payoffs earned
along the unique history that leads to it.

The above six components define a game in extensive form. Often, we shall rely on
a graphical description of matters where

- the unfolding events $x \in K$ induced by players' actions are represented
  through a tree structure of the sort illustrated in Figure 1.2;
- intermediate nodes are labeled with the index $i \in N$ of the player who takes
  a decision at that point;
- the edges departing from intermediate nodes $x \in K$ are labeled with the
  respective actions $a \in A(x)$ leading to each of its different successors in
  $P^{-1}(x)$;
- the intermediate nodes $\{x \in h\}$ that belong to the same information set h
  are joined by a dashed line;
- the final nodes $z \in Z$ have real vectors $\pi(z)$ associated with them, express-
  ing the payoffs attained by each player in that game play.

A simple illustration of such a graphical way of describing a game in extensive
form is displayed in Figure 1.3.

1.2.2   *Examples*

1.2.2.1 *A simple entry game.* Consider two firms, 1 and 2, involved in the following
game. Firm 1 is considering whether to enter the market originally occupied by a
single incumbent, firm 2. In deciding what to do (enter (E) or not (N)), firm 1
must anticipate what will be the reaction of the incumbent (fight (F) or concede
(C)), a decision the latter will implement only after it learns that firm 1 has entered
the market. Assume that the monopoly (or collusive) profits to be derived from the
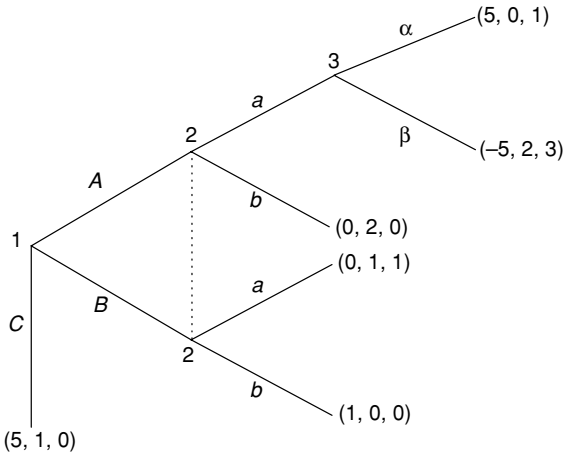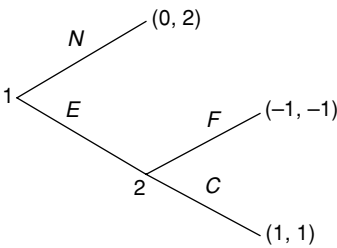

Figure 1.3: A game in extensive form.

Figure 1.4: A simple entry game, extensive form.

market are given by two million dollars, which firm 2 either can enjoy alone if it remains the sole firm or must share with firm 1 if it concedes entry. On the other hand, if firm 2 fights entry, both firms are assumed to incur a *net* loss of one million dollars because of the reciprocal predatory policies then pursued.

The extensive-form representation of the entry game considered is described in Figure 1.4. In this simple extensive form, each firm has just one information set consisting of only one node. Thus, in both of these information sets, the corresponding firm is fully informed of what has happened at preceding points in the game. With this information at hand, each firm has two possible actions to choose from (N or E for firm 1; F or C for firm 2).

1.2.2.2 *A matching-pennies game.* Consider the following game. Two players simultaneously choose "heads" or "tails." If their choices coincide (i.e., both select heads, or both select tails) player 2 pays a dollar to player 1; in the opposite cases, player 1 pays this amount to player 2.

As explained above, the extensive form is to be conceived as the most basic and complete way of representing a game. However, since an extensive-form representation displays, by construction, a sequential decision structure (i.e., any decision node can belong to only a single agent), one might be tempted to think that it is inherently unsuited to model any simultaneity of choices such as the one proposed here. To resolve this puzzle, the key step is to grasp the appropriate interpretation of the notion of "simultaneity" in a strategic context. In any given game, the fact that certain actions are described as "simultaneous" does not necessarily reflect the idea that they are chosen at the same moment in *real time*. Rather, the only essential requirement in this respect is that at the time when one of the players takes her decision, she does not know any of the "simultaneous" decisions taken by the other players.

To formalize such a notion of simultaneity, we rely on the concept of information set, as formulated in Subsection 1.2.1. This allows us to model the matching-pennies game through *any* of the two extensive-form representations displayed in Figures 1.5 and 1.6 (recall the graphical conventions illustrated in Figure 1.3). In either of these alternative representations, each player has just one information set and two possible actions (heads (H) or tails (T)). However, while in the first representation it is player 1 who "fictitiously" starts the game and then player 2 follows, the second representation has the formal roles of the players reversed. Clearly, both of these
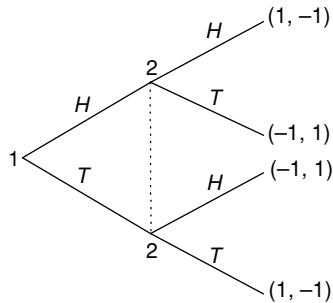
Figure 1.5: A matching-pennies game in extensive form, alternative 1.
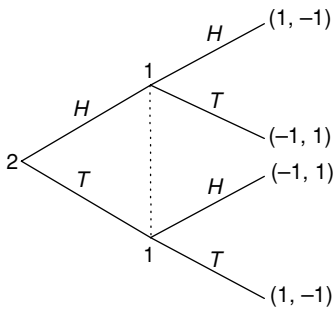

Figure 1.6: A matching-pennies game in extensive form, alternative 2.

alternative extensive-form representations of the game should be viewed by the players as *strategically equivalent*. In both of them, no player is informed of the action played by the opponent, either because she moves first or because she is unable to distinguish between the possible "prior" moves of the other player.

1.2.2.3 *Battle of the sexes.* Along the lines pursued for the previous example, we may return to the battle of the sexes introduced in Section 1.1 and describe the extensive-form representation of its simultaneous version as displayed in Figure 1.7.

Again, the alternative representation of the simultaneous battle of the sexes where the formal roles of the boy and the girl are reversed is strategically equivalent to the one described in Figure 1.7. Of course, this is no longer the case if we consider instead the sequential version of the game where the girl moves first. Such a game has the extensive-form representation described in Figure 1.1. In it, the girl still has only one information set (she moves without knowing the decision her partner will make), but the boy has two information sets (he already knows the decision adopted by the girl at the time he makes his own decision). As explained in our informal discussion of Section 1.1, this sequential version of the game leads to a rather strong strategic position for the girl. It is obvious, however, that the relative strength of the strategic positions is reversed if the roles of the players (i.e., their order of move) is permuted. Thus, in contrast with the simultaneous version, such