Part II

Sequential Games

In this part we extend the analysis of strategic thinking to interactive decision situations where players move sequentially. As the play unfolds, players obtain new information. We consider a world inhabited by players who maximize their subjective expected payoff.[60] When a player receives a piece of information that was possible (had positive probability) according to his previous beliefs, then he simply updates his beliefs according to the standard rules of conditional probabilities. When the new piece of information is completely unexpected, the player forms "brand new" subjective beliefs, unconstrained by the rules of conditional probabilities.

The observation of some previous moves by the co-players may provide information about the strategies they are implementing and/or their type. How such information is interpreted depends, for example, on beliefs about the co-players' rationality, or whether they are carrying out their plan. This introduces a new fascinating dimension to strategic thinking.

As in Part I, we first analyze games with complete information and then move on to games with incomplete information. Chapter 9 provides a mathematical description of **multistage games with observed actions**, i.e., games where at the beginning of each stage all the actions taken in previous stages are publicly observed and hence are common knowledge. We then define the central concepts of "strategy" and "strategic form." Chapter 10 analyzes rational planning for given beliefs about the behavior of the co-players. For games with finite horizon (i.e., games that end in bounded, finite time), we represent rational planning with the "folding-back" procedure. It turns out that a strategy is consistent with folding back if and only if it is "sequentially optimal," and if and only if it satisfies the "one-step optimality property." Sequential optimality and the one-step optimality property, which are equivalent in finite-horizon games, can be defined also for games with infinite horizon, and the equivalence between sequential optimality and the one-step optimality property—called **one-deviation principle**—extends to a very large class of sufficiently regular infinite-horizon games. Chapter 11 puts forward notions of rationalizability for multistage games justified by different assumptions about strategic reasoning. In many games of interest, these solution concepts admit useful strategic-form characterizations, by means of weak dominance. Chapter 12 analyzes equilibrium concepts. A first definition of equilibrium is

[60]Economic theorists often say "Bayesian players" to mean this. See our comments on language in Chapter 8.

given by the Nash equilibrium of the strategic form. This, however, is unsatisfactory. The more demanding concept of **subgame perfect equilibrium** is based on the details of the sequential representation of the multistage game: it requires that the strategy of each player be sequentially optimal given correct conjectures about the strategies of the co-players. The one-deviation principle implies that, in finite games where only one player is active at each stage and payoffs are "generic," there is only one subgame perfect equilibrium that can be computed with the so called "backward-induction" algorithm, a kind of inter-personal folding-back procedure. In games with finite horizon, one can use a "case-by-case backward induction" method to find the subgame perfect equilibria. Chapters 13 and 14 apply the one-deviation principle to the analysis of repeated games and bargaining games. Chapters 15 and 16 extend the analysis to sequential games with incomplete information and with imperfectly observed actions.

9

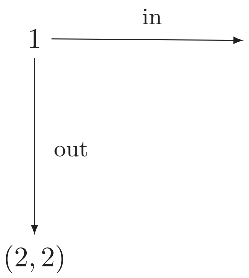
Multistage Games with Complete Information

We start by providing a mathematical definition of multistage games whereby, at the beginning of each stage, all the actions taken in previous stages are publicly observed, hence are common knowledge (Section 9.1). In a single stage, actions are chosen simultaneously, but some players may be inactive. Inactive players will be formally represented as players with only one feasible action. It is possible that in a given stage all the players but one are inactive. The set of feasible actions in a given stage may be endogenous, that is, it may depend on the actions taken in previous stages. The length of the game may also be endogenous. Armed with such formal notation, we move on to define "strategies" and the (reduced) "strategic form" of a game (Section 9.3). We can interpret a strategy as (i) the complete description of the information-dependent behavior of a player, and (ii) the contingent plan of a player. The strategic form of a game is based on the first interpretation. The analysis of rational planning (formally developed in Chapter 10) considers the second interpretation. For a player who would always carry out his plan, (i) and (ii) coincide. Any solution concept for static games can be applied to the strategic form of a multistage game and thus yields a "candidate solution." In Section 9.3.1, we will discuss whether some solutions concepts for static games—such as rationalizability or Nash equilibrium—make sense as solutions of a multistage game, when applied to its strategic form. Finally, we extend the analysis by introducing two different notions of randomized strategic behavior, "mixed strategies" and "behavior strategies"; the latter are better suited to define subgame perfect equilibrium in randomized strategies, as we will see in Chapter 12.

9.1 Preliminary Definitions

It is useful to start with a preliminary example before we plunge into the mathematical representation of multistage games with observed actions.

Example 41. Ann (player 1) and Bob (player 2) play the **Battle of the Sexes (BoS) with an Outside Option**. Ann first chooses between playing the BoS with Bob (action in) or not (action out). If she chooses in, this choice becomes common knowledge and then the simultaneous-move "subgame" BoS is played. If she chooses out the game ends. Payoffs are displayed in Figure 9.1 below.



1\2	B_2	S_2
B_1	$3, 1$	$0, 0$
S_1	$0, 0$	$1, 3$

Figure 9.1: Battle of the Sexes with an Outside Option.

The set of feasible actions of a player depends on the position reached: Ann's feasible set is $\{\text{in}, \text{out}\}$ at the beginning of the game, and it is $\{B_1, S_1\}$ if she moves in. At the beginning of the game, Bob can only "wait and see" what Ann does; therefore his feasible set is the singleton $\{\text{wait}\}$. If Ann moves in, Bob's feasible set is $\{B_2, S_2\}$. There are 5 possible plays of the game: either the pair of actions $(\text{out}, \text{wait})$ is played and the game ends, or one of the following 4 sequences of actions pairs is played: $((\text{in}, \text{wait}), (B_1, B_2))$, $((\text{in}, \text{wait}), (B_1, S_2))$, $((\text{in}, \text{wait}), (S_1, B_2))$ and $((\text{in}, \text{wait}), (S_1, S_2))$. A pair of payoffs $(u_1(z), u_2(z))$ is associated with each possible play z. ▲

A possible play of the game is called **terminal history**. Possible partial plays like (in, wait) are called **nonterminal** (or **partial**) histories. The rules of the game specify what sequences of action profiles are terminal or nonterminal histories. For each terminal history, viz., z,[1] the rules of the game specify an outcome (consequence) $y = g(z)$, e.g., a distribution of money among the players. Each player i assigns to each outcome y utility $v_i(y)$. As in the first part, v_i represents i's preferences over lotteries of outcomes (consequences) $\lambda \in \Delta(Y)$ by way of expected utility calculations. Given the rules and the utility functions, each terminal history z is mapped to a profile of "payoffs" (induced utilities) $(u_i(z))_{i\in I} = (v_i(g(z)))_{i\in I}$. For example, the payoff pair attached to terminal history $z = ((\text{in}, \text{wait}), (\text{S}_1, \text{S}_2))$ is $(u_1(z), u_2(z)) = (1, 3)$. Under the complete information assumption, the rules of the game (including the outcome function g) and players' preferences over lotteries of consequences (represented by the utility functions v_i, $i \in I$) are common knowledge. Thus, also the payoff functions $z \mapsto u_i(z)$ are common knowledge. In this chapter, we assume complete information and we directly specify the payoff functions. But, as in the analysis of static games, it is important to keep in mind that such payoff functions are derived from preferences and the outcome function g.

When discussing a *specific example* like the BoS with an Outside Option it is simpler to represent the possible plays as sequences like (out), $((\text{in}), (\text{B}_1, \text{B}_2))$, $((\text{in}), (\text{B}_1, \text{S}_2))$, etc. But we use this awkward notation for a reason. Having each inactive player automatically choose "wait" simplifies the *abstract* notation for *general* games: possible plays of the game are just sequences of action profiles $(a^1, a^2, ...)$ where each element a^k of the sequence is a profile $(a_i^k)_{i\in I}$ and there is no need to keep track in the formal notation of who is active. Thus, if A denotes the set of all action profiles, histories are just sequences of elements of set A. To allow for the theoretical possibility of games that can go on forever, we also consider infinite sequences of elements of A. The rules of the game specify which sequences are possible, i.e., they specify the set of "**histories**."

Since sequences of elements from a given domain are a crucial ingredient of the formal representation of a game, it is useful to introduce some preliminary concepts and notation about sequences.

[1]Since z is the last (i.e., terminal) letter of the alphabet, it seems like a good choice as a symbol to denote terminal histories.

9.1.1 Sequences and Trees

Fix a set X. The set of finite sequences of length ℓ of elements from set X is the ℓ-fold Cartesian product

$$X^\ell = \underbrace{X \times ... \times X}_{\ell \text{ times}}.$$

Such sequences can be thought of as functions from the set of the first ℓ positive integers, $\{1, ..., \ell\}$ to X; thus, X^ℓ is isomorphic to the set of functions $X^{\{1,...,\ell\}}$. Now consider the union of all such sets, i.e., the set of all finite sequences of elements from X, plus the singleton $X^0 = \{\varnothing\}$ containing only the **empty sequence** $\varnothing$ (a convenient mathematical object analogous to the empty set $\emptyset$).[2] With a descriptive notation, such union is represented by the symbol $X^{<\mathbb{N}_0}$, i.e.,

$$X^{<\mathbb{N}_0} = \bigcup_{\ell \in \mathbb{N}_0} X^\ell$$

where $\mathbb{N}_0 = \{0, 1, 2, ...\}$ is the set of natural numbers including 0. It may help to think of X as an *alphabet* for some language Λ. Then, the set of *words* in language Λ is a subset of $X^{<\mathbb{N}_0}$. The *rules* of language Λ determine which elements of $X^{<\mathbb{N}_0}$ are words of Λ.

Similarly, the rules of a given game Γ determine which sequences of action profiles are feasible. But, unlike words in natural languages, some histories of games may have an infinite length. This means that the game does not necessarily end in a finite number of stages. This is a useful abstraction. Consider, for example, a bargaining game where the parties go on making offers and counteroffers until they reach a binding agreement: it is possible that they never agree and hence bargain forever. Countably infinite sequences of elements from X are like functions from $\mathbb{N} = \{1, 2, ...\}$ to X. Therefore the set of all such sequences is denoted $X^{\mathbb{N}}$:

$$X^{\mathbb{N}} = \{(x^k)_{k=1}^{\infty} : \forall k \in \mathbb{N}, x^k \in X\}.$$

The set of all finite and infinite sequences from X (empty sequence included) is

$$X^{\leq \mathbb{N}_0} = X^{<\mathbb{N}_0} \cup X^{\mathbb{N}}.$$

[2]Similarly, the power set of X contains the empty set, that is, $\emptyset \in 2^X$, or equivalently $\{\emptyset\} \subseteq 2^X$.

The histories of a game form a **rooted tree**, i.e., a partially ordered subset[3] $\bar{H} \subseteq X^{\leq \mathbb{N}_0}$ where (i) the order relation $x \prec y$ is "x is a **prefix** (initial subsequence) of y" (by convention, $\varnothing$ is a prefix of every sequence $\left(x^k\right)_{k=1}^{\ell}$, $\ell = 1, 2, ..., \infty$), (ii) every prefix of a sequence in $\bar{H}$ (including the empty sequence) is also a sequence in $\bar{H}$, so that the empty sequence is the root of the tree, and (iii) (in games that may never end) if $z \in X^{\mathbb{N}}$ and every (finite) prefix of z is in $\bar{H}$ then also z belongs to $\bar{H}$.[4]

9.2 Multistage Games with Observed Actions

Consider a game that proceeds through stages. At each stage there is a subset of active players, and this set may depend on previous moves; each active player chooses an action in a feasible set with two or more alternatives, and the choices of the active players are simultaneous. The rules of the game are such that at the beginning of each stage the action profiles chosen in previous stages become public information. The rules determine which actions are feasible for each player according to previous choices of every player. The only feasible action of an inactive player is to wait, hence his feasible set is a singleton. When the game ends, players have empty feasible sets: not only they cannot choose between alternatives, they also stop waiting. The consequences for the players implied by the rules of the game, in general, may depend on the actions taken since the game started, and on the order in which they were taken, that is, the consequences depend on the sequence of action profiles that led to the end of the game. Game forms with simultaneous moves, or static game forms, are a special case whereby the first action profile ends the game.

As we did in the case of simultaneous moves games, we do not represent the rules directly, but rather we describe the players' set, their feasible sets of actions (including how they depend on previous moves) and the consequences of their actions. We do this by providing an

[3]A set P is partially ordered by the binary relation $\prec \subseteq P \times P$ if $\prec$ is transitive and asymmetric, that is, for all $p, q, r \in P$, $(p \prec q \wedge q \prec r)$ implies $p \prec r$, and $p \prec q$ implies $\neg (q \prec p)$. The reflexive closure of $\prec$ is the relation $\preceq$ given by $(\prec \cup =)$, that is, $p \preceq q$ if either $p \prec q$ or $p = q$. If $\prec$ is transitive and asymmetric, then $\preceq$ is transitive, reflexive, and antisymmetric (which means that $p \preceq q \wedge q \preceq p$ implies $p = q$). A set P can be equivalently said to be partially ordered by $\preceq$ if $\preceq$ is transitive, reflexive, and antisymmetric. See [64, A.1.4].

[4]Note that (ii) and (iii) do not hold for the words of natural languages.

abstract representation of the possible sequences of action profiles, called "histories;" from this we obtain a game tree with a set of terminal histories; then we introduce a consequence map from terminal histories to collective consequences, thus obtaining a multistage game form, which is a mathematical description of the rules of the game. Finally, we introduce players' preferences and obtain a multistage game. If the rules of the game and players' preferences are common knowledge, this provides a complete mathematical description of those aspects of the game that we deem relevant for the analysis of strategic interaction.

A **multistage game tree** with observed actions is a structure

$$\langle I, (A_i, \mathcal{A}_i(\cdot))_{i \in I} \rangle$$

given by the following components:

- For each $i \in I$, A_i is a nonempty set of **potentially feasible actions**.
- Let $A = \times_{i \in I} A_i$ and consider the set $A^{<\mathbb{N}_0}$ of finite sequences of action profiles; then, for each $i \in I$, $\mathcal{A}_i(\cdot) : A^{<\mathbb{N}_0} \rightrightarrows A_i$ is a **feasibility correspondence** that assigns to each finite sequence of action profiles $h^\ell = (a^k)_{k=1}^{\ell}$ the set $\mathcal{A}_i(h^\ell)$ of actions of i that are feasible immediately after h^ℓ. It is assumed that $\mathcal{A}_i(\varnothing) \neq \emptyset$ and, for all $h \in A^{<\mathbb{N}_0}$, $\mathcal{A}_i(h) = \emptyset$ if and only if $\mathcal{A}_j(h) = \emptyset$ for every $j \in I$ (the reason for this assumption will be explained below).

The reason why we say that this structure has observed actions is that this is the intended interpretation. In particular, if the feasible set $\mathcal{A}_i(h)$ depends on h, the assumption that h is observed as soon as it occurs implies that player i always knows what his feasible set of action is. Indeed, we assume more, i.e., that when a history h occurs it becomes common knowledge that h has occurred. If this were not the case then we would have to describe what players know about what other players may have observed.

Let $\mathcal{A}(h) = \times_{i \in I} \mathcal{A}_i(h)$ denote the set of feasible action profiles given $h \in A^{<\mathbb{N}_0}$. A sequence $(a^k)_{k=1}^{\ell}$ $(\ell = 1, 2, ..., \infty)$ is a (feasible) **history** if $a^1 \in \mathcal{A}(\varnothing)$ and $a^{k+1} \in \mathcal{A}(a^1, a^2, ..., a^k)$ for each positive integer $k < \ell$. Thus, a history is a sequence of action profiles whereby each action profile is feasible given the previous ones. By convention, the empty sequence

$\varnothing$ is a history. Let $\bar{H} \subseteq A^{\leq\mathbb{N}_0}$ denote the set of histories. A history $z = (a^k)_{k=1}^{\ell} \in \bar{H}$ is **terminal** if either $z \in A^{\mathbb{N}}$ (i.e., z is an infinite history) or $\mathcal{A}(z) = \emptyset$. Let

$$Z = \left\{ z \in \bar{H} : z \in A^{\mathbb{N}} \text{ or } \mathcal{A}(z) = \emptyset \right\}$$

denote the set of terminal histories, and let

$$H = \bar{H} \backslash Z$$

denote the set of **nonterminal** (or **partial**) histories.[5]

Consider the restriction to $\bar{H}$ of the prefix-of relation $\prec$ on $A^{\leq\mathbb{N}_0}$: let $h = \left(a^1, ..., a^k\right)$ and $\bar{h} = \left(\bar{a}^1, ..., \bar{a}^{\ell}\right)$, then $h \prec \bar{h}$ if $k < \ell$ and $\left(a^1, ..., a^k\right) = \left(\bar{a}^1, ..., \bar{a}^k\right)$, that is, $\bar{h} = \left(a^1, ..., a^k, \bar{a}^{k+1}, ...\right)$; in this case we say that h **precedes** $\bar{h}$, or—equivalently—that $\bar{h}$ **follows** h. We write $h \preceq \bar{h}$ if either $h \prec \bar{h}$ or $h = \bar{h}$, and we say that history h **weakly precedes** history $\bar{h}$ (or that $\bar{h}$ **weakly follows** h). We also write $\bar{h} \succ h$ ($\bar{h} \succeq h$) to mean that $\bar{h}$ (weakly) follows h. The following result implies that $\bar{H}$ ordered by the precedence relation $\prec$ (restricted to $\bar{H}$) is a tree with distinguished root $\varnothing$:

Remark 30. *The set of histories $\bar{H}$ has the following properties:*
(1) $\varnothing \in \bar{H}$ and, for each $h \in \bar{H} \backslash \{\varnothing\}$, $\varnothing \prec h$;
(2) for each sequence $h \in A^{<\mathbb{N}_0}$ and each history $h' \in \bar{H}$, if $h \prec h'$ then $h \in \bar{H}$;
(3) for each infinite sequence $z \in A^{\mathbb{N}}$, if every predecessor (prefix) of z is in $\bar{H}$, then $z \in \bar{H}$.

Proof. (1) holds by convention. To verify (2), fix arbitrarily a history $h' = \left(a^t\right)_{t=1}^{\ell} \in \bar{H}$. Then, by definition, $a^1 \in \mathcal{A}(\varnothing)$ and $a^{t+1} \in \mathcal{A}(a^1, a^2, ..., a^t)$ for each positive integer $t < \ell$. If $h \prec h'$ then either $h = \varnothing \in \bar{H}$, or $h = \left(a^k\right)_{k=1}^{t}$ for some positive integer $t < \ell$. In the latter case, we must have $a^1 \in \mathcal{A}(\varnothing)$ and $a^{k+1} \in \mathcal{A}(a^1, a^2, ..., a^k)$ for each $k \in \{1, ..., t-1\}$. Therefore h must be a history as well, i.e., $h \in \bar{H}$. To verify (3), let $z = \left(a^k\right)_{k=1}^{\infty}$ be such that $h = \left(a^k\right)_{k=1}^{\ell} \in \bar{H}$ for each $\ell \in \mathbb{N}$, i.e., for each prefix of z. Then $a^1 \in \mathcal{A}(\varnothing)$ and $a^{t+1} \in \mathcal{A}(a^1, a^2, ..., a^t)$ for each $t \in \mathbb{N}$, which implies that $z \in \bar{H}$. ■

[5] It makes sense to assume that the effective range of the feasibility correspondence of a player is his action set, that is $\cup_{h \in H} \mathcal{A}_i(h) = A_i$. But this is not necessary for our analysis.

A comment on language We interpret histories $h = \left(a^1, ..., a^\ell\right) \in \bar{H}$ *events*, i.e., propositions stating that actions in profile $a^1 = \left(a_i^1\right)_{i\in I}$ are/were taken in stage 1, and so on. Yet, Remark 30 shows that the elements of $\bar{H}$ correspond to *nodes* in a tree, which can be represented as a set of points and edges connecting them in a plane. When we refer to histories as events, we use phrases like "given that h occurred," "given h," or "conditional on h." When we refer to histories as nodes in a tree, we use a spatial/geometrical language, e.g., we write "at (history/node) h." Both kinds of language are acceptable, and we will opt for the one that seems more appropriate given the context of our descriptions and explanations.

Next we add to our description a set Y of outcomes, or consequences (e.g., monetary payoffs for all the players), and an **outcome** (or consequence) **function** $g : Z \to Y$, thus obtaining a **multistage game form** with observed actions

$$\langle I, Y, g, (A_i, \mathcal{A}_i(\cdot))_{i\in I}\rangle .$$

This structure describes the rules of the game, but not how each player would like this game to end. Even when $Y \subseteq \mathbb{R}^I$ and $g : Z \to Y$ describes monetary payoffs, it may be the case that players do not just care about their own money, and even in this case we still lack a specification of risk attitudes. Thus, in order to analyze interaction between a specific group of individuals for given rules of the game, we add their personal preferences over lotteries, represented by a profile of von Neumann-Morgenstern utility functions $(v_i : Y \to \mathbb{R})_{i\in I}$ according to expected utility calculations.

With this, we obtain a **multistage game** with observed actions

$$\Gamma = \langle I, Y, g, (A_i, \mathcal{A}_i(\cdot), v_i)_{i\in I}\rangle .$$

For each player $i \in I$, the composition

$$u_i = v_i \circ g : Z \to \mathbb{R}$$

is called the **payoff function** of i. As we did for static games, we will often neglect the specification of the outcome function and utility functions, and look at a simplified, reduced representation $\langle I, (A_i, \mathcal{A}_i(\cdot), u_i)_{i\in I}\rangle$ showing only the payoff functions. In this chapter, we maintain the informal assumption of *complete information*: the rules of the game and players' preferences over lotteries are common knowledge. As we have seen for static games, if there is incomplete information we have to enrich our description of the game, and this matters for the analysis.

We consider both games that end within a given finite number of stages and games that can last for an arbitrarily many stages:

Definition 47. *Game Γ has **finite horizon** if $\bar{H} \subseteq \bigcup_{k=1}^{L} A^k$ for some $L \in \mathbb{N}$, otherwise Γ has **infinite horizon**. Game Γ is **finite** if $\bar{H}$ is finite, otherwise Γ is **infinite**. Game Γ is a **static**, or simultaneous-move game if $Z = \mathcal{A}(\varnothing)$. Game Γ is **compact-continuous** if A is a compact subset of a Euclidean space, $\bar{H}$ is compact*[6] *and u_i is continuous for each $i \in I$.*

Remark 31. *If Γ is finite, then Γ has finite horizon.*

Proof. We prove the result by contraposition, showing that if Γ has infinite horizon, then Γ is infinite. Suppose that, for each $L \in \mathbb{N}$, there is some history $\left(a^k\right)_{k=1}^{\ell} \in \bar{H}$ with length $\ell > L$, where $\ell \in \mathbb{N} \cup \{\infty\}$. Then, there are two (not mutually exclusive) cases: (i) $\bar{H}$ is not finite, which means that Γ is not finite, or (ii) $\bar{H}$ contains an infinite (hence, necessarily terminal) history, $z = \left(a^k\right)_{k=1}^{\infty}$. In the latter case, by Remark 30, the countably infinite set of predecessors $\left(a^k\right)_{k=1}^{\ell}$ ($\ell \in \mathbb{N}$) of z is included in $\bar{H}$ as well; hence Γ is not finite. ■

For each $h \in \bar{H}$, we let $H(h)$ and $Z(h)$ respectively denote the set of nonterminal and terminal histories that weakly follow h:[7]

$$H(h) = \{h' \in H : h \preceq h'\},$$

$$Z(h) = \{z \in Z : h \preceq z\}.$$

Note that the set of histories $H(h) \cup Z(h)$ and the restriction of the precedence relation $\preceq$ on $H(h) \cup Z(h)$ form a sub-tree with root h.

For any nonterminal history $h \in H$, the sub-tree given by $H(h) \cup Z(h)$ and the restriction of g (hence, of each u_i) to sub-domain $Z(h)$ determine

---

[6]For any $\ell \in \mathbb{N} \cup \{\infty\}$, we say that $h_n = \left(a^{k,n}\right)_{k=1}^{\ell}$ converges to $h = \left(a^k\right)_{k=1}^{\ell}$ (in the product topology), written $h_n \to h$, if $a^{k,n} \to a^k$ for each k. In the finite-dimentional case, this is the standard notion of convergence. The set $\bar{H}$ is closed if, for every $\ell \in \mathbb{N} \cup \{\infty\}$ and $(h_n)_{n=1}^{\infty} \in \left(A^\ell\right)^{\mathbb{N}}$ with $\{h_n\}_{n=1}^{\infty} \subseteq \bar{H}$, $\lim_{n\to\infty} h_n = h$ implies $h \in \bar{H}$.

[7]If $h \in Z$, then $H(h) = \emptyset$ and $Z(h) = \{h\}$.

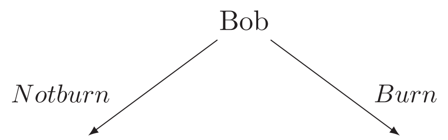
the **subgame with root** h, denoted by $\Gamma(h)$. The maximal length of histories in this subgame, called **height** of $\Gamma(h)$, is

$$L(\Gamma(h)) := \max_{z \in Z(h)} \ell(z) - \ell(h).$$

For notational convenience, we extend $L(\Gamma(\cdot))$ to terminal histories as well: $L(\Gamma(z)) = 0$ for each $z \in Z$.

We illustrate our notation and the definitions with a leading example: the "**Battle of the Sexes (BoS) with Dissipative Action.**"

Example 42. The game depicted in the picture below is the Battle of the Sexes preceded by an observed "money-burning" (dissipative) move by Bob:[8] if Bob chooses the dissipative action *Burn*, he burns 2 dollars.[9]



Ann\Bob	L	R
U	$4, 1$	$0, 0$
D	$0, 0$	$1, 4$

Ann\Bob	l	r
u	$4, -1$	$0, -2$
d	$0, -2$	$1, 2$

The game tree is described by the following elements:

– $\mathcal{A}(\varnothing) = \{w\} \times \{N, B\}$, $\mathcal{A}((w, N)) = \{U, D\} \times \{L, R\}$, $\mathcal{A}((w, B)) = \{u, d\} \times \{l, r\}$, where $\mathcal{A}(h) = \mathcal{A}_a(h) \times \mathcal{A}_b(h)$ for each history/node h, w is the "wait" action of an inactive player, and we omit one order of parentheses when this causes no confusion (e.g., we write $\mathcal{A}((w, B))$ instead of $\mathcal{A}(((w, B)))$ for the feasible action pairs at $h = ((w, B))$, a sequence of action pairs of length one);

– $H = \{\varnothing, ((w, N)), ((w, B))\}$;

– $Z = \{(w, N)\} \times \mathcal{A}((w, N)) \cup \{(w, B)\} \times \mathcal{A}((w, B))$, $\bar{H} = H \cup Z$;

– there are eight maximal chains[10] of histories/nodes starting with the root

[8]See, e.g., Ben-Porath and Dekel [25].

[9]Utility is measured in dollars. In the figure, payoff pairs follow the alphabetical order: the first number is Ann's payoff.

[10]A chain in a partially ordered set $(X, \preceq)$ is a subset $C \subseteq X$ that is totally ordered by $\preceq$.

and ending with a terminal history, the following is an example:

$$\varnothing \prec ((w,B)) \prec ((w,B),(u,l)) \in Z;$$

also note that

$$((w,N)) \nprec ((w,B),(u,l)).$$

▲

Sometimes the definition of games with observed actions takes the set $\bar{H}$ as primitive and posits properties (1)-(3) of Remark 30 as maintained axioms (cf. Osborne and Rubinstein [65, pp 89-90, 102.]). The most common definition of game with sequential moves takes the tree structure as primitive, but deals with simultaneous actions in a somewhat arbitrary and un-intuitive way (see, e.g., Fudenberg and Tirole [46, pp 80-81]). We will take advantage of the tree structure to provide graphical representations of some games, in particular the games where at each stage k at most one player has at least two feasible actions; such games are said to have "perfect information":

Definition 48. *Player i is **active** at history $h \in H$ (that is, immediately after the occurrence of h) if i has at least two feasible actions at h: $|\mathcal{A}_i(h)| \geq 2$. A multistage game with observed actions has **perfect information** if, for each $h \in H$, at most one player is active at h.*

Note, perfect information is a formal property of multistage games, and it should not be confused with complete information, an assumption with a different meaning that here we do not express formally.[11]

9.2.1 Comments

Active and Inactive Players

Our definition of game tree allows for the possibility that all players are inactive at a given history h. We do not even exclude that along some path z all players become inactive after some stage k; such path z may even be infinite. In this case, the path can be truncated at the stage where all players become inactive, and this may transform a formally infinite

[11] Recall that a game features complete information is there is common knowledge of the rules of the game and of players' preferences over lotteries of outcomes.

game into a finite one. Under standard assumptions about preferences and strategic thinking (such as those considered in this textbook), such transformations are innocuous.

According to the present notation, when $\mathcal{A}_i(h) = \emptyset$ for some player i, it means that the game is over. Therefore it is assumed that $\mathcal{A}_i(h) = \emptyset$ for *some* i if and only if $\mathcal{A}_i(h) = \emptyset$ for *all* i. An **inactive** player in an ongoing game is a player with only *one* feasible action, say action "*wait*."

One could give a more elegant definition in which there is a set $\mathcal{I}(h)$ of active players at each partial history h and feasible actions are specified for active players only. This would have the advantage of simplifying the notation for perfect information games and also for specific examples of games. But such a definition would be more complex[12] without essential gains in generality.

Memory and the interpretation of observed actions

The narrowly intended interpretation of the formalism introduced above is that, as soon as a nonterminal history h materializes, it is publicly observed, hence, it becomes common knowledge. But applications where this assumption is literally true are rare. Consider, for example, the game of Chess as usually played in friendly matches: The action a^k taken by the active player at stage k is commonly observed by both players. Furthermore, the resulting positions of the pieces on the board, call it "state P_k" of the play at the end of stage k, is also commonly observed. Can we claim that there is common knowledge of h as soon as it has materialized? No, unless some strong assumptions related to memory hold. In particular, suppose that, just as we assumed that players' preferences over (lotteries of) outcomes are commonly known, another personal feature of players is commonly known, i.e., that they have "perfect memory." This implies that each realized history h becomes common knowledge.

[12] With this alternative definition, action profiles have the form $a_J = (a_j)_{j \in J}$ with $\emptyset \neq J \subseteq I$; and the set of such profiles is

$$\bar{A} := \bigcup_{\emptyset \neq J \subseteq I} (\bigtimes_{i \in J} A_i);$$

an active players correspondence $\mathcal{I}(\cdot) : \bar{A}^{<\mathbb{N}} \rightrightarrows 2^I$ has to be introduced among the primitive elements defining the game; the feasibility correspondence $\mathcal{A}_i(\cdot)$ is defined on the subset $H_i = \{h \in \bar{A}^{<\mathbb{N}} : i \in \mathcal{I}(h)\}$.

Common knowledge of "perfect memory" is a quite strong assumption when we consider games like Chess (although one may argue that it is relatively innocuous given that the continuation-subgame after stage k depends only on the state P_k which *is* common knowledge). Let us note that such assumptions about (knowledge of) players' personal cognitive features are not relevant to the issue we are discussing if the rules of the game provide "cumulative information," that is, the information made public at the end of any stage k is made public again, together with new information, at the end of later stages $\ell > k$. This situation is well approximated in Chess competitions, where a ledger listing past moves is progressively updated. So, there are game situations where the intended interpretation of the formalism applies literally, and others where it applies only under—often tacit—assumptions concerning players' personal cognitive abilities and knowledge about such abilities. It is desirable to clearly separate, in the formalism used to represent interactive strategic situations, the part that just describes the rules of the game from the part that describes players' personal features and knowledge about such features, just as we did for static games. We postpone the implementation of this methodological "separation principle" to Chapter 16, where we analyze games with imperfectly observed actions.[13] For now, we just flag this conceptual issue and note that the "common-knowledge-of-h" assumption should be assessed in the context of applications of the theory.

Streams of Outcomes and Intertemporal Preferences

As a matter of interpretation, $y = g(z)$ need not be an outcome that realizes at the end of the game. Indeed, this cannot be literally true when z is an infinite history (i.e., $z \in Z \cap A^{\mathbb{N}}$). In many applications, $y = g(z)$ is a sequence, or *stream*, of outcomes that realize at the end of each time period with periods being composed by one or more stages. Suppose, for example, that each stage coincides with a time period; furthermore, to ease notation suppose that the game has a fixed—finite or infinite—duration $T \in \mathbb{N} \cup \{\infty\}$. In this case, g may be derived from a sequence of functions $\left(g_t : A^t \to Y_t\right)_{t=1}^{T}$, where $g_t\left(a^1, ..., a^t\right)$ is the period-t outcome, which in

[13]On the "separation principle," the objective representation of how information reaches players according to the game rules, and the representation of players' memory see Battigalli and Generoso [14].

general may depend on the whole history up to period t. With this, we let $Y = \bigtimes_{t=1}^{T} Y_t$ and

$$\begin{array}{rlcl} g: & Z & \to & Y, \\ & \left(a^t\right)_{t=1}^{T} & \mapsto & \left(g_t\left(\left(a^\tau\right)_{\tau=1}^{t}\right)\right)_{t=1}^{T}, \end{array}$$

that is, $y = g(z)$ is the sequence of end-of-period outcomes realized through terminal history z. For example, $g_t\left(\left(a^\tau\right)_{\tau=1}^{t}\right) \in \mathbb{R}^I$ could be the profile of period-t profits in an oligopoly with a set of firms I. If $y = (y_t)_{t=1}^{T}$ is a stream of outcomes, $v_i : Y \to \mathbb{R}$ aggregates the outcomes of different periods and represents the intertemporal preferences of player i. The most common intertemporal aggregator in economic applications satisfies time separability and exponential discounting, that is, there are a discount factor $\delta_i \in (0,1)$ and a sequence of utility functions $(v_{i,t} : Y_t \to \mathbb{R})_{t=1}^{T}$ such that

$$\forall y \in Y,\ v_i(y) = \sum_{t=1}^{T} \delta_i^{t-1} v_{i,t}(y_t).$$

When periods comprise several stages the explicit description of histories and outcome functions requires a more complex notation. We do not pursue this issue here.

9.2.2 Graphical Representation

It is traditional to represent games of *perfect information* (see Definition 48) as **game trees** in the sense of graph theory: each $h \in \bar{H}$ is a **node**; each $z \in Z$ is a **terminal node** (or leaf); each $h \in H$ is a **decision node**; each terminal history/node z is associated with payoff profile $(u_i(z))_{i \in I}$; each partial history/decision node h is associated with the only active player, denoted by $i = \iota(h)$; if $h' = (h, a)$ (the concatenation of h and a), then there is a **directed arc** (arrow) from h to h', this directed arc is associated with action $a_{\iota(h)}$, the action of the active player in profile $a = (a_{\iota(h)}, \text{wait}_{-\iota(h)})$. Sometimes it is notationally useful to distinguish between "physically identical" actions if they are taken after different histories. In this case, there is a 1-1 correspondence between arcs and actions in perfect information games. The following example illustrates such graphical representation.

Example 43. Consider the following game **Take-it-Or-Leave-it** game of length L between player 1 and player 2. A referee (or mechanical device) puts a dollar on the table. Player 1 can take it or leave it, player 2 just observes and waits. If player 1 takes the dollar the game is over, otherwise the referee puts another dollar on the table, player 2 can take the two dollars or leave them on the table, player 1 observes and waits. If player 2 takes the dollars the game is over, otherwise the referee puts another dollar on the table. The game goes on like this until the referee has exhausted his L dollars or someone has taken the dollars on the table. If there are L dollars on the table and the active player (player 1 if L is odd) leaves them, they go to the other player. It is common knowledge that players only care about money and are risk neutral. The action set of each player i is $A_i = \{\text{Take}, \text{Leave}, \text{Wait}\}$. If $L = 4$, the game is represented as in Figure 9.2.

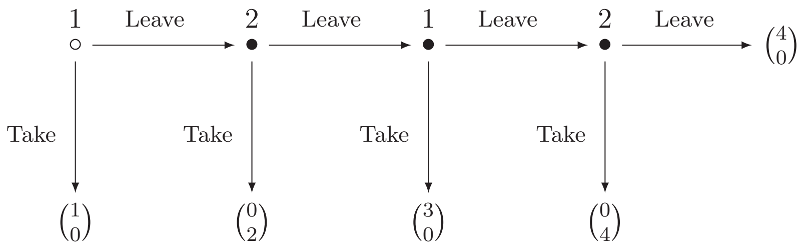


Figure 9.2: Game for $L = 4$.

We have

$$H = \{\varnothing, ((\text{Leave}, \text{Wait})), ((\text{Leave}, \text{Wait}),(\text{Wait}, \text{Leave})), \\ ((\text{Leave}, \text{Wait}),(\text{Wait}, \text{Leave}),(\text{Leave}, \text{Wait}))\},$$

$\mathcal{A}_i(h) = \{\text{Take}, \text{Leave}\}$ if i is active at $h \in H$ and $\mathcal{A}_i(h) = \{\text{Wait}\}$ otherwise. The vertical (respectively, horizontal) directed arcs are labeled by the action Take (respectively Leave) of the active player. ▲

In Example 43, we explicitly included the pseudo-action Wait only to clarify the abstract, general notation introduced above; but, from now on, we will identify histories with sequences of actions by *active* *players only.* For example, (Leave, Wait) will be simply written as (Leave), ((Leave, Wait),(Wait, Leave)) will be written as (Leave, Leave), etc.

More generally, *all games with observed actions can be graphically represented as game trees.* Consider the BoS with an Outside Option of Example 41: if history (in) occurs, then the four action pairs—that is, (B_1, B_2), (B_1, S_2), (S_1, B_2), and (S_1, S_2)—can be represented by directed arcs (arrows) from (in) to the terminal nodes, where each directed arc is associated with the corresponding action pair $(a_1, a_2) \in \mathcal{A}((\text{in}))$. See Figure 9.3 below.

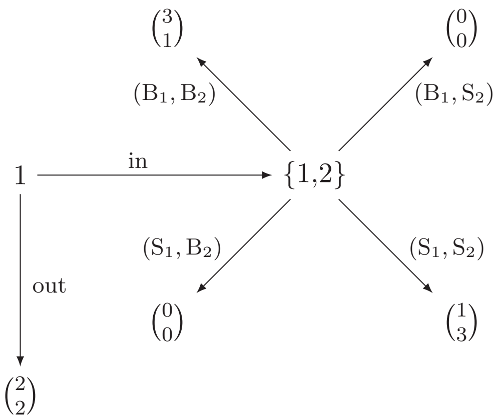


Figure 9.3: BoS with an Outside Option as game tree.

Pictures with trees such as in Figure 9.3 are perfectly legitimate and they faithfully represent the simultaneity of moves at some stages of games *without* perfect information, but they are not very common in the game-theoretic literature. The reason is twofold. First, they are not easily readable. Second, they are not traditional. Indeed, the traditional approach to representing multistage games without perfect information is to pretend that simultaneous moves are, instead, sequential, and use so called "information sets" to make such misrepresentation innocuous.[14] Consider again the BoS with an Outside Option of Example 41. If Ann (player 1) goes in, then the simultaneous-move subgame BoS is played. The approach via "information sets" is to pretend that, after history (in),

[14]See, in particular, von Neumann and Morgenstern [84] and Kuhn [55], who set the stage for the following literature on dynamic games.

one of the players moves first, and the co-player moves second, but does not observe the action chosen by the first mover. In Figure 9.4, the representation of the BoS "subgame" starting after history (in) is that Ann is the first mover. To represent the assumption that Bob (player 2) cannot observe Ann's choice, i.e., that he does not know whether he is at history/node (in, B_1) or (in, S_1), a dashed line joins the two decision nodes of Bob. A (maximal) set of nodes that a player cannot distinguish is called **information set.** That said, we do not follow such approach for the representation of games with observed actions. Information sets will be useful for the analysis of multistage games with incomplete information (Chapter 15) or imperfectly observed actions (Chapter 16).[15]

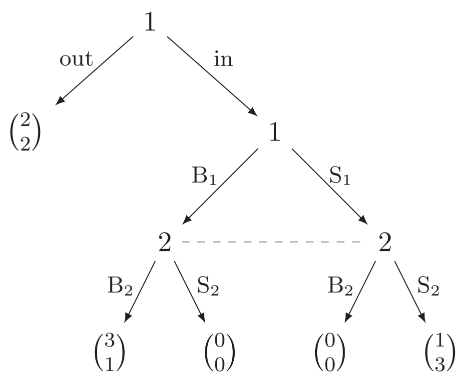


Figure 9.4: Alternative representation of BoS with an Outside Option.

9.3 Strategies, Plans, and Strategic Forms

A strategy is a complete contingent plan which includes "instructions" contingent on every nonterminal history, including histories where the opponents make unexpected moves as well as histories inconsistent with the plan itself. More formally:

[15]Recall that—unfortunately—"(im)perfect information" and "(in)complete information" have very different meanings in the language of game theory (see Chapters 1.4.3, 8, and Definition 48), despite the fact that they are essentially synonymous in the natural language.

Definition 49. *A* ***strategy*** *for player i is an element of the product set $\bigtimes_{h\in H}\mathcal{A}_i(h)$, that is, a function $s_i : H \to A_i$ such that $s_i(h) \in \mathcal{A}_i(h)$ for each $h \in H$. The set of strategies for player i is denoted S_i, that is, $S_i = \bigtimes_{h\in H}\mathcal{A}_i(h)$. We let $S = \bigtimes_{i\in I} S_i$ and $S_{-i} = \bigtimes_{j\neq i} S_j$.*

Remark 32. *In a finite game the number of strategies of each player $i \in I$ is $|S_i| = \prod_{h\in H}|\mathcal{A}_i(h)|$.*

A typical element of S_i is denoted by $s_i = (s_i(h))_{h\in H}$. Note that we can equivalently define the set of strategies neglecting the histories where player i is not active: let

$$H_i = \{h \in H : |\mathcal{A}_i(h)| \geq 2\}$$

denote the set of histories (nodes of the game tree) where player i is active. Then $\bigtimes_{h\in H}\mathcal{A}_i(h)$ is isomorphic to $\bigtimes_{h\in H_i}\mathcal{A}_i(h)$ and it makes sense to write $S_i = \bigtimes_{h\in H_i}\mathcal{A}_i(h)$.

The set of **sub-strategies in the sub-tree with root** $h \in H$ is denoted by $S_i^{\succeq h}$, that is,

$$S_i^{\succeq h} = \bigtimes_{h'\in H(h)} \mathcal{A}_i(h').$$

A generic element of $S_i^{\succeq h}$ is denoted by $s_i^{\succeq h}$. The sub-strategy induced by $s_i \in S_i$ in the sub-tree with root h is denoted by

$$(s_i|h) = (s_i(h'))_{h'\in H(h)} \in S_i^{\succeq h}.$$

Given S, it is possible to derive a fictitious static game $G = \mathcal{N}(\Gamma)$, called the "strategic form," or "normal form" of Γ, that captures some important aspects of the game Γ.[16] The idea is the following: each player instructs a trustworthy agent, or codes a computer program for a machine, so that the agent/machine implements a particular strategy. Thus, each player commits to (*chooses*) a strategy, rather than just planning it in his mind. Each profile of strategies $s = (s_i)_{i\in I}$ induces a terminal history and an associated profile of payoffs. Of course, the players' choices of strategies are simultaneous (or, at least, no player can observe anything about the strategies chosen by the opponents before he chooses his own strategy).

[16] According to some theorists—but not ourselves—all the relevant aspects of a game are captured by this derived representation.

The strategic form of the Take-it-Or-Leave-it game of length $T = 4$ of Figure 9.2 (henceforth, TOL4) is represented in Figure 9.5. Strategies are labelled by sequences of letters separated by dots, with the following convention: the first letter denotes the action to be taken at the first history where the player is active ($\varnothing$ for player 1 and h = (Leave) for player 2), the second letter denotes the action to be taken at the second history where the player is active (h = (Leave, Leave) for player 1, h = (Leave, Leave, Leave) for player 2). For example, $s_1 = T.T$ is the strategy [Take; Take if (Leave, Leave)] of player 1.

$1\backslash 2$	$T.T$	$T.L$	$L.T$	$L.L$
$T.T$	$1,0$	$1,0$	$1,0$	$1,0$
$T.L$	$1,0$	$1,0$	$1,0$	$1,0$
$L.T$	$0,2$	$0,2$	$3,0$	$3,0$
$L.L$	$0,2$	$0,2$	$0,4$	$4,0$

Figure 9.5: Strategic form of ToL4.

Definition 50. *The* ***path function*** $\zeta : S \to Z$ *associates each strategy profile* $s \in S$ *with the induced terminal history* $\zeta(s) = (a^k)_{k=1}^{\ell}$ ($\ell \in \mathbb{N} \cup \{\infty\}$) *defined recursively as follows:*

$$\begin{aligned} a^1 &= (s_i(\varnothing))_{i \in I}, \\ a^{k+1} &= \left(s_i\left((a^j)_{j=1}^k\right)\right)_{i \in I} \end{aligned}$$

for every $k < \ell$ *such that* $(a^j)_{j=1}^k$ *is not terminal. The* ***strategic*** *or* ***normal form*** *of* Γ *is the static game*

$$\mathcal{N}(\Gamma) = \langle I, (S_i, U_i)_{i \in I} \rangle$$

where $U_i = u_i \circ \zeta : S \to \mathbb{R}$ *for every* $i \in I$.

As stated in Definition 50, "normal form" and "strategic form" are (for us) synonyms. We use mostly the latter because it is more self-explanatory, yet symbol $\mathcal{N}(\Gamma)$ refers to the former.

Consider the strategies $T.T$ and $T.L$ of player 1 in TOL4. Both of them prescribe to take one dollar immediately, an action that terminates the game. They differ only for the instruction corresponding to the history $h = (\text{Leave}, \text{Leave})$, that is, a history prevented by both strategies. These two strategies yield the same terminal history, (Take), independently of the strategy adopted by player 2. A similar observation holds for player 2: the two strategies $T.T$ and $T.L$ differ only for the instruction corresponding to the history $h = (\text{Leave}, \text{Leave}, \text{Leave})$, that both of them prevent. Which terminal history is reached depends on the strategy of player 1, but does not depend on which of these two strategies is implemented by player 2. This is the reason why these strategies correspond to identical rows (player 1) or columns (player 2).

These considerations motivate the following definitions. For every pair of histories $h, \bar{h} \in \bar{H}$ with $h \prec \bar{h}$, let $\alpha_i(h, \bar{h})$ denote the (unique) **action of i at h that allows $\bar{h}$**, that is, for each $a_i \in \mathcal{A}_i(h)$,

$$a_i = \alpha_i(h, \bar{h}) \Leftrightarrow (\exists a_{-i} \in \mathcal{A}_{-i}(h), (h, (a_i, a_{-i})) \preceq \bar{h}).$$

With this, the set of nonterminal **histories not prevented by** s_i can be defined as follows:

$$H_i(s_i) = \{\bar{h} \in H : \forall h \in H, h \prec \bar{h} \Rightarrow s_i(h) = \alpha_i(h, \bar{h})\}.$$

In words, $H_i(s_i)$ is the set of nonterminal histories $\bar{h}$ such that s_i selects the action leading to $\bar{h}$ at each predecessor $h \prec \bar{h}$.

For each strategy s_i, we can characterize the set $H_i(s_i)$ as follows. Formally, history h is not prevented, or is allowed by s_i if there is a profile of strategies of the opponents s_{-i} such that (s_i, s_{-i}) induces h. Since (s_i, s_{-i}) induces h if and only if it induces a terminal history preceded by h, we obtain the following characterization of the set of nonterminal histories allowed by s_i.

Lemma 25. $H_i(s_i) = \{\bar{h} \in H : \exists s_{-i} \in S_{-i}, \bar{h} \prec \zeta(s_i, s_{-i})\}$.

Proof. Fix $\bar{h}$ and suppose that $\bar{h} \prec \zeta(s_i, s_{-i})$ for some s_{-i}. Then, by definition of ζ, $\alpha_i(h, \bar{h}) = s_i(h)$ for each $h \prec \bar{h}$, which implies $\bar{h} \in H_i(s_i)$. This shows

$$\{\bar{h} \in H : \exists s_{-i} \in S_{-i}, \bar{h} \prec \zeta(s_i, s_{-i})\} \subseteq H_i(s_i).$$

To show that

$$H_i(s_i) \subseteq \{\bar{h} \in H : \exists s_{-i} \in S_{-i}, \bar{h} \prec \zeta(s_i, s_{-i})\},$$

pick $\bar{h} \in H_i(s_i)$ arbitrarily and construct a strategy profile $s_{-i} \in S_{-i}$ as follows: $s_j(h) = \alpha_j(h, \bar{h})$ for each $h \prec \bar{h}$ and $j \neq i$, otherwise $s_j(h)$ is arbitrarily chosen in $\mathcal{A}_j(h)$. Since $\bar{h} \in H_i(s_i)$, we have $\alpha_i(h, \bar{h}) = s_i(h)$ for each $h \prec \bar{h}$. Hence, by construction, $(s_j(h))_{j \in I} = (\alpha_j(h, \bar{h}))_{j \in I}$ for each $h \prec \bar{h}$. Therefore $\bar{h} \prec \zeta(s_i, s_{-i})$. ■

Definition 51. *Two strategies s_i and s_i' are **realization equivalent** if*

$$\forall s_{-i} \in S_{-i},\ \zeta(s_i, s_{-i}) = \zeta(s_i', s_{-i});$$

*they are **behaviorally equivalent** if*

$$H_i(s_i) = H_i(s_i') \text{ and } \forall h \in H_i(s_i), s_i(h) = s_i'(h).$$

In words, two strategies s_i and s_i' are realization-equivalent if, for every given strategy profile of the co-players, they induce the same terminal history. Strategies s_i and s_i' are behaviorally equivalent if they allow the same nonterminal histories and prescribe the same actions at such histories.

Remark 33. *For any fixed strategy s_i, the cardinality of the set of strategies behaviorally equivalent to s_i is $\prod_{h \in H \backslash H_i(s_i)} |\mathcal{A}_i(h)|$.*

Proof. By inspection of Definition 51, all the strategies s_i' behaviorally equivalent to s_i yield the same set of possible histories $H_i(s_i)$ and select action $s_i(h)$ at every history $h \in H_i(s_i)$, whereas the actions selected at all the other histories $h \in H \backslash H_i(s_i)$ can be chosen arbitrarily. The number of such arbitrary selections is $\prod_{h \in H \backslash H_i(s_i)} |\mathcal{A}_i(h)|$. ■

Lemma 26. *Two strategies are realization-equivalent if and only if they are behaviorally equivalent.*

Proof. We prove both directions of the statement by contraposition. Suppose first that s_i and s_i' are *not* realization equivalent, then there is some s_{-i} such that $\zeta(s_i, s_{-i}) \neq \zeta(s_i', s_{-i})$. We must show that s_i and s_i' are not behaviorally equivalent. Let $\bar{h}$ denote the maximal element of the chain of common predecessors of $\zeta(s_i, s_{-i})$ and $\zeta(s_i', s_{-i})$ (the chain is not empty because it contains the root $\varnothing$, and it is finite, hence it has a maximal element); this is the node at which the two paths "bifurcate." By Lemma 25, $\bar{h} \in H_i(s_i) \cap H_i(s'_i)$. The two strategy profiles (s_i, s_{-i}) and (s'_i, s_{-i}) specify the same action $s_j(\bar{h})$ at $\bar{h}$ for each co-player $j \neq i$; hence, it must be the case that $s_i(\bar{h}) \neq s'_i(\bar{h})$ (otherwise, there would be no "bifurcation" at $\bar{h}$). Therefore, strategies s_i and s'_i differ at a history not precluded by either one of them, which implies that they are not behaviorally equivalent.

Now suppose that s_i and s'_i are not behaviorally equivalent. To show that they are not realization-equivalent, we consider the following exhaustive cases, and we show that in both cases Lemma 25 implies that $\zeta\left(s_i, s_{-i}\right) \neq \zeta\left(s'_i, s_{-i}\right)$ for some s_{-i}:

1. $H_i(s_i) \neq H_i(s'_i)$,

2. $H_i(s_i) = H_i(s'_i)$ and $s_i(h) \neq s'_i(h)$ for some $h \in H_i(s_i)$.

In **case 1**, there is some $h \in (H_i(s_i) \backslash H_i(s'_i)) \cup (H_i(s'_i) \backslash H_i(s_i))$. By Lemma 25, if $h \in H_i(s_i) \backslash H_i(s'_i)$, there is some s_{-i} such that $h \prec \zeta(s_i, s_{-i})$ and $h \not\prec \zeta(s'_i, s_{-i})$, that is, $\zeta(s_i, s_{-i}) \in Z\left(h\right)$ and $\zeta(s'_i, s_{-i}) \in Z \backslash Z\left(h\right)$. Case $h \in H_i(s'_i) \backslash H_i(s_i)$ is analogous. Whatever the case, $\zeta(s_i, s_{-i}) \neq \zeta(s'_i, s_{-i})$ for some s_{-i}.

In **case 2**, there is $h \in H_i(s_i) = H_i(s'_i)$ such that $s_i(h) \neq s'_i(h)$. By Lemma 25 there is some s_{-i} such that $h \prec \zeta(s_i, s_{-i})$ and $h \prec \zeta(s'_i, s_{-i})$. Since $s_i(h) \neq s'_i(h)$, we have $\zeta(s_i, s_{-i}) \neq \zeta(s'_i, s_{-i})$. ■

In particular, the proof of Lemma 26 **(case 1**) shows that if s_i and s'_i are realization-equivalent they allow for the same histories, that is, $H_i\left(s_i\right) = H_i\left(s'_i\right) = H_i(s_i) \cap H_i(s'_i)$. Furthermore, from the first part of the proof we obtain another characterization of behavioral equivalence. We leave the proof as an exercise.

Remark 34. *Two strategies s_i and s'_i are behaviorally equivalent if and only if $s_i\left(h\right) = s'_i\left(h\right)$ for every $h \in H_i(s_i) \cap H_i(s'_i)$.*

The behavioral equivalence relation for player i corresponds to a partition of S_i (the quotient set of S_i with respect to the equivalence relation): two strategies belong to the same cell of the partition if and only if they are behaviorally equivalent.[17]

Definition 52. *A **reduced strategy** for player i is an element of the partition of S_i induced by the behavioral equivalence relation.*

We let $\mathbf{S}_i^r$ denote the set of reduced strategies of player i, and we let $\mathbf{S}^r = \times_{i\in I} \mathbf{S}_i^r$ denote the set of reduced strategy profiles. Note, each $\mathbf{s}_i^r \in \mathbf{S}_i^r$ is a *subset* $\mathbf{s}_i^r \subseteq S_i$ of behaviorally equivalent strategies. Note that, by Remark 33 (and Lemma 26), the cardinality of any reduced strategy $\mathbf{s}_i^r$ is

$$|\mathbf{s}_i^r| = \prod_{h\in H\setminus H_i(\mathbf{s}_i^r)} |\mathcal{A}_i(h)|,$$

where $H_i(\mathbf{s}_i^r) = H_i(s_i)$ for every $s_i \in \mathbf{s}_i^r$. By Lemma 26, we can define a reduced strategic-form path function $\zeta^r : \mathbf{S}^r \to Z$ as follows:

$$\zeta^r((\mathbf{s}_i^r)_{i\in I}) = \zeta((s_i)_{i\in I}) \text{ if } s_i \in \mathbf{s}_i^r \text{ for each } i \in I.$$

Given the payoff functions $(u_i : Z \to \mathbb{R})_{i\in I}$, we obtain the reduced-form payoff functions $(U_i^r = u_i \circ \zeta^r : \mathbf{S}^r \to \mathbb{R})_{i\in I}$, where $U_i^r(\mathbf{s}^r) = U_i(s)$ for each $\mathbf{s}^r \in \times_{j\in I} \mathbf{S}_j^r$ and $s \in \mathbf{s}^r$.

Definition 53. *The **reduced strategic** (or normal) **form** of a game Γ is the static game $\mathcal{N}^r(\Gamma) = \langle I, (\mathbf{S}_i^r, U_i^r)_{i\in I}\rangle$.*

The reduced strategic form of the TOL4 game is represented in Figure 9.6.

$1\backslash 2$	T	$L.T$	$L.L$
T	$1,0$	$1,0$	$1,0$
$L.T$	$0,2$	$3,0$	$3,0$
$L.L$	$0,2$	$0,4$	$4,0$

Figure 9.6: Reduced strategic form of ToL4.

[17] The quotient of a set S with respect to an equivalence relation $\approx$ on S is the collection of equivalence classes:

$$S/\approx := \{\mathbf{s} \subseteq S : \forall (s,t) \in S\times S, \{s,t\} \subseteq \mathbf{s} \Leftrightarrow s \approx t\}.$$

See Ok [64, A.1.3].

We illustrate the previous concepts and results with the BoS with a Dissipative Action of Example 42.

Example 44. We denote strategies as lists of actions separated by dots, we denote strategy sets (profiles) as lists (ordered lists) of strategies separated by commas:

$$\begin{aligned} S_a &= \{U, D\} \times \{u, d\} = \{U.u, U.d, D.u, D.d\}, \\ S_b &= \{N, B\} \times \{L, R\} \times \{l, r\} \\ &= \{N.L.l, N.R.l, N.L.r, N.R.r, B.L.l, B.R.l, B.L.r, B.R.r\}. \end{aligned}$$

The following are examples of actions leading from a history h to a successor $\bar{h}$:

$$\alpha_b(\varnothing, (B, (u, l))) = B,\ \alpha_a((B), (B, (u, l))) = u.$$

The set of nonterminal histories is

$$H = \{\varnothing, (N), (B)\},$$

and the set of nonterminal histories allowed by $s_b = N.L.r$ is

$$H_b(N.L.r) = \{\varnothing, (N)\}\,.$$

The sub-strategy induced by $s_b = N.L.r$ in the subgame with root $h = (B)$ is $s_b^{\succeq h} = r$;
The terminal history induced by strategy pair $(U.d, B.L.r)$ is

$$\zeta(U.d, B.L.r) = (B, (d, r)).$$

Finally, the following pairs of strategies of Bob are realization-equivalent:

$N.L.l$ and $N.L.r$ (reduced strategy $N.L$),
$N.R.l$ and $N.R.r$ (reduced strategy $N.R$),
$B.L.l$ and $B.R.l$ (reduced strategy $B.l$),
$B.L.r$ and $B.R.r$ (reduced strategy $B.r$);

therefore the reduced strategic form of the game is

$a\backslash b$	$N.L$	$N.R$	$B.l$	$B.r$
$U.u$	4,1	0,0	4,-1	0,-2
$U.d$	4,1	0,0	0,-2	1,2
$D.u$	0,0	1,4	4,-1	0,-2
$D.d$	0,0	1,4	0,-2	1,2

▲

A reduced strategy is sometimes called "plan of action" (see Rubinstein [74]). The reason is that "plan of action" corresponds to the intuitive idea that a player has to plan for all *external* contingencies, i.e., contingencies that not depend on his behavior, but he does not have to plan for contingences that cannot occur if he follows his plan. For example, in the TOL4 game, $T.T$ and $T.L$ are realization-equivalent and they correspond to the reduced strategy, or "plan" T, that is, "take at the first opportunity."

We avoid this terminology, and stick to the more neutral "reduced strategy," because there is a perfectly meaningful notion of "planning" that yields the specification of a whole strategy, not just a reduced strategy. To see this, consider the BoS with an Outside Option of Example 41. Suppose Ann, player 1, believes that, conditional on choosing in, the probability that Bob chooses B_2 is $\mu^1(\mathrm{B}_2|\mathrm{in}) = q$. Then Ann can compute a dynamically optimal plan with the following **folding back** procedure. First, she computes the values of choosing, respectively, B_1 and S_1 given in:

$$\begin{aligned} V_1^q(\mathrm{B}_1|\mathrm{in}) &= 3q, \\ V_1^q(\mathrm{S}_1|\mathrm{in}) &= 1-q. \end{aligned}$$

Next, assuming that she would maximize her expected payoff in the subgame if she chose in, Ann computes the value of in:

$$V_1^q(\mathrm{in}) = \max\{3q, 1-q\} = \begin{cases} 3q, & \text{if } q \geq \frac{1}{4}, \\ 1-q, & \text{if } q < \frac{1}{4}. \end{cases}$$

This means that Ann "plans" to choose B_1 (resp. S_1) in the subgame if $q > 1/4$ (resp. $q < 1/4$).[18] Finally, Ann compares $V_1^q(\mathrm{in})$ with the value

[18] If $q = 1/4$, then Ann is indifferent, and the optimal plan for the subgame is arbitrary.

of the outside option $V_1(\text{out}) = 2$ and "plans" to choose in (resp. out) if $V_1^q(\text{in}) > 2$ (resp. $V_1^q(\text{in}) < 2$), that is, if $q > 2/3$ (resp. $q < 2/3$). The upshot of all this is that, given her subjective beliefs q, Ann can use the folding back procedure to form a dynamically optimal plan, which corresponds to a *full* strategy, not a reduced strategy:[19]

$$s_1 = \begin{cases} \text{out.S}_1, & \text{if } q < \frac{1}{4}, \\ \text{out.B}_1, & \text{if } \frac{1}{4} < q < \frac{2}{3}, \\ \text{in.B}_1, & \text{if } q > \frac{2}{3}. \end{cases}$$

On the other hand, one may object that specifying a complete strategy is not necessary for rational planning: if q is not high enough, i.e., if $\max\{3q, 1-q\} < 2$, Ann has no need to plan ahead for the subgame, as she can see that it is not worth her while to reach it. Specifically, she can look forward for the payoff consequences of choosing out, in and B_1, or in and S_1, and her best "forward plan" is just out, a reduced strategy. Since both notions of rational planning—backward and forward—are meaningful and intuitive, we are not going to endorse only the second one by calling "plan of action" the reduced strategies. The concepts presented in Chapters 10, 11, and 12 further clarify why the notion of full strategy (as opposed to reduced strategy) is important in game theory. We are going to elaborate on the folding back procedure of dynamic programming in Chapter 10. While the folding back procedure refers to a *single* player and his arbitrarily given *subjective* beliefs, the standard equilibrium theory of multistage games looks for *"interpersonal"* elaborations of the folding back procedure, see Chapter 12. Yet, as we did in the part on static games, before moving on to standard equilibrium theory we analyze multistage versions of the rationalizability solution concept in Chapter 11. The distinction between full and reduced strategies will be dicussed again in that context.

We considered above two notions of equivalence between (full) strategies—behavioral and realization equivalence—that depend only on the game tree, and we observed in Lemma 26 that they are congruent, that is, for each $i \in I$, they correspond to the same subset of $S_i \times S_i$. Now we consider a weaker equivalence relation that depends on the whole game Γ as it does not rely on induced behavior, but rather on induced payoffs.

Definition 54. *Two strategies s_i and s_i' are **payoff equivalent**, written $s_i \sim_i s_i'$ if, for each strategy profile of the co-players, they yield the same*

[19] Again, we ignore ties.

profile of payoffs:

$$\forall s_{-i} \in S_{-i},\ (s_i \sim_i s_i') \Longleftrightarrow (\forall j \in I, U_j(s_i, s_{-i}) = U_j(s_i', s_{-i})).$$

Since $U_j = u_j \circ \zeta$ ($j \in I$), it is clear that realization equivalence implies payoff equivalence. Thus, Lemma 26 implies the following:

Remark 35. *If two strategies are behaviorally equivalent, then they are payoff-equivalent.*

In the games of Examples 41 and 43, as well as in many other games, there is no difference between payoff-equivalence and behavioral equivalence. A difference may arise only if the game features some ties between payoffs at distinct terminal histories, that is, if there are $z, z' \in Z$ such that $z \neq z'$ and yet $(u_i(z))_{i\in I} = (u_i(z'))_{i\in I}$. Such ties are "structural" when z and z' yield the same outcome, $g(z) = g(z')$, e.g., the same allocation of resources; otherwise, they are due to "non generic" ties between utility profiles. This happens if, for some $z, z' \in Z$, $g(z) \neq g(z')$ and $(v_i(g(z)))_{i\in I} = (v_i(g(z')))_{i\in I}$. Unfortunately, the outcome function g is mostly overlooked in game theory and all ties between payoffs at distinct terminal histories are called "non generic." Therefore, this terminology has to be taken with a grain of salt.

Comment on reduced strategic forms The most common definition of "reduced strategic form" refers to classes of *payoff*-equivalent strategies, rather than realization-equivalent strategies. This is fine if one is only interested in computing equilibria (or other solutions) of the strategic form. Instead, we are interested in the strategic form $\mathcal{N}(\Gamma)$ only as an auxiliary tool that (sometimes) helps analyzing Γ itself (as we will see in Section 9.3.1). Therefore we chose to emphasize the concept of realization-equivalence and reduced strategy, which is based on a meaningful notion of plan, and the corresponding concept of reduced strategic form.

9.3.1 Old Wine in New Bottles

Any solution concept for static games can be applied to the strategic (or normal) form $\mathcal{N}(\Gamma)$ of a multistage game Γ and thus yields a "candidate solution" for Γ. For example, we can find the rationalizable strategies or the Nash equilibria of $\mathcal{N}(\Gamma)$ and ask ourselves if they make sense as solutions of Γ.

Consider first the following very stylized **Entry Game**: a firm, player 1, has the opportunity to enter a (so far) monopolistic market. The incumbent, player 2, may fight the entry with a price war that damages both, or it may "acquiesce." The game is (summarily) represented in Figure 9.7.

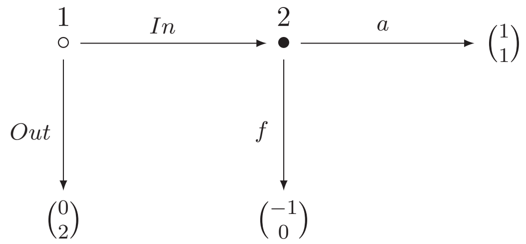


Figure 9.7: An "entry game."

The strategic form of this game is represented in the following table:

1\2	f	a
Out	0, 2	0, 2
In	-1, 0	1, 1

It is easily checked that every strategy profile of the strategic form of this game is rationalizable, and there are two Nash equilibria, (Out, f) and (In, a) (there is also a continuum of mixed equilibria where player 1 goes Out and 2 plays f with probability larger than $\frac{1}{2}$). Yet, it should be noted that—under complete information—if the potential entrant believes that the incumbent is rational (in an obvious sense), then she should expect acquiescence and therefore she should enter. So, the only "reasonable" solution should be (In, a).

Consider now the strategic form of the TOL4 game (Figure 9.5). Again, it is easily checked that every strategy profile is rationalizable, and there are multiple Nash equilibria, namely, all the pairs in the set $\{T.T, T.L\} \times \{T.T, T.L\}$. Note that all Nash equilibria "correspond" to the unique Nash equilibrium (T, T) of the reduced strategic form (Figure

9.6), and they induce the terminal history (Take). This should not be surprising, as strategies $T.T$ and $T.L$ are realization-equivalent. Indeed:

Observation 9. Payoff-equivalences (hence, also realization-equivalences) induce a trivial multiplicity of Nash equilibria*: if a strategy profile s^* is a Nash equilibrium, every profile $\bar{s}^*$ obtained from s^* by replacing some strategy s_i^* with an equivalent strategy $\bar{s}_i^*$ is also a Nash equilibrium. In other words, to find Nash equilibria we may just look at the equilibria $\mathbf{s}^{*,r}$ of reduced strategic form and then take into account that any profile s^* of the non-reduced form that corresponds to $\mathbf{s}^{*,r}$ is a Nash equilibrium of the multistage game. Similar considerations apply to strategic-form rationalizability: if a strategy s_i^* is rationalizable in $\mathcal{N}(\Gamma)$ every payoff-equivalent (hence, every realization-equivalent) strategy $\bar{s}_i^*$ is also rationalizable in $\mathcal{N}(\Gamma)$.*

The problem with rationalizability and Nash equilibrium of the strategic form is that such solution concepts require that players maximize their expected payoff on the path of play, i.e., along the play induced by the given strategy profile, but they also allow players to "plan" non maximizing actions at "off-path" histories. For instance, in the Nash equilibrium (Out, f) of the Entry Game player 2 plans to choose a non maximizing action given "off-path" history (In). Does this make sense?

In a multistage game, a player has initial beliefs which he updates as the play unfolds. Sometimes what he observes may be unexpected (i.e., have zero probability) according to his previous beliefs. In this case, he will not be able to pin down his new beliefs by updating his previous ones according to the rules of conditional probability, but he will still form new beliefs about his opponents' behavior. For example, suppose that in the TOL4 game of Figure 9.2 player 2 initially expects that player 1 will immediately take one dollar. If instead player 1 leaves the dollar on the table, player 2 will revise his beliefs about player 1, because he now knows that player 1 is implementing one of the following strategies: $L.T$ (Leave, then Take if also player 2 leaves) or $L.L$ (Leave, then Leave again if also player 2 leaves). The revised beliefs must assign zero probability to every other strategy of player 1.[20]

[20] In this informal discussion we are taking for granted that players actually implement the strategies they have in mind, and that this is "transparent." We will come back to this point in our analysis of solution concepts for multistage games.

We say that a player in a multistage game is *rational* if he would make expected utility maximizing choices given his (updated) beliefs *for every possible history of observed choices of his opponents.* Consider ToL4. Can one say that player 1 is irrational if he leaves three dollars on the table? No. Player 1 may hope that then player 2 will be "generous" and leave him four dollars. We might argue that such a belief is not very reasonable, indeed it is inconsistent with the rationality of player 2; yet, if player 1 had this belief, leaving three dollars on the table would be rational, i.e., expected utility maximizing.

In the analysis of static games, one can characterize with a solution concept the behavioral implications of the following assumptions: all players are rational and there is common belief of rationality. The solution concept is rationalizability. Furthermore, an action is rationalizable if and only if it is iteratively undominated.

What assumptions about rationality and strategic reasoning are worth considering in the case of multistage games? How should we extend the notion of rationalizability from static to multistage games so as to characterize the behavioral implications of such assumptions? How does such extension relate to rationalizability in the strategic form?

In Chapter 10 we analyze rationality, i.e., the implementation of strategies obtained by rational planning given conjectures about the behavior of other players. In Chapter 11 we introduce different forms of rationalizability for multistage games justified by different assumptions about strategic reasoning. Of course, these solution concepts have to coincide with rationalizability in the strategic form $\mathcal{N}(\Gamma)$ when the given game Γ has only one stage, and therefore is static. But in games with two or more stages they yield refinements of rationalizability in the strategic form, that is, some strategies that are rationalizable in the strategic form $\mathcal{N}(\Gamma)$ are deleted by the appropriate version of the rationalizability concept for multistage game Γ. Does this mean that the strategic form is useless? Not exactly: in many games of interest, notions of rationalizability for multistage games admit a useful characterization that relies only on the strategic form.

Consider first an elementary example, the Entry Game. It is pretty clear that if the players are rational in the sense specified above and if the first mover believes that also the opponent is rational, then the "reasonable" solution (In, a) obtains. Furthermore, (In, a) can be obtained in the strategic form of the game by first eliminating the *weakly* dominated strategies (just f in this case) and then eliminating the (strictly) dominated strategies (just *Out* in this case) of the residual strategic form. This procedure works in all two-stage games with perfect information with "no relevant ties."

Now consider more complex games—such as the BoS with an Outside Option—and the associated strategic forms. We will argue in Chapters 10 and 11 that, in many games of interest, **iterated admissibility**, i.e., the iterated deletion of all weakly dominated strategies,[21] coincides with **strong rationalizability**. The latter is a solution concept for multistage games which captures the following **best rationalization principle**: every player always ascribes to his opponents the highest degree of "strategic sophistication" consistent with their observed behavior. To illustrate this point, consider the BoS with an Outside Option (Figure 9.1). The highest degree of strategic sophistication that Bob can ascribe to Ann if he observes action "in" is that Ann is rational, because strategy in.B_1 is justifiable. On the other hand, strategy in.S_1 is not justifiable (indeed, it is dominated by "out"). Thus, at history (in) Bob should believe that Ann is rational and that she is going to choose B_1 in the subgame. The best reply to such belief is B_2. Anticipating this, Ann implements strategy in.B_1. Hence, the unique solution consistent with the best rationalization principle is (in.B_1, B_2).

It can be easily checked that (in.B_1, B_2) can be obtained *in the strategic form* of the game by iterated admissibility. Chapter 11 will provide a formal definition of strong rationalizability for multistage games with observed actions. Here we just note that iterated admissibility on the strategic form is a useful solution concept that yields in many games "reasonable" solutions capturing the best rationalization principle.

In Chapter 12 we show how to extend Nash's classical idea that players best respond to correct conjectures. This will be a rather direct application of the analysis of rational planning of Chapter 10. The resulting solution concept is called **subgame perfect** (Nash) **equilibrium (SPE)**, because it is characterized by the property of inducing a Nash equilibrium in every "subgame." For example, the Entry Game has only one SPE, (In, a), because this is the only strategy pair that is both a Nash equilibrium

[21] Iterated admissibility is a **maximal** iterated elimination procedure, that is, in each step *all* the weakly dominated strategies of *all* players are eliminated.

of the game and also induces an equilibrium of the (trivial) one-person game following history (In). The BoS with an Outside Option instead has two SPEs, $(\text{in.B}_1, \text{B}_2)$ and $(\text{out.S}_1, \text{S}_2)$, each one of them induces a Nash equilibrium of the BoS subgame following history (in). The standard way to compute SPEs relies on all the details of the description of the multistage game Γ.

What about self-confirming equilibrium (SCE)? Intuitively, a strategy profile s^* is a (pure) SCE if each s_i^* can be justified as a best reply to a conjecture about the other players that is confirmed by the evidence obtained ex post by i if everybody plays according to s^*. Of course, to give a definition of SCE we have to specify players' feedback, which can be described, for each player i, by a function $f_i : Z \to M_i$. Given the profile of feedback functions $f = (f_i)_{i \in I}$, we obtain a multistage game with feedback (Γ, f). Assume that players recall whatever they observe. In the context of multistage games with observed actions, it makes sense to assume that either f_i is one-to-one, because players observe *ex post* each other actions also at the end of the last stage, or—at least—player i remembers all the actions played in all stages but the last one, and remembers what he did in the last stage. Be that as it may, the key observation for SCE is that, unlike SPE, *we can analyze the essential aspects of SCE for multistage games with feedback* (Γ, f) *by looking at the strategic (or normal) form*

$$\mathcal{N}(\Gamma, f) = \left\langle I, (U_i, F_i)_{i \in I} \right\rangle = \left\langle I, (u_i \circ \zeta, f_i \circ \zeta)_{i \in I} \right\rangle,$$

where, for each player $i \in I$, $F_i : S \to M_i$ is the strategic-form feedback function of i. Specifically, any fixed strategy profile s^* is an SCE of (Γ, f) if and only if there is an SCE $\bar{s}^*$ of $\mathcal{N}(\Gamma, f)$ inducing the same terminal history, i.e., such that $\zeta(s^*) = \zeta(\bar{s}^*)$.

To get some intuition for this result, let Γ be the Entry Game and assume (reasonably in this case) that players observe ex post the terminal history; in particular, if player 1 goes In, he observes ex post whether 2 acquiesces or fights. We argued that, whatever the conjecture of player 2 about 1, the only reasonable "best reply" should be a. Thus, if player 1—knowing 2's payoff function—anticipates this, we obtain (In, a). Yet, *SCE does not rely on the idea that players reason strategically given common knowledge of the game*, they only have to best respond to confirmed justifying conjectures, given that each i knows his own payoff function u_i and feedback function f_i. Therefore, player 1 may be afraid of a fight and go Out, which makes his fears unverifiable. Hence, there are two (pure) SCE strategy pairs, (In, a) and (Out, a). The latter is supported by every conjecture of player 1 assigning at least 50% probability to a fight, because staying Out is a best reply and, given this, the conjecture—even if inaccurate—is trivially confirmed. The strategic form feedback is such that $F_i(Out, a) = Out = F_i(Out, f)$ for each i. Hence, there is a third (pure) SCE pair in the strategic form, the (imperfect) Nash equilibrium (Out, f) (player 2 is *ex ante* indifferent if he is certain of Out), but this additional strategic-form SCE yields the same terminal history as (Out, a). Note that the equality between Nash and SCE outcomes in this game is not a coincidence: indeed, one can show that, *in every two-person game with feedback where players observe ex post the terminal history, pure SCE and Nash outcomes coincide.*[22]

9.4 Randomized Strategies

In this section we introduce randomization in multistage games with observed actions. To simplify the probabilistic analysis we focus on *finite* games. In the context of multistage games (and more generally of all games with a sequential structure) one can think of two types of randomization:

(1) player i implements a pure strategy at random according to a probability measure $\sigma_i \in \Delta(S_i)$, which is called a **mixed strategy**;

(2) for each non terminal history $h \in H$, player i would choose an action at random according to a probability measure $\beta_i(\cdot|h) \in \Delta(\mathcal{A}_i(h))$; the array of probability measures $\beta_i = (\beta_i(\cdot|h))_{h \in H}$ is called **behavior strategy**.[23]

Note that *we can always regard a pure strategy as a degenerate randomized strategy*: s_i can be identified with the mixed strategy σ_i such that $\sigma_i(s_i) = 1$ and with the behavior strategy β_i such that $\beta_i(s_i(h)|h) = 1$ for every $h \in H$.

[22]The result extends to the mixed SCEs such that all pure strategies in the support can be justified by the same confirmed conjecture. See, e.g., Battigalli et al. [19].

[23]We use Kuhn's [55] original terminology. Some authors, e.g., Osborne and Rubinstein [65], say "behavioral strategy." Kuhn introduced the concept and the terminology in his analysis of general games with a sequential structure. Obviously, $\beta_i(\cdot|h)$ is non-trivial only at histories where i is active.

9.4.1 Equivalence of Mixed and Behavior Strategies

The literal interpretation of randomized strategies is that players spin roulette wheels or toss coins and let their actions be decided by the outcomes of such randomization devices. But this seems a bit farfetched, especially in the case of mixed strategies. Furthermore, agents who maximize expected utility have no use for randomization, as pure choices always do at least as well as randomized choices.

As in static games, a randomized strategy of i can be interpreted as a representation of the opponents' beliefs about the pure strategy of i. For example, suppose that game Γ is played by agents drawn at random from large populations; let $\sigma_i(s_i)$ be the fraction of agents in population i (the population of agents that may play in role i) that would implement strategy s_i. If j knows the statistical distribution σ_i, this will also represent the belief of j about the strategy of i. On the other hand, $\beta_i(a_i|h)$ may be interpreted as the conditional probability assigned by j to action a_i given history h.

This belief interpretation suggests a relationship between mixed and behavior strategies. Suppose that h has just realized; what does anyone learn about the pure strategy implemented by i? She learns that i is implementing a strategy that allows (does not prevent) h. Let $S_i(h)$ be the set of such strategies. Formally,[24]

$$S_i(h) = \{s_i \in S_i : \exists s_{-i} \in S_{-i}, h \prec \zeta(s_i, s_{-i})\}.$$

Next we define the set of strategies of i that allow h *and* select a_i at h:

$$S_i(h, a_i) = \{s_i \in S_i(h) : s_i(h) = a_i\}.$$

Now suppose that β_i is derived from the statistical distribution σ_i under the "*independence-across-players*" assumption that what is observed about the behavior of other players does not affect the beliefs about the strategy of i. Then the probability of a_i given h is just the fraction of agents in population i implementing a strategy that allows h and selects a_i at h, divided by the fraction of agents in population i implementing a strategy that allows h (if positive), that is,

$$\forall h \in H,\ \sigma_i(S_i(h)) > 0 \Rightarrow \beta_i(a_i|h) = \frac{\sigma_i(S_i(h, a_i))}{\sigma_i(S_i(h))} \qquad (9.4.1)$$

[24] If the game has chance moves the definition in the text must be adapted by including in s_{-i} also s_0, the "strategy" of the chance player.

(for every subset $X \subseteq S_i$, we write $\sigma_i(X) = \sum_{s_i \in X} \sigma_i(s_i)$). If $\sigma_i(S_i(h)) = 0$, $\beta_i(\cdot|h)$ can be specified *arbitrarily*. Note that (9.4.1) can be written in a more compact, but slightly less transparent form:

$$\forall h \in H,\ \beta_i(a_i|h)\sigma_i(S_i(h)) = \sigma_i(S_i(h, a_i)).$$

The formula above, or (9.4.1), says that σ_i and β_i are mutually consistent in the sense that they jointly satisfy a kind of chain rule for conditional probabilities.

Definition 55. *A mixed strategy $\sigma_i \in \Delta(S_i)$ and a behavior strategy $\beta_i = (\beta_i(\cdot|h))_{h\in H} \in \times_{h\in H} \Delta(\mathcal{A}_i(h))$ are mutually consistent if they satisfy (9.4.1).*

Observation 10. *If mixed strategy σ_i is such that $\sigma_i(S_i(h)) = 0$ for some h where player i is active, then there is a continuum of behavior strategies β_i consistent with σ_i. If $\sigma_i(S_i(h)) > 0$ for every h where i is active, then there is a unique β_i consistent with σ_i. If i is active at more than one history, then there is a continuum of mixed strategies σ_i consistent with any given behavior strategy β_i.*

The last statement in the observation can be understood with a counting-dimensionality argument. Let $H_i = \{h \in H : |\mathcal{A}_i(h)| \geq 2\}$ denote the set of histories where i is active. In a finite game, the number of elements of S_i is $|S_i| = \prod_{h\in H_i} |\mathcal{A}_i(h)|$, thus the dimensionality of $\Delta(S_i)$ is $|S_i| - 1 = \prod_{h\in H_i} |\mathcal{A}_i(h)| - 1$. On the other hand, the dimensionality of the set of behavior strategies is $\sum_{h\in H_i}(|\mathcal{A}_i(h)| - 1) = \sum_{h\in H_i} |\mathcal{A}_i(h)| - |H_i|$. It can be shown (by induction on the cardinality of H_i) that $\prod_{h\in H_i} |\mathcal{A}_i(h)| \geq \sum_{h\in H_i} |\mathcal{A}_i(h)|$ because $|\mathcal{A}_i(h)| \geq 2$ for each $h \in H_i$.[25] Thus the dimensionality of $\Delta(S_i)$ is higher than the dimensionality of $\times_{h\in H} \Delta(\mathcal{A}_i(h))$ if $|H_i| \geq 2$.

[25] First note that for every integer $L \geq 1$, we have $2^{L-1} \geq L$. (This can be easily proved by induction: it is trivially true for $L = 1$; suppose it is true for some $L \geq 1$, then $2^L = 2 \times 2^{L-1} \geq 2 \times L = L + L \geq L + 1$.)

Let $n_k \geq 2$ for each $k = 1, 2, ...$ We show that $\prod_{k=1}^{L} n_k \geq \sum_{k=1}^{L} n_k$ for each $L = 1, 2, ...$ Let $n^* = \max\{n_1, ..., n_\ell\}$. Then

$$\prod_{k=1}^{L} n_k \geq n^* \times 2^{L-1} \geq n^* \times L \geq \sum_{k=1}^{L} n_k.$$

Example 45. In the game "tree" in Figure 9.8, Rowena (player 1) has 4 strategies, $S_1 = \{\text{out}.u, \text{out}.d, \text{in}.u, \text{in}.d\}$; two of them, out.$u$ and out.d, are realization-equivalent and correspond to the reduced strategy out. The figure labels terminal histories: $v = (\text{out})$, $w = (\text{in}, (u, l))$ etc. The set of nonterminal histories is $H = \{\varnothing, (\text{in})\}$ (recall that $\varnothing$ denotes the initial history) and $S_1(\varnothing) = S_1$, $S_1(\text{in}) = \{\text{in}.u, \text{in}.d\}$.

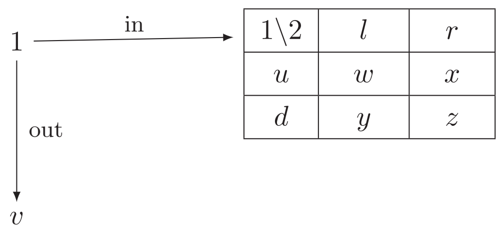


Figure 9.8: A game "tree."

Consider the following mixed strategy, parameterized by $p \in [0, 1]$:

$$\begin{aligned} \sigma_1^p(\text{out}.u) &= \frac{p}{2},\ \sigma_1^p(\text{out}.d) = \frac{1-p}{2}, \\ \sigma_1^p(\text{in}.u) &= \frac{1}{6},\ \sigma_1^p(\text{in}.d) = \frac{2}{6}. \end{aligned}$$

Since no history in H is ruled out by σ_1, there is only one behavior strategy consistent with σ_1:

$$\begin{aligned} \beta_1(\text{in}|\varnothing) &= \frac{\sigma_1(S_1(\text{in}))}{\sigma_1(S_1)} = \frac{1}{6} + \frac{2}{6} = \frac{1}{2}, \\ \beta_1(u|\text{in}) &= \frac{\sigma_1(S_1(\text{in}.u))}{\sigma_i\,(S_1(\text{in}))} = \frac{\frac{1}{6}}{\frac{1}{6}+\frac{2}{6}} = \frac{1}{3}. \end{aligned}$$

Note also that β_1 does not depend on the distribution $p : 1-p$ of the probability mass $\sigma_1^p(S_1(\text{out})) = \frac{1}{2}$ between the realization-equivalent strategies out.u and out.d. The reason is that the only difference between these two strategies is the choice between u and d contingent on in, a counterfactual contingency if Rowena plans to go out. Now consider the mixed strategy $\overline{\sigma}_1^p$ with $\overline{\sigma}_1^p(\text{out}.u) = p$, $\overline{\sigma}_1^p(\text{out}.d) = 1-p$, which amounts to the deterministic plan of going out. Since $\overline{\sigma}_1^p(S_1(\text{in})) = 0$, there is a continuum of behavior strategies consistent with $\overline{\sigma}_1^p$, all those in the set

$$\{\overline{\beta}_1 \in \Delta(\{\text{out}, \text{in}\}) \times \Delta(\{u, d\}) : \overline{\beta}_1(\text{out}|\varnothing) = 1\}.$$

▲

Next suppose that player i does indeed randomize according to behavior strategy β_i. Is there a method to obtain a mixed strategy σ_i consistent with β_i? Here is an answer: It is natural, although not entirely obvious, to consider the case where the randomization devices that i would use at different histories are stochastically independent. We call this assumption "*independence across agents*," as one can regard the contingent choice of i at each $h \in H$ as carried out by an agent (i, h) of i who only operates under contingency h. A pure strategy s_i is just a collection $(s_i(h))_{h \in H}$ of contingent choices of the agents (i, h), $h \in H$. If the random choices at different histories are mutually independent, then the probability of pure strategy s_i is the product of the probabilities of the contingent choices $s_i(h)$, $h \in H$. Therefore one can derive from β_i the mixed strategy σ_i that satisfies

$$\forall s_i \in S_i,\ \sigma_i(s_i) = \prod_{h \in H} \beta_i(s_i(h)|h). \tag{9.4.2}$$

It is boring, but rather straightforward to show that such σ_i is consistent with β_i:[26]

Lemma 27. *For all $i \in I$, $\sigma_i \in \Delta(S_i)$, $\beta_i \in \times_{h \in H} \Delta(\mathcal{A}_i(h))$, if (9.4.2) holds, then also (9.4.1) holds, that is, σ_i and β_i are mutually consistent.*

Example 46. In the game "tree" of Figure 9.8, the mixed strategy associated with $\beta_1(\text{in}|\varnothing) = \beta_1(\text{out}|\varnothing) = \frac{1}{2}$, $\beta_1(u|\text{in}) = \frac{1}{3}$, $\beta_1(d|\text{in}) = \frac{2}{3}$ under the "independence-across-agents" assumption is σ_1:

$$\begin{aligned} \sigma_1(\text{out}.u) &= \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6},\ \sigma_1(\text{out}.d) = \frac{1}{2} \cdot \frac{2}{3} = \frac{2}{6}, \\ \sigma_1(\text{in}.u) &= \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6},\ \sigma_1(\text{in}.d) = \frac{1}{2} \cdot \frac{2}{3} = \frac{2}{6}. \end{aligned}$$

[26] This is one part of what is known as Kuhn [55] theorem for mixed and behavior strategies. For a proof, see the appendix.

Note that β_1 was derived in Example 45 from an arbitrary mixed strategy with the parametric representation σ_1^p, whereas here we derived exactly $\sigma_1 = \sigma_1^{1/3}$. The reason is that only σ_1^p with $p = \frac{1}{3}$ is consistent with "independence across agents." ▲

Observation 10 says that many behavior strategies may be consistent with a given mixed strategy and that many mixed strategies may be consistent with a given behavior strategy, although only one satisfies the assumption of "independence across agents." Then why is consistency between mixed and behavior strategies important? Intuitively, the reason is that mutually consistent mixed and behavior strategies yield the same probabilities of histories, which is all that matter to expected payoff maximizing players. To make the intuition precise, note that a profile of randomized strategies ξ ($\xi = \sigma = (\sigma_i)_{i\in I}$ or $\xi = \beta = (\beta_i)_{i\in I}$) induces a probability measure $\hat{\zeta}(\cdot|\xi) \in \Delta(Z)$ on terminal histories, which is determined by the following formulas (which implicitly assume "independence across players"):

$$\forall z \in Z,\, \hat{\zeta}(z|\sigma) = \sum_{s:\zeta(s)=z} \left(\prod_{i\in I} \sigma_i(s_i)\right),$$

$$\forall z \in Z,\, \hat{\zeta}(z|\beta) = \prod_{k=1}^{\ell(z)} \prod_{i\in I} \beta_i(\mathbf{a}_i^k(z)|\mathbf{h}^{k-1}(z)).$$

(In the second formula, $\mathbf{a}_i^k(z)$ is the action played by i at stage k in history z, and $\mathbf{h}^{k-1}(z)$ is the prefix of length $k-1$ of z.)

Theorem 37. *(Kuhn)*[27] *For any two profiles of mixed and behavior strategies σ and β the following statements hold:*
(a) Fix any $i \in I$; if σ_i and β_i satisfy either (9.4.1) or (9.4.2), then

$$\forall s_{-i} \in S_{-i},\, \hat{\zeta}(\cdot|\sigma_i, s_{-i}) = \hat{\zeta}(\cdot|\beta_i, s_{-i}).$$

[27] In his seminal article ([55]), Harold Kuhn proved two important results about general games with a sequential structure (so called "extensive-form games"), one concerns the realization-equivalence of mixed and behavior strategies under the assumption of perfect recall (a generalization of the observed actions assumption), the other concerns the existence of equilibria.

(b) If, for all $i \in I$, σ_i and β_i satisfy either (9.4.1) or (9.4.2), then

$$\hat{\zeta}(\cdot|\sigma) = \hat{\zeta}(\cdot|\beta).$$

Example 47. Again, the example of Figure 9.8 illustrates. For Rowena (player 1) consider the randomized strategies σ_1^p and β_1 of Example 45. Colin (player 2) is active at only one history, hence there is an obvious isomorphism between his mixed and behavior strategies. Let $\sigma_2(l) = \beta_2(l|\text{in}) = q$. Then $\hat{\zeta}(\cdot|\beta_1, l) = \hat{\zeta}(\cdot|\sigma_1, l)$, $\hat{\zeta}(\cdot|\beta_1, r) = \hat{\zeta}(\cdot|\sigma_1, r)$ and $\hat{\zeta}(\cdot|\beta) = \hat{\zeta}(\cdot|\sigma)$. In particular

$$\begin{aligned}
\hat{\zeta}(v|\beta) &= \frac{1}{2} = \hat{\zeta}(v|\sigma),\ \hat{\zeta}(w|\beta) = \frac{q}{6} = \hat{\zeta}(w|\sigma),\ \hat{\zeta}(x|\beta) = \frac{1-q}{6} = \hat{\zeta}(x|\sigma), \\
\hat{\zeta}(y|\beta) &= \frac{q}{3} = \hat{\zeta}(y|\sigma),\ \hat{\zeta}(z|\beta) = \frac{1-q}{3} = \hat{\zeta}(z|\sigma).
\end{aligned}$$

▲

9.5 Appendix

Proof of Lemma 27. Fix i, σ_i, β_i, $\hat{h} \in H$ and $\hat{a}_i \in \mathcal{A}_i(\hat{h})$ (it is notationally convenient to "put a hat" on the fixed history and action). Suppose that (9.4.2) holds; it must be shown that $\sigma_i(S_i(\hat{h})) > 0$ implies $\beta_i(\hat{a}_i|\hat{h}) = \sigma_i(S_i(\hat{h}, \hat{a}_i))/\sigma_i(S_i(\hat{h}))$. For every $h \prec \hat{h}$, let $\hat{\mathbf{a}}_i(h)$ denote the action taken by i at h to reach $\hat{h}$ from h (in other words, $\hat{\mathbf{a}}(h) = (\hat{\mathbf{a}}_j(h))_{j\in I}$ is the unique action profile such that $(h, \hat{\mathbf{a}}(h)) \preceq \hat{h}$). Then $S_i(\hat{h}) = \{s_i : \forall h \prec \hat{h}, s_i(h) = \hat{\mathbf{a}}_i(h)\}$ and $S_i(\hat{h}, \hat{a}_i) = \{s_i : s_i(\hat{h}) = \hat{a}_i, \forall h \prec \hat{h}, s_i(h) = \hat{\mathbf{a}}_i(h)\}$, so that, for each $s_i \in S_i(\hat{h})$,

$$\prod_{h\in H} \beta_i(s_i(h)|h) = \left(\prod_{h\in H: h\prec\hat{h}} \beta_i(\hat{\mathbf{a}}_i(h)|h)\right) \cdot \left(\prod_{h\in H: h\not\prec\hat{h}} \beta_i(s_i(h)|h)\right),$$

and for each $s_i \in S_i(\hat{h}, \hat{a}_i)$,

$$\prod_{h\in H} \beta_i(s_i(h)|h) = \left(\prod_{h\in H: h\prec\hat{h}} \beta_i(\hat{\mathbf{a}}_i(h)|h)\right) \cdot \beta_i(\hat{a}_i|\hat{h}) \cdot \left(\prod_{h\in H: h\not\preceq\hat{h}} \beta_i(s_i(h)|h)\right).$$

Therefore,

$$\begin{aligned}\sigma_i(S_i(\hat{h})) &= \sum_{s_i\in S_i(\hat{h})}\sigma_i(s_i) = \sum_{s_i\in S_i(\hat{h})}\prod_{h\in H}\beta_i(s_i(h)|h)\\ &= \left(\prod_{h\in H:h\prec\hat{h}}\beta_i(\hat{\mathbf{a}}_i(h)|h)\right)\cdot\left(\sum_{s_i\in S_i(\hat{h})}\prod_{h\in H:h\not\preceq\hat{h}}\beta_i(s_i(h)|h)\right),\end{aligned}$$

$$\begin{aligned}\sigma_i(S_i(\hat{h},\hat{a}_i)) &= \sum_{s_i\in S_i(\hat{h},a_i)}\sigma_i(s_i) = \sum_{s_i\in S_i(\hat{h},a_i)}\prod_{h\in H}\beta_i(s_i(h)|h)\\ &= \left(\prod_{h\in H:h\prec\hat{h}}\beta_i(\hat{\mathbf{a}}_i(h)|h)\right)\cdot\beta_i(\hat{a}_i|\hat{h})\cdot\left(\sum_{s_i\in S_i(\hat{h},\hat{a}_i)}\prod_{h\in H:h\not\preceq\hat{h}}\beta_i(s_i(h)|h)\right).\end{aligned}$$

Count the partial histories that do not weakly precede $\hat{h}$ in any order, e.g., $\{h\in H : h\not\preceq\hat{h}\} = \{h_1,...,h_L\}$. Now note that there is a canonical bijection between $S_i(\hat{h})$ and $\mathcal{A}_i(\hat{h})\times\left(\bigtimes_{k=1}^{L}A_i(h_k)\right)$, and between $S_i(\hat{h},\hat{a}_i)$ and $\bigtimes_{k=1}^{L}\mathcal{A}_i(h_k)$. Given this, we can re-order terms as follows:

$$\begin{aligned}\sum_{s_i\in S_i(\hat{h},\hat{a}_i)}\prod_{h\in H:h\not\preceq\hat{h}}\beta_i(s_i(h)|h) &= \sum_{k=1}^{L}\sum_{a_{i,k}\in\mathcal{A}_i(h_k)}\beta_i(a_{i,k}|h_k) = 1,\\ \sum_{s_i\in S_i(\hat{h})}\prod_{h\in H:h\not\preceq\hat{h}}\beta_i(s_i(h)|h) &= \sum_{a_i\in\mathcal{A}_i(\hat{h})}\beta_i(a_i|\hat{h})\sum_{k=1}^{L}\sum_{a_{i,k}\in\mathcal{A}_i(h_k)}\beta_i(a_{i,k}|h_k) = 1.\end{aligned}$$

Then

$$\begin{aligned}\sigma_i(S_i(\hat{h})) &= \prod_{h\in H:h\prec\hat{h}}\beta_i(\hat{\mathbf{a}}_i(h)|h),\\ \sigma_i(S_i(\hat{h},\hat{a}_i)) &= \left(\prod_{h\in H:h\prec\hat{h}}\beta_i(\hat{\mathbf{a}}_i(h)|h)\right)\cdot\beta_i(\hat{a}_i|\hat{h}),\end{aligned}$$

and

$$\beta_i(\hat{a}_i|\hat{h})\sigma_i(S_i(\hat{h})) = \beta_i(\hat{a}_i|\hat{h})\left(\prod_{h\in H:h\prec\hat{h}}\beta_i(\hat{\mathbf{a}}_i(h)|h)\right) = \sigma_i(S_i(\hat{h},\hat{a}_i)).$$

■

10

Rational Planning

In this chapter we analyze rational planning from the perspective of a single player with a subjective probabilistic conjecture about the behavior of co-players. We introduce and analyze several dynamic programming properties for strategies in finite multistage games with observed actions. In particular, we focus on **folding-back optimality** as a representation of rational planning: the decision maker computes his subjectively optimal strategy starting from the last stage of the game and factoring into the decision problems of earlier stages the expected payoffs computed for the later stages. One-Step Optimality is a property of a strategy given the conjecture: at every history, the decision maker has no incentive to change the prescribed action, keeping fixed the actions prescribed for the future stages. Sequential Optimality takes a continuation-plan approach: at every history, the continuation-strategy maximizes the player's expected payoff conditional on reaching the history. We show that folding-back optimality is equivalent to One-Step Optimality (*Folding-Back Principle*) and to Sequential Optimality (*Optimality Principle*). It follows that Sequential Optimality is equivalent to One-Step Optimality (*One-Deviation Principle*). In light of these equivalence results, we adopt Sequential Optimality for at least one conjecture as a notion of justifiability for multistage games. Extending the analogous result obtained for static games (Lemma 2 of Chapter 3), we characterize justifiability with the notion of **conditional dominance**, according to which a strategy s_i is conditionally dominated if there exists a history h (consistent with s_i) such that s_i is dominated conditional on observing h. To simplify the analysis, we assume that *the game tree is finite*, that is, there is finite horizon and the cardinality of each feasible set of actions is finite. This maintained assumption is not explicitly mentioned in the results below. The finiteness assumption is removed in Section 10.5, where we show that the One-Deviation Principle also holds for all games with *infinite horizon* that satisfy a regularity property called "continuity at infinity."

10.1 Conditional Beliefs and Decision Trees

In the analysis of best replies in static games (Chapter 3), we modeled players' conjectures as probability measures over the actions of the co-players. There are at least two conceivable ways to extend the definition of conjecture to multistage games: (i) Define conjectures as (possibly correlated) probability measures over the strategies of the co-players, where such strategies are interpreted as descriptions of how co-players (would) behave at each nonterminal history. Note that conjectures have to be updated, or revised, as the play unfolds. (ii) Define conjectures as arrays of (possibly correlated) probability measures over co-players' feasible actions sets, one such measure for each nonterminal history. These two approaches are essentially equivalent. To see this, consider the easier case of two-person games, where there is only one co-player. Return to Section 9.4: An array of probability measures $\beta^i \in \bigtimes_{h\in H} \Delta(\mathcal{A}_{-i}(h))$ can also be interpreted as a behavior strategy of co-player $-i$. By Theorem 37, β^i is realization equivalent to a probability measure $\mu^i \in \Delta(S_{-i})$. Now consider any $\mu^i \in \Delta(S_{-i})$ such that every nonterminal history can be reached with positive probability, that is, $\mu^i(S_{-i}(h)) > 0$ for every $h \in H$. Then we can recover a realization-equivalent β^i letting

$$\beta^i(a_{-i}|h) = \frac{\mu^i(S_{-i}(h, a_{-i}))}{\mu^i(S_{-i}(h))}$$

for all $h \in H$ and $a_{-i} \in \mathcal{A}_{-i}(h)$.[1] If instead $\mu^i(S_{-i}(h)) = 0$ for some h, we assume that upon observing h player i forms a new conjecture, which can be used to derive $\beta^i(\cdot|h)$. Thus, we obtain $\beta^i(\cdot|h)$ for every $h \in H$.

[1] Recall from Section 9.4 of Chapter 9 that $S_j(h, a_j)$ is the set of strategies of player j consistent with h that select action a_j at h. Here, player j is i's opponent, that is, $j = -i$.

We omit the details.[2]

To analyze rational planning, it is easier to model conjectures as arrays of probability measures over co-players' feasible actions. Arrays of probability measures over strategies will be more convenient to analyze how conjectures are formed through strategic reasoning—we will introduce them in Section 10.4 before using them in Chapter 11. From now on, we reserve the term **conjecture** for a system of conditional beliefs about co-players' actions

$$\beta^i = (\beta^i(\cdot|h))_{h\in H} \in \bigtimes_{h\in H} \Delta(\mathcal{A}_{-i}(h)),$$

where $\beta^i(a_{-i}|h)$ denotes the subjective probability of a_{-i} conditional on h. In this section, *we focus on the subjectively optimal behavior of player i given his conjectures*, hence we *fix* β^i. Nonetheless, we make the dependence on β^i explicit, because conjectures will be required to be consistent with the strategic analysis of the game. With multiple co-players, β^i is a kind of "correlated behavior strategy."

Let $\beta_i^{s_i}$ denote the degenerate behavior strategy determined by pure strategy s_i according to the obvious rule:

$$\forall h \in H, \forall a_i \in \mathcal{A}_i(h),\ \beta_i^{s_i}(a_i|h) = 1 \Leftrightarrow a_i = s_i(h).$$

The probability of action profile $a = (a_i, a_{-i})$ conditional on h given strategy s_i and conjecture β^i is

$$\mathbb{P}_{s_i,\beta^i}(a|h) := \beta_i^{s_i}(a_i|h)\beta^i(a_{-i}|h) = \begin{cases} 0, & \text{if } a_i \neq s_i(h), \\ \beta^i(a_{-i}|h), & \text{if } a_i = s_i(h). \end{cases} \tag{10.1.1}$$

Equation (10.1.1) can be interpreted in two ways. The first is that player i will certainly implement his plan s_i. The second is that s_i is an objective description of how player i would behave conditional on reaching each history h. In this chapter we adopt the first interpretation, from the viewpoint of player i: he is certain that he is going to implement his plan s_i; therefore, for the continuation of the game, he assigns probability 1 to the

[2] The key assumption is that player i updates or revises his beliefs in compliance with the chain rule of conditional probabilities. See below.

actions precribed by s_i. Using the **chain rule** of conditional probabilities,[3] we can define the subjective probability of h' conditional on prefix h ($h \prec h'$), given that s_i is played from h onward: let $h = (a^1, ..., a^{\ell(h)})$, $h' = (a^1, ..., a^{\ell(h)}, ..., a^{\ell(h')})$; then[4]

$$\mathbb{P}_{s_i,\beta^i}(h'|h) = \prod_{t=\ell(h)+1}^{\ell(h')} \beta_i^{s_i}(a_i^t|h, ..., a^{t-1})\beta^i(a_{-i}^t|h, ..., a^{t-1}). \qquad (10.1.2)$$

Recall from Section 9.3 that $S_i^{\succeq h}$ is the set of continuation-strategies in the subgame with root h, and that $(s_i|h) \in S_i^{\succeq h}$ is the continuation-strategy implied by s_i (that is, the projection of s_i onto $S_i^{\succeq h}$). With this, note that the above conditional probability is well defined and meaningful *even if s_i precludes h*, because $\mathbb{P}^{s_i,\beta^i}(h'|h)$ only depends on $(s_i|h) \in S_i^{\succeq h}$, i.e., it depends only on how s_i behaves in the sub-tree with root h. The subjective unconditional probability of reaching h' given s_i is

$$\mathbb{P}_{s_i,\beta^i}(h') = \mathbb{P}_{s_i,\beta^i}(h'|\varnothing).$$

Comment on conjectures If $-i$ is just one player, then β^i—as a mathematical object—is a behavior strategy. For example, $-i = 0$ may be the Chance player,[5] in this case $\beta^i = \pi_0$ represents the objective probabilities of chance moves in a decision tree; or $-i = 0$ may be Nature[6] and then $\beta^i = \beta_0$ represents the subjective conditional probabilities assigned by i to moves by Nature. Such subjective probabilities may be obtained *via* updating from some prior $\mu_0 \in \Delta\left(\times_{h\in H} \Delta(\mathcal{A}_0(h))\right)$ on objective probability models: for example, if supp$\mu_0 = \{\pi_0^1, ..., \pi_0^K\}$ and

[3] The chain rule of conditional probabilities states that, given a chain of three events $D \subseteq E \subseteq F$,

$$\mathbb{P}(D|F) = \mathbb{P}(D|E)\mathbb{P}(E|F).$$

Here we can look at chains of events in Z corresponding to chains of histories: if $h \prec h' \prec h''$, then $Z(h'') \subseteq Z(h') \subseteq Z(h)$.

[4] Equation (10.1.2) and other equations in the continuation of this chapter contain the following abuse of notation: When $t = \ell(h) + 1$, we have $a^{t-1} = a^{\ell(h)}$, which is already included in h, therefore $h, ..., a^{t-1}$ shall be read as just h.

[5] A pseudo-player choosing some moves with known objective probabilities.

[6] A pseudo-player selecting a state of nature with unknown objective probabilities.

no h is ruled out by μ_0,[7] then

$$\forall h \in H,\ \forall a_0 \in \mathcal{A}_0(h),\ \beta_0(a_0|h) = \frac{\sum_{k=1}^K \mathbb{P}_{s_i,\pi_0^k}(h,(s_i(h),a_0))\mu_0(\pi_0^k)}{\sum_{k=1}^K \mathbb{P}_{s_i,\pi_0^k}(h)\mu_0(\pi_0^k)}$$

for every s_i consistent with h.[8]

The structure $\langle(\bar{H},\preceq),u_i,\beta^i\rangle$ forms a **subjective decision tree for player** i: $(\bar{H},\preceq)$ is the tree, $u_i : Z \to \mathbb{R}$ is the payoff function of i, and β^i is a subjective assessment of the probabilities of actions not controlled by i.

Example 48. The BoS with Dissipative Action is a two-person game; therefore, a conjecture of player i—as a mathematical object—corresponds to a behavior strategy of the co-player $-i$. We will focus on the following examples of conditional beliefs (they are not meant to form an equilibrium):

$$\begin{aligned}
\beta^b &: \quad \beta^b(U|N) = \frac{3}{4},\ \beta^b(u|B) = \frac{1}{4};\\
\beta^a &: \quad \beta^a(N|\varnothing) = 1,\ \beta^a(L|N) = 0,\ \beta^a(l|B) = 1;
\end{aligned}$$

hence β^a corresponds to the pure strategy $s_b = N.R.l$ of Bob. Next we give a few examples of realization probabilities, where s_b, β^b and β^a are specified above:

$$\begin{aligned}
\mathbb{P}_{s_b,\beta^b}(B,(u,l)) &= 0 = \mathbb{P}_{s_b,\beta^b}(N,(U,L)),\\
\mathbb{P}_{s_b,\beta^b}(N,(U,R)) &= \frac{3}{4},\\
\mathbb{P}_{s_b,\beta^b}(B,(u,l)) &= 0,\ \mathbb{P}_{D.u,\beta^a}(B,(u,l)|B) = 1.
\end{aligned}$$

▲

[7] That is, $\sum_{k=1}^K \mathbb{P}^{s_i,\pi_0^k}(h)\mu_0(\pi_0^k) > 0$ for every $h \in H$ and s_i consistent with h. This is true, for example, if there is at least one model $\pi_0^k \in$suppμ_0 such that $\pi_0^k(a_0|\bar{h}) > 0$ for every $\bar{h} \in H$ and $a_0 \in \mathcal{A}_0(\bar{h})$.

[8] Recall that strategy s_i is consistent with history h if $h \preceq \zeta(s_i,s_{-i})$ for some $s_{-i} \in S_{-i}$. The set of such strategies is denoted by $S_i(h)$ (see Section 9.4).

10.2 Subjective Values

For all $s_i \in S_i$, $\beta^i \in \times_{h\in H} \Delta(\mathcal{A}_{-i}(h))$ and $h \in H$, we can determine the subjective **value of** h given that s_i is followed from h onward:

$$V_i^{s_i,\beta^i}(h) := \sum_{z\in Z(h)} u_i(z)\mathbb{P}_{s_i,\beta^i}(z|h).$$

Like $\mathbb{P}_{s_i,\beta^i}(z|h)$, also $V_i^{s_i,\beta^i}(h)$ is well defined and meaningful even if s_i precludes h, because it depends only on $(s_i|h)$, the sub-strategy induced by s_i in the sub-tree with root h. Therefore, it makes sense to write (with a slight abuse of notation) $V_i^{s_i,\beta^i}(h) = V_i^{(s_i|h),\beta^i}(h)$, and to write $V_i^{t_i,\beta^i}(h)$ for any $t_i \in S_i^{\succeq h}$; that is, $V_i^{t_i,\beta^i}(h)$ is the value of playing sub-strategy t_i in the decision sub-tree with root h.

Similarly, for every $a_i \in \mathcal{A}_i(h)$ we can define the **value of** taking **action** a_i at h given that s_i will be followed from the next stage:

$$V_i^{s_i,\beta^i}(h,a_i) := \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \beta^i(a_{-i}|h)V_i^{s_i,\beta^i}(h,(a_i,a_{-i})).$$

Finally, the **ex ante value** of playing s_i is

$$V_i^{s_i,\beta^i}(\varnothing) := \sum_{z\in Z} u_i(z)\mathbb{P}_{s_i,\beta^i}(z).$$

Note: since $\beta_i^{s_i}(a_i|h) = 1$ if $a_i = s_i(h)$ and $\beta_i^{s_i}(a_i|h) = 0$ otherwise, by inspection of the definitions we have

$$\begin{aligned}
& V_i^{s_i,\beta^i}(h,s_i(h)) \\
= & \sum_{(a_i,a_{-i})\in\mathcal{A}(h)} V_i^{s_i,\beta^i}(h,(a_i,a_{-i}))\beta_i^{s_i}(a_i|h)\beta^i(a_{-i}|h) \\
= & \sum_{(a_i,a_{-i})\in\mathcal{A}(h)} \beta_i^{s_i}(a_i|h)\beta^i(a_{-i}|h) \sum_{z\in Z(h,(a_i,a_{-i}))} u_i(z)\mathbb{P}_{s_i,\beta^i}(z|h,(a_i,a_{-i})) \\
= & \sum_{z\in Z(h)} u_i(z)\mathbb{P}_{s_i,\beta^i}(z|h) = V_i^{s_i,\beta^i}(h)
\end{aligned}$$

(the second to last equality follows from the fact that $Z(h) = \bigcup_{a\in\mathcal{A}(h)} Z(h,a)$ and from the chain rule). Thus we obtain:

Remark 36. *For every $s_i \in S_i$ and $h \in H$, $V_i^{s_i,\beta^i}(h,s_i(h)) = V_i^{s_i,\beta^i}(h)$.*

10.3 Rational Planning

The values defined above take a particular strategy as given. Next we define recursively the subjective value of reaching a history h and of taking an action a_i at h, *under the presumption that the behavior of player i will be subjectively rational in the following stages.* We use the symbol $\hat{V}_i^{\beta^i}$ to denote such values to emphasize that they are optimal given conjecture β^i.

The recursion is based on $L(\Gamma(h))$, the height of the subgame starting at $h \in \bar{H}$:[9]

- If $L(\Gamma(h)) = 0$ (that is, $h \in Z$) let

$$\hat{V}_i^{\beta^i}(h) = u_i(h).$$

- Suppose $\hat{V}_i^{\beta^i}(h')$ has been defined for every h' with $L(\Gamma(h')) \leq k$. Then if $L(\Gamma(h)) = k + 1$ let

$$\hat{V}_i^{\beta^i}(h, a_i) = \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \hat{V}_i^{\beta^i}(h, (a_i, a_{-i}))\beta^i(a_{-i}|h),$$

$$\hat{V}_i^{\beta^i}(h) = \max_{a_i \in \mathcal{A}_i(h)} \hat{V}_i^{\beta^i}(h, a_i). \quad (10.3.1)$$

This determination of optimal values by backward recursion is often called "**folding back**" in the literature on dynamic programming. We can interpret folding back as a method (an algorithm) to compute an optimal plan, which "collects" actions that solve the maximization problems of equation (10.3.1) at every history. We call a strategy obtained in this way "folding-back optimal."

Definition 56. *A strategy $\bar{s}_i \in S_i$ is **folding-back optimal** given β^i if*

$$\forall h \in H,\ \hat{V}_i^{\beta^i}(h, \bar{s}_i(h)) = \max_{a_i \in \mathcal{A}_i(h)} \hat{V}_i^{\beta^i}(h, a_i) = \hat{V}_i^{\beta^i}(h).$$

For every conjecture there is a folding-back optimal strategy: by finiteness of the game, the maximization problem of equation (10.3.1) has a solution at every history. Thus, we obtain the following:

---

[9]Recall, $L(\Gamma(h)) := \max_{z \in Z(h)} \ell(z) - \ell(h)$.

Remark 37. *There exists at least one folding-back optimal strategy for any given* β^i.

Of course, there can be more than one folding-back optimal strategy given the same conjecture. This happens whenever the maximazion problem of equation (10.3.1) has more than one solution at a history. When there are multiple solutions at more than one history, a natural question arises: is every combination of optimal actions a folding-back optimal strategy? The answer is affirmative, because folding back only depends on the maximal expected payoff computed at future histories, not on the specific maximizers.

Remark 38. *For every two folding-back optimal strategies* $\bar{s}_i, \bar{s}'_i$ *given* β^i, *and for every strategy* $\bar{s}''_i$ *such that* $\bar{s}''_i(h) \in \{\bar{s}_i(h), \bar{s}'_i(h)\}$ *for each* $h \in H$, $\bar{s}''_i$ *is folding-back optimal given* β^i.

Remark 38 contrasts with Example 52, where we will show that the set of strategies that are "justifiable"—i.e., optimal for at least one conjecture—does not feature every combination of actions that are prescribed by some strategy in the set.[10]

Next we move to alternative notions of optimal planning. **One step-optimality** takes the strategy $\bar{s}_i$ as given and makes sure that at every history h player i has no incentive to deviate to an action $a_i \neq \bar{s}_i(h)$ given β^i. According to **sequential optimality**, conditional on every history, the continuation-strategy must maximize the expected payoff given β^i. The weaker notion of **weak sequential optimality** only considers the histories that can be reached given the candidate plan. Finally, we introduce **ex-ante optimality**, which only requires optimality from the viewpoint of the initial history.

Definition 57. *A strategy* $\bar{s}_i \in S_i$ *is*

- ***one-step optimal*** *given* β^i *if*

$$\forall h \in H,\ V_i^{\bar{s}_i,\beta^i}(h, \bar{s}_i(h)) = \max_{a_i \in \mathcal{A}_i(h)} V_i^{\bar{s}_i,\beta^i}(h, a_i); \tag{10.3.2}$$

[10] By Remark 38, the set of folding-back optimal strategies given β^i can be written as the intersection over all histories h of the sets of strategies that prescribe a folding-back optimal action at h given β^i, and an intersection of Cartesian sets is Cartesian. By contrast, the union over all conjectures β^i of the Cartesian sets of folding-back optimal strategies given β^i need not be Cartesian.

- ***sequentially optimal*** *given* β^i *if*

$$\forall h \in H,\ V_i^{\bar{s}_i,\beta^i}(h) = \max_{t_i \in S_i^{\succeq h}} V_i^{t_i,\beta^i}(h); \tag{10.3.3}$$

- ***weakly sequentially optimal*** *given* β^i *if*[11]

$$\forall h \in H_i(\bar{s}_i),\ V_i^{\bar{s}_i,\beta^i}(h) = \max_{t_i \in S_i^{\succeq h}} V_i^{t_i,\beta^i}(h); \tag{10.3.4}$$

- ***ex ante optimal*** *given* β^i *if*

$$\bar{s}_i \in \arg\max_{s_i \in S_i} V_i^{s_i,\beta^i}(\varnothing). \tag{10.3.5}$$

We will sometimes omit the phrase "given β^i" when the reference to a specific conjecture is clear from the context. We will sometimes refer to the maximization property at h for a strategy $\bar{s}_i$ that satisfies one-step optimality as the **local optimality of** $\bar{s}_i$ **at** h.

Ex-ante optimality implies that the strategy is optimal not just from the viewpoint of the initial history, but also at every history that player i initially expects to reach with positive probability.

Theorem 38. *A strategy* $\bar{s}_i$ *is ex ante optimal if and only if*

$$\forall h \in H,\ \mathbb{P}_{\bar{s}_i,\beta^i}(h) > 0 \Rightarrow (\bar{s}_i|h) \in \arg\max_{t_i \in S_i^{\succeq h}} V_i^{t_i,\beta^i}(h). \tag{10.3.6}$$

We will state a more general result for beliefs over strategies in the next section (Lemma 28), thus we refer the reader to the proof of that result for an intuition of how to prove Theorem 38.

In the rest of this section, we investigate the relations among the notions of optimal plan introduced in Definitions 56 and 57. By inspection of Definition 57 and Remark 36, we immediately obtain the following relation between sequential optimality and one-step optimality (and ex-ante optimality as well).

Remark 39. *If a strategy is sequentially optimal given* β^i*, then it is one-step optimal and ex ante optimal given* β^i*.*

[11] Recall from Chapter 9 that $H_i(\bar{s}_i)$ is the set of nonterminal histories not precluded by $\bar{s}_i$.

The following results show that folding-back optimality and sequential optimality are equivalent to one-step optimality. First we establish the equivalence between the two properties defined by local optimization conditions, that is, folding-back optimality and one-step optimality; this is called the "Folding Back Principle." Next we show that folding-back optimality is equivalent to sequential optimality; this is the so called "Optimality Principle." As a corollary we obtain the "One-Deviation Principle" which states the equivalence between one-step optimality and sequential optimality.

Theorem 39. (Folding Back Principle)
(I) A strategy $\bar{s}_i$ is one-step optimal given β^i if and only if

$$\begin{aligned} V_i^{\bar{s}_i,\beta^i}(h,a_i) &= \hat{V}_i^{\beta^i}(h,a_i), \\ V_i^{\bar{s}_i,\beta^i}(h) &= \hat{V}_i^{\beta^i}(h) \end{aligned} \tag{10.3.7}$$

for every $h \in H$ and $a_i \in \mathcal{A}_i(h)$.
(II) A strategy $\bar{s}_i$ is folding-back optimal given β^i if and only if it is one-step optimal given β^i.

Proof. We prove by induction on the height of subgames that if $\bar{s}_i$ is one-step optimal given β^i then (10.3.7) holds for every $h \in H$ and $a_i \in \mathcal{A}_i(h)$. It is easy to show that also the converse of this statement holds. Thus (I) holds and (I) is equivalent to (II). We leave the proof of these steps as an exercise.

Suppose that $\bar{s}_i$ is one-step optimal given β^i.

Basis step. Consider any $h \in H$ such that $L(\Gamma(h)) = 1$. Then $(h,a) \in Z$ for every $a \in \mathcal{A}(h)$; therefore

$$V_i^{\bar{s}_i,\beta^i}(h,a_i) = \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} u_i(h,(a_i,a_{-i}))\beta^i(a_{-i}|h) = \hat{V}_i^{\beta^i}(h,a_i)$$

for every $a_i \in \mathcal{A}_i(h)$; hence

$$\begin{aligned} V_i^{\bar{s}_i,\beta^i}(h) &= V_i^{\bar{s}_i,\beta^i}(h,\bar{s}_i(h)) \overset{\text{(loc.opt.)}}{=} \max_{a_i\in\mathcal{A}_i(h)} V_i^{\bar{s}_i,\beta^i}(h,a_i) \\ &= \max_{a_i\in\mathcal{A}_i(h)} \hat{V}_i^{\beta^i}(h,a_i) = \hat{V}_i^{\beta^i}(h), \end{aligned}$$

where the first equality holds by Remark 36 and the second equality holds because $\bar{s}_i$ is locally optimal (loc.opt.) at h.

Inductive step. Suppose that (10.3.7) holds for each $h \in H$ with $L(\Gamma(h)) \leq k$. Now, fix h with $L(\Gamma(h)) = k+1$. Then $L(\Gamma(h,a)) \leq k$ for each $a \in \mathcal{A}(h)$. Therefore the inductive hypothesis (I.H.) implies

$$V_i^{\bar{s}_i,\beta^i}(h,a_i) = \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} V_i^{\bar{s}_i,\beta^i}(h,(a_i,a_{-i}))\beta^i(a_{-i}|h)$$

$$\overset{\text{(I.H.)}}{=} \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \hat{V}_i^{\beta^i}(h,(a_i,a_{-i}))\beta^i(a_{-i}|h) =: \hat{V}_i^{\beta^i}(h,a_i)$$

for every $a_i \in \mathcal{A}_i(h)$; hence

$$\begin{aligned} V_i^{\bar{s}_i,\beta^i}(h) &= V_i^{\bar{s}_i,\beta^i}(h,\bar{s}_i(h)) \overset{\text{(loc.opt.)}}{=} \max_{a_i\in\mathcal{A}_i(h)} V_i^{\bar{s}_i,\beta^i}(h,a_i) \\ &= \max_{a_i\in\mathcal{A}_i(h)} \hat{V}_i^{\beta^i}(h,a_i) = \hat{V}_i^{\beta^i}(h), \end{aligned}$$

where the first equality holds by Remark 36 and the second equality holds because $\bar{s}_i$ is locally optimal at h. ■

With this, we can prove the Optimality Principle.

Theorem 40. (Optimality Principle) *A strategy of player i is sequentially optimal given β^i if and only if it is folding-back optimal given β^i.*

Proof. To ease notation, in the proof we use the symbols t_i, t_i' to denote generic *sub*-strategies.

(If) Let $\bar{s}_i$ be folding-back optimal given β^i. We will prove by induction on the height of subgames that

$$\forall h \in H,\ \hat{V}_i^{\beta^i}(h) \geq \max_{t_i\in S_i^{\succeq h}} V_i^{t_i,\beta^i}(h). \tag{10.3.8}$$

Since $\bar{s}_i$ is folding-back optimal, it satisfies one-step optimality (Theorem 39 II), therefore $V_i^{\bar{s}_i,\beta^i}(h) = \hat{V}_i^{\beta^i}(h)$ (Theorem 39 I); this means that $\bar{s}_i$ is sequentially optimal given β^i.

Basis step. Let $L(\Gamma(h)) = 1$. Then $(h,a) \in Z$ for every $a \in \mathcal{A}(h)$; therefore

$$\max_{t_i\in S_i^{\succeq h}} V_i^{t_i,\beta^i}(h) = \max_{a_i\in\mathcal{A}_i(h)} \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} u_i(h,(a_i,a_{-i}))\beta^i(a_{-i}|h);$$

furthermore

$$\hat{V}_i^{\beta^i}(h) \geq \hat{V}_i^{\beta^i}(h, a_i) = \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} u_i(h, (a_i, a_{-i}))\beta^i(a_{-i}|h)$$

for every $a_i \in \mathcal{A}_i(h)$, where both the inequality and the equality hold by definition. Hence (10.3.8) holds at h.

Inductive step. Suppose that $\hat{V}_i^{\beta^i}(h) \geq \max_{t_i \in S_i^{\succeq h}} V_i^{t_i,\beta^i}(h)$ for every $h \in H$ with $L(\Gamma(h)) \leq k$. Now fix h with $L(\Gamma(h)) = k+1$. Then $L(\Gamma(h,a)) \leq k$ for each $a \in \mathcal{A}(h)$, and the inductive assumption (I.H.) yields

$$\begin{aligned} \hat{V}_i^{\beta^i}(h) \quad &\geq \quad \hat{V}_i^{\beta^i}(h, a_i) = \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \hat{V}_i^{\beta^i}(h, (a_i, a_{-i}))\beta^i(a_{-i}|h) \\ &\overset{(I.H.)}{\geq} \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \max_{t_i \in S_i^{\succeq (h,(a_i,a_{-i}))}} V_i^{t_i,\beta^i}(h, (a_i, a_{-i}))\beta^i(a_{-i}|h) \end{aligned}$$

for every $a_i \in \mathcal{A}_i(h)$. Therefore, for every sub-strategy $t_i' \in S_i^{\succeq h}$,

$$\begin{aligned} \hat{V}_i^{\beta^i}(h) \quad &\geq \quad \max_{a_i \in \mathcal{A}_i(h)} \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \max_{t_i \in S_i^{\succeq (h,(a_i,a_{-i}))}} V_i^{t_i,\beta^i}(h, (a_i, a_{-i}))\beta^i(a_{-i}|h) \\ &\geq \quad \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \max_{t_i \in S_i^{\succeq (h,(t_i'(h),a_{-i}))}} V_i^{t_i,\beta^i}(h, (t_i'(h), a_{-i}))\beta^i(a_{-i}|h) \\ &\geq \quad \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} V_i^{t_i',\beta^i}(h, (t_i'(h), a_{-i}))\beta^i(a_{-i}|h) = V_i^{t_i',\beta^i}(h). \end{aligned}$$

The latter set of inequalities yield

$$\hat{V}_i^{\beta^i}(h) \geq \max_{t_i' \in S_i^{\succeq h}} V_i^{t_i',\beta^i}(h)$$

as desired.

(Only if) If $\bar{s}_i$ is sequentially optimal, then it satisfies one-step optimality (Remark 39), which in turn implies that it is folding-back optimal (Theorem 39 II). ∎

Theorems 39 and 40 yield the following result, known as the **One-Deviation (OD) Principle**:

Corollary 6. (One-Deviation Principle) *A strategy of player i is sequentially optimal given β^i if and only if it is one-step optimal given β^i.*

Since folding-back optimality, sequential optimality, and one-step optimality are equivalent, we can generically refer to any of them as "**rational planning.**" The following result relates rational planning to weak sequential optimality.

Proposition 3. *A strategy $\bar{s}_i$ is weakly sequentially optimal given β^i if and only if there is a strategy s_i that is behaviorally equivalent to $\bar{s}_i$ and is sequentially optimal given β^i.*

Proof. Suppose that $\bar{s}_i$ is weakly sequentially optimal. Let s_i^* be any strategy that is folding-back optimal given β^i (by finiteness, there is at least one such strategy). Define s_i as follows:

$$s_i(h) = \begin{cases} s_i^*(h), & \text{if } h \in H \backslash H_i(\bar{s}_i), \\ \bar{s}_i(h), & \text{if } h \in H_i(\bar{s}_i). \end{cases}$$

By construction, s_i is behaviorally equivalent to $\bar{s}_i$ and satisfies $V_i^{s_i,\beta^i}(h) = \hat{V}_i^{\beta^i}(h)$ for every $h \in H \backslash H_i(\bar{s}_i)$. Since $\bar{s}_i$ is weakly sequentially optimal and behaviorally equivalent to s_i,

$$V_i^{s_i,\beta^i}(h) = V_i^{\bar{s}_i,\beta^i}(h) = \max_{t_i \in S_i^{\succeq h}} V_i^{t_i,\beta^i}(h) = \hat{V}_i^{\beta^i}(h)$$

for every $h \in H_i(\bar{s}_i)$, where the first equality follows from behavioral equivalence, which implies realization-equivalence, hence, value-equivalence, and the third equality follows from the proof of the optimality principle (Theorem 40). Thus, s_i is folding-back optimal, because $V_i^{s_i,\beta^i}(h) = \hat{V}_i^{\beta^i}(h)$ for every $h \in H$. By the optimality principle, s_i is also sequentially optimal.

The converse follows by inspection of the definitions. ∎

The following example illustrates all the notions of optimal plan introduced in Definitions 56 and 57.

Example 49. Using the conditional beliefs of Example 48 we obtain the following values and optimal strategies. First consider Ann:

$$\begin{aligned}\hat{V}_a^{\beta^a}(N,U) &= 0,\ \hat{V}_a^{\beta^a}(N,D)=1,\\ \hat{V}_a^{\beta^a}(B,u) &= 4,\ \hat{V}_a^{\beta^a}(B,d)=0,\\ \hat{V}_a^{\beta^a}(N) &= 1=\hat{V}_a^{\beta^a}(\varnothing),\ \hat{V}_a^{\beta^a}(B)=4.\end{aligned}$$

Strategy $D.u$ of Ann is folding-back, sequentially, and ex ante optimal, and it satisfies one-step optimality; strategy $D.d$ is ex ante optimal, but it does not satisfy one-step optimality and is not folding-back, nor sequentially optimal. The difference between ex ante optimality and the three kinds of dynamic optimality depends on the fact that Ann initially does not expect that the subgame with root B can be reached. Since Ann cannot move more than once, there is no difference between sequential and weakly sequential optimality in her case. Next consider Bob:

$$\begin{aligned}\hat{V}_b^{\beta^b}(N,L) &= \frac{3}{4},\ \hat{V}_b^{\beta^b}(N,R)=1,\\ \hat{V}_b^{\beta^b}(B,l) &= -\frac{7}{4},\ \hat{V}_b^{\beta^b}(B,r)=1,\\ \hat{V}_b^{\beta^b}(B) &= \hat{V}_b^{\beta^b}(N)=\hat{V}_b^{\beta^b}(\varnothing)=1;\end{aligned}$$

strategies $N.R.r$ and $B.R.r$ of Bob are folding-back, sequentially, and ex ante optimal, and they satisfy one-step optimality; strategies $N.R.l$ (realization-equivalent to $N.R.r$) and $B.L.r$ (realization-equivalent to $B.R.r$) are ex ante optimal and weakly sequentially optimal, but do not satisfy one-step optimality and are not folding-back, nor sequentially optimal. ▲

10.4 Conditional Probability and Dominance

We have seen how, in order to plan rationally, player i must form conjectures about the behavior of the co-players at each nonterminal history. A deterministic conjecture specifies an action profile $a_{-i} \in \mathcal{A}_{-i}(h)$ for each $h \in H$. Formally, such specification is an element of $S_{-i} = \times_{h\in H}\mathcal{A}_{-i}(h)$, the set of functions $s_{-i} : H \to A_{-i}$ such that $s_{-i}(h) \in \mathcal{A}_{-i}(h)$ for each $h \in H$. Elements of S_{-i} may also be interpreted as profiles of plans in the minds of the co-players, but this is *not* the interpretation used here, because the payoff of i is affected by co-players' behavior, not by their plans. (Of course, if we assume that plans are necessarily carried out, the two interpretations are equivalent.) Thus, a probabilistic conjecture can be expressed by assigning probabilities to elements of S_{-i}. Such assignments must take into account the information about s_{-i} revealed by the play as it unfolds, and condition on such information. This representation in terms of conditional beliefs allows us to characterize rational planning with a notion of dominance.

At any given history $h \in H$, a probability measure over the elements of S_{-i} that are consistent with h suffices for our player to solve his decision problem at h. Then, rational planning is possible if the beliefs at different histories satisfy natural consistency rules. Recall from Chapter 9 that $S_{-i}(h)$ is the set of s_{-i} that do not prevent h from occurring, that is, the s_{-i} such that $h \prec \zeta(s_i, s_{-i})$ for some s_i. Thus, let

$$\mathcal{H}_{-i} = \{C_{-i} \subseteq S_{-i} : \exists h \in H, C_{-i} = S_{-i}(h)\}$$

denote the collection of "relevant hypotheses" or "conditioning events" about co-players' behavior.

Definition 58. *A* ***Conditional Probability System*** *is an array of probability measures* $\mu^i = (\mu^i(\cdot|C_{-i}))_{C_{-i}\in\mathcal{H}_{-i}} \in (\Delta(S_{-i}))^{\mathcal{H}_{-i}}$ *such that:*

1. *for every* $C_{-i} \in \mathcal{H}_{-i}$, $\mu^i(C_{-i}|C_{-i}) = 1$*;*
2. *for all* $C_{-i}, D_{-i} \in \mathcal{H}_{-i}$ *with* $D_{-i} \subseteq C_{-i}$*, and for every* $E_{-i} \subseteq D_{-i}$*,*

$$\mu^i(E_{-i}|C_{-i}) = \mu^i(E_{-i}|D_{-i}) \cdot \mu^i(D_{-i}|C_{-i}).$$

A Conditional Probability System (henceforth, CPS) is an array of probability measures over S_{-i} that conditions on information (property 1) and satisfies the *chain rule of conditional probabilities* (property 2). The chain rule imposes coherency between the belief of player i at a history h and the belief at a future history h' that player i deems possible at h: player i should not "change his mind" from h to h', but just update his belief. To see this clearly, note that the equality in property 2 is equivalent to the following condition:

$$\mu^i(D_{-i}|C_{-i}) > 0 \Rightarrow \mu^i(E_{-i}|D_{-i}) = \frac{\mu^i(E_{-i}|C_{-i})}{\mu^i(D_{-i}|C_{-i})}.$$

Considering that C_{-i} and D_{-i} are elements of $\mathcal{H}_{-i}$, this can also be written as follows. Let $h \prec h'$ and $s_{-i} \in S_{-i}(h')$, then

$$\mu^i(S_{-i}(h')|S_{-i}(h)) > 0 \Rightarrow \mu^i(s_{-i}|S_{-i}(h')) = \frac{\mu^i(s_{-i}|S_{-i}(h))}{\mu^i(S_{-i}(h')|S_{-i}(h))},$$

because $S_{-i}(h') \subseteq S_{-i}(h)$. Given a CPS μ^i, at every history h that player i initially deems possible (i.e., $\mu^i(S_{-i}(h)|S_{-i}) > 0$, where $S_{-i} = S_{-i}(\varnothing)$), player i simply updates the initial belief $\mu^i(\cdot|S_{-i})$ using the standard rule of conditional probabilities. Instead, at a history h that was deemed impossible (i.e., $\mu^i(S_{-i}(h)|S_{-i}) = 0$), player i has to come up with a new belief $\mu^i(\cdot|S_{-i}(h))$. Then, player i starts updating this new belief exactly as he was doing with the initial one, until also this new belief is "falsified" by the observed behavior of the co-players, and so on. To ease notation, from now on we write

$$\mu^i(\cdot|S_{-i}(h)) = \mu^i(\cdot|h),$$

and we let $\Delta^H(S_{-i})$ denote the set of all CPSs of player i on S_{-i}.[12]

We now redefine rational planning with respect to a CPS. Of the different representations of rational planning provided for conjectures, we seek to adapt to CPSs the definition of weak sequential optimality, for two reasons. First, sequential optimality is easier to define than folding-back optimality with respect to a CPS, because a sequentially optimal strategy is defined via a comparison of continuation-plans at each history, and the belief specified by the CPS for the history suffices for this comparison. Second, we look for just *weakly* sequentially optimal strategies because, *under the assumption that players always execute their plans*, they are indistinguishable from the realization-equivalent sequentially optimal strategies, based on the observation of co-players' behavior. (We will elaborate on this issue in the next chapter.)

Given the belief $\mu^i(\cdot|h)$ that player i holds at history h, and given a strategy s_i that is consistent with h, we can compute the probability of reaching a terminal history $z \in Z(h)$ as follows:

$$\mathbb{P}_{s_i,\mu^i}(z|h) = \begin{cases} 0, & \text{if } s_i \notin S_i(z), \\ \mu^i(S_{-i}(z)|h), & \text{if } s_i \in S_i(z). \end{cases}$$

[12] Note that, for all $\mu^i \in \Delta^H(S_{-i})$ and $h', h'' \in H$, $S_{-i}(h') = S_{-i}(h'')$ implies $\mu^i(\cdot|h') = \mu^i(\cdot|h'')$. In other words, beliefs about co-players depend only on what the history reveals about their behavior, and do not depend on what it reveals about own behavior.

With this, we can recompute the subjective value of reaching h when s_i is followed from h onwards as follows:

$$V_i^{s_i,\mu^i}(h) = \sum_{z\in Z(h)} u_i(z)\mathbb{P}_{s_i,\mu^i}(z|h).$$

Then, we can adapt to CPSs the definition of weak sequential optimality provided for conjectures. Recall that $H_i(\bar{s}_i)$ is the set of nonterminal histories consistent with (not precluded by) playing strategy $\bar{s}_i$.

Definition 59. *A strategy $\bar{s}_i$ is **weakly sequentially optimal given** μ^i if*[13]

$$\forall h \in H_i(\bar{s}_i),\ V_i^{\bar{s}_i,\mu^i}(h) = \max_{s_i\in S_i(h)} V_i^{s_i,\mu^i}(h). \tag{10.4.1}$$

Condition (10.4.1) must hold at every history not precluded by $\bar{s}_i$, but given the coherency between beliefs at different histories imposed by the chain rule, it is sufficient to check that it holds at the histories that are deemed impossible by player i until they are actually reached.

Lemma 28. *Fix any strategy $\bar{s}_i$. For all $h, h' \in H_i(\bar{s}_i)$ such that $h \prec h'$ and $\mu^i(S_{-i}(h')|S_{-i}(h)) > 0$, if*

$$V_i^{\bar{s}_i,\mu^i}(h) = \max_{s_i\in S_i(h)} V_i^{s_i,\mu^i}(h) \tag{10.4.2}$$

then

$$V_i^{\bar{s}_i,\mu^i}(h') = \max_{s_i\in S_i(h')} V_i^{s_i,\mu^i}(h').$$

Proof. Fix $s_i \in S_i(h')$. Define $s'_i \in S_i(h')$ as $s'_i(\widetilde{h}) = \bar{s}_i(\widetilde{h})$ for each $\widetilde{h} \not\succeq h'$, and $s'_i(\widetilde{h}) = s_i(\widetilde{h})$ for each $\widetilde{h} \succeq h'$. Then, we have $\zeta(s'_i, s_{-i}) = \zeta(s_i, s_{-i})$ for each $s_{-i} \in S_{-i}(h')$, and $\zeta(s'_i, s_{-i}) = \zeta(\bar{s}_i, s_{-i})$

[13] We have previously defined the continuation value at h given a conjecture β^i considering the *continuation* strategies in $S_i^{\succeq h}$. To ease notation, here we define the continuation value at h given a CPS μ^i for the strategies that are consistent with h, that is, $S_i(h)$. The two representations are equivalent because $\mu^i(S_{-i}(h)|h) = 1$, therefore only the actions prescribed by each $s_i \in S_i(h)$ from h onwards matter for the expected payoff.

for each $s_{-i} \notin S_{-i}(h')$. Hence,

$$\sum_{s_{-i}\in S_{-i}(h)\backslash S_{-i}(h')} u_i(\zeta(s_i', s_{-i}))\mu^i(s_{-i}|h) \tag{10.4.3}$$
$$= \sum_{s_{-i}\in S_{-i}(h)\backslash S_{-i}(h')} u_i(\zeta(\bar{s}_i, s_{-i}))\mu^i(s_{-i}|h),$$

$$\sum_{s_{-i}\in S_{-i}(h')} u_i(\zeta(s_i', s_{-i}))\mu^i(s_{-i}|h) \tag{10.4.4}$$
$$= \sum_{s_{-i}\in S_{-i}(h')} u_i(\zeta(s_i, s_{-i}))\mu^i(s_{-i}|h).$$

By (10.4.2) and (10.4.3), we get

$$\sum_{s_{-i}\in S_{-i}(\bar{h}')} u_i(\zeta(s_i', s_{-i}))\mu^i(s_{-i}|h) \leq \sum_{s_{-i}\in S_{-i}(\bar{h}')} u_i(\zeta(\bar{s}_i, s_{-i}))\mu^i(s_{-i}|h),$$

so by (10.4.4),

$$\sum_{s_{-i}\in S_{-i}(\bar{h}')} u_i(\zeta(s_i, s_{-i}))\mu^i(s_{-i}|h) \leq \sum_{s_{-i}\in S_{-i}(\bar{h}')} u_i(\zeta(\bar{s}_i, s_{-i}))\mu^i(s_{-i}|h).$$

Finally, dividing both sides by $\mu^i(S_{-i}(h')|h)$, we get

$$\sum_{s_{-i}\in S_{-i}(\bar{h}')} u_i(\zeta(s_i, s_{-i}))\mu^i(s_{-i}|h') \leq \sum_{s_{-i}\in S_{-i}(\bar{h}')} u_i(\zeta(\bar{s}_i, s_{-i}))\mu^i(s_{-i}|h'),$$

as desired. ■

Now we state and prove formally the equivalence between weak sequential optimality given a CPS and weak sequential optimality given the conjecture derived from the CPS. Recall that, for all $h \in H$ and $a_{-i} \in \mathcal{A}_{-i}(h)$, we let $S_{-i}(h, a_{-i})$ denote the set of all $s_{-i} \in S_{-i}(h)$ such that $s_{-i}(h) = a_{-i}$.

Proposition 4. *For each CPS $\mu^i \in \Delta^H(S_{-i})$, define the corresponding conjecture $\beta^i \in \times_{h\in H} \Delta(\mathcal{A}_{-i}(h))$ as follows:*

$$\forall h \in H,\ \forall a_{-i} \in \mathcal{A}_{-i}(h),\ \beta^i(a_{-i}|h) = \mu^i\left(S_{-i}\left(h, a_{-i}\right)|h\right).$$

Then, a strategy is weakly sequentially optimal given μ^i if and only if it is weakly sequentially optimal given β^i.

Proof. A strategy $\bar{s}_i \in S_i$ is weakly sequentially optimal given β^i if

$$\forall h \in H_i(\bar{s}_i),\ V_i^{\bar{s}_i,\beta^i}(h) = \max_{s_i \in S_i(h)} V_i^{s_i,\beta^i}(h),$$

where

$$V_i^{s_i,\beta^i}(h) = \sum_{z \in Z(h)} u_i(z)\mathbb{P}_{s_i,\beta^i}(z|h).$$

Thus, we have only to show that, for all $s_i \in S_i$, $h \in H_i(s_i)$, and $z \in Z(h)$,

$$\mathbb{P}_{s_i,\mu^i}(z|h) = \mathbb{P}_{s_i,\beta^i}(z|h).$$

Since h is a prefix of every $z \in Z(h)$, there is some $n \geq 1$ and some $(a^1, ..., a^n)$ such that $(h, a^1, ..., a^n)$. Case $n = 1$ is trivial. Thus, suppose that $n \geq 2$. Previously, we defined

$$\mathbb{P}_{s_i,\beta^i}(z|h) = \mathbb{P}_{s_i,\beta^i}\left(a^1|h\right) \prod_{t=2}^{n} \mathbb{P}_{s_i,\beta^i}\left(a^t|(h, a^1, ..., a^{t-1})\right),$$

where, for each $\widetilde{h} \in H$ and $a = (a_i, a_{-i}) \in \mathcal{A}(\widetilde{h})$,

$$\mathbb{P}_{s_i,\beta^i}\left(a|\widetilde{h}\right) = \begin{cases} 0, & \text{if } a_i \neq s_i(\widetilde{h}), \\ \beta^i(a_{-i}|\widetilde{h}), & \text{if } a_i = s_i(\widetilde{h}). \end{cases}$$

Suppose first that $s_i \notin S_i(z)$. Since $s_i \in S_i(h)$, there exists t such that $a_i^t \neq s_i(h)$. Then we have

$$\mathbb{P}_{s_i,\beta^i}(z|h) = 0 = \mathbb{P}_{s_i,\mu^i}(z|h).$$

If $s_i \in S_i(z)$, then $a_i^t = s_i(h)$ for every $t = 1, ..., n$. So we have

$$\mathbb{P}_{s_i,\beta^i}(z|h) = \beta^i(a_{-i}^1|h) \prod_{t=2}^{n} \beta^i\left(a_{-i}^t|(h, a^1, ..., a^{t-1})\right). \tag{10.4.5}$$

By the chain rule, we can write

$$\begin{aligned}
\mathbb{P}_{s_i,\mu^i}(z|h) &= \mu^i(S_{-i}(z)|h) \\
&= \mu^i(S_{-i}((h,a^1))|h) \cdot \mu^i(S_{-i}(z)|(h,a^1)) \\
&= \mu^i(S_{-i}((h,a^1))|h) \cdot \mu^i(S_{-i}((h,a^1,a^2))|(h,a^1)) \cdot \mu^i(S_{-i}(z)|(h,a^1,a^2)) \\
&= ... \\
&= \mu^i(S_{-i}((h,a^1))|h) \cdot \prod_{t=2}^{n} \mu^i(S_{-i}((h,a^1,...,a^t))|(h,a^1,...,a^{t-1})). \quad (10.4.6)
\end{aligned}$$

For each $\widetilde{h} \in H$ and $a = (a_i, a_{-i}) \in \mathcal{A}(\widetilde{h})$, by definition we have $\beta^i(a_{-i}|\widetilde{h}) = \mu^i(S_{-i}(\widetilde{h}, a_{-i})|\widetilde{h})$, and obviously $S_{-i}(\widetilde{h}, a_{-i}) = S_{-i}(\widetilde{h}, a)$. Hence, (10.4.5) and (10.4.6) coincide, as desired. ∎

Weak sequential optimality given a conjecture is weaker than sequential optimality, which in turn is equivalent to folding-back optimality by the optimality principle (Theorem 40). Hence, the existence of a folding-back optimal strategy guarantees the existence of a weakly sequentially optimal strategy given the conjecture derived from the CPS; thus, by Proposition 4, we obtain the existence of a weakly sequentially optimal strategy given the CPS. Therefore, we can adopt weak sequential optimality for at least one CPS as our extension of the notion of justifiability of Chapter 3 to multistage games.

Definition 60. *A strategy $\bar{s}_i$ is* ***justified by*** *a CPS μ^i if it is weakly sequentially optimal given μ^i. We will say that a strategy is* ***justifiable*** *if it is justified by some CPS.*

We let

$$r_i : \Delta^H(S_{-i}) \rightrightarrows S_i$$

denote the correspondence that assigns to each CPS μ^i the set of strategies justified by μ^i. With this, the set of justifiable strategies is

$$r_i\left(\Delta^H(S_{-i})\right) = \bigcup_{\mu^i \in \Delta^H(S_{-i})} r_i\left(\mu^i\right).$$

The set of strategies justified by a CPS satisfies the following property, which we will use in the next chapter. Fix a CPS μ^i. If a history h is consistent with some weakly sequentially optimal strategy and with a strategy $\bar{s}_i$ that satisfies Condition (10.4.1) at h, then there is a weakly sequentially optimal strategy that is consistent with h and prescribes $\bar{s}_i(h)$. (We omit the proof.)

Lemma 29. *For each $h \in \bigcup_{s_i \in r_i(\mu^i)} H_i(s_i)$ and $\bar{s}_i \in S_i(h)$ such that*

$$V_i^{\bar{s}_i,\mu^i}(h) = \max_{s_i \in S_i(h)} V_i^{s_i,\mu^i}(h),$$

there exists $s_i' \in r_i(\mu^i) \cap S_i(h)$ such that $s_i'(h) = \bar{s}_i(h)$.

In Chapter 3, we stated and proved an equivalence result relating (un)justifiability and strict dominance (Lemma 2). Here, we extended the notion of justifiability to multistage games. Can we extend the notion of strict dominance to multistage games in such a way to preserve the equivalence between justifiability and not being dominated? The answer is yes.

Definition 61. *A strategy $\bar{s}_i$ is* ***conditionally dominated*** *if there exist a history $h \in H_i(\bar{s}_i)$ and a mixed strategy $\sigma_i \in \Delta(S_i)$ with $\sigma_i(S_i(h)) = 1$ such that*

$$\forall s_{-i} \in S_{-i}(h), \sum_{s_i \in S_i(h)} \sigma_i(s_i) u_i(\zeta(s_i, s_{-i})) > u_i(\zeta(\bar{s}_i, s_{-i})). \tag{10.4.7}$$

The set of strategies of agent i that are not conditionally dominated *is denoted by NCD_i.*

A strategy $\bar{s}_i$ is conditionally dominated if there exists a mixed strategy that yields a higher expected payoff conditional on reaching some history consistent with $\bar{s}_i$, no matter what the co-players do. We restrict attention to histories that are consistent with the strategy itself, because we want to relate conditional dominance to *weak* sequential optimality.[14] We can state the equivalence result.

Lemma 30. *A strategy is justifiable if and only if it is not conditionally dominated, that is, $r_i\left(\Delta^H(S_{-i})\right) = NCD_i$.*

Proof. Fix a strategy $\bar{s}_i$. Suppose first that $\bar{s}_i$ is justifiable. Then there exists $\mu^i \in \Delta^H(S_{-i})$ such that

$$\forall h \in H_i(\bar{s}_i), V_i^{\bar{s}_i, \mu^i}(h) = \max_{s_i \in S_i(h)} V_i^{s_i, \mu^i}(h),$$

which means that

$$\forall h \in H_i(\bar{s}_i), \bar{s}_i \in \arg\max_{s_i \in S_i(h)} \sum_{s_{-i} \in S_{-i}(h)} u_i(\zeta(s_i, s_{-i})) \mu^i(s_{-i}|h). \tag{10.4.8}$$

[14] One can establish an analogous equivalence between sequential optimality and a notion of conditional dominance that also considers the histories that are precluded by the dominated strategy (focusing on the continuation strategy).

Next, fix any $h \in H_i(\bar{s}_i)$. Consider an auxiliary simultaneous-move game with, for each $j \in I$, action set $A_j^h = S_j(h)$ and payoff function

$$\begin{array}{rcl} v_j : \times_{k \in I} A_k^h & \to & \mathbb{R} \\ s & \mapsto & u_j(\zeta(s)). \end{array}$$

By (10.4.8), $\bar{s}_i$ is a best reply to the conjecture $\nu^i \in \Delta(S_{-i}(h))$ such that $\nu^i(s_{-i}) = \mu^i(s_{-i}|h)$ for each $s_{-i} \in S_{-i}(h)$. Then, by Lemma 2, no mixed action $\sigma_i \in \Delta(S_i(h))$ dominates $\bar{s}_i$. Since this is true for every $h \in H_i(\bar{s}_i)$, it follows that $\bar{s}_i$ is not conditionally dominated.

Conversely, suppose that $\bar{s}_i$ is not conditionally dominated. Fix any $h \in H_i(\bar{s}_i)$ and construct an auxiliary simultaneous-move game as above. Thus, $\bar{s}_i$ is not dominated by any mixed action in this game. Then, by Lemma 2, $\bar{s}_i$ is a best reply among the strategies in $S_i(h)$ to some conjecture $\nu^{i,h} \in \Delta(S_{-i}(h))$. Order the conjectures $(\nu^{i,h})_{h \in H_i(\bar{s}_i)}$ in such a way that if history h precedes history h' in the game, $\nu^{i,h}$ comes before $\nu^{i,h'}$ in the ordering. Next, fix some residual measure $\nu \in \Delta(S_{-i})$ such that $\nu(s_{-i}) > 0$ for every $s_{-i} \in S_{-i}$. With this, define $\mu^i = (\mu^i(\cdot|h))_{h \in H} \in (\Delta(S_{-i}))^H$ as follows. For each $h \in H$, derive $\mu^i(\cdot|h)$ by considering prefixes $\bar{h}$ of h and conditioning the first $\nu^{i,\bar{h}}$ in the ordering such that $\nu^{i,\bar{h}}(S_{-i}(h)) > 0$, *if any*; otherwise, derive $\mu^i(\cdot|h)$ by conditioning the residual measure ν.

First we check that μ^i is a CPS. Fix $C_{-i}, D_{-i} \in \mathcal{H}_{-i}$ and $h, h' \in H$ such that $D_{-i} \subseteq C_{-i}$, $C_{-i} = S_{-i}(h)$, $D_{-i} = S_{-i}(h')$, and $h \neq h'$. Suppose that $\mu^i(D_{-i}|h) > 0$. If $\mu^i(\cdot|h)$ was derived by conditioning ν, so was $\mu^i(\cdot|h')$, therefore condition 2 of a CPS is satisfied. If $\mu^i(\cdot|h)$ was derived by conditioning some $\nu^{i,\bar{h}}$, $\mu^i(D_{-i}|h) > 0$ implies $\nu^{i,\bar{h}}(D_{-i}) > 0$. Moreover, for each $\nu^{i,\widetilde{h}}$ that precedes $\nu^{i,\bar{h}}$ in the ordering, we must have $\nu^{i,\widetilde{h}}(S_{-i}(h)) = 0$, otherwise $\mu^i(\cdot|h)$ could not be derived from $\nu^{i,\bar{h}}$. Hence, $\mu^i(\cdot|h')$ was derived by conditioning $\nu^{i,\bar{h}}$ as well, and condition 2 of a CPS is satisfied.

Now we want to show that $\bar{s}_i$ is weakly sequentially optimal under μ_i. We want to show that, for each $h \in H_i(\bar{s}_i)$, $\bar{s}_i$ satisfies (10.4.1). For every $\widetilde{h} \in H_i(\bar{s}_i)$ with $\widetilde{h} \not\preceq h$ and $h \not\preceq \widetilde{h}$, there is no $s_{-i} \in S_{-i}$ such that $\widetilde{h} \prec \zeta(\bar{s}_i, s_{-i})$ and $h \prec \zeta(\bar{s}_i, s_{-i})$. Hence, $S_{-i}(\widetilde{h}) \cap S_{-i}(h) = \emptyset$ and $\mu^i(\cdot|h)$ cannot be derived by conditioning $\nu^{i,\widetilde{h}}$. So, since $\nu^{i,h}(S_{-i}(h)) = 1$, $\mu^i(\cdot|h)$ is derived by conditioning some $\nu^{i,\bar{h}}$ with $\bar{h} \preceq h$ and $\nu^{i,\bar{h}}(S_{-i}(h)) > 0$. Note that $\mu^i(\cdot|\bar{h}) = \nu^{i,\bar{h}}$, therefore $\bar{s}_i$ satisfies (10.4.1) at $\bar{h}$. Then, by Lemma 28, $\bar{s}_i$ satisfies (10.4.1) at h. ■

To better understand conditional dominance, it is useful to compare it to strict dominance and weak dominance applied to the *strategic form* of the game. Strict dominance in the strategic form clearly implies conditional dominance. Indeed, if a strategy is strictly dominated, then it is conditionally dominated given the empty (initial) history $h = \varnothing$, because $S_{-i}(\varnothing) = S_{-i}$.

Remark 40. *If a strategy is strictly dominated in the strategic form of the game, then it is conditionally dominated.*

Furthermore, conditional dominance implies weak dominance: if $\bar{s}_i$ is worse than some mixed strategy σ_i from some history h onwards, then we can construct a mixed strategy that imitates σ_i from h onwards and $\bar{s}_i$ everywhere else, so that it is equivalent to $\bar{s}_i$ if h is not reached.

Proposition 5. *If a strategy is conditionally dominated, then it is weakly dominated in the strategic form of the game.*

Proof. We use Lemma 5: a strategy is weakly dominated if (and only if) it is not a best response to any conjecture assigning *strictly positive* probability to every profile s_{-i}. Suppose that $\bar{s}_i$ is conditionally dominated at some history $\bar{h}$. Pick any conjecture $\mu^i \in \Delta(S_{-i})$ such that $\mu^i(s_{-i}) > 0$ for each $s_{-i} \in S_{-i}$, and let ν^i be the conjecture over $S_{-i}(\bar{h})$ obtained from μ^i by conditioning. Now, focus on the auxiliary simultaneous-move game where the action set of each player j coincides with the subset of strategies of the multistage game $S_j(\bar{h})$, and for each profile of actions s the payoff of j is given by $u_j(\zeta(s))$. We can view ν^i as a conjecture of player i in this game. Since $\bar{s}_i$ is conditionally dominated at $\bar{h}$ in the multistage game, it is dominated in this simultaneous-move game, therefore by Lemma 2 it is not a best reply to ν^i. It follows that there exists $s_i \in S_i(\bar{h})$ such that

$$\sum_{s_{-i} \in S_{-i}(\bar{h})} u_i(\zeta(\bar{s}_i, s_{-i}))\nu^i(s_{-i}) < \sum_{s_{-i} \in S_{-i}(\bar{h})} u_i(\zeta(s_i, s_{-i}))\nu^i(s_{-i}).$$

Now go back to the multistage game. Define $\widetilde{s}_i$ as $\widetilde{s}_i(h) = s_i(h)$ for all $h \succeq \bar{h}$ and $\widetilde{s}_i(h) = \bar{s}_i(h)$ for all $h \not\succeq \bar{h}$. Hence,

$$\sum_{s_{-i} \in S_{-i}(\bar{h})} u_i(\zeta(\widetilde{s}_i, s_{-i}))\nu^i(s_{-i}) = \sum_{s_{-i} \in S_{-i}(\bar{h})} u_i(\zeta(s_i, s_{-i}))\nu^i(s_{-i}),$$

$$\forall s_{-i} \in S_{-i} \backslash S_{-i}(\bar{h}), \quad \zeta(\widetilde{s}_i, s_{-i}) = \zeta(\bar{s}_i, s_{-i}).$$

Thus, we obtain

$$\begin{aligned}
& \sum_{s_{-i} \in S_{-i}} u_i(\zeta(\bar{s}_i, s_{-i}))\mu^i(s_{-i}) \\
= \; & \mu^i(S_{-i}(\bar{h})) \sum_{s_{-i} \in S_{-i}(\bar{h})} u_i(\zeta(\bar{s}_i, s_{-i}))\nu^i(s_{-i}) + \sum_{s_{-i} \in S_{-i} \backslash S_{-i}(\bar{h})} u_i(\zeta(\bar{s}_i, s_{-i}))\mu^i(s_{-i}) \\
< \; & \mu^i(S_{-i}(\bar{h})) \sum_{s_{-i} \in S_{-i}(\bar{h})} u_i(\zeta(\widetilde{s}_i, s_{-i}))\nu^i(s_{-i}) + \sum_{s_{-i} \in S_{-i} \backslash S_{-i}(\bar{h})} u_i(\zeta(\widetilde{s}_i, s_{-i}))\mu^i(s_{-i}) \\
= \; & \sum_{s_{-i} \in S_{-i}} u_i(\zeta(\widetilde{s}_i, s_{-i}))\mu^i(s_{-i}).
\end{aligned}$$

This argument proves that $\bar{s}_i$ is not a best reply to any strictly positive conjecture in the strategic form of the multistage game. Therefore, by Lemma 5, s_i is weakly dominated. ■

The reason why weak dominance does not imply conditional dominance can be seen in a simultaneous-move game, where conditional dominance coincides with strict dominance. However, it is often the case that all the weak dominance relations in the strategic form of a multistage game imply a strict dominance relation from some history h onwards, as in the definition of conditional dominance. For example, in all leader-follower games[15] a strategy of the follower is weakly dominated in the strategic form of the game if and only if it does not select a best reply for at least one action of the leader, and it is therefore conditionally dominated given this action. Indeed, in a large class of games, weak dominance in the strategic form coincides with conditional dominance. Next we quantify how large this class is.

Suppose that a mixed strategy $\bar{\sigma}_i$ weakly dominates strategy $\bar{s}_i$ in the strategic form of the game. Fix a history $h \in H_i(\bar{s}_i)$ such that (i) $\bar{\sigma}_i(S_i(h)) = 1$ and (ii) $\bar{\sigma}_i(s_i) > 0$ for some s_i with $s_i(h) \neq \bar{s}_i(h)$. By weak dominance, $\bar{\sigma}_i$ does weakly better than $\bar{s}_i$ against every $s_{-i} \in S_{-i}(h)$. Now decrease by $\varepsilon > 0$ the (only) payoffs of player i associated with terminal histories that are consistent with $\bar{s}_i$. Then, by (ii), $\bar{\sigma}_i$ now

[15] Formally, a leader-follower game is a two-person, two-stage game with perfect information such that the first-mover cannot move again in the second stage.

does *strictly* better than $\bar{s}_i$ against every $s_{-i} \in S_{-i}(h)$, and thus $\bar{s}_i$ is conditionally dominated. So, if $\bar{s}_i$ is not conditionally dominated in the original game, it is conditionally dominated in "neighboring games" with the same tree and slightly different payoffs. The bottom line is that games where weak dominance in the strategic form does not imply conditional dominance represent "an exception" from a mathematical perspective, as long as it makes sense to slightly modify the payoffs of some terminal histories and not others, which is certainly the case when the outcome function $g : Z \to Y$ is one-to-one. We formalize this concept using the notion of *Lebesgue measure.* In $\mathbb{R}^n$, the Lebesgue measure is the generalization of the notions of length, area, and volume for $\mathbb{R}^1$, $\mathbb{R}^2$, and $\mathbb{R}^3$, respectively. Recall that the set $\mathbb{R}^Z$ of all payoff functions $u_i : Z \to \mathbb{R}$ is essentially the same as the Euclidean space $\mathbb{R}^{|Z|}$: fix arbitrarily a bijection $n : Z \to \{1, ..., |Z|\}$, then $u_i \in \mathbb{R}^Z$ corresponds to the vector $\left(u_i\left(n^{-1}(1)\right), ..., u_i\left(n^{-1}(|Z|)\right)\right) \in \mathbb{R}^{|Z|}$.

Lemma 31. *Fix a finite multistage game tree. The closure*[16] *of the set of payoff vectors $(u_i(z))_{z \in Z}$ such that some strategy of player i is weakly dominated in the strategic form but not conditionally dominated has Lebesgue measure zero in $\mathbb{R}^Z$.*

The proof of Lemma 31 is in the appendix of Chapter 11, where we provide a more general version of this result.

In words, we will say that a statement (such as the equivalence between conditional dominance and weak dominance) holds **generically** when, given every multistage game tree, it is true for all payoff vectors except possibly for a subset with zero-measure closure. Now we provide the intuition for Lemma 31 with an example.

Example 50. Consider a modified version of the Battle of the Sexes of Example 41, where Bob's payoff at terminal history $((\text{in}), (\text{B}_1, \text{B}_2))$ is substituted by a parameter $v \in \mathbb{R}$. Strategy B_2 of Bob is conditionally dominated if and only if $v < 0$. If $v < 0$, then B_2 is conditionally dominated by S_2 at history (in). If $v \geq 0$, then B_2 is not conditionally dominated: for every mixed strategy $\sigma_2 \in \Delta(\{\text{S}_2, \text{B}_2\})$, we have

$$\begin{aligned} \sigma_2(\text{S}_2)u_2(\zeta(\text{in.B}_1, \text{S}_2)) + \sigma_2(\text{B}_2)u_2(\zeta(\text{in.B}_1, \text{B}_2)) &= \sigma_2(\text{B}_2) \cdot v \qquad (10.4.9) \\ &\leq v = u_2(\zeta(\text{in.B}_1, \text{B}_2)), \end{aligned}$$

[16] The closure of a set is the smallest closed set that contains it.

therefore condition (10.4.7) is not satisfied for s_1 = in.B$_1$ (at either nonterminal history).

Strategy B$_2$ is weakly dominated in the strategic form of the game if and only if $v \leq 0$. If $v \leq 0$, then B$_2$ is weakly dominated by S$_2$—note that B$_2$ is not strictly dominated because the two strategies yield the same payoff if Ann plays out. If $v > 0$, then B$_2$ is not weakly dominated, because then inequality (10.4.9) is strict.

So, for $v \neq 0$, B$_2$ is weakly dominated if and only if it is conditionally dominated. There is only one value of v in the entire real line, $v = 0$, such that B$_2$ is weakly dominated but not conditionally dominated. ▲

10.5 Infinite Horizon and Continuity at Infinity

Although the previous analysis is focused on finite games, it applies more generally to games with finite horizon.[17] In games with infinite horizon, folding-back planning does not apply, because there is no last stage from which the backward computation can start. However, the OD Principle still holds under a mild regularity property, called "continuity at infinity," implied by compactness-continuity of the game.

For any history $h = (a^1, a^2, ...) \in \bar{H}$ and number of stages $\ell \in \mathbb{N}$, let h^ℓ denote the prefix of h of length ℓ if $\ell(h) > t$ (recall that $\ell(h)$ is the length of h, and $\ell(h) = \infty$ if h is an infinite history), otherwise let $h^\ell = h$; that is, if $h = \left(a^k\right)_{k=1}^{\ell(h)}$, then $h^\ell = \left(a^k\right)_{k=1}^{\min\{\ell(h),\ell\}}$.

Definition 62. *Game Γ is* ***continuous at infinity*** *if, for every player $i \in I$,*

$$\lim_{\ell\to\infty}\left[\sup\left\{\left|u_i(h) - u_i(\widetilde{h})\right| : h, \widetilde{h} \in Z, h^\ell = \widetilde{h}^\ell\right\}\right] = 0$$

where we let $\sup\emptyset = 0$ *by convention.*

Note that every game with finite horizon is trivially continuous at infinity. Indeed, let $L < \infty$ denote the horizon, then for every $\ell \geq L$, $h^\ell = \widetilde{h}^\ell$ implies $h = \widetilde{h}$, hence the limsup must be 0. Thus, Γ is continuous

[17] In infinite games with finite horizon, the optimization problems defining the dynamic programming properties studied above may not have a solution, but the main equivalence results (Folding-Back Principle, Optimality Principle, OD Principle) still hold. Other results, such as Proposition 3, do not hold for all infinite games, even if they have finite horizon.

at infinity if and only if either Γ has finite horizon, or Γ has infinite horizon and for all $\epsilon > 0$ there is some sufficiently large positive integer L_ϵ such that, for all $\ell \geq L_\epsilon$ and all pairs of histories $h, \widetilde{h} \in Z$ satisfying $h^\ell = \widetilde{h}^\ell$, we have $\left| u_i(h) - u_i(\widetilde{h}) \right| < \epsilon$. This means that what happens in the distant future has little impact on overall payoffs.

Remark 41. *Every compact-continuous game is continuous at infinity.*

Proof. Let Γ be compact-continuous and fix a terminal history $h \in Z$ arbitrarily; we must show that

$$\lim_{\ell \to \infty} \left[\sup \left\{ \left| u_i(h) - u_i(\widetilde{h}) \right| : \widetilde{h} \in Z, h^\ell = \widetilde{h}^\ell \right\} \right] = 0.$$

The equality holds trivially if h is finite. Suppose that h is infinite. For every $\ell \in \mathbb{N}$, the set $\left\{ \widetilde{h} \in Z : h^\ell = \widetilde{h}^\ell \right\}$ of terminal histories with prefix h^ℓ is compact; therefore, by continuity of u_i, there is some terminal history $\widetilde{h}_\ell \in \left\{ \widetilde{h} \in Z : h^\ell = \widetilde{h}^\ell \right\}$ such that the supremum is attained:

$$\sup \left\{ \left| u_i(h) - u_i(\widetilde{h}) \right| : \widetilde{h} \in Z, h^\ell = \widetilde{h}^\ell \right\} = \left| u_i(h) - u_i(\widetilde{h}_\ell) \right| < \infty.$$

By construction, the sequence of terminal histories $\left(\widetilde{h}_\ell \right)_{\ell=1}^{\infty}$ converges to h, because $\widetilde{h}_\ell^k = h^k$ for $k \leq \ell$, therefore $\lim_{\ell \to \infty} \widetilde{h}_\ell^k = h^k$ for each $k \in \mathbb{N}$, which implies $\lim_{\ell \to \infty} \widetilde{h}_\ell = h$. By continuity of u_i, $\lim_{\ell \to \infty} u_i \left(\widetilde{h}_\ell \right) = u_i(h)$, which implies the claim. ■

Example 51. Suppose that the game has infinite horizon and, at every stage, all the action profiles in the Cartesian set $A = \times_{i \in I} A_i$ are feasible, that is, $H = A^{<\mathbb{N}_0}$ and $Z = A^{\mathbb{N}}$. *Stages coincide with time periods*, denoted by t, and consequences materialize at the end of each period t. For every $i \in I$ and $t = 1, 2, ...$, there is a period-t (instantaneous, or flow) payoff function $u_{i,t} : A^t \to \mathbb{R}$.[18] Furthermore, there is some upper bound $\overline{v} \in \mathbb{R}$ such that, for all $i \in I$,

$$\sup_{h^t \in A^t} |u_{i,t}(h^t)| \leq \overline{v}.$$

[18]The period-t payoff functions $u_{i,t}$ may be the composition of period-t outcome functions $g_t : A^t \to Y_t$ and period-t utility functions $v_{i,t} : Y_t \to \mathbb{R}$, that is, $u_{i,t} = v_{i,t} \circ g_t : A^t \to \mathbb{R}$. See the discussion in Section 9.2.1.

Payoffs are assigned to terminal histories as follows: for each player i there is a discount factor $\delta_i \in (0,1)$ such that

$$u_i(h^\infty) = (1-\delta_i)\sum_{t=1}^{\infty}(\delta_i)^{t-1}u_{i,t}(h^t)$$

for all $h^\infty \in Z$, where h^t denotes the length-t prefix of h^∞. One can verify that such games are continuous at infinity. Infinitely repeated games with discounting are a special case where $u_{i,t}(a^1,...,a^{t-1},a^t) = u_i(a^t)$ is independent of $\left(a^1,...,a^{t-1}\right)$. ▲

Theorem 41. (OD Principle with continuity at infinity)**.** *Suppose that* Γ *is continuous at infinity. Then, for every* $s_i \in S_i$ *and* $\beta^i \in \times_{h\in H}\Delta(\mathcal{A}_{-i}(h))$*, the following conditions are equivalent:*
(1) s_i *is sequentially optimal given* β^i*;*
(2) s_i *is one-step optimal given* β^i*.*

Proof. (1) implies (2) by Remark 39. To show that (2) implies (1), we prove the contrapositive: if s_i is *not* sequentially optimal given β^i, then there must be incentives for one-step deviations. If β^i is non-deterministic, the proof requires some measure-theoretic technicalities. Therefore we provide the proof for the *deterministic case*: for some $s_{-i} \in S_{-i}$ and every $h \in H$, $\beta^i\left(s_{-i}\left(h\right)|h\right) = 1$.[19] Thus, we replace β^i with s_{-i} in the formulas, and we have $V_i^{s_i,s_{-i}}(h) = u_i\left(\zeta\left(s_i, s_{-i}|h\right)\right)$, where $\zeta\left(s_i, s_{-i}|h\right)$ denotes the terminal history induced by (s_i, s_{-i}) starting from h, which depends only on $(s_i|h) \in S_i^{\succeq h}$ and $(s_{-i}|h) \in S_{-i}^{\succeq h}$.

Suppose that s_i is *not* sequentially optimal given some $s_{-i} \in S_{-i}$, that is, there are $h \in H$, $\widehat{s}_i \in S_i$ and $\epsilon > 0$ such that

$$u_i\left(\zeta\left(s_i, s_{-i}|h\right)\right) = u_i\left(\zeta\left(\widehat{s}_i, s_{-i}|h\right)\right) - \epsilon. \tag{10.5.1}$$

For every positive integer L and for every strategy profile $s \in S$, let us define the auxiliary "truncated" game $\Gamma^{L,s} = \left\langle I, (\mathcal{A}_i^L(\cdot), u_i^{L,s})_{i\in I}\right\rangle$ with horizon L derived from Γ as follows:

- $\mathcal{A}_i^L(h') = \mathcal{A}_i(h')$ if $\ell(h') < L$, and $\mathcal{A}_i^L(h') = \emptyset$ otherwise, so that $\bar{H}^L = \{h' \in \bar{H} : \ell(h') \leq L\}$ and $Z^L = \{h' \in \bar{H}^L$: either $\ell(h') = L$, or $\ell(h') < L$ and $h' \in Z\}$,

[19] The proof of this special case is sufficient for the analysis of pure equilibria in games without chance moves.

- for all $\bar{h} \in Z \cap Z^L$, $u_i^{L,s}(\bar{h}) = u_i(\bar{h})$,
- for all $\bar{h} \in Z^L \backslash Z$, $u_i^{L,s}(\bar{h}) = u_i\left(\zeta\left(s_i, s_{-i}|\bar{h}\right)\right)$.

Intuitively, the game is truncated so that it has at most L stages. If a terminal history h' of a truncated game $\Gamma^{L,s}$ is not terminal for the original game Γ, then the imputed payoffs are those that would obtain in Γ if the players followed the given strategy profile s after history h'.

By continuity at infinity, there is a sufficiently large L (which depends on ϵ) such that, for all $\widetilde{h}, \widehat{h} \in Z$, if $\widetilde{h}^L = \widehat{h}^L$ (i.e., if these two histories coincide in the first L stages) then

$$|u_i(\widetilde{h}) - u_i(\widehat{h})| < \epsilon. \tag{10.5.2}$$

Let $\widetilde{s}_i$ be the strategy that behaves as s_i after L and as $\widehat{s}_i$ before L. More precisely:

$$\widetilde{s}_i(h') = \begin{cases} s_i(h'), & \text{if } \ell(h') \geq L, \\ \widehat{s}_i(h'), & \text{if } \ell(h') < T. \end{cases}$$

Let $\widehat{h} = \zeta\left(\widehat{s}_i, s_{-i}|h\right)$ and $\widetilde{h} = \zeta\left(\widetilde{s}_i, s_{-i}|h\right)$, where s_{-i} and h are given above in (10.5.1). Then

$$u_i\left(\zeta\left(\widetilde{s}_i, s_{-i}|h\right)\right) = u_i(\widetilde{h}) > u_i(\widehat{h}) - \epsilon = u_i\left(\zeta\left(\widehat{s}_i, s_{-i}|h\right)\right) - \epsilon,$$

where the equalities hold by definition and the inequality follows from (10.5.2). Taking (10.5.1) into account, we obtain

$$V_i^{\widetilde{s}_i, s_{-i}}(h) = u_i\left(\zeta\left(\widetilde{s}_i, s_{-i}|h\right)\right) > u_i\left(\zeta\left(s_i, s_{-i}|h\right)\right) = V_i^{s_i, s_{-i}}(h).$$

This means that the restriction of s_i to $\bar{H}^L$ is not sequentially optimal given s_{-i} in the finite horizon game $\Gamma^{L,s}$. Thus, by Corollary 6,[20] strategy s_i is not one-step optimal given s_{-i}. By definition of $\Gamma^{L,s}$, this implies that also in game Γ player i has an incentive to make one-step deviations after history h. ■

[20] Corollary 6 pertains to finite games, but the result can be extended to games with finite horizon.

11

Rationalizability in Multistage Games

In this chapter we analyze the behavior of rational players who form their subjective beliefs about the behavior of co-players by means of strategic thinking. As we did in Chapter 4 for simultaneous-move games, the cornerstone of players' reasoning is the belief of a rational player that the co-players are rational as well. Therefore, we seek to extend the idea of rationalizability from simultaneous-move games to multistage games. However, there is no obvious way to do it. The reason is that, as the game unfolds, the initial belief about the co-players may be falsified by their observed behavior. Depending on how players revise their beliefs after unexpected moves, we will define three notions of belief in the rationality of co-players, and, correspondingly, three notions of rationalizability for multistage games. **Initial rationalizability** captures the idea that players believe in the rationality of co-players at the beginning of the game and form a belief about their (present and future) behavior accordingly; yet, if this belief is falsified, they are free to entertain the possibility that their co-players are not rational, even if their past behavior does not contradict their rationality. **Strong rationalizability**, instead, captures a principle of **best rationalization**: players always ascribe to their co-players the highest level of strategic sophistication that is consistent with their observed behavior. In particular, they believe in the rationality of the co-players as long as the evidence is consistent with this hypothesis. A similar assumption holds for higher levels of strategic sophistication.

This idea is used to rationalize unexpected moves of co-players and predict their future moves accordingly—an instance of so-called "forward-induction reasoning." **Continuation-rationalizability** is based instead on the idea that, when a player is surprised by a move of a co-player, she considers the possibility that the co-player made a mistake in the execution of his plan, but continues to believe that the co-player is rational and will not make other mistakes in the continuation of the game. In this way, continuation-rationalizability generalizes to multistage games the **backward-induction** algorithm for perfect-information games, a kind of "inter-personal folding-back" procedure, which we present at the end of the chapter. We assume throughout this chapter that the given multistage game is *finite*.[1]

11.1 Rationality in Multistage Games

Different versions of rationalizability characterize the behavioral implications of rationality and of different forms of common belief in rationality. It is therefore important to be clear about the notion of rationality we adopt. We want "rationality" to mean two things: (i) players "plan rationallly"; and (ii) they actually implement what they planned. In Chapter 10 we introduced different, but equivalent notions of rational planning for multistage games: sequential optimality, folding-back optimality, and one-step optimality. We also introduced the behaviorally equivalent notion of weak sequential optimality, and we adapted its definition to the case in which players' beliefs about co-players' behavior are represented by conditional probability systems (CPSs). As anticipated, in this chapter we adopt CPSs over co-players' strategies to describe how beliefs about co-players' (history-dependent) behavior are updated, or revised as the play unfolds. Furthermore, in the analysis of initial and strong rationalizability (Sections 11.2 and 11.3) it is convenient to use *weak* sequential optimality, whereas in the analysis of continuation-rationalizability (Section 11.4) we use sequential optimality.

The reason for choosing *weak* sequential optimality in Sections 11.2 and

[1] The analysis can be extended to all **simple** multistage games, whereby the set of feasible actions of each player is finite for all histories except, possibly, those of height 1, for which the set of feasible actions must be, if not finite, at least compact. This implies that the set H of nonterminal histories is finite or at least countable. See Battigalli [8].

11.3 is the following. There, we assume to be transparent (i.e., true and always commonly believed) that each player i always executes his plan. Thus, when a co-player j observes a deviation from the expected behavior of i, he does *not* believe that player i committed a mistake in execution and will revert to his original plan, but rather that player i has a different plan in his mind and is executing this different plan. Accordingly, there is no need to specify the moves that a player's strategy prescribes after deviations from the strategy itself. In fact, the whole analysis of Sections 11.2 and 11.3 of this chapter goes through verbatim with reduced strategies in place of strategies, since the elimination procedures we consider there do not make any distinction between any two behaviorally-equivalent (hence, realization-equivalent) strategies. As anticipated, continuation-rationalizability is instead based on the opposite idea: upon observing a deviation from the expected plan of i, the co-players believe that player i committed a mistake and will revert to his original plan; if i is rational, this plan is sequentially optimal given i system of beliefs.

The reason for working with CPSs over co-players' strategies—rather than conditional probabilities of co-players' actions—is that a player who reasons about the game is interested in the history-dependent behavior of the co-players, which is conveniently described by strategies (see the discussion in Section 10.4 of Chapter 10). Rationality of the co-players yields "across-histories restrictions" on their behavior. Thus, just keeping track of the actions the co-players could play at different histories taken in isolations and just considering conjectures over such actions would ultimately expand the set of behaviors that a player may deem possible. The following example illustrates this point.

Example 52. Consider the common interest game with perfect information in Figure 11.1 (the numbers at terminal nodes represent the common payoff).

We want to compute the set of justifiable strategies of Ann. (As usual, we omit action "wait" from the description of players' strategies.)

▶ First note that any $s_a \in S_a$ such that $s_a(\varnothing) = out$ is justified by all CPSs μ^a such that $\mu^a(\{m\} \times S_c|\varnothing) = 1$.

▶ Strategy $in.b.b'$ is justified by a CPS μ^a such that $\mu^a((\ell, d.d')|\varnothing) = 1$ and $\mu^a((r, d.d')|(in, r)) = 1$.

▶ Strategy $in.b.a'$ is justified by a CPS μ^a such that $\mu^a((\ell, d.d')|\varnothing) = 1$ and $\mu^a((r, d.c')|(in, r)) = 1$.

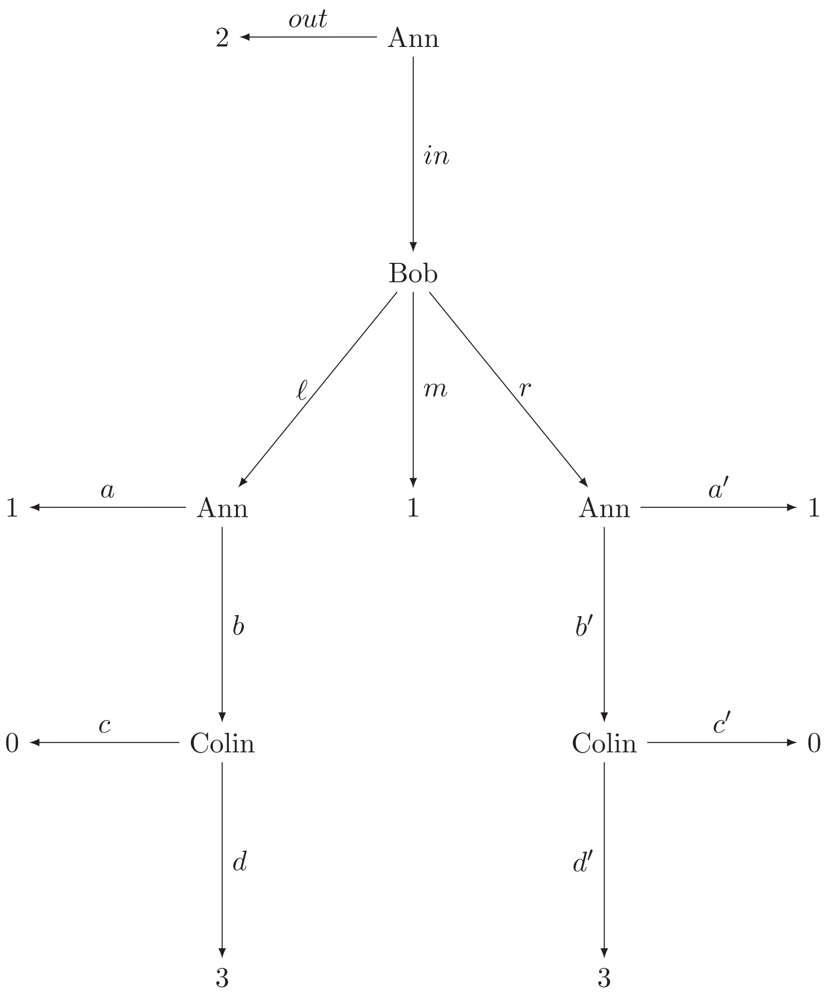


Figure 11.1: A common interest game.

▶ Strategy $in.a.b'$ is justified by a CPS μ^a such that $\mu^a((r, d.d')|\varnothing) = 1$ and $\mu^a((\ell, c.c')|(in, \ell)) = 1$.
▶ Strategy $in.a.a'$ is *not justifiable*. To see this, observe that $in.a.a'$ is dominated by *out* at the root of the game: with $in.a.a'$ Ann obtains payoff 1 for sure, with *out* she obtains 2. Therefore, by Lemma 30, $in.a.a'$ is not justifiable.

It follows that every action of Ann is prescribed by at least one justifiable strategy, but some combination of these actions, such as $in.a.a'$, is *not* a justifiable strategy.

Take now the viewpoint of Bob at history (in) with a belief over the justifiable strategies of Ann and Colin that are consistent with (in). Bob is certain that Colin plays $d.d'$ and, furthermore, he assigns probability 1 to the set of strategies of Ann that prescribe either b, or b', or both. With this, we obtain a conjecture of Bob that assigns probability at least 1/2 to either b (given ℓ) or b' (given r). Therefore, his optimal action can be ℓ or r, but never m. Consider now instead a conjecture of Bob that, at each history of the game, assigns probability 1 to the actions of Ann and Colin that are prescribed by some justifiable strategy consistent with the history. Such a conjecture can assign probability 1 to a at (in, ℓ) and to a' at (in, r). So, Bob could optimally choose m. The problem just described arises because, if we pick for each history h an action $a_{i,h}$ consistent with i's rationality (the action "wait" if i is not active) and we form the strategy $s_i = (a_{i,h})_{h \in H}$, it is possible that s_i is *in*consistent with i's rationality. This is the case for strategy $in.a.a'$ of Ann.[2] ▲

Finally, we provide a characterization of weak sequential optimality that will come in handy later in this chapter.

Lemma 32. *Fix a player $i \in I$ and a CPS $\mu^i \in \Delta^H(S_{-i})$. Consider a nonempty subset of strategies $C_i \subseteq S_i$ such that*

$$r_i(\mu^i) \subseteq C_i. \tag{11.1.1}$$

A strategy $\bar{s}_i \in C_i$ is justified by μ^i if and only if

$$\forall h \in H_i(\bar{s}_i),\ V_i^{\bar{s}_i,\mu^i}(h) = \max_{s_i \in C_i \cap S_i(h)} V_i^{s_i,\mu^i}(h). \tag{11.1.2}$$

[2]Formally, the set of justifiable strategies

$$r_i\left(\Delta^H(S_{-i})\right) \subseteq S_i = \bigtimes_{h \in H} \mathcal{A}_i(h)$$

is not the Cartesian product of its projections on the sets $\mathcal{A}_i(h)$ for $h \in H$. Strategy $in.a.a'$ belongs to the Cartesian product of the projections of $r_a\left(\Delta^H(S_{-a})\right)$, but $in.a.a' \notin r_a\left(\Delta^H(S_{-a})\right)$.

Proof. The "only if" part is obvious. The proof of the "if" part is by induction on the length of histories. Fix any $\bar{s}_i \in C_i$ such that (11.1.2) holds. We want to show that, for every $h \in H_i(\bar{s}_i)$,

$$V_i^{\bar{s}_i,\mu^i}(h) = \max_{s_i \in S_i(h)} V_i^{s_i,\mu^i}(h). \tag{11.1.3}$$

At the initial history we have

$$V_i^{\bar{s}_i,\mu^i}(\varnothing) \overset{(11.1.2)}{=} \max_{s_i \in C_i} V_i^{s_i,\mu^i}(\varnothing) \overset{(11.1.1)}{=} \max_{s_i \in S_i} V_i^{s_i,\mu^i}(\varnothing).$$

Now, fix $n > 0$ and $h' \in H_i(\bar{s}_i)$ of length n. Suppose by way of induction that for every $h \in H_i(\bar{s}_i)$ of length smaller than n, $r_i(\mu^i) \cap S_i(h) \neq \emptyset$ and condition (11.1.3) holds. Now let h be the immediate predecessor of h'. Then, by Lemma 29, there exists $\widetilde{s}_i \in r_i(\mu^i) \cap S_i(h)$ such that $\widetilde{s}_i(h) = \bar{s}_i(h)$, thus $\bar{s}_i \in S_i(h')$. Hence,

$$V_i^{\widetilde{s}_i,\mu^i}(h') = \max_{s_i \in S_i(h')} V_i^{s_i,\mu^i}(h').$$

By (11.1.1), we have $\widetilde{s}_i \in C_i$, therefore

$$\max_{s_i \in S_i(h')} V_i^{s_i,\mu^i}(h') = \max_{s_i \in C_i \cap S_i(h')} V_i^{s_i,\mu^i}(h'). \tag{11.1.4}$$

So

$$V_i^{\bar{s}_i,\mu^i}(h') \overset{(11.1.2)}{=} \max_{s_i \in C_i \cap S_i(h')} V_i^{s_i,\mu^i}(h') \overset{(11.1.4)}{=} \max_{s_i \in S_i(h')} V_i^{s_i,\mu^i}(h').$$

■

11.2 Initial Rationalizability

The first natural extension of common belief in rationality from simultaneous-move games to multistage games is common *initial* belief in rationality. At the beginning of the game, players believe that co-players are rational; believe that co-players believe that everybody else is rational; and so on. To derive the behavioral implications of rationality and common initial belief in rationality, as we did in Chapter 4, we introduce a justification operator that characterizes the behavioral implications for rational players of each step of this reasoning procedure. Consider the collection $\mathcal{C}$ of all Cartesian subsets of S. For each $C = \times_{i \in I} C_i \in \mathcal{C}$, and for each $i \in I$, we define

$$\Delta^H_\varnothing(C_{-i}) = \left\{\mu^i \in \Delta^H(S_{-i}) : \mu^i(C_{-i}|\varnothing) = 1\right\},$$

$$\rho_i(C_{-i}) = \left\{s_i \in S_i : \exists \mu^i \in \Delta^H_\varnothing(C_{-i}), s_i \in r_i(\mu^i)\right\} = r_i(\Delta^H_\varnothing(C_{-i})),$$

$$\rho(C) = \bigtimes_{i \in I} \rho_i(C_{-i}).$$

The set $\Delta^H_\varnothing(C_{-i})$ is the subset of CPSs such that player i initially believes C_{-i}, i.e., C_{-i} is assigned probability one at the initial history. Given this, $\rho_i(C_{-i})$ is the set of strategies of i that are justified by a CPS in $\Delta^H_\varnothing(C_{-i})$. Finally, $\rho(C)$ is the set of strategy profiles that can be justified by a CPS with initial belief in C_{-i} for each player $i \in I$. We use the same symbols r_i and ρ_i used for best replies and justification operator in simultaneous-move games, because we are considering an extension of these notions to multistage games: when the game has simultaneous moves, $\Delta^H_\varnothing(C_{-i})$ coincides with $\Delta(C_{-i})$, and weakly sequentially optimal strategies coincide with best replies.

Equipped with the operator $\rho : \mathcal{C} \to \mathcal{C}$, we define "**initial rationalizability**" with the standard iteration of a self-map as in Chapter 4. Let $\rho^0(S) = S$, and for each $k > 0$, let $\rho^k(S) = \rho(\rho^{k-1}(S))$. Note that whenever $C' \subseteq C$, we also have $\Delta^H_\varnothing(C'_{-i}) \subseteq \Delta^H_\varnothing(C_{-i})$ for each $i \in I$. Then, ρ is *monotone*. Therefore, since $\rho^0(S) = S$, $\left(\rho^k(S)\right)_{k=0}^{\infty}$ must be a weakly decreasing sequence, and it makes sense to define:[3]

$$\rho^\infty(S) = \bigcap_{k \geq 1} \rho^k(S).$$

Definition 63. *A strategy profile $s \in S$ is* ***initially rationalizable*** *if $s \in \rho^\infty(S)$.*

At every step of reasoning, initial rationalizability preserves the strategies of each player that are justified by a CPS assigning probability 1 at the initial history to the co-players' strategies that survived the previous steps. So, for instance, the strategies that survive step 2 are

[3] Compare with Chapter 4.

justified by a CPS that, at the initial history, assigns probability 1 to the justifiable strategies of the co-players. Then, the belief in the rationality of the co-players is maintained at all histories that a player deems possible given the initial belief. However, at histories that are deemed subjectively impossible, initial rationalizability does not impose any restriction on the beliefs over co-players' strategies, even if such histories are consistent with some or all steps of the algorithm—Example 53 below is case in point.

Leveraging on the monotonicity of operator ρ, one can prove that extensions of Theorems 2 and 3 of Chapter 4 hold for initial rationalizability in multistage games. Also, in Chapter 4 we reformulated rationalizability as a *reduction* procedure, by iterating a reduction operator $\overline{\rho}$ that only considers the strategies in the reduced game as candidate best replies. Can we do the same with initial rationalizability? The answer is again yes, and a possible way to construct the reduction procedure is identical to that in Chapter 4, once the set of conjectures over the reduced set of co-players' action is replaced by the set of CPSs that initially believe in the reduced set of co-players' strategies. We leave these proofs as exercises.

In Chapter 10, Lemma 30, we have shown that weak sequential optimality for some CPS is characterized by the notion of *conditional dominance*. With this, it is relatively easy to see that initial rationalizability can be given the following dominance characterization. The first step coincides with deletion of all conditionally dominated strategies. The following steps coincide with the iterated deletion of strategies that are strictly dominated at the initial history, or equivalently, with iterated strict dominance in the residual strategic form obtained after the deletion of conditionally dominated strategies. To formalize, let $NCD = \times_{i\in I} NCD_i$ denote the set of profiles of conditionally undominated (not conditionally dominated) strategies. Also, let $\mathrm{ND}(C)$ denote the set of profile of strategies that are not dominated (by mixed strategies) in the restricted Cartesian set $C \subseteq S$, i.e., the undominated profiles in the restricted strategic form $\langle I, (U_i|_C, C_i)_{i\in I}\rangle$. As in Chapter 4, $\mathrm{ND} : \mathcal{C} \to \mathcal{C}$ is a restriction operator: $\mathrm{ND}(C) \subseteq C$ for every $C \in \mathcal{C}$. As such, it yields a decresing sequence of subsets $\left(\mathrm{ND}^k\left(\bar{S}\right)\right)_{k\in\mathbb{N}}$ starting from any subset $\bar{S}$. With this, we can state our result.

Theorem 42. $\rho^\infty(S) = \mathrm{ND}^\infty(NCD)$.

We omit the proof. Note the order of elimination: *first* one has to delete all the conditionally dominated strategies of the multistage game Γ, obtaining the subset $NCD \subseteq S$, *then* one continues with the iterated deletion of dominated strategies starting from NCD.

Finally, we illustrate initial rationalizability in the BoS with an outside option (Example 41).

Example 53. Consider the BoS with an Outside Option. We omit action "wait" from the description of player 2's strategies, and we illustrate the steps of initial rationalizability. For player 1, we have $r_1(\mu^1) = \{\text{out.B}_1, \text{out.S}_1\}$ for each CPS μ^1 such that $\mu^1(\text{S}_2|\varnothing) = 1$, and $r_1(\mu^1) = \{\text{in.B}_1\}$ for each $\mu^1 \in \Delta^H(S_2)$ such that $\mu^1(\text{B}_2|\varnothing) = 1$. For player 2, we have $r_2(\mu^2) = \{\text{B}_2\}$ for each CPS μ^2 such that $\mu^2(\text{in.B}_1|\varnothing) = 1$, and $r_2(\mu^2) = \{\text{S}_2\}$ for each CPS μ^2 such that $\mu^2(\text{in.S}_1|\varnothing) = 1$. Thus,

$$\rho^1(S) = \{\text{out.B}_1, \text{out.S}_1, \text{in.B}_1\} \times \{\text{B}_2, \text{S}_2\}.$$

At second step, no strategy for player 1 is deleted. For player 2, we have $r_2(\mu^2) = \{\text{B}_2\}$ for each CPS μ^2 such that $\mu^2(\text{in.B}_1|\varnothing) = 1$, and $r_2(\mu^2) = \{\text{S}_2\}$ for each CPS μ^2 such that $\mu^2(\{\text{out.B}_1, \text{out.S}_1\}|\varnothing) = 1$ and $\mu^2(\text{in.S}_1|(\text{in})) = 1$. Hence,

$$\rho^2(S) = \{\text{out.B}_1, \text{out.S}_1, \text{in.B}_1\} \times \{\text{B}_2, \text{S}_2\}.$$

No strategy has been eliminated with the second step of the procedure. Therefore

$$\rho^\infty(S) = \{\text{out.B}_1, \text{out.S}_1, \text{in.B}_1\} \times \{\text{B}_2, \text{S}_2\}.$$

▲

Note that, in Example 53, the set of initially rationalizable strategies of player 1 includes in.B_1 but not in.S_1. Despite this, the set of initially rationalizable strategies of player 2 includes not only B_2 but also S_2. This is so because player 2 can assign probability 1 to the initially rationalizable strategies of player 1 that prescribe action out; if this belief is contradicted by the observation of in, then player 2 is free to think that player 1 is not carrying out a rational plan, and that player 1 will then play S_1. The consequence is that both strategies of player 2 remain justifiable under initial belief in rationality of player 1. Thus, in this game, initial rationalizability yields the same strategies as rationalizability in the strategic form of the game (cf. Chapter 9, Section 9.3.1),[4] where both B_2 and S_2 survive simply because the strategy of player 2 is irrelevant for the outcome when player 1 plays out. In the next section, we will see what happens if instead player 2 believes that in was part of a rational plan (and player 1 believes that player 2 reasons in this way).

11.3 Strong Rationalizability

When the initial belief about co-players' behavior is falsified by observation, a player may still try to interpret unexpected moves as part of a rational plan. Obviously, this is not always possible as there may be observable behavior that is not part of any rational plan of the co-player. But when possible, this rationalization may narrow down what a player can expect from the co-players in the continuation of the game. This way of looking at past moves to predict future moves is called "forward-induction reasoning." Forward-induction reasoning is enabled precisely by the hypothesis that past and future moves of a co-player must be part of the same rational plan, which the co-player does not fail to carry out. An alternative way of reasoning is to think that the unexpected moves were mistakes that disrupted the implementation of co-players' plans; in this case, the future moves can at best be predicted from their optimality in the continuation of the game, a form of "backward-induction reasoning" that will be analyzed in Section 11.4.

Thus, forward-induction reasoning rests on maintaining the belief that the co-players are rational (i.e., they have a rational plan and implement it) at all histories that are consistent with their rationality. This persistency of the belief in an event as long as not contradicted by observation is called "strong belief."

Definition 64. *Fix a player $i \in I$, a CPS μ^i and a subset of co-players' strategy profiles $E_{-i} \subseteq S_{-i}$.*

We say that μ^i ***strongly believes*** *E_{-i} if μ^i* initially believes *E_{-i} ($\mu^i \in \Delta^H_\varnothing(E_{-i})$) and $\mu^i(E_{-i}|h) = 1$ for every $h \in H \backslash \{\varnothing\}$ such that $E_{-i} \cap S_{-i}(h) \neq \emptyset$.*

[4] Rationalizability in the strategic form of the game coincides with the iterated elimination of strategies that are not ex-ante optimal for some conjecture consistent with the previous steps—cf. Chapter 10.

We let $\Delta^H_{\mathrm{sb}}(E_{-i})$ denote the set of CPSs of player i that strongly believe E_{-i}.

Note that, by definition, strong belief implies initial belief, that is, $\Delta^H_{\mathrm{sb}}(E_{-i}) \subseteq \Delta^H_{\varnothing}(E_{-i})$ for every $E_{-i} \subseteq S_{-i}$. Strong belief in rationality is the cornerstone of strategic reasoning. Here we represent this as strong belief in the set of co-players' profiles of justifiable strategies.[5] More generally, we impose a **best rationalization principle**: players always ascribe to co-players the highest level of "strategic sophistication" that is consistent with the observed behavior. For example, on top of strong belief in rationality, we also require that each player strongly believes that *his co-players are rational and* that *they strongly believe in rationality.* The latter is the second order of strategic sophistication of co-players. As we consider further iterations of this form of strategic thinking, we want to study the behavioral implications of *rationality and common strong belief in rationality.* Such behavioral implications are characterized by a solution concept called **strong rationalizability**.

Definition 65. *Consider the following elimination procedure.*

(Step $n = 0$**)** *For each $i \in I$, let $S^0_i = S_i$. Also, let $S^0_{-i} = \times_{j \neq i} S_j$ and $S^0 = S$.*

(Step $n > 0$**)** *For each $i \in I$, let*

$$\begin{aligned} \Delta^n_i &= \bigcap\nolimits_{m=0}^{n-1} \Delta^H_{\mathrm{sb}}(S^m_{-i}); \\ S^n_i &= \left\{ s_i \in S_i : \exists \mu^i \in \Delta^n_i, s_i \in r_i(\mu^i) \right\}. \end{aligned}$$

Also, let $S^n_{-i} = \times_{j \neq i} S^n_j$ and $S^n = \times_{i \in I} S^n_i$.

Finally, for each $i \in I$, let $S^\infty_i = \bigcap_{n>0} S^n_i$, and $S^\infty = \times_{i \in I} S^\infty_i$. For each $i \in I$, the strategies in S^∞_i are called ***strongly rationalizable****.*

Let us compare strong and initial rationalizability. To ease the comparison, write the sequence of subsets obtained with the initial

[5] Recall that rationality is a relationship between beliefs and behavior. Yet, here we represent formally only beliefs about behavior, not beliefs about others' beliefs. Thus, we represent belief in rationality as belief in behavior consistent with rationality, i.e., belief in justifiable behavior.

rationalizability procedure as $\left(S^n_\varnothing\right)_{n\in\mathbb{N}} = (\rho^n(S))_{n\in\mathbb{N}}$ with $S^n_\varnothing = \times_{i\in I} S^n_{i,\varnothing}$. Since $\Delta^1_i = \Delta^H(S_{-i})$, the first step ($n = 1$) of strong rationalizability eliminates only the strategies of player i that are not justifiable, exactly like initial rationalizability: $S^1_i = S^1_{i,\varnothing}$. At the second step of reasoning ($n = 2$), Δ^2_i captures strong belief in the co-players' rationality, i.e., strong belief in S^1_{-i}: $\mu^i(S^1_{-i}|h) = 1$ for every $h \in H$ such that $S^1_{-i} \cap S_{-i}(h) \neq \emptyset$. Strong belief in an event E_{-i} clearly implies initial belief in E_{-i}, as the initial history is always consistent with the believed event. Hence,

$$\Delta^2_i = \Delta^H_{\text{sb}}\left(S^1_{-i}\right) \subseteq \Delta^H_\varnothing\left(S^1_{-i}\right) = \Delta^H_\varnothing\left(S^1_{-i,\varnothing}\right)$$

($i \in I$), so that $S^2 \subseteq S^2_\varnothing = \rho^2(S)$. Iterating this argument, we obtain the following result.

Remark 42. *Initial rationalizability is weaker than strong rationalizability, that is,*

$$S^n \subseteq S^n_\varnothing = \rho^n(S)$$

for all $n \in \mathbb{N}\cup\{\infty\}$.

Proof. We already know that the inclusion holds for the first two steps. Suppose, by way of induction, that $S^m \subseteq S^m_\varnothing$ for each $m < n$ (IH). Then, for each $i \in I$,

$$\Delta^n_i \overset{\text{(def)}}{=} \bigcap_{m=0}^{n-1} \Delta^H_{\text{sb}}\left(S^m_{-i}\right) \overset{\text{(sb}-\text{ib)}}{\subseteq} \bigcap_{m=0}^{n-1} \Delta^H_\varnothing\left(S^m_{-i}\right) \overset{\text{(IH,mon)}}{\subseteq} \Delta^H_\varnothing\left(S^{n-1}_{-i,\varnothing}\right),$$

where the equality holds by definition, the first inclusion holds because strong belief implies initial belief, and the second holds by the inductive hypothesis and the monotonicity of initial belief. Thus, for each $i \in I$,

$$S^n_i = r_i\left(\Delta^n_i\right) \subseteq r_i\left(\Delta^H_\varnothing\left(S^{n-1}_{-i,\varnothing}\right)\right) = S^n_{i,\varnothing},$$

which implies $S^n \subseteq \rho^n(S)$ for all $n \in \mathbb{N}$. The latter implies $S^\infty \subseteq \rho^\infty(S)$. ■

Example 54. In Example 53, we have seen that in the BoS with an outside option the initially rationalizable strategies coincide with the justifiable strategies. In particular, Bob could play S_2 even though in.S_1 is not justifiable for Ann because Bob can initially believe that Ann will rationally play out, and then think that in has not been played as part of a rational plan. Strong rationalizability, instead, requires Bob to believe that in has been played as part of the rational plan in.B_1. Therefore, as anticipated in Chapter 9, the observation of in leads Bob to believe, by forward induction, that Ann will play B_1 next. Thus, we have $S_2^2 = \{B_2\} = S_2^\infty$. Moreover, Ann anticipates this, because she (strongly) believes that (i) Bob is rational, and (ii) Bob strongly believes that she is rational. Thus, we have $S_1^3 = \{\text{in.B}_1\} = S_1^\infty$. ▲

A more involved example of forward-induction reasoning captured by strong rationalizability is provided by the BoS with Dissipative Action.

Example 55. Consider the game in Example 42, Chapter 9, where Bob (b) is the first-mover. All strategies of Ann (a) are justifiable. Note that for Bob, given any CPS μ^b, the weakly sequentially optimal strategies are entirely determined by $\mu^b(\cdot|\varnothing)$; this is because Bob never plays after observing an action of Ann, hence Bob never has to revise $\mu^b(\cdot|\varnothing)$ before the game is over. Now, for any initial belief, it is not optimal for Bob to burn and then aim for the least preferred coordination outcome. So, we have

$$S_b^1 = \{B.L.r, B.R.r\} \cup \{s_b \in S_b : s_b(\varnothing) = N\}.$$

No strategy of Ann can be eliminated at the first step. Hence, no strategy of Bob can be eliminated at the second step, and so on; therefore we can focus on just one player at each step of reasoning. For Ann, strong belief in S_b^1 entails that playing u at history (B) is suboptimal. So, we have

$$\begin{aligned} \Delta_a^2 &= \left\{\mu^a \in \Delta^H(S_b) : \mu^a(\{B.L.r, B.R.r\} \,|(B)) = 1\right\}, \\ S_a^2 &= \{U.d, D.d\}. \end{aligned}$$

Then, for Bob, (initial) belief in S_a^2 guarantees that burning will effectively signal the intention to coordinate on his preferred outcome. Then, not burning makes sense for him only if he assigns sufficiently high probability to Ann playing d also after not burning. Hence, we have

$$\begin{aligned} \Delta_b^3 &= \left\{\mu^b \in \Delta^H(S_a) : \mu^b(\{U.d, D.d\} \,|\varnothing) = 1\right\}, \\ S_b^3 &= \{N.R.l, N.R.r, B.L.r, B.R.r\}. \end{aligned}$$

So, for Ann, strong belief in S_b^3 translates into certainty of R at history (N), on top of certainty of r at history (B). Thus, we have

$$\begin{aligned}\Delta_a^4 &= \left\{ \begin{array}{ll} \mu^a \in \Delta^H(S_b): & \mu^a(\{B.L.r, B.R.r\}\,|(B)) = \\ & \mu^a(\{N.R.l, N.R.r\}\,|(N)) = 1 \end{array} \right\}, \\ S_a^4 &= \{D.d\} = S_a^\infty.\end{aligned}$$

For Bob, (initial) belief in S_a^4 guarantees that his preferred coordination outcome can be obtained regardless of his initial move, and then, obviously, it is optimal for him not to burn:

$$\begin{aligned}\Delta_b^5 &= \left\{\mu^b \in \Delta^H(S_a) : \mu^b(\{D.d\}\,|\varnothing) = 1\right\}, \\ S_b^5 &= \{N.R.l, N.R.r\}.\end{aligned}$$

Forward-induction reasoning, as captured by strong rationalizability, pins down a unique path: $(N, (R, D))$. The mere possibility of burning allows Bob to induce coordination on his preferred outcome even without actually using the burning option. ▲

11.3.1 Strong Rationalizability as a Reduction Procedure

An attentive reader probably noticed that we have not defined strong rationalizability with the iterated application of a justification operator, as we did for rationalizability in simultaneous-move games, and for initial rationalizability. Why is it so? The reason is that the best rationalization principle requires each player i, at any step of reasoning $n > 2$, to strongly believe in all the sets/events $S_{-i}^{n-1}, ..., S_{-i}^1$, not just S_{-i}^{n-1}. If instead we were to impose only strong belief in S_{-i}^{n-1}, then Δ_i^n would not necessarily be a subset of Δ_i^{n-1}, and thus we would not have a well-defined elimination procedure. The point is that strong belief in S_{-i}^{n-1} is not a stronger condition than strong belief in S_{-i}^{n-2}: although $S_{-i}^{n-1} \subseteq S_{-i}^{n-2}$, the set $\Delta_{\text{sb}}^H(S_{-i}^{n-1})$ is not necessarily a subset of $\Delta_{\text{sb}}^H(S_{-i}^{n-2})$, because the set of histories consistent with S_{-i}^{n-1} can be a strict subset of those consistent with S_{-i}^{n-2}, and thus strong belief in S_{-i}^{n-1} may restrict i's conditional beliefs at fewer histories than strong belief in S_{-i}^{n-2}. (In general, the smaller an event, the fewer the contingencies where one can believe the event is true.)

An alternative route is to reformulate strong rationalizability as a *reduction* procedure. By focusing on the strategies that survived the previous steps of elimination, we can impose strong belief in S_{-i}^{n-1} only, and rely on the fact that the strategies in S_i^{n-1} already capture the best rationalization principle for the lower orders of belief, i.e., they already capture strong belief in $S_{-i}^{n-2}, ..., S_{-i}^{1}$. In this way, we overcome the nonmononicity of strong belief and we can write the reduction procedure with a *constrained* justification operator. To this end, fix any $C = \times_{i \in I} C_i \in \mathcal{C}$ and let

$$H(C) = \{h \in H : \exists s \in C, h \prec \zeta(s)\} = \{h \in H : S(h) \cap C \neq \emptyset\}$$

denote the set of nonterminal histories consistent with some strategy profile in C. We define a constrained optimality correspondence as follows: for each $i \in I$ and $\mu^i \in \Delta_{\text{sb}}^H(C_{-i})$, let

$$r_i(\mu^i|C) = \left\{ \overline{s}_i \in C_i : \forall h \in H_i(\overline{s}_i) \cap H(C), V_i^{\overline{s}_i,\mu^i}(h) = \max_{s_i \in C_i \cap S_i(h)} V_i^{s_i,\mu^i}(h) \right\}.$$

Thus, $r_i(\mu^i|C)$ is the set of strategies that, among the strategies in C_i, maximize the continuation-value given $\mu^i(\cdot|h)$ at every history h consistent with C (and with the strategy itself). The correspondence $r_i(\cdot|C)$ would be well defined also if the domain contained all CPSs, but conceptually (and for our purpose) the CPSs that strongly believe C_{-i} are the interesting ones: it is under these CPSs that player i can focus exclusively on the reduced set C, as long as no deviation from the paths in $\zeta(C)$ is observed.

With this, we introduce the reduction operator $\overline{\rho}_{\text{sb}} : \mathcal{C} \to \mathcal{C}$ as follows: for every $C \in \mathcal{C}$,

$$\begin{aligned} \overline{\rho}_{i,\text{sb}}(C) &= r_i(\Delta_{\text{sb}}^H(C_{-i})|C), \\ \overline{\rho}_{\text{sb}}(C) &= \bigtimes_{i \in I} \overline{\rho}_{i,\text{sb}}(C). \end{aligned}$$

Finally, we can iterate $\overline{\rho}_{\text{sb}}$ starting from S in the usual way:

$$\begin{aligned} & \overline{\rho}_{\text{sb}}^0(S) = S, \\ \forall k > 0, \quad & \overline{\rho}_{\text{sb}}^k(S) = \overline{\rho}_{\text{sb}}(\overline{\rho}_{\text{sb}}^{k-1}(S)), \\ & \overline{\rho}_{\text{sb}}^\infty(S) = \textstyle\bigcap_{k \geq 1} \overline{\rho}_{\text{sb}}^k(S). \end{aligned}$$

We want to prove that the reduction procedure $(\overline{\rho}_{\text{sb}}^k(S))_{k=0}^\infty$ coincides with strong rationalizability. Compared to the analogous result for rationalizability in Chapter 4, there is here a difficulty: even if $\overline{\rho}_{\text{sb}}^{k-1}(S) = S^{k-1}$, the set of CPSs $\Delta_{\text{sb}}^{H}(\overline{\rho}_{-i,\text{sb}}^{k-1}(S))$ is larger than the set Δ_i^k we use in the definition of strong rationalizability, which is, the set of CPSs that strongly believe $S_{-i}^{k-1}, ..., S_{-i}^{0}$. Nonetheless, the expected equivalence holds.

Lemma 33. *For each* $k \in \mathbb{N} \cup \{\infty\}$, $\overline{\rho}_{\text{sb}}^{k}(S) = S^k$.

Proof. Obviously, $\overline{\rho}_{\text{sb}}^{0}(S) = S^0$. So, fix $k \geq 0$ and suppose by way of induction that $\overline{\rho}_{\text{sb}}^{k}(S) = S^k$. We need to show that, for each $i \in I$,

$$r_i(\Delta_{\text{sb}}^{H}(\overline{\rho}_{-i,\text{sb}}^{k}(S))|\overline{\rho}_{\text{sb}}^{k}(S)) = r_i(\Delta_i^{k+1});$$

then, $\overline{\rho}_{i,\text{sb}}^{k+1}(S) = S_i^{k+1}$ follows by definition.

We first show that $r_i(\Delta_i^{k+1}) \subseteq r_i(\Delta_{\text{sb}}^{H}(\overline{\rho}_{-i,\text{sb}}^{k}(S))|\overline{\rho}_{\text{sb}}^{k}(S))$. Fix any $s_i \in r_i(\Delta_i^{k+1})$. Then, by definition, there exists μ^i that strongly believes S_{-i}^k such that $s_i \in r_i(\mu^i)$. By the induction hypothesis, $S_{-i}^k = \overline{\rho}_{-i,\text{sb}}^{k}(S)$. Hence $\mu^i \in \Delta_{\text{sb}}^{H}(\overline{\rho}_{-i,\text{sb}}^{k}(S)) = \Delta_{\text{sb}}^{H}(S_{-i}^k)$, which implies $s_i \in S_i^{k+1} \subseteq S_i^k = \overline{\rho}_{i,\text{sb}}^{k}(S)$. With this, we obtain

$$s_i \in r_i(\Delta_{\text{sb}}^{H}(\overline{\rho}_{-i,\text{sb}}^{k}(S))|\overline{\rho}_{\text{sb}}^{k}(S)).$$

We now show the opposite inclusion. Fix any $\overline{s}_i \in r_i(\Delta_{\text{sb}}^{H}(\overline{\rho}_{-i,\text{sb}}^{k}(S))|\overline{\rho}_{\text{sb}}^{k}(S))$. Then, by definition, there exists μ^i that strongly believes $\overline{\rho}_{-i,\text{sb}}^{k}(S)$ such that $\overline{s}_i \in r_i(\mu^i|\overline{\rho}_{\text{sb}}^{k}(S))$. By the induction hypothesis, $S_{-i}^k = \overline{\rho}_{-i,\text{sb}}^{k}(S)$, hence μ^i strongly believes S_{-i}^k. By definition of $r_i(\cdot|\overline{\rho}_{\text{sb}}^{k}(S))$ and the induction hypothesis, $\overline{s}_i \in \overline{\rho}_{i,\text{sb}}^{k}(S) = S_i^k$. Therefore, there is $\hat{\mu}^i$ that strongly believes $S_{-i}^{k-1}, ..., S_{-i}^{0}$ such that $\overline{s}_i \in r_i(\hat{\mu}^i)$. Then, we can construct a CPS $\bar{\mu}^i \in \Delta_i^{k+1}$ as follows: for every $h \in H$,

$$\bar{\mu}^i(\cdot|h) = \begin{cases} \mu^i(\cdot|h), & \text{if } S_{-i}^k \cap S_{-i}(h) \neq \emptyset, \\ \hat{\mu}^i(\cdot|h), & \text{otherwise.} \end{cases}$$

As $\bar{\mu}^i$ is a well-defined CPS,[6] it remains to show that $\overline{s}_i \in r_i(\bar{\mu}^i)$. For each $h \in H_i(\overline{s}_i)$, we have

$$V_i^{\overline{s}_i,\bar{\mu}^i}(h) = \max_{s_i \in \overline{\rho}_{i,\text{sb}}^{k}(S) \cap S_i(h)} V_i^{s_i,\bar{\mu}^i}(h).$$

[6]The chain rule holds: to see this, pick any $h' \in H$ such that $S_{-i}^k \cap S_{-i}(h') = \emptyset$ and $h \in H$ such that $S_{-i}^k \cap S_{-i}(h) \neq \emptyset$. Then we have $\mu^i(S_{-i}(h')|h) = 0$, as μ^i strongly believes S_{-i}^k.

Specifically, if $S^k_{-i} \cap S_{-i}(h) = \emptyset$, this follows from the fact that $\overline{s}_i \in r_i(\hat{\mu}^i)$ and $\bar{\mu}^i(\cdot|h) = \hat{\mu}^i(\cdot|h)$; if instead $S^k_{-i} \cap S_{-i}(h) \neq \emptyset$, then the above conclusion still holds, because (i) $\overline{\rho}^k_{-i,\text{sb}} \cap S_{-i}(h) \neq \emptyset$ (by the inductive hypothesis), (ii) $\overline{s}_i \in r_i(\mu^i | \overline{\rho}^k_{\text{sb}}(S))$ (by the inductive hypothesis), and (iii) $\bar{\mu}^i(\cdot|h) = \mu^i(\cdot|h)$. Using again the inductive hypothesis, we obtain $\overline{\rho}^k_{i,\text{sb}}(S) = S^k_i$. Moreover, $\bar{\mu}^i \in \Delta^k_i$. Therefore, $r_i(\bar{\mu}^i) \subseteq \overline{\rho}^k_{i,\text{sb}}(S)$. Then, by Lemma 32, $\overline{s}_i \in r_i(\bar{\mu}^i)$. ■

11.3.2 Independent Rationalization

A natural variant of strong rationalizability arises from the observation that, under the baseline definition, player i is free to believe that *all* his co-players are not rational once just one co-player chooses an action that is not prescribed by any justifiable strategy. To capture the hypothesis that each player i ascribes to each individual co-player j the highest level of strategic sophistication that is consistent with j's behavior, we introduce the following variation of strong rationalizability, which we call **independent strong rationalizability**.

Definition 66. *For each $i \in I$, let $S^0_{i,\text{ir}} = S_i$. For each $n > 0$, let $s_i \in S^n_{i,\text{ir}}$ if and only if there exists $\mu^i \in \Delta^H(S_{-i})$ such that:*

1. *for all $j \neq i$, $m = 0, ..., n-1$, and $h \in H(S^m_{j,\text{ir}} \times S_{-j})$,*
$$\mu^i \left(S^m_{j,\text{ir}} \times \left(\bigtimes_{k \neq i,j} S_k \right) \middle| h \right) = 1;$$
2. *$s_i \in r_i(\mu^i)$.*

   Finally, let $S^\infty_{i,\text{ir}} = \bigcap_{n>0} S^n_{i,\text{ir}}$.

A natural question is whether independent strong rationalizability is equivalent to strong rationalizability. The answer is Yes and no. In terms of strategies, this version of strong rationalizability can yield a different result, but the induced set of possible terminal histories does not change. In this case, we say that the two procedures are path-equivalent. (The proof is in the appendix.)

Theorem 43. *For each $n \in \mathbb{N} \cup \{\infty\}$, $\zeta(S^n_{\text{ir}}) = \zeta(S^n)$.*

The intuition behind Theorem 43 is that *independent rationalization* (i.e, the belief that each individual co-player has the highest level of strategic sophistication that is consistent with his own behavior) only has bite at histories that are not compatible with strategic reasoning for one or more co-players (although they are consistent with strategic reasoning for at least one co-player).

So far, we allowed for the possibility that players have correlated beliefs about the behavior of different co-players. Sometimes, it makes sense for a player to assume that, according to his subjective beliefs, the strategies of different co-players are mutually independent. This means that what player i believes about co-player j does not depend on the observed behavior of co-player $k \neq j$. This assumption on players' beliefs is called *strategic independence*, and it is naturally complemented by independent rationalization. Here we do not provide the formal definition of strong rationalizability under strategic independence, but we do observe that, differently from mere independent rationalization, strategic independence refines the set of strongly rationalizable paths.[7]

11.3.3 Iterated Conditional Dominance

The reformulation of strong rationalizability as a reduction procedure opens up the opportunity to characterize it with a notion of iterated dominance. The appropriate notion of dominance is conditional dominance. In Chapter 10 we have shown that a strategy is conditionally undominated if and only if it is weakly sequentially optimal under a CPS, absent any restrictions on players' CPSs given by strategic reasoning. To characterize strong rationalizability with *iterated* conditional dominance, we need to extend the notion of conditional dominance and its relation with weak sequential optimality to a "reduced game" $C = \times_{i \in I} C_i$, where C_{-i} describes the contingent behavior that each player i expects from the co-players. For each $i \in I$, say that a strategy $\bar{s}_i \in C_i$ is **conditionally dominated in** C if there exist a history $h \in H_i(\bar{s}_i) \cap H(C)$ and a mixed strategy $\sigma_i \in \Delta(C_i)$, with $\sigma_i(S_i(h)) = 1$, such that

$$\forall s_{-i} \in C_{-i} \cap S_{-i}(h), \quad \sum_{s_i \in C_i \cap S_i(h)} \sigma_i(s_i) u_i(\zeta(s_i, s_{-i})) > u_i(\zeta(\bar{s}_i, s_{-i})).$$

[7] See Battigalli [7].

We say that $\bar{s}_i \in C_i$ is **conditionally undominated in** C if it is *not* conditionally dominated in C. The set of strategies of player i that are *conditionally undominated* (Not Conditionally Dominated) *in* C is denoted by $\mathrm{NCD}_i(C)$.

Lemma 34. *Fix $C = \times_{i\in I} C_i \in \mathcal{C}$. A strategy $\bar{s}_i \in C_i$ is conditionally undominated in C if and only if there exists $\mu^i \in \Delta^H_{\mathrm{sb}}(C_{-i})$ such that $\bar{s}_i \in r_i(\mu^i|C)$.*

Proof. The proof follows the same lines as the proof of Lemma 30 in Chapter 10. Here we adapt the key steps, and the reader can consult the proof of Lemma 30 for further details.

Suppose there exists $\mu^i \in \Delta^H_{\mathrm{sb}}(C_{-i})$ such that $\bar{s}_i \in r_i(\mu^i|C)$. Then, for every $h \in H_i(\bar{s}_i) \cap H(C)$,

$$\bar{s}_i \in \arg \max_{s_i \in C_i \cap S_i(h)} \sum_{s_{-i}\in S_{-i}(h)} u_i(\zeta(s_i, s_{-i}))\mu^i(s_{-i}|h),$$

and since μ^i strongly believes C_{-i},

$$\mu^i(C_{-i} \cap S_{-i}(h)|h) = 1.$$

Then, by Lemma 2,[8] no mixed action $\sigma_i \in \Delta(C_i \cap S_i(h))$ dominates $\bar{s}_i$ in $(C_i \cap S_i(h)) \times (C_{-i} \cap S_{-i}(h))$. Thus, $\bar{s}_i$ is not conditionally dominated in C.

Conversely, suppose now that $\bar{s}_i$ is not conditionally dominated in C. Thus, for each $h \in H_i(\bar{s}_i)$, $\bar{s}_i$ is not dominated in $(C_i \cap S_i(h)) \times (C_{-i} \cap S_{-i}(h))$. Then, by Lemma 2, there exists $\nu^{i,h} \in \Delta(C_{-i} \cap S_{-i}(h))$

$$\bar{s}_i \in \arg \max_{s_i \in C_i \cap S_i(h)} \sum_{s_{-i}\in C_{-i}\cap S_{-i}(h)} u_i(\zeta(s_i, s_{-i}))\nu^{i,h}(s_{-i}|h)$$

For each $h \notin H_i(\bar{s}_i)$, fix $\nu^{i,h} \in \Delta(S_{-i}(h))$ such that, if $h \in H(C_{-i})$, $\nu^{i,h}(C_{-i}) = 1$. Then, construct a CPS $\mu^i \in \Delta^H_{\mathrm{sb}}(C_{-i})$ and check that $\bar{s}_i \in r_i(\mu^i|C)$ as in the proof of Lemma 30. ∎

We are now in position to state the equivalence between strong rationalizability and iterated conditional dominance. Let $\mathrm{NCD}(C) =$

[8] Lemma 2 is applied here to the strategic form of the game reduced to C.

$\times_{i\in I} \mathrm{NCD}_i(C) \subseteq S$ denote the set of conditionally undominated strategy profiles in $C = \times_{i\in I} C_i \in \mathcal{C}$. This defines a restriction operator, that is, an operator $\mathrm{NCD} : \mathcal{C} \to \mathcal{C}$ such that $\mathrm{NCD}\,(C) \subseteq C$ for every $C \in \mathcal{C}$. As we did for iterated dominance in Chapter 4, we can represent iterated conditional dominance through the following sequence of subsets $\left(\mathrm{NCD}^k(S)\right)_{k\in\mathbb{N}}$, where—as usual—$\mathrm{NCD}^k = \mathrm{NCD} \circ \mathrm{NCD}^{k-1}$ (with $\mathrm{NCD}^0\,(C) = C$ for every C).

Definition 67. *A profile of strategies $s \in S$ survives iterated conditional dominance if $s \in \mathrm{NCD}^\infty(S) = \bigcap_{k\geq 1} \mathrm{NCD}^k(S)$.*

Theorem 44. *For every $k = 1, 2, ...,$ we have $S^k = \mathrm{NCD}^k(S)$. Therefore, a strategy profile is strongly rationalizable if and only if it survives iterated conditional dominance.*

Proof. For each $k \in \mathbb{N} \cup \{\infty\}$, it follows from Lemma 33 that $\overline{\rho}^k_{\mathrm{sb}}(S) = S^k$. By definition, for every $C \in \mathcal{C}$ and $i \in I$, we have $\overline{\rho}_{i,\mathrm{sb}}(C) = r_i(\Delta^H_{\mathrm{sb}}(C_{-i})|C)$. Then, by Lemma 34, $\overline{\rho}^1_{\mathrm{sb}}(S) = \mathrm{NCD}^1(S)$, and by a simple inductive argument, $S^k = \mathrm{NCD}^k(S)$ for every $k \in \mathbb{N}$. ■

In Chapter 10 we have shown that conditional dominance and weak dominance in the strategic form of the game are generically equivalent. The intuition is that, generically, there are no ties between a player's payoffs at different terminal histories, and in this case the additional "caution" captured by weak dominance but not by conditional dominance has no bite. In light of this result, strong rationalizability is generically equivalent to **iterated admissibility**, i.e., the iterated deletion of all weakly dominated strategies. To prove this result, we need first to extend Lemma 31 to conditional dominance within a set of strategy profiles C—the proof is in the appendix of this chapter. For any Cartesian set C of strategy profiles, let $\mathrm{NWD}_i\,(C)$ denote the set of strategies of player i that are not weakly dominated in C, that is, are not weakly dominated in the strategic form restricted to C, $\left\langle I, (U_i|_C, C_i)_{i\in I}\right\rangle$, and let $\mathrm{NWD}\,(C) = \times_{i\in I} \mathrm{NWD}_i\,(C)$.

Lemma 35. *Fix arbitrarily a Cartesian subset of strategy profiles $C = \times_{i\in I} C_i$. Generically,*[9] $\mathrm{NCD}(C) = \mathrm{NWD}\,(C)$.

Theorem 45. *Generically, $S^\infty = \mathrm{NWD}^\infty(S)$.*

[9] For a precise definition of genericity, see Chapter 10.

Proof. Fix $k \geq 0$ and suppose by induction that, generically, $S^k = \mathrm{NWD}^k(S)$ (it is trivially true for $n = 0$). By Theorem 44, $S^{k+1} = \mathrm{NCD}^{k+1}(S)$. Let $C = S^k$. Thus, $\mathrm{NCD}^{k+1}(S) = \mathrm{NCD}(C)$. By Lemma 35, we have $\mathrm{NCD}(C) = \mathrm{NWD}(C)$, except for a subset of payoff vectors W with zero-measure closure. By the inductive hypothesis, we also have

$$\mathrm{NCD}(C) = \mathrm{NCD}(\mathrm{NWD}^k(S)) = \mathrm{NWD}^{k+1}(S),$$

except for a subset of payoff vectors W' with zero-measure closure. The union of the closures of W and W' has zero measure. Moreover, it is a closed set and contains $W \cup W'$. Hence, it contains the closure of $W \cup W'$. It follows that the closure of $W \cup W'$ has zero measure. We conclude that, generically, $S^{k+1} = \mathrm{NWD}^{k+1}(S)$. Since the game is finite, there exists $M \in \mathbb{N}$ such that $S^\infty = S^M$ and $\mathrm{NWD}^\infty(S) = \mathrm{NWD}^M(S)$, and we have proven that, generically, $S^M = \mathrm{NWD}^M(S)$. ■

In some games, different terminal histories are associated with the same outcome, or at least with outcomes that are payoff-equivalent for some players. This is an important, albeit non-generic, class of games; for instance, the (often sequential) games used by social planners to achieve certain goals typically assign to players the same outcome (e.g., the losing outcome in an ascending auction, or a particular match in a school admissions mechanism). A natural conjecture is that, when there are ties between some players' payoffs at different terminal histories, or for more general forms of non-genericity leading to violations of the equivalence stateed in Theorem 45, the terminal histories that are consistent with iterated admissibility would still be a subset of those consistent with strong rationalizability, because of the additional bite of the form of caution captured by weak dominance. It turns out that this conjecture is wrong, as shown by the following example due to Catonini [33].

Example 56. Consider the two-player, two-stage game depicted below. In the first stage (left matrix), Ann and Bob simultaneously choose an action. The game ends under all action pairs except for (U, R) and (M, R), after which the game moves to a second stage. Both after (U, R) (middle matrix) and after (M, R) (right matrix), Ann and Bob move simultaneously (Ann chooses the row and Bob chooses the column), and then the game ends.

$A \backslash B$	L	C	R
U	$2,4$	$2,0$	Γ_1
M	$0,2$	$0,2$	Γ_2
D	$0,0$	$0,4$	$2,2$

Γ_1	W	E
N	$1,2$	$5,5$
S	$1,2$	$5,0$

Γ_2	F	G	H
T	$4,1$	$0,0$	$0,0$
B	$0,1$	$1,5$	$5,0$

To carry out the iterated deletion of weakly dominated strategies, it is useful to visualize the reduced strategic form of the game.

$A \backslash B$	L	C	$R.W.F$	$R.W.G$	$R.W.H$	$R.E.F$	$R.E.G$	$R.E.H$
$U.N$	$2,4$	$2,0$	$1,2$	$1,2$	$1,2$	$5,5$	$5,5$	$5,5$
$U.S$	$2,4$	$2,0$	$1,2$	$1,2$	$1,2$	$5,0$	$5,0$	$5,0$
$M.T$	$0,2$	$0,2$	$4,1$	$0,0$	$0,0$	$4,1$	$0,0$	$0,0$
$M.B$	$0,2$	$0,2$	$0,1$	$1,5$	$5,0$	$0,1$	$1,5$	$5,0$
D	$0,0$	$0,4$	$2,2$	$2,2$	$2,2$	$2,2$	$2,2$	$2,2$

We leave as an exercise to verify the following steps of elimination:

$\mathrm{NWD}_i^n(S)$	$i = 1$ (Ann)	$i = 2$ (Bob)
$n = 1$	$\{U.N, U.S, M.T, M.B, D\}$	$\{L, C, R.W.G, R.E.F, R.E.G\}$
$n = 2$	$\{U.N, U.S, D\}$	$\{L, C, R.W.G, R.E.F, R.E.G\}$
$n = \infty$	$\{U.N, U.S, D\}$	$\{L, C, R.W.G, R.E.F, R.E.G\}$

Strong Rationalizability requires more steps of elimination, as illustrated by the next table:

S_i^n	$i = 1$ (Ann)	$i = 2$ (Bob)
$n = 1$	S_1	$S_2 \backslash \{R.W.H, R.E.H\}$
$n = 2$	$S_1 \backslash \{M.B\}$	$S_2 \backslash \{R.W.H, R.E.H\}$
$n = 3$	$S_1 \backslash \{M.B\}$	$\{L, C, R.W.F, R.E.F\}$
$n = 4$	$\{U.N, U.S, M.T\}$	$\{L, C, R.W.F, R.E.F\}$
$n = 5$	$\{U.N, U.S, M.T\}$	$\{L, C, R.E.F\}$
$n = 6$	$\{U.N, U.S\}$	$\{L, C, R.E.F\}$
$n = 7$	$\{U.N, U.S\}$	$\{L, R.E.F\}$
$n = 8$	$\{U.N, U.S\}$	$\{L, R.E.F\}$
$n = \infty$	$\{U.N, U.S\}$	$\{L, R.E.F\}$

We leave as an exercise to verify the steps, which are explained in Catonini [33]; here we only discuss the crucial difference with the iterated elimination of weakly dominated strategies. Strategy $R.W.F$ of Bob is weakly dominated by the mixed strategy that gives probability 0.5 to L and C; however, $R.W.F$ is justified by the CPS that initially assigns probability 0.5 to U and D, and probability 1 to $M.T$ at history (M, R). Then, while both $M.T$ and $M.B$ can be eliminated at the second step of iterated weak dominance, $M.T$ survives the second step of strong rationalizability, as it is optimal for Ann against $R.W.F$. This implies that, at the third step of strong rationalizability, Bob must still assign probability 1 to S_1^2 at history (M, R), so in particular he must concentrate his belief on $M.T$. This refinement of beliefs at a history that is ruled out by iterated weak dominance but not (yet) by strong rationalizability induces the further elimination steps, which eventually lead strong rationalizability to refine the path-predictions of iterated weak dominance. In the following figure, the payoffs in **bold** indicate the terminal histories that are consistent with iterated weak dominance but not with strong rationalizability.

$A\backslash B$	L	C	R
U	$2,4$	$\mathbf{2,0}$	Γ_1
M	$0,2$	$0,2$	Γ_2
D	$\mathbf{0,0}$	$\mathbf{0,4}$	$\mathbf{2,2}$

Γ_1	W	E
N	$\mathbf{1,2}$	$5,5$
S	$\mathbf{1,2}$	$5,0$

Γ_2	F	G	H
T	$4,1$	$0,0$	$0,0$
B	$0,1$	$1,5$	$5,0$

▲

11.3.4 Elimination Orders

From an algorithmic viewpoint, it is interesting to understand whether "slowing down" the elimination of conditionally dominated strategies may change the final output of the procedure. For instance, in a two-player game, it could be convenient to alternate the eliminations for the two players in the following way: first compute the justifiable strategies S_1^1 of player 1 only; then compute directly the strategies S_2^2 of player 2 that are justified by a CPS that strongly believes S_1^1; then let $\widetilde{S}_1^3$ be the set of strategies of player 1 that are justified by a CPS that strongly believes S_2^2, and so on. In Chapter 4, we have seen that—in simultaneous-move games—the order of elimination of strictly dominated strategies is immaterial for the final output of iterated strict dominance. Here, the problem is much complicated by the non-monotonicity of strong belief. In the procedure above, we have $\widetilde{S}_1^3 \supseteq S_1^3$, because the set of CPSs that strongly believe S_2^2 are clearly a superset of those that strongly believe S_2^2 *and* S_2^1. However, the set of CPSs that strongly believe $\widetilde{S}_1^3$ is not a superset of those that strongly believe S_1^3, and at this point $\widetilde{S}_2^4$ need not be a superset of S_2^4. Where does this lead to? Fortunately, although the final output will look different from strong rationalizability in terms of strategies, the induced terminal histories will be exactly the same. We omit the details.[10]

The iterated elimination of *weakly* dominated strategies from the strategic form of the game, instead, does not have the same order independence property. Note that, in a game in which the iterated elimination of weakly dominated strategies is indeed order-independent in terms of induced terminal histories, such terminal histories will all be consistent with strong rationalizability. This is because the iterated elimination of conditionally dominated strategies can be seen as a slow, unfinished order of elimination of weakly dominated strategies: by Proposition 5, a conditionally dominated strategy is also weakly dominated. We omit the formalization of this result.

11.4 Continuation-Rationalizability and Backward Induction

In this section, we introduce another version of the rationalizability idea for multistage games, called "continuation-rationalizability." After exploring its relation with the other notions of rationalizability introduced in this chapter, we will show that—in *finite* games—the predictions of continuation-rationalizability can also be computed by means of a convenient backward procedure, which starts from the end of the game and procedes backward. For this reason, the solution concept is also called "backward rationalizability." In games with perfect information, this procedure coincides with the well-known backward-induction algorithm,

[10] The outcome-equivalence of different orders of elimination of conditionally dominated strategies was first proven by Chen and Micali [36]. Perea [70] directly proves the order independence (in terms of outcomes) of the strong belief operator. Catonini [32] proves outcome inclusions and equivalences between different notions of rationalizability for sequential games.

which we introduce at the end of the section.[11]

11.4.1 Continuation-Rationalizability

The distinctive feature of continuation-rationalizability is that each player i, whenever surprised by an action of a co-player j, may believe that j—instead of carrying out his strategy (plan)—chose this action by mistake; yet, i deems it impossible that j will make other mistakes as he carries out his strategy in the future. So, once surprised, player i keeps assigning probability 1 to strategies of j that are consistent with strategic reasoning (differently from initial rationalizability), but not necessarily consistent with his past behavior (differently from strong rationalizability).

Given this interpretation of unexpected behavior, a player needs to reason about what moves the co-players subjectively deem optimal at future histories that are *in*consistent with their plans. Thus, it is convenient to describe players' rational planning by means of *sequentially optimal strategies*, rather than *weakly* sequentially optimal strategies. To introduce the notion of sequential optimality under a CPS, we recall (in a sligthly different, but equivalent way) the notion of the subjective value of a history h given a strategy and a CPS of player i. Fix a strategy $\bar{s}_i$ and a CPS μ^i. For each $h \in H \backslash H_i(\bar{s}_i)$, define

$$V_i^{\bar{s}_i,\mu^i}(h) = V_i^{s_i,\mu^i}(h)$$

for any $s_i \in S_i(h)$ that *coincides* with $\bar{s}_i$ at histories $h' \succeq h$, that is those that *weakly follow* h. We say that $\bar{s}_i$ is **sequentially optimal given** μ^i if

$$\forall h \in H,\, V_i^{\bar{s}_i,\mu^i}(h) = \max_{s_i \in S_i} V_i^{s_i,\mu^i}(h).$$

Let $\hat{r}_i(\mu^i)$ denote the set of all sequentially optimal strategies given μ^i. Sequential optimality given a CPS is equivalent to sequential optimality given the conjecture β^i derived from CPS μ^i, defined in Chapter 10 on rational planning.

Proposition 6. *For each CPS* $\mu^i \in \Delta^H(S_{-i})$*, define the corresponding conjecture* $\beta^i \in \bigtimes_{h \in H} \Delta(\mathcal{A}_{-i}(h))$ *as follows:*

$$\forall h \in H,\, \forall a_{-i} \in \mathcal{A}_{-i}(h),\; \beta^i(a_{-i}|h) = \mu^i\left(S_{-i}\left(h, a_{-i}\right)|h\right).$$

[11] For an in-depth analysis and foundation of the concepts presented in this section, see Perea [70] and Battigalli and De Vito [11].

Then, a strategy is sequentially optimal given μ^i if and only if it is sequentially optimal given β^i.

The proof of this result is analogous to the one given for weak sequential optimality, so we omit it (see Proposition 4 and its proof).

Before presenting the mathematical definition of this version of the rationalizability concept, it is important to recall our interpretations of the mathematical objects called "strategies" in our analysis of the behavioral implications of rationality and strategic reasoning. The strategies $s_j \in S_j = \times_{h\in H} \mathcal{A}_j(h)$ of any player j (e.g., a co-player of i) have a dual interpretation: they represent conjunctions of *conditional statements about behavior* of the form "if h occurred, j's action would be $s_j(h)$," and they also represent *plans* made of the conjunction of statements of the form "j plans to take action $s_j(h)$ should h occur." The natural connection between these two intepretations is that *rational players carry out their plans*. As we allow players to think that some observed actions of their co-players were "mistakes," that is, they were not planned, we have to refer to this dual intepretations in a rather subtle way. At any history h, player i—thinking of co-players' strategies as descriptions of behavior—assigns probability 0 to all the profiles s_{-i} that are inconsistent with h, that is, those outside $S_{-i}(h)$. On top of this—thinking of co-players' strategies as descriptions of plans—player i predicts that, starting from h, each co-player will continue according to one of the plans that i deems possible (e.g., consistent with some level of rationality and strategic reasoning) *even if carrying out such plans would have prevented h.* In other words, differently from initial and strong rationalizability, observed actions are not necessarily interpreted as evidence of co-players' plans; nonetheless, i believes that such plans will be carried out in the continuation of the game.

Recall that we let $s_j|h$ denote the continuation-strategy of player j in subgame $\Gamma(h)$ that selects the same actions as s_j at histories $h' \succeq h$ (see Chapter 9.2, Section 9.3 and Chapter 10). For any subset $\overline{S}_{-i}$ and nonterminal history h, we introduce the set

$$\chi^h_{-i}(\overline{S}_{-i}) = \left\{s_{-i} \in S_{-i}(h) : \exists s'_{-i} \in \overline{S}_{-i}, \forall j \in I\setminus\{i\}, s_j|h = s'_j|h\right\}.$$

In words, given a subset $\overline{S}_{-i}$ of co-players' profiles of plans that i deems possible, $\chi^h_{-i}(\overline{S}_{-i})$ is the set of all descriptions of the co-players' behavior that are consistent with h and coincide with some profile of plans in $\overline{S}_{-i}$ in the continuation-game[12] starting with h. With this, for any (nonempty) $\overline{S}_{-i}$, consider the following subset of CPSs of player i:

$$\Delta^H_{\text{fb}}(\overline{S}_{-i}) = \left\{\mu^i \in \Delta^H(S_{-i}) : \forall h \in H,\ \mu^i(\chi^h_{-i}(\overline{S}_{-i})|h) = 1\right\}.$$

In words, $\Delta^H_{\text{fb}}(\overline{S}_{-i})$ is the set of CPSs with "**future belief**" in $\overline{S}_{-i}$: at every history h, the CPS assigns probability 1 to the event that the behavior of the co-players from h onwards is correctly described by $\overline{S}_{-i}$ (interpreted as a set of profiles of co-players' plans). Note that whenever $\overline{S}'_{-i} \subseteq \overline{S}_{-i}$, we also have $\Delta^H_{\text{fb}}(\overline{S}'_{-i}) \subseteq \Delta^H_{\text{fb}}(\overline{S}_{-i})$ for each $i \in I$.

When $\overline{S}_{-i}$ consists of the strategies of the co-players that are sequentially optimal under some CPS, $\Delta^H_{\text{fb}}(\overline{S}_{-i})$ represents "belief in future rationality," i.e., the belief that the co-players, given their own beliefs, will behave as planned making subjectively rational choices in the future, even if they may have deviated from their rational plans in the past. Continuation-rationalizability characterizes the *behavioral implications of "rationality and common future belief in rationality"* as follows.

Definition 68. *Consider the following elimination procedure.*

(Step $n = 0$**)** *For each $i \in I$, let $\hat{S}^0_i = S_i$. Also, let $\hat{S}^0_{-i} = \times_{j\neq i} S_j$ and $\hat{S}^0 = S$.*

(Step $n > 0$**)** *For each $i \in I$, let*

$$\hat{S}^n_i = \left\{s_i \in S_i : \exists \mu^i \in \Delta^H_{\text{fb}}(\hat{S}^{n-1}_{-i}), s_i \in \hat{r}_i(\mu^i)\right\}.$$

Also, let $\hat{S}^n_{-i} = \times_{j\neq i} \hat{S}^n_j$ and $\hat{S}^n = \times_{i\in I} \hat{S}^n_i$.

*Finally, for each $i \in I$, let $\hat{S}^\infty_i = \bigcap_{n>0} \hat{S}_i$, and $\hat{S}^\infty = \times_{i\in I} \hat{S}^\infty_i$. For each $i \in I$, the strategies in $\hat{S}^\infty_i$ are called **continuation-rationalizable**.*

We will illustrate continuation-rationalizability by example in the next subsection. For now, note the following. At every step n of continuation-rationalizability, at every history h, each player i assigns probability 1 to co-players' strategies that, from h onwards, coincide with

[12] Symbol χ is the Greek letter Chi (pronounced "kai") for *c*ontinuation.

the strategies in $\hat{S}_{-i}^{n-1}$, i.e., those that survived the previous steps of elimination, *without any distinction between those consistent with h and those inconsistent with h*, since the co-players may have taken some of the actions that show up in h by mistake. Hence, differently from strong rationalizability, continuation-rationalizability does not embody the best rationalization principle. Consequently, continuation-rationalizability does not capture forward-induction reasoning, but just a form of backward-induction reasoning. In light of this, one may expect two things. First, that continuation-rationalizability yields weaker predictions about the path of play than strong rationalizability. Second, that the predictions of continuation-rationalizability can also be computed with a "backward procedure," which generalizes to games with observed actions the well-known backward-induction algorithm for games with perfect information. Both facts are true, as we are going to elaborate in the next two subsections.

11.4.2 Relationship with Initial Rationalizability and Strong Rationalizability

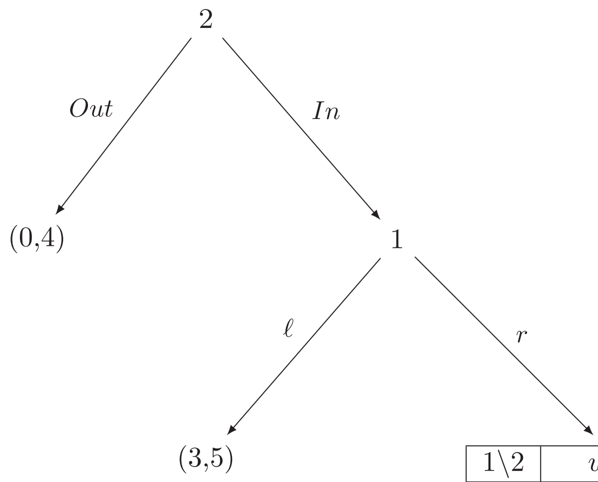
The following example illustrates continuation-rationalizability and its relation with initial and strong rationalizability.

Example 57. Player 2 moves first, choosing between In and Out. After Out, the game ends, after In, Player 1 chooses between ℓ and r. After ℓ the game ends, after r players move simultaneously for the last time. This game is illustrated in Figure 11.2.

Continuation-rationalizability works as follows. For player 2, every strategy is sequentially optimal under some CPS, while for player 1 strategies $r.n$ and $\ell.s$ are not. In particular, $r.n$ is conditionally dominated at history (In) by any strategy that prescribes ℓ, while s is not a continuation-best reply at history (In, r) to any belief that justifies ℓ at history (In), and player 1's belief is the same at the two histories by the chain rule, as he does not observe a new move by player 2. Hence,

$$\hat{S}_2^1 = S_2, \quad \hat{S}_1^1 = \{r.c, r.s, \ell.n, \ell.c\}\,.$$

At the second step of continuation-rationalizability, every strategy of player 2 that prescribes In is sequentially optimal under some $\mu^2 \in \Delta_{\text{fb}}^H(\hat{S}_1^1)$. In



1\2	w	m	e
n	2, 3	2, 2	2, 0
c	1, 2	8, 0	8, 1
s	0, 0	9, 2	9, 3

Figure 11.2: A 3-stage game.

particular, μ^2 can assign probability 1 to $\{\ell.n, \ell.c\}$ at the initial history, and this justifies In. Then, at history (In, r), player 2 must revise his belief, and μ^2 can assign positive probability to every $s_1 \in S_1(In, r)$, hence justifying either w, m, or e. This is because every $a_1 \in \mathcal{A}_1(In, r)$ is prescribed at (In, r) by some $s_1 \in \hat{S}_1^1$ — the fact that n is prescribed only by $\ell.n \notin S_1(In, r)$ is irrelevant for continuation-rationalizability. Instead, strategy $Out.m$ is not sequentially optimal under any $\mu^2 \in \Delta_{\text{fb}}^H(\hat{S}_1^1)$. The reason is that Out is optimal only if $\mu^2(\cdot|\varnothing)$ assigns positive probability to r, hence $\mu^2(\cdot|(In, r))$ must derived from $\mu^2(\cdot|\varnothing)$ by conditioning, but then $\mu^2(r.n|(In, r)) = \mu^2(r.n|\varnothing) = 0$ (because $r.n \notin \hat{S}_1^1$), so e does better than m at h. Thus,

$$\hat{S}_1^2 = \hat{S}_1^1, \quad \hat{S}_2^2 = S_2 \backslash \{Out.m\}.$$

The elimination of $Out.m$ does not change the set of viable beliefs of player 1 at history (In), because $In.m \in \hat{S}_2^2$. So, in conclusion,

$$\hat{S}_1^\infty = \{r.c, r.s, \ell.n, \ell.c\}, \quad \hat{S}_2^\infty = S_2 \backslash \{Out.m\}.$$

We defined initial and strong rationalizability by means of weak sequential optimality, therefore the first step of both procedures eliminates only $r.n$ — strategy $\ell.s$ is behaviorally equivalent to $\ell.n$ and $\ell.c$ and therefore it is weakly sequentially optimal under the same CPSs. By the same token, the second step of both procedures does not eliminate $Out.m$. Initial rationalizability does not eliminate any other strategy, while strong rationalizability eliminates $In.m$. This is because, even if player 2 initially assigns probability 1 to ℓ, after being surprised by r player 2 must still assign probability 0 to strategy $r.n$, by strong belief in the rationality of player 1.

In conclusion, initial rationalizability yields a superset of the continuation-rationalizable strategy profiles, while strong rationalizability yields a subset of the continuation-rationalizable *paths*, as the paths in which the last move of player 2 is m are consistent with continuation-rationalizability, but not with strong rationalizability. ▲

The conclusions of the example on the relationship between continuation-rationalizability and the other notions of rationalizability can be generalized. To better compare continuation-rationalizability with initial rationalizability, it is useful to rewrite also the former by means of the iteration of a justification operator. Consider the collection $\mathcal{C}$ of all Cartesian subsets of S, and for all $C = \times_{i\in I} C_i \in \mathcal{C}$, for each $i \in I$, define

$$\rho_{i,\text{fb}}(C_{-i}) = \left\{s_i \in S_i : \exists \mu^i \in \Delta^H_{\text{fb}}(C_{-i}), s_i \in \hat{r}_i(\mu^i)\right\} = \hat{r}_i(\Delta^H_{\text{fb}}(C_{-i})),$$

$$\rho_{\text{fb}}(C) = \bigtimes_{i\in I} \rho_{i,\text{fb}}(C_{-i}).$$

Thus, $\rho_{\text{fb}}(C)$ is the set of strategy profiles that can be justified by a CPS with future belief in C_{-i} for each player $i \in I$. Like r_i and ρ_i as defined in Section 11.2, also $\hat{r}_i$ and ρ_{fb} are extensions to multistage games of the sets of best replies and the justification operator defined for simultaneous-move games. We can now rewrite continuation-rationalizability as the standard iteration of the self-map $\rho_{\text{fb}} : \mathcal{C} \to \mathcal{C}$. Let $\rho^0_{\text{fb}}(S) = S$ and $\rho^k_{\text{fb}}(S) = \rho_{\text{fb}}(\rho^{k-1}_{\text{fb}}(S))$ for each integer $k > 0$. Since $\Delta^H_{\text{fb}}(C'_{-i}) \subseteq \Delta^H_{\text{fb}}(C_{-i})$ whenever $C'_{-i} \subseteq C_{-i}$, ρ_{fb} is monotone. Therefore, since $\rho^0_{\text{fb}}(S) = S$, $\left(\rho^k_{\text{fb}}(S)\right)^\infty_{k=0}$ is a weakly decreasing sequence, and we can define:

$$\rho^\infty_{\text{fb}}(S) = \bigcap_{k\geq 1} \rho_{\text{fb}}(S).$$

Remark 43. *For all steps $k \in \mathbb{N}$, $\hat{S}^k = \rho^k_{\text{fb}}(S)$. Therefore, a strategy profile $s \in S$ is continuation-rationalizable if and only if $s \in \rho^{\infty}_{\text{fb}}(S)$.*

Recall that ρ is the justification operator used to define initial rationalizability, the simplest extension of the rationalizability idea from simultaneous-move games to multistage games: for all $C \in \mathcal{C}$,

$$\rho(C) = \bigtimes_{i \in I} r_i \left(\Delta^H_{\varnothing}(C_{-i}) \right)$$

Note that, for all $C \in \mathcal{C}$, $\rho_{\text{fb}}(C) \subseteq \rho(C)$, because $\Delta^H_{\text{fb}}(C_{-i}) \subseteq \Delta^H_{\varnothing}(C_{-i})$ (future belief in C_{-i} entails initial belief in C_{-i}) and $\hat{r}_i \left(\mu^i \right) \subseteq r_i \left(\mu^i \right)$ for every $i \in I$ and every CPS μ^i. By monotonicity and the previous observation, we also have $\rho_{\text{fb}}(C) \subseteq \rho_{\text{fb}}(C') \subseteq \rho(C')$ whenever $C \subseteq C'$. Then, with a simple induction argument, we can conclude that continuation-rationalizability refines initial rationalizability.

Remark 44. *For all steps $k \in \mathbb{N}$, $\rho^k_{\text{fb}}(S) \subseteq \rho^k(S)$. Therefore, every continuation-rationalizable strategy profile is initially rationalizable.*

Furthermore, as pointed out for initial rationalizability, leveraging on the monotonicity of operator ρ_{fb} one can prove that extensions of Theorems 2 and 3 of Chapter 4 hold for continuation-rationalizability in multistage games.

The relationship between continuation-rationalizability and strong rationalizability is more complex. Strong rationalizability does not yield a subset of the set of continuation-rationalizable strategy profiles. In Example 57, this is simply due to the use of weak sequential optimality in place of sequential optimality. Indeed, the only strategies that are strongly rationalizable but not continuation-rationalizable are behaviorally equivalent to some continuation-rationalizable strategy. However, this is not always the case: *strong rationalizability can yield strategies that are not behaviorally equivalent to any continuation-rationalizable strategy.* One reason is that strong belief in C_{-i} has no bite at histories that are inconsistent with C_{-i}, whereas future belief in C_{-i} does. Furthermore, after some steps of reasoning, continuation and strong rationalizability may require, respectively, future and strong belief in *different* sets C_{-i} and C'_{-i}, which are not even behaviorally equivalent. Then, future belief in C_{-i} may have different implications from strong belief in C'_{-i} regarding co-players' behavior at histories that *are* consistent with C_{-i}. (See Example 60 for an instance of this.) May this difference induce player i, under future belief in C_{-i}, to always find it profitable to leave a strongly rationalizable path? The answer is no: *continuation-rationalizability never eliminates all the strategies of player i that are consistent with some strongly rationalizable path.* This answer is based on rather sophisticated arguments, so we skip the details. The consequence is that the set of *paths* induced by the strongly rationalizable strategy profiles is a subset of those induced by continuation-rationalizability.

Theorem 46. *Every strongly rationalizable path is continuation-rationalizable:* $\zeta(S^\infty) \subseteq \zeta(\hat{S}^\infty)$.

Here we only provide a sketch of the proof. Theorem 46 can be proved in two steps. First, one can exploit the monotonicity of ρ_{fb} to compute continuation-rationalizability through the following, slower elimination order $((\widetilde{S}_i^n)_{i\in I})_{n>0}$. Start by iteratively eliminating, for each player i, the strategies s_i that, at step n, are not *weakly* sequentially optimal given some CPS μ^i such that i has future belief in $\widetilde{S}_{-i}^{n-1}$ *at the histories that are consistent with* $\widetilde{S}_{-i}^{n-1}$, that is,

$$\begin{aligned}\mu^i &\in \left\{\mu^i \in \Delta^H(S_{-i}) : \forall h \in H(\widetilde{S}_{-i}^{n-1}),\ \mu^i(\chi_{-i}^h(\widetilde{S}_{-i}^{n-1})|h) = 1\right\} \\ &\supseteq \Delta_{\text{fb}}^H(\widetilde{S}_{-i}^{n-1}) \cup \Delta_{\text{sb}}^H(\widetilde{S}_{-i}^{n-1}).\end{aligned}$$

Since the game is finite, after some step K no more strategies can eliminated in this way. Next, for each $n \geq K$, let $\widetilde{S}^n = \rho_{\text{fb}}(\widetilde{S}^{n-1})$. Via an argument of independence of the final set on the order of elimination, one can show that $\widetilde{S}^\infty = \hat{S}^\infty$. Since steps $n \geq K$ only refine behavior at histories that are inconsistent either with one's own strategy, or with the strategies of the co-players, it can be checked that $\zeta(\widetilde{S}^K) = \zeta(\widetilde{S}^\infty) = \zeta(\hat{S}^\infty)$. Next, note that until step K every set $\widetilde{S}_i^n$ contains all the strategies that are *weakly* sequentially optimal under a CPS that strongly believes $\widetilde{S}_{-i}^{n-1}$. Hence, $((\widetilde{S}_i^n)_{i\in I})_{n>0}^K$ can be seen as a slow, possibly unfinished, order of elimination under the strong justification operator $\overline{\rho}_{\text{sb}}$. At this point, one can invoke the outcome-equivalence of different elimination orders based on $\overline{\rho}_{\text{sb}}$: see Section 11.3.

11.4.3 The Backward Procedure

In this section, we introduce a "backward procedure" that can be used to compute the predictions of continuation-rationalizability starting from the preterminal histories and moving backward.

For each $\ell = 1, ..., L$ (where L is the game horizon), we consider the nonterminal histories h such that $L(\Gamma(h)) = \ell$, where $L(\Gamma(h))$ denotes the height of the subgame starting at h, i.e., the histories such that the longest continuation (suffix) has length ℓ.[13] For each $h \in H$, let $H^1(h)$ denote the **set of *nonterminal* histories that *immediately* follow** h. In particular, $H^1(h)$ is empty if h is preterminal, i.e., if $(h, a) \in Z$ for every $a \in \mathcal{A}(h)$. In Chapter 10 we let $s_i^{\succeq h} \in S_i^{\succeq h}$ denote the continuation-strategies of player i in the subgame $\Gamma(h)$ with root h. Although such notation is expressive, to avoid cumbersome formulas here we *simplify* it *as follows*: $s_i^h \in S_i^h$, where $S_i^h = \times_{h' \succeq h} \mathcal{A}_i(h')$. With this, for each $s_i^h \in S_i^h$ and $\mu_i^h \in \Delta(S_{-i}^h)$, let $U_i^h(s_i^h, \mu_i^h)$ denote the expected payoff of player i in subgame $\Gamma(h)$, starting from history h and continuing with s_i^h, given the probability measure μ_i^h on the continuation-strategies of the co-players. Furthermore, for $h \prec h' \in H$ we let $s_i^h|h' \in S^{h'}$ denote the continuation induced by s_i^h in subgame $\Gamma(h')$.

Intuitively, the elimination procedure starts finding the rationalizable *actions* of every *last-stage* simultaneous-move subgame $\Gamma(h)$, i.e., for every h with $L(\Gamma(h)) = 1$. For "pre-preterminal" histories h with $L(\Gamma(h)) = 2$, the procedure iteratively eliminates strategies of $\Gamma(h)$ that are not best replies to beliefs taking into account that (1) the last-stage continuations must be rationalizable in the last-stage games and (2) strategies of $\Gamma(h)$ already eliminated in previous steps must have probability 0. The iterated elimination procedure for strategies in subgames with height $L(\Gamma(h)) > 2$ is similar.

Definition 69. *For all h with $L(\Gamma(h)) = 1$ and $i \in I$, let $\hat{S}_i^{h,\infty}$ be the set of rationalizable actions in the simultaneous-move subgame with root h.*

Now fix $\ell = 2, ..., L$ and suppose by way of induction that, for all h' with $L(\Gamma(h')) = \ell - 1$ and $i \in I$, $\hat{S}_i^{h',\infty}$ has been defined. Then, for each h with $L(\Gamma(h)) = \ell$, define the following elimination procedure:

[13] Recall that the height of h, viewed as a node of the game tree, is $L(\Gamma(h)) := \max_{z:h \prec z}[\ell(z) - \ell(h)]$.

Step 0 *For each $i \in I$, let*

$$\hat{S}_i^{h,0} = \left\{ s_i \in S_i^h : \forall h' \in H^1(h), s_i^h | h' \in \hat{S}_i^{h',\infty} \right\}.$$

Step k *For each $i \in I$ and $s_i^h \in S_i^h$, let $s_i^h \in \hat{S}_i^{h,k}$ if there exists $\mu_i^h \in \Delta(S_{-i}^h)$ such that:*

(BP1): $\mu_i^h \left(\hat{S}_{-i}^{h,k-1} \right) = 1$.

(BP2): for all $\widetilde{s}_i^h \in S_i^h$, $U_i \left(s_i^h, \mu_i^h \right) \geq U_i \left(\widetilde{s}_i^h, \mu_i^h \right)$.

Finally, for each $i \in I$, let $\hat{S}_i^{h,\infty} = \bigcap_{k>0} \hat{S}_i^{h,k}$.

For each history h, the backward procedure performs the iterated elimination of never best replies in an appropriately trimmed strategic form of the subgame with root h. Such trimming consists of eliminating every continuation-strategy s_i^h whose projection in some subgame with root $h' \in H^1(h)$ did not survive the iteration of the backward procedure for that subgame i.e., $s_i^h | h' \notin \hat{S}_i^{h',\infty}$. Using strategic-form best replies, the backward procedure eliminates continuation-strategies only based on their behavioral-equivalence class. For this reason, the backward procedure is not able to eliminate every strategy that is not continuation-rationalizable. We illustrate this fact in the game of Example 57.

Example 58. Consider the game of Figure 11.2. Start from the preterminal history (In, r). In the strategic form of the subgame, every action of both players is justifiable, therefore the first iteration of the backward procedure does not eliminate any action. Hence, we move to the subgame with root (In) and we consider its entire strategic form. Strategy $r.n$ is dominated by any strategy that prescribes ℓ. Every other strategy of player 1 is undominated, and all the continuation-strategies of player 2, trivially, best reply to a measure that assigns probability 1 to $\{\ell.n, \ell.c, \ell.s\}$.

Thus, we move to the initial history and initialize the elimination procedure as follows:

$$\hat{S}_1^{\varnothing,0} = S_1 \backslash \{r.n\}, \quad \hat{S}_2^{\varnothing,0} = S_2.$$

The strategies of player 2 that prescribe Out (resp., In) are strategic-form best replies to a measure that assigns probability 0 (resp., 1) to $\{\ell.n, \ell.c, \ell.s\}$. The strategies of player 1, trivially, are strategic form best replies to a measure that assigns probability 1 to $\{Out.w, Out.m.Out.e\}$. So, in conclusion,

$$\hat{S}_1^{\varnothing,\infty} = S_1 \backslash \{r.n\}, \quad \hat{S}_2^{\varnothing,\infty} = S_2.$$

Compared to continuation-rationalizability, the backward procedure does not eliminate $\ell.s$ and $Out.m$. This is because s and m are rationalizable in the simultanous-move subgame with root (In, r), and they do not affect the payoffs of strategies $\ell.s$ and $Out.m$ in the strategic forms of the larger subgames, as they are prescribed at a history that is inconsistent with the strategies themselves. ▲

In the example, although the backward procedure yields a superset of the continuation-rationalizable strategies, it yields the same *reduced* strategies. This is a general result: the backward procedure can be used to compute the predictions of continuation-rationalizability. For any subset $\bar{S}_i \subseteq S_i$, let $[\bar{S}_i] \subseteq S_i$ denote the **set of strategies of** i **that are behaviorally equivalent**[14] **to at least one strategy in** $\bar{S}_i$. With this, we have:

Theorem 47. *For every* $i \in I$, $\hat{S}_i^{\infty} \subseteq \hat{S}_i^{\varnothing,\infty} \subseteq \left[\hat{S}_i^{\infty}\right]$.

The key intution behind Theorem 47 is that, for each player i, at every history h, every continuation-strategy that is sequentially optimal given a CPS in the subgame $\Gamma(h)$ is also induced by some strategy that is sequentially optimal given a CPS for the entire game. Given this, concentrating beliefs on the co-players' continuation-strategies that are induced by continuation-rationalizability is not different from focusing on the continuation-strategies that cannot be iteratively eliminated by simply "looking ahead" in $\Gamma(h)$ (as in the backward procedure).

11.4.4 Backward Induction in Games with Perfect Information

The backward procedure is a generalization to all multistage games with observed actions of the well-known backward-induction algorithm for games with perfect information.

[14] Or realization-equivalent, see Lemma 26.

Recall that a game Γ has **perfect information (PI)** if, for each $h \in H$, only one player—denoted $\iota(h)$—is active at h:

$$\forall h \in H, \exists! i \in I, |\mathcal{A}_i(h)| > 1.$$

Thus, for each $h \in H$,

$$|\mathcal{A}_i(h)| > 1 \Leftrightarrow i = \iota(h),$$

and we can write *histories* as *sequences* of *actions of active players*.

Now fix two terminal histories $z', z'' \in Z$; if $z' \neq z''$, there is a player who is "decisive," or "pivotal," for reaching z' rather than z'', that is, the player who is active at the longest prefix (last common predecessor) of z' and z''. We let $\pi(z', z'')$ denote this player; formally,

$$\pi\left(z', z''\right) = \iota\left(\arg \max_{h \in H: h \prec z', h \prec z''} \ell(h)\right).$$

With this, we introduce the condition that a player who is pivotal between two paths is never indifferent between them.

Definition 70. *A PI game Γ has **no relevant ties** if the pivotal player is never indifferent between distinct continuation-paths; formally,*

$$\forall z', z'' \in Z,\ z' \neq z'' \Rightarrow u_{\pi(z',z'')}\left(z'\right) \neq u_{\pi(z',z'')}\left(z''\right).$$

In games with perfect information and no relevant ties, the backward-induction algorithm is well-defined and yields a unique strategy s_i^* for each player i. The algorithm can be seen as a kind of "inter-personal" version of the folding-back optimality algorithm of Chapter 10. So, for each $h \in H$, we are going to compute the action $s_{\iota(h)}^*(h)$ and the value $V_j^*(h)$ for each $j \in I$ of reaching h by induction on $L(\Gamma(h))$, the height of the subgame starting at h. Of course, the value for a player of reaching a terminal history is just the payoff of that history; therefore, $V_j^*(z) = u_j(z)$ for all $j \in I$ and $z \in Z$. With this, **backward induction** (**BI**) is the following recursive procedure:

- If $L(\Gamma(h)) = 1$, then $\left(h, a_{\iota(h)}\right) \in Z$ for every $a_{\iota(h)} \in \mathcal{A}_{\iota(h)}(h)$; thus,
  - $i = \iota(h) \Rightarrow s_i^*(h) = \arg\max_{a_i \in \mathcal{A}_i(h)} V_i^*(h, a_i) = \arg\max_{a_i \in \mathcal{A}_i(h)} u_i(h, a_i)$,

– $\forall j \in I$, $V_j^*(h) = V_j^*\left(h, s_{\iota(h)}^*(h)\right)$.

- Let $s_{\iota(h)}^*(h)$ and $V_j^*(h)$ be defined for every $h \in \bar{H}$ such that $L(\Gamma(h)) \leq k$ and every $j \in I$; if $L(\Gamma(h)) = k+1$, then $L\left(\Gamma\left(h, a_{\iota(h)}\right)\right) \leq k$ for every $a_{\iota(h)} \in \mathcal{A}_{\iota(h)}(h)$; thus,

  – $i = \iota(h) \Rightarrow s_i^*(h) = \arg\max_{a_i \in \mathcal{A}_i(h)} V_i^*(h, a_i)$,

  – $\forall j \in I$, $V_j^*(h) = V_j^*\left(h, s_{\iota(h)}^*(h)\right)$.

Note that $s_{\iota(h)}^*(h)$ is well defined because the given perfect information game is assumed to be finite and with no relevant ties, which implies that—at each step of the backward algorithm—the active player has *only one maximizing action*. In words, we start from the *last* stage of the game: h is such that all feasible actions terminate the game, that is $L(\Gamma(h)) = 1$.[15] According to the algorithm, the active player selects the payoff maximizing action. This determines a profile of payoffs for all players, denoted $(V_i^*(h))_{i\in I}$. Then we go backward to the second-to-last stage, or—more precisely—we consider histories of height 2: $L(\Gamma(h)) = 2$. The value $V_i^*\left(h, a_{\iota(h)}\right)$ has already been computed for all histories $(h, a_{\iota(h)}) \in \bar{H}$, because such histories correspond to the last stage of the game, or are terminal. According to the algorithm, the active player $\iota(h)$ chooses the feasible action $a_{\iota(h)} = s_{\iota(h)}^*(h)$ that maximizes $V_{\iota(h)}^*\left(h, a_{\iota(h)}\right)$, the reason is that the active player expects that every following player (possibly himself) would maximize his own payoff in the last stage. The algorithm continues to go backward in this way until it reaches the first stage ($h = \varnothing$).

Now we illustrate the BI algorithm by example.

Example 59. In the ToL4 of Figure 9.2, let T_k (L_k) denote the action of taking (leaving) k euros. It can be verified that there are no relevant ties. We obtain the strategy pair $s^* = (\mathrm{T}_1.\mathrm{T}_3, \mathrm{T}_2.\mathrm{T}_4)$ by backward induction as follows.

► $L(\Gamma(h)) = 1$: The only history h with $L(\Gamma(h)) = 1$ is $h = (\mathrm{L}_1, \mathrm{L}_2, \mathrm{L}_3)$ and $\iota(\mathrm{L}_1, \mathrm{L}_2, \mathrm{L}_3) = 2$. Player 2 maximizes by taking 4 euros; thus, $s_2^*(\mathrm{L}_1, \mathrm{L}_2, \mathrm{L}_3) = \mathrm{T}_4$ and $V_1^*(\mathrm{L}_1, \mathrm{L}_2, \mathrm{L}_3) = 0$.

[15] Recall that which stage is the last, second-to-last and so on may be endogenous. For example, we may have a game that lasts for two stages if the first mover chooses Left and three stages if the first mover chooses Right.

▶ $L(\Gamma(h)) = 2$: The only history h with $L(\Gamma(h)) = 2$ is $h = (\mathrm{L}_1, \mathrm{L}_2)$ and $\iota(\mathrm{L}_1, \mathrm{L}_2) = 1$. Player 1 maximizes by taking 3 euros, because $V_1^*(\mathrm{L}_1, \mathrm{L}_2, \mathrm{T}_3) = 3 > 0 = V_1^*(\mathrm{L}_1, \mathrm{L}_2, \mathrm{L}_3)$; thus, $s_1^*(\mathrm{L}_1, \mathrm{L}_2) = \mathrm{T}_3$ and $V_2^*(\mathrm{L}_1, \mathrm{L}_2) = 0$.
▶ $L(\Gamma(h)) = 3$: The only history h with $L(\Gamma(h)) = 3$ is $h = (\mathrm{L}_1)$ and $\iota(\mathrm{L}_1) = 2$. Player 2 maximizes by taking 2 euros, because $V_2^*(\mathrm{L}_1, \mathrm{T}_2) = 2 > 0 = V_2^*(\mathrm{L}_1, \mathrm{L}_2)$; thus, $s_2^*(\mathrm{L}_1) = \mathrm{T}_2$ and $V_1^*(\mathrm{L}_1) = 0$.
▶ $L(\Gamma(h)) = 4$: The only history h with $L(\Gamma(h)) = 4$ is the root $\varnothing$ and $\iota(\varnothing) = 1$. Player 1 maximizes by taking 1 euros, because $V_1^*(\mathrm{T}_1) = 1 > 0 = V_2^*(\mathrm{L}_1)$; thus, $s_1^*(\varnothing) = \mathrm{T}_1$. ▲

It is easy to observe that, in PI games with no relevant ties, the backward procedure boils down to the BI algorithm. For every history h, assuming by induction that the previous iterations of the two procedures have pinned down the same continuation-strategies from each $h' \in H^1(h)$ onwards, the backward procedure first reduces the strategic form of the subgame to the unique continuation-strategy for each non-active player, then immediately eliminates for the active player all the remaining continuation-strategies except the one that prescribes the BI move at h.

Remark 45. *In PI games with no relevant ties, the backward procedure and the BI algorithm coincide.*

In light of this, the BI algorithm inherits from Theorem 47 the following property: there is only one continuation-rationalizable profile and it coincides with the BI strategy profile. In turn, by Theorem 46, the induced path coincides with the only strongly rationalizable path.

Theorem 48. *In every finite game with perfect information and no relevant ties, all the strongly rationalizable strategy profiles yield the same terminal history as the unique backward induction strategy profile:* $\zeta(S^\infty) = \{\zeta(s^*)\}$.

Theorem 48 is a striking result because it holds although, even in PI game with no relevant ties, the strongly rationalizable profiles may not include the unique BI profile. That is, the BI strategies may be inconsistent with rationality and common strong belief in rationality. We conclude the section with a classic example of this fact.[16]

[16] See Reny [71].

Example 60. Consider the PI game between Ann (a) and Bob (b) depicted in Figure 11.3. It can be verified that there are no relevant ties.

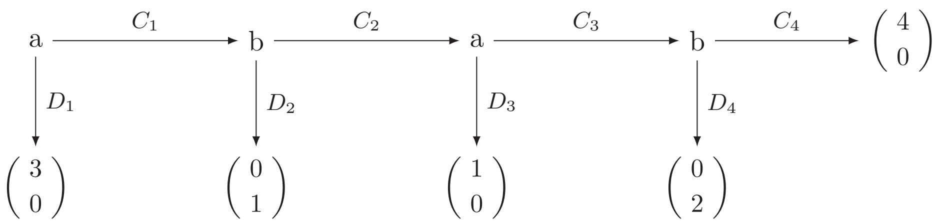


Figure 11.3: A PI game with no relevant ties.

The *BI profile* is $s^* = (D_1.D_3, D_2.D_4)$. Strategies $C_1.D_3$ and $C_2.C_4$ are conditionally dominated. With this,

$$S^1 = \{D_1.D_3, D_1.C_3, C_1.C_3\} \times \{C_2.D_4, D_2.D_4, D_2.C_4\}$$

(recall that strong rationalizability does not distinguish between behaviorally equivalent strategies). The *strongly rationalizable set* is $S^\infty = S^2 = \{D_1.D_3, D_1.C_3\} \times \{C_2.D_4\}$. Thus, $s^* \notin S^\infty$, because $s_2^* \neq C_2.D_4$. Indeed, strong belief in rationality requires that, upon observing C_1, Bob would believe that Ann is implementing the only justifiable (i.e., conditionally undominated) strategy that selects C_1 at the root, that is, $C_1.C_3$. The weakly sequentially optimal strategy given conjecture $C_1.C_3$ is $C_2.D_4$. ▲

11.5 Appendix

Proof of Lemma 35. Fix a strategy $\bar{s}_i$ and let $W \subset \mathbb{R}^Z$ denote the set of player i's payoff vectors $(y_z)_{z\in Z}$ such that $\bar{s}_i$ is *weakly dominated* in the strategic form with strategy profiles C, *but not conditionally dominated* within C. We show that the closure of W, denoted by $\mathrm{cl}\,(W)$, has zero measure. Since the number of strategies is finite, this yields the result.

Let $\bar{Z}$ be the set of all $z \in Z$ such that $\bar{s}_i \in S_i(z)$. Fix any $\bar{z} \in \bar{Z}$. For each payoff vector $(y_z)_{z\in Z}$, there is a unique vector $(x_z)_{z\in Z}$ such that

$$\begin{cases} y_z = x_z, & \text{if } z \notin \bar{Z}, \\ y_z = x_z + x_{\bar{z}}, & \text{if } z \in \bar{Z}; \end{cases} \tag{11.5.1}$$

that is, vector $(x_z)_{z\in Z}$ such that $x_{\bar{z}} = y_{\bar{z}}/2$, $x_z = y_z - x_{\bar{z}}$ for all $z \in \bar{Z}$, and $x_z = y_z$ for all $z \notin \bar{Z}$. Let $f : \mathbb{R}^Z \to \mathbb{R}^Z$ denote the transformation defined by (11.5.1).

We are going to study for which vectors $(x_z)_{z\in Z}$ the corresponding payoff vector $(y_z)_{z\in Z} = f((x_z)_{z\in Z})$ belongs to $\mathrm{cl}\,(W)$. In particular, we now show that for each $(x_z)_{z\neq\bar{z}} \in \mathbb{R}^{Z\setminus\{\bar{z}\}}$, there is at most one value of $x_{\bar{z}}$ such that $(y_z)_{z\in Z} \in \mathrm{cl}\,(W)$. Fix two values $x < x'$ of $x_{\bar{z}}$, with corresponding payoff vectors $(y_z)_{z\in Z}$ and $(y'_z)_{z\in Z}$. Thus[17]

$$(y'_z)_{z\in\bar{Z}} \;>\; (y_z)_{z\in\bar{Z}}, \tag{11.5.2}$$
$$(y'_z)_{z\notin\bar{Z}} \;=\; (x_z)_{z\notin\bar{Z}}, \tag{11.5.3}$$
$$(y_z)_{z\notin\bar{Z}} \;=\; (x_z)_{z\notin\bar{Z}}. \tag{11.5.4}$$

Suppose that $(y'_z)_{z\in Z} \in \mathrm{cl}\,(W)$. We are going to show that $(y_z)_{z\in Z} \notin \mathrm{cl}\,(W)$, because there are payoff vectors arbitrarily close to $(y_z)_{z\in Z}$ such that $\bar{s}_i$ is conditionally dominated. Since $(y'_z)_{z\in Z} \in \mathrm{cl}\,(W)$, there exists some $(\widetilde{y}'_z)_{z\in Z} \in W$ that preserves (11.5.2) and (11.5.3):[18]

$$(\widetilde{y}'_z)_{z\in\bar{Z}} \;>\; (y_z)_{z\in\bar{Z}}, \tag{11.5.5}$$
$$(\widetilde{y}'_z)_{z\notin\bar{Z}} \;=\; (x_z)_{z\notin\bar{Z}}. \tag{11.5.6}$$

Since $(\widetilde{y}'_z)_{z\in Z} \in W$, there exists $\bar{\sigma}_i \in \Delta(S_i\setminus\{\bar{s}_i\})$ with $\bar{\sigma}_i(C_i) = 1$ such that

$$\forall s_{-i} \in C_{-i}, \quad \sum_{s_i\in\mathrm{supp}\bar{\sigma}_i} \bar{\sigma}_i(s_i)\widetilde{y}'_{\zeta(s_i,s_{-i})} \geq \widetilde{y}'_{\zeta(\bar{s}_i,s_{-i})}, \tag{11.5.7}$$
$$\exists s_{-i} \in C_{-i}, \quad \sum_{s_i\in\mathrm{supp}\bar{\sigma}_i} \bar{\sigma}_i(s_i)\widetilde{y}'_{\zeta(s_i,s_{-i})} > \widetilde{y}'_{\zeta(\bar{s}_i,s_{-i})}, \tag{11.5.8}$$

i.e., $\bar{s}_i$ is weakly dominated in C. Condition (11.5.8) requires that there exist histories $h \in H(\bar{s}_i)$ such that (i) $C_{-i} \cap S_{-i}(h) \neq \emptyset$ and (i) $\hat{s}_i(h) \neq \bar{s}_i(h)$ for some $\hat{s}_i \in \mathrm{supp}\bar{\sigma}_i$. Among these histories, fix one history $\bar{h}$ such that no $h \prec \bar{h}$ has property (i). This means that, for every $h \prec \bar{h}$ and $\hat{s}_i \in \mathrm{supp}\bar{\sigma}_i$, $\hat{s}_i(h) = \bar{s}_i(\bar{h})$, that is, every $\hat{s}_i$ prescribes the same action.

[17] Given any two vectors $b = (b_1, ..., b_n), b' = (b'_1, ..., b'_n)$ of real numbers of the same length, we write $b' > b$ when $b'_j > b_j$ for each $j = 1, ..., n$.

[18] Equation (11.5.3) can be preserved because it pertains to terminal histories that are inconsistent with $\bar{s}_i$.

Hence, $\bar{\sigma}_i(S_i(\bar{h})) = 1$. Now fix $s_{-i} \in C_{-i} \cap S_{-i}(\bar{h})$. For each $s_i \in \text{supp}\bar{\sigma}_i$, we have

$$\zeta(s_i, s_{-i}) = \zeta(\bar{s}_i, s_{-i}) \;\Leftrightarrow\; \zeta(s_i, s_{-i}) \in \bar{Z}. \tag{11.5.9}$$

Direction $\Rightarrow$ is obvious; to see $\Leftarrow$, reason by contraposition: $\zeta(s_i, s_{-i}) \neq \zeta(\bar{s}_i, s_{-i})$ requires $\hat{s}_i(h) \neq \bar{s}_i(h)$ for some $h \prec \zeta(s_i, s_{-i})$, and thus $\zeta(s_i, s_{-i})$ cannot be reached with $\bar{s}_i$ under any $s'_{-i} \in S_{-i}$. At this point, we can conclude that

$$\begin{aligned}
&\sum_{s_i \in \text{supp}\bar{\sigma}_i} \bar{\sigma}_i(s_i) \left(y_{\zeta(s_i,s_{-i})} - y_{\zeta(\bar{s}_i,s_{-i})}\right) \overset{(11.5.9)}{=} \\
&\sum_{s_i \in \text{supp}\bar{\sigma}_i : \zeta(s_i,s_{-i}) \notin \bar{Z}} \bar{\sigma}_i(s_i) \left(y_{\zeta(s_i,s_{-i})} - y_{\zeta(\bar{s}_i,s_{-i})}\right) \overset{(11.5.4)}{=} \\
&\sum_{s_i \in \text{supp}\bar{\sigma}_i : \zeta(s_i,s_{-i}) \notin \bar{Z}} \bar{\sigma}_i(s_i) \left(x_{\zeta(s_i,s_{-i})} - y_{\zeta(\bar{s}_i,s_{-i})}\right) \overset{(11.5.5)}{>} \\
&\sum_{s_i \in \text{supp}\bar{\sigma}_i : \zeta(s_i,s_{-i}) \notin \bar{Z}} \bar{\sigma}_i(s_i) \left(x_{\zeta(s_i,s_{-i})} - \widetilde{y}'_{\zeta(\bar{s}_i,s_{-i})}\right) \overset{(11.5.6)}{=} \\
&\sum_{s_i \in \text{supp}\bar{\sigma}_i : \zeta(s_i,s_{-i}) \notin \bar{Z}} \bar{\sigma}_i(s_i) \left(\widetilde{y}'_{\zeta(s_i,s_{-i})} - \widetilde{y}'_{\zeta(\bar{s}_i,s_{-i})}\right) \overset{(11.5.9)}{=} \\
&\sum_{s_i \in \text{supp}\bar{\sigma}_i} \bar{\sigma}_i(s_i) \left(\widetilde{y}'_{\zeta(s_i,s_{-i})} - \widetilde{y}'_{\zeta(\bar{s}_i,s_{-i})}\right) \overset{(11.5.7)}{\geq} 0.
\end{aligned}$$

Since this is true for every $s_{-i} \in C_{-i} \cap S_{-i}(\bar{h})$, $\bar{s}_i$ is conditionally dominated by $\bar{\sigma}_i$ in C when $x_{\bar{z}} = x$. Thus, $(y_z)_{z\in Z} \notin W$. Moreover, conditional dominance is preserved in a neighborhood of $(y_z)_{z\in Z}$, because it is defined by strict inequalities. This shows that $(y_z)_{z\in Z} \notin \text{cl}\,(W)$.

Now, the Lebesgue measure of $\text{cl}\,(W)$ is defined as

$$\lambda(\text{cl}\,(W)) = \int \mathbf{1}_{\text{cl}(W)} \left((y_z)_{z\in Z}\right)_{z\in Z} \mathrm{d}y_z.$$

We evaluate this integral after a change of variables, from $(y_z)_{z\in Z}$ to $(x_z)_{z\in Z}$. The change of variables yields

$$\int \mathbf{1}_{\text{cl}(W)} \left((y_z)_{z\in Z}\right)_{z\in Z} \mathrm{d}y_z = \int \mathbf{1}_{\text{cl}(W)} \left((x_z)_{z\in Z\setminus\bar{Z}}, (x_z + x_{\bar{z}})_{z\in\bar{Z}}\right) k_{z\in Z} \mathrm{d}x_z,$$

where k denotes the determinant of the Jacobian matrix of f. By linearity of f, k is a constant. Using the fact that $\mathbf{1}_{\mathrm{cl}(W)}\left((y_z)_{z\in Z}\right)$ is not zero for at most one value of $x_{\bar{z}}$, we obtain

$$\begin{aligned}\lambda(\mathrm{cl}\,(W)) &= \int\left(\int \mathbf{1}_{\mathrm{cl}(W)}\left((x_z)_{z\in Z\setminus\bar{Z}}\,,(x_z+x_{\bar{z}})_{z\in\bar{Z}}\right)k\mathrm{d}x_{\bar{z}}\right)_{z\neq\bar{z}}\mathrm{d}x_z\\ &= \int 0_{z\neq\bar{z}}\mathrm{d}x_z = 0.\end{aligned}$$

■

Lemma 31 is a corollary of Lemma 35: just set $C = S$.

Proof of Theorem 43. The proof is by induction, but we need a stronger induction hypothesis than the claim we aim to prove. Recall that $H(C) = \{h \in H : S\,(h) \cap C \neq \emptyset\}$.

Inductive hypothesis (n): $H(S^n) = H(S^n_{\mathrm{ir}})$; furthermore, for each $i \in I$ and $s_i \in S^n_i$ (resp., $\widehat{s}_i \in S^n_{i,\mathrm{ir}}$), there exists $\widehat{s}_i \in S^n_{i,\mathrm{ir}}$ (resp., $s_i \in S^n_i$) such that $s_i(h) = \widehat{s}_i(h)$ for each $h \in H(S^n) = H(S^n_{\mathrm{ir}})$.

Basis step ($n = 0$): $S^0_{i,\mathrm{ir}} = S^0_i = S_i$ for each $i \in I$.

Inductive step $(n+1)$**:** Let $H^* := H(S^n) = H(S^n_{\mathrm{ir}})$ and $Z^* := \zeta(S^n) = \zeta(S^n_{\mathrm{ir}})$, where the equalities follow from the inductive hypothesis (n). We prove that for all $i \in I$ and $s_i \in S^{n+1}_i$, there exists $\widehat{s}_i \in S^{n+1}_{i,\mathrm{ir}}$ such that $s_i(h) = \widehat{s}_i(h)$ for each $h \in H^*$; the proof of the converse is identical. Fix any μ^i that strongly believes S^n_{-i} and satisfies $s_i \in r_i(\mu^i)$. By the inductive hypothesis, we can construct a map $\eta : S^n_{-i} \to S^n_{-i,\mathrm{ir}}$ that associates each $s_{-i} \in S^n_{-i}$ with some $\widehat{s}_{-i} \in S^n_{-i,\mathrm{ir}}$ such that $s_{-i}(h) = \widehat{s}_{-i}(h)$ for every $h \in H^*$. For each $h \in H^*$, since $\mu^i(S^n_{-i}|h) = 1$, we obtain an associated belief $\hat{\mu}^i(\cdot|h) \in \Delta(S^n_{-i,\mathrm{ir}})$ defined as $\hat{\mu}^i(s_{-i}|h) = \mu^i(\eta^{-1}(s_{-i})|h)$ for all $s_{-i} \in S^n_{-i,\mathrm{ir}}$. It is easy to see that, by construction, $\hat{\mu}^i(S_{-i}(h)|h) = 1$. Moreover, for each $h' \in H^*$ with $S_{-i}(h') \subseteq S_{-i}(h)$, and for each $E_{-i} \subseteq S_{-i}(h')$,

$$\begin{aligned}\hat{\mu}^i(E_{-i}|S_{-i}(h)) &= \mu^i(\eta^{-1}(E_{-i})|S_{-i}(h))\\ &= \mu^i(\eta^{-1}(E_{-i})|S_{-i}(h'))\cdot\mu^i(S_{-i}(h')|S_{-i}(h))\\ &= \hat{\mu}^i(E_{-i}|S_{-i}(h'))\cdot\hat{\mu}^i(S_{-i}(h')|S_{-i}(h)),\end{aligned}$$

where the second equality follows from the fact that $\eta^{-1}(E_{-i}) \subseteq S_{-i}(h')$, while the third equality holds because, for each $s_{-i} \in S^n_{-i,\mathrm{ir}}$, we have

$s_{-i} \in S_{-i}(h')$ if and only if $\eta^{-1}(s_{-i}) \in S_{-i}(h')$. For each $h' \notin H^*$ such that $\hat{\mu}^i(S_{-i}(h')|h) > 0$ for some $h \in H^*$, derive $\hat{\mu}^i(\cdot|h')$ by conditioning. For every other $h' \notin H^*$, let $\hat{\mu}^i(\cdot|h') = \widetilde{\mu}^i(\cdot|h')$ for any CPS $\widetilde{\mu}^i$ that satisfies condition 1 in Definition 66. Thus, $\hat{\mu}^i$ satisfies the chain rule and condition 1 in Definition 66.

Now we are going to construct $\widehat{s}_i \in r_i(\hat{\mu}^i) \subseteq S^n_{i,\text{ir}}$ such that $\widehat{s}_i(h) = s_i(h)$ for all $h \in H^*$. By the order-extension principle (e.g., Davey and Priestley [37], p. 32) there is a linear (i.e., complete) extension $\prec^\ell$ of the strict partial order $\prec$ on the set $H^* \cap H_i(s_i)$; that is, for all $h, h' \in H^* \cap H_i(s_i)$, $h \prec h'$ implies $h \prec^\ell h'$.

We start by fixing any $\widehat{s}_i \in r_i(\hat{\mu}^i) \subseteq S^n_{i,\text{ir}}$, and we are going to modify it as desired with an inductive procedure based on $\prec^\ell$. Thus, fix $h \in H^* \cap H_i(s_i)$ and suppose by way of induction that there exists $\widehat{s}_i \in r_i(\hat{\mu}^i) \subseteq S^n_{i,\text{ir}}$ such that $\widehat{s}_i(h') = s_i(h')$ for every $h' \prec^\ell h$; this is vacuously true for $h = \varnothing$. Since $s_i \in S_i(h)$ and $\widehat{s}_i$ coincides with s_i on every predecessor of h, it is the case that $\widehat{s}_i \in S_i(h)$ as well. Since $s_i \in S^n_i$ (resp., $\widehat{s}_i \in S^n_{i,\text{ir}}$) and $\mu^i(S^n_{-i}|h) = 1$ (resp., $\hat{\mu}^i(S^n_{-i,\text{ir}}|h) = 1$), we have $\zeta(s_i, s_{-i}) \in Z^*$ (resp., $\zeta(\widehat{s}_i, s_{-i}) \in Z^*$) for each s_{-i} with $\mu^i(s_{-i}|h) > 0$ (resp., $\hat{\mu}^i(s_{-i}|h) > 0$). By construction, we have $\mu^i(S_{-i}(z)|h) = \hat{\mu}^i(S_{-i}(z)|h)$ for each $z \in Z^*$. Hence, the pairs $(s_i, \mu^i(\cdot|h))$ and $(s_i, \hat{\mu}^i(\cdot|h))$ (resp., $(\widehat{s}_i, \mu^i(\cdot|h))$ and $(\widehat{s}_i, \hat{\mu}^i(\cdot|h))$) induce the same probability mass function over terminal histories. With this,

$$V_i^{s_i,\hat{\mu}^i}(h) = V_i^{s_i,\mu^i}(h) \geq V_i^{\widehat{s}_i,\mu^i}(h) = V_i^{\widehat{s}_i,\hat{\mu}^i}(h) = \max_{s'_i \in S_i(h)} V_i^{s'_i,\hat{\mu}^i}(h),$$

where the inequality holds because $s_i \in r_i(\mu^i) \cap S_i(h)$, and the last equality follows from the fact that $\widehat{s}_i \in r_i(\hat{\mu}^i) \cap S_i(h)$. But then, there exists $\widehat{s}'_i \in r_i(\hat{\mu}^i)$ such that $\widehat{s}'_i(h) = s_i(h)$ and $\widehat{s}'_i(h') = \widehat{s}_i(h')$ for every $h' \prec^\ell h$. ■

12

Equilibrium in Multistage Games

In this chapter we analyze subgame perfect equilibrium, a solution concept that—unlike Nash equilibrium, or iterated admissibility—is based on the details of the multistage game Γ rather than just its strategic form $\mathcal{N}(\Gamma)$. Somewhat informally, a **subgame perfect equilibrium** is a profile of strategies $s^* = (s_i^*)_{i \in I}$ such that each strategy s_i^* is sequentially optimal given the conjecture that other players behave as specified by s_{-i}^*. In other words, the *ex ante* optimality condition of Nash equilibrium is replaced by the more demanding *sequential* optimality condition. Indeed, as in the analysis of rationalizability of Chapter 11, it is *not* assumed that strategies are implemented by a machine or a trustworthy agent. Rather, it is informally assumed that each player makes a plan in advance and, as the play unfolds, he is free to either choose actions as planned or to deviate from the plan; therefore, players must not have incentives to deviate at any history, given that—as is typically assumed in equilibrium analysis—they have correct conjectures about each other. In finite games where only one player is active at each stage and payoffs are "generic," there is only one subgame perfect equilibrium that can be computed with the backward-induction algorithm we introduced in Chapter 11. In more general games with finite horizon, one can use a "case-by-case backward induction" method to find the subgame perfect equilibria. Finally, we extend the analysis to account for randomized strategies and chance moves.

12.1 Subgame Perfect Equilibrium

In Section 9.3.1 we noticed that every solution concept for games with simultaneous moves can be applied to the strategic form $\mathcal{N}(\Gamma)$ of a multistage game Γ. This may be a useful first step. Indeed, a strategy is sequentially optimal given a conjecture only if it is an ex ante best reply to this conjecture (see Remark 39 and Theorem 38 in Chapter 10). This implies that strategies that are not rationalizable in $\mathcal{N}(\Gamma)$ cannot be rationalizable in Γ for whatever extension of the rationalizability concept from simultaneous to multistage games. Yet, we noticed in Chapter 11 that, in games with at least two stages, rationality entails a stronger form of justifiability than just being an ex ante best reply to some conjecture and this yields notions of rationalizability in Γ stronger than mere rationalizability in the strategic form $\mathcal{N}(\Gamma)$. Similar considerations apply to traditional equilibrium analysis: there may be Nash equilibria s^* of $\mathcal{N}(\Gamma)$ such that the strategy s_i^* of some player i is not sequentially optimal given that i "correctly" expects the co-players to behave as prescribed by strategy profile s_{-i}^*. The strategy of fighting entry in the Entry Game of Figure 9.7 is a case in point (see Section 9.3.1). The credibility of such strategies is questionable. Such considerations induced traditional game theorists to endorse a refinement[1] of the Nash equilibrium concept called "subgame perfect (Nash) equilibrium" for reasons that will soon be made clear. As in the case of the Nash equilibrium concept, no formal and general justification is offered for the "correct-conjectures" assumption underlying this equilibrium concept. We first focus on equilibria in pure strategies, then move to randomized equilibria.

12.1.1 Pure Strategies

For any player i and conjecture $\beta^i \in \times_{h \in H} \Delta(\mathcal{A}_{-i}(h))$, let $\hat{r}_i(\beta^i)$ denote the set of sequentially optimal strategies given β^i, or **sequential best replies** to β^i.[2] Suppose that conjecture β^i is deterministic because it

[1] A "refinement" of a solution concept is a different, more demanding solution concept that, for some games, allows for a strictly smaller set of strategy profiles. For example, initial rationalizability is a refinement of rationalizability in the strategic form, and strong rationalizability is a refinement of initial rationalizability.

[2] Recall that $r_i(\beta^i)$ instead denotes the set of *weakly* sequential optimal strategies given β^i. Thus, $\hat{r}_i(\beta^i) \subseteq r_i(\beta^i)$ and the inclusion may be strict if player i moves (is

corresponds to a pure strategy profile s_{-i}, that is, $\beta^i(s_{-i}(h)|h) = 1$ for all $h \in H$; then, with an abuse of notation similar to the one adopted for static games, we write $\hat{r}_i(s_{-i})$.

Definition 71. *A pure strategy profile $s^* = (s_i^*)_{i\in I}$ is a **subgame perfect (Nash) equilibrium (SPE)** if, for each $i \in I$, s_i^* is sequentially optimal given the deterministic conjecture corresponding to s_{-i}^*, that is,*

$$\forall i \in I, s_i^* \in \hat{r}_i\left(s_{-i}^*\right).$$

By inspection of the definition above and of the backward justification operator ρ_{fb} of Chapter 11, one can show that every SPE s^* satisfies $\{s^*\} \subseteq \rho_{\text{fb}}(\{s^*\}) \subseteq \rho_{\text{fb}}^{\infty}(S)$. Therefore,

Remark 46. *Every SPE is continuation-rationalizable.*

For every strategy profile $s \in S$ and history $h \in H$, let $\zeta(s|h)$ denote the terminal history $z = (h, s(h), ...)$ induced by s *starting from* h. Note, if s induces h, that is $h \prec \zeta(s)$, then $\zeta(s|h) = \zeta(s)$. But $\zeta(s|h)$ is well defined even if s does not induce h, because $\zeta(s|h)$ depends only on the actions prescribed by s at h and the nonterminal histories following h (if any). For example, in the BoS with an Outside Option of Figure 9.1, $\zeta((\text{Out.B}_1, \text{B}_2)\,|\,(\text{In})) = (\text{In}, (\text{B}_1, \text{B}_2))$. The following remark, which holds by inspection of Definition 71, explains the name "subgame perfect (Nash) equilibrium":

Remark 47. *A strategy profile $s^* = (s_i^*)_{i\in I}$ is an SPE if and only if*

$$\forall i \in I, \forall h \in H, \forall s_i \in S_i,\ u_i(\zeta(s^*|h)) \geq u_i(\zeta(s_i, s_{-i}^*|h)).$$

Thus, for every $h \in H$, an SPE yields a Nash equilibrium in the subgame $\Gamma(h)$ starting at h.

Example 61. In the BoS with an Outside Option there are two (pure) SPEs. Indeed, the BoS subgame starting at (In) has two pure equilibria, (B_1, B_2) and (S_1, S_2). Thus, a SPE must have the form $(a_1.\text{B}_1, \text{B}_2)$ or $(a_1.\text{S}_1, \text{S}_2)$ with $a_1 \in \{\text{In}, \text{Out}\}$. Furthermore, $a_1.\text{X}_1$ must be a best reply to X_2 (with X= B or X = S) at the root, because the strategy pair must be a Nash equilibrium. Only In.B_1 is a best reply to B_2, while Out.S_1 is

active) at leat twice along some given path.

the only best reply to S_2 of the required form (also $\mathrm{Out.B}_1$ is a best reply to S_2 at the root, but it does not have the required form, hence, it is not a sequential best reply to S_2). Thus, we obtain two SPEs: $(\mathrm{In.B}_1, \mathrm{B}_2)$ and $(\mathrm{Out.S}_1, \mathrm{S}_2)$. ▲

We will present a method to find all the SPEs of (finite) two-stage games. Most methods to compute the SPEs of a game are based on an equilibrium version of the OD principle. Say that **strategy profile** $s^* = (s_i^*)_{i \in I}$ satisfies **one-step optimality with correct conjectures** if, for each $i \in I$, strategy s_i^* is one-step optimal given the deterministic conjecture corresponding to s_{-i}^*. With this, Theorem 41 in Section 10.5 yields:

Corollary 7. (OD principle for pure SPE) *Suppose that* Γ *is continuous at infinity. Then a strategy profile is an SPE if and only if it satisfies one-step optimality with correct conjectures.*

By Remark 41 in Section 10.5, it follows that, for every *compact-continuous* game Γ (hence, in particular, for every finite game) a strategy profile is an SPE if and only if it satisfies one-step optimality with correct conjectures.

Backward Induction

Recall that in Chapter 11 we introduced the backward induction procedure to solve perfect-information (PI) games with no relevant ties as a special case of the backward procedure, which computes continuation-rationalizable strategies. It can be checked that this procedure computes the unique SPE in such games.

Remark 48. *For every finite PI game* Γ *with no relevant ties, the strategy profile* s^* *obtained with the backward induction algorithm is the unique SPE of* Γ.

"Case-by-Case" Backward Induction

For finite games that do not have a unique SPE computable by backward induction, Corollary 7 still offers a method to compute all the SPEs. We describe it for *two-stage games*. Consider a *two-stage* game Γ. Fix any first-stage, nonterminal (hence, preterminal) action profile $a \in \mathcal{A}(\varnothing) \backslash Z$.[3] Consider the second-stage subgame

$$G^2(a) = \left\langle I, (\mathcal{A}_i(a), u_i(a, \cdot))_{i \in I} \right\rangle .$$

In other words, the set of feasible actions of player i in $G^2(a)$ is $\mathcal{A}_i(a)$ and his payoff function in $G^2(a)$ is

$$a' \mapsto u_{i,a}(a') = u_i(a, a') .$$

For example, in the BoS with an Outside Option $\mathcal{A}(\varnothing) \backslash Z = \{\text{In}\}$ and $G^2(\text{In})$ is the BoS subgame.

Stage 2. Let $NE^2(a)$ denote the set of (pure) Nash equilibria of $G^2(a)$. If $NE^2(a) = \emptyset$ for some $a \in \mathcal{A}(\varnothing) \backslash Z$, then Γ cannot have any (pure) SPE. Suppose that $NE^2(a) \neq \emptyset$ for all $a \in \mathcal{A}(\varnothing) \backslash Z$, and consider all possible selections from the second-stage equilibrium correspondence $a \mapsto NE^2(a)$, that is, all the functions

$$s^2 : (\mathcal{A}(\varnothing) \backslash Z) \to \bigcup_{a \in \mathcal{A}(\varnothing) \backslash Z} \mathcal{A}(a)$$

such that $s^2(a) \in NE^2(a)$ for each $a \in \mathcal{A}(\varnothing) \backslash Z$. Note, the set of such selections is the Cartesian product $\times_{a \in \mathcal{A}(\varnothing) \backslash Z} NE^2(a)$, which has cardinality $\prod_{a \in \mathcal{A}(\varnothing) \backslash Z} \left| NE^2(a) \right|$.

Stage 1. Each selection $s^2 \in \times_{a \in \mathcal{A}(\varnothing) \backslash Z} NE^2(a)$ is a "**case**" to which we apply a backward induction calculation. Specifically, for each selection s^2, define the *auxiliary* simultaneous-move *game*

$$G^1(s^2) = \left\langle I, \left(\mathcal{A}_i(\varnothing), u_i^1(\cdot, s^2)\right)_{i \in I} \right\rangle ,$$

where

$$u_i^1(a, s^2) = \begin{cases} u_i\left(a, s^2(a)\right), & \text{if } a \in \mathcal{A}(\varnothing) \backslash Z, \\ u_i(a), & \text{if } a \in Z. \end{cases}$$

In words, $G^1(s^2)$ specifies the payoffs of each first-stage action profile a under the hypothesis (commonly believed by the players) that, if a is nonterminal , the following second-stage profile will be $s^2(a)$. Since

[3] With a small abuse of notation, we do not distinguish between an action profile a and the sequence of length one (a). Therefore, if $(a) \in Z$, we write $a \in Z$.

$s^2(a)$ is a Nash equilibrium of $G^2(a)$ for each $a \in \mathcal{A}(\varnothing)\backslash Z$, every Nash equilibrium $s^1 \in \mathcal{A}(\varnothing)$ of $G^1(s^2)$ yields an SPE $s = (s^1, s^2)$: specifically, for each $i \in I$, $s_i(\varnothing) = s_i^1$ and $s_i(a) = s_i^2(a)$ for all $a \in \mathcal{A}(\varnothing)\backslash Z$.

Let $NE^1(s^2)$ denote the set of (pure) equilibria of the auxiliary $G^1(s^2)$. Then the number of SPEs of Γ is the summation over "cases" s^2 of number of equilibria in the auxiliary games $G^1(s^2)$:

$$\sum_{s^2 \in \times_{a\in\mathcal{A}(\varnothing)\backslash Z} NE^2(a)} \left|NE^1\left(s^2\right)\right|.$$

Hence, there is only one SPE of Γ *if* (i) the second-stage equilibrium correspondence $a \mapsto NE^2(a)$ is actually a function s^2 (hence $\times_{a\in\mathcal{A}(\varnothing)\backslash Z} NE^2(a)$ is a singleton) and (ii) $G^1(s^2)$ has a unique equilibrium.

Example 62. The BoS with a dissipative move of Example 42 has two strategically equivalent second-stage BoS subgames, $G^2(B)$ and $G^2(N)$. The second-stage equilibrium correspondence is $B \mapsto \{(u,l),(d,r)\}$ and $N \mapsto \{(U,L),(D,R)\}$. Therefore we obtain $2 \times 2 = 4$ cases/selections, i.e., 4 auxiliary games:

$G^1((U,L),(u,l))$	N	B
	4, **1**	4, -1

$G^1((U,L),(d,r))$	N	B
	4, 1	1, **2**

$G^1((D,R),(d,r))$	N	B
	1, **4**	1, 2

$G^1((D,R),(u,l))$	N	B
	1, **4**	4, -1

By inspection of the equilibria of these auxiliary games, we obtain the set of SPEs (see the payoffs in bold):

$$SPE = \{(U.u, N.L.l), (U.d, B.L.r), (D.d, N.R.r), (D.u, N.R.l)\}.$$

We may interpret SPE $(U.d, B.L.r)$ as follows: the dissipative action B is regarded as a signal that Bob is going for his favorite equilibrium in subgame $G^2(B)$, whereas action N is interpreted as a signal that Bob "concedes" and just aims at equilibrium (U,L) of $G^2(N)$. ▲

Properties (i)-(ii) give a sufficient, but not necessary condition for uniqueness, as the following example shows.

Example 63. Ann and Bob play the following two-stage game Γ, where every first-stage action pair but (D,R) is terminal, whereas (D,R) leads to a sort of Battle-of-the-Sexes subgame G^2:

$a\backslash b$	L	R
U	2, 1	1, 2
D	1, 2	G^2

with

G^2	l	r
u	**2, 1**	0, 0
d	0, 0	**1/2, 2**

According to the subgame equilibrium of G^2 we obtain two auxiliary games:

$G^1(u,l)$	L	R
U	2, 1	1, 2
D	1, 2	2, 1

$G^1(d,r)$	L	R
U	2, 1	**1, 2**
D	1, 2	1/2, 2

Auxiliary game $G^1(u,l)$ is a kind of "Matching Pennies" and has no equilibrium; thus, there is no SPE where (u,l) is selected in subgame G^2. Auxiliary game $G^1(d,r)$ instead is dominance solvable and its unique equilibrium is (U,R). Therefore, Γ has a *unique* SPE, $(U.d, R.r)$. In this equilibrium, Ann does not deviate in the first stage because she expects to be "punished" by the subgame equilibrium of G^2—in which she gets only 1/2—if she chooses D. ▲

12.1.2 Randomized Strategies

The definition of the continuation-value of strategies and actions can be extended to locally randomized strategies, i.e., behavior strategies. To simplify the probabilistic analysis we focus throughout on *finite* games.

For any behavior strategy profile $\beta = (\beta_i)_{i\in I}$ and histories $h = (a^1, ..., a^{\ell(h)})$, $h' = (h, a^{\ell(h)+1}, ..., a^{\ell(h')})$ let

$$\mathbb{P}_\beta(h'|h) = \prod_{t=\ell(h)+1}^{\ell(h')} \prod_{i\in I} \beta_i(a_i^t|h, ..., a^{t-1})$$

denote the probability of reaching h' from h. Then

$$V_i^\beta(h) = \sum_{z\in Z(h)} u_i(z)\mathbb{P}_\beta(z|h),$$

$$V_i^\beta(h, a_i) = \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} V_i^\beta(h, (a_i, a_{-i}))\beta_{-i}(a_{-i}|h),$$

where $\beta_{-i}(a_{-i}|h) := \prod_{j\neq i}\beta_j(a_j|h)$. Obviously, $V_i^{\beta}(h)$ and $V_i^{\beta}(h,a_i)$ depend only on the behavior of β in the sub-tree with root h.

Definition 72. *A* ***behavior strategy profile*** $\beta^* = (\beta_i^*)_{i\in I}$ *is an SPE if, for each* $i \in I$*,*

$$\forall h \in H,\ V_i^{\beta^*}(h) \geq \max_{\beta_i} V_i^{\beta_i,\beta^*_{-i}}(h);$$

β^* *satisfies* ***one-step optimality with correct conjectures*** *if, for each* $i \in I$*,*

$$\forall h \in H,\ \text{supp}\beta_i^*(\cdot|h) \subseteq \arg\max_{a_i\in\mathcal{A}_i(h)} V_i^{\beta^*}(h,a_i).$$

The results about dynamic optimality can be extended to behavior strategies, and in particular they imply the OD Principle for randomized SPE:

OD Principle for randomized SPE: *A behavior strategy profile is an SPE if and only if it satisfies one-step optimality with correct conjectures.*

As with pure strategies, there is a "case-by-case" backward induction algorithm to compute the set of SPEs in behavior strategies, which can be easily spelled out for two-stage games and is very similar to the algorithm explained above: Just replace the stage-2 pure equilibrium selection s^2 with a stage-2 mixed equilibrium selection $\beta^2 \in \times_{a\in\mathcal{A}(\varnothing)\setminus Z} MNE^2(a)$, where $MNE^2(a)$ denotes the set of mixed Nash equilibria of $G^2(a)$. There are two important differences with respect to pure strategy equilibria. First, since every finite simultaneous-move game has always at least one mixed Nash equilibrium, we have $\times_{a\in\mathcal{A}(\varnothing)\setminus Z} MNE^2(a) \neq \emptyset$ and $MNE\left(G^1\left(\beta^2\right)\right) \neq \emptyset$ for each β^2; therefore, there is a unique SPE in behavior strategies *if and only if* (i) $\times_{a\in\mathcal{A}(\varnothing)\setminus Z} MNE^2(a) = \{\beta^2\}$ is a singleton and (ii) $G^1\left(\beta^2\right)$ has a unique mixed equilibrium (that is, (i)-(ii) give a *necessary* and sufficient condition for uniqueness of randomized equilibria). Second, in some rare "nongeneric" games $\times_{a\in\mathcal{A}(\varnothing)\setminus Z} MNE^2(a)$ has the power of the continuum and there is a continuum of "cases" to deal with in the stage-1 analysis.

Example 64. The BoS subgames of Example 42 have a completely mixed equilibrium on top of the two pure equilibria, that is,

$\left(\frac{4}{5}\delta_U + \frac{1}{5}\delta_D, \frac{1}{5}\delta_L + \frac{4}{5}\delta_R\right)$ in $G^2(N)$ and $\left(\frac{4}{5}\delta_u + \frac{1}{5}\delta_d, \frac{1}{5}\delta_l + \frac{4}{5}\delta_r\right)$ in $G^2(B)$. Therefore there are $3 \times 3 = 9$ second-stage selections β^2 with 9 corresponding auxiliary games $G^1(\beta^2)$. We have already computed the 4 pure selections and corresponding pure SPE's in Example 62. Without any actual computation, we can claim that there are at least 5 partially or totally *randomized* additional equilibria, that is, at least one for each non pure selection $\beta^2 = s^2$.[4] For example, the selection given by $\beta^2(N) = \left(\frac{4}{5}\delta_U + \frac{1}{5}\delta_D, \frac{1}{5}\delta_L + \frac{4}{5}\delta_R\right)$ in $G^2(N)$ and $\beta^2(B) = (d, r)$ in $G^2(B)$ yields the auxiliary game

$G^1(\beta^2)$	N	B
	$\frac{4}{5}$, $\frac{4}{5}$	1, **2**

and the partially randomized equilibrium

$$\left(\left(\frac{4}{5}\delta_U + \frac{1}{5}\delta_D\right).d, B.\left(\frac{1}{5}\delta_L + \frac{4}{5}\delta_R\right).r\right),$$

where $\beta_a = \left(\frac{4}{5}\delta_U + \frac{1}{5}\delta_D\right).d$ is the partially randomized strategy of Ann that would play $\left(\frac{4}{5}\delta_U + \frac{1}{5}\delta_D\right)$ upon observing N and plays d after B, and $\beta_b = B.\left(\frac{1}{5}\delta_L + \frac{4}{5}\delta_R\right).r$ is the partially randomized strategy of Bob that plays B in the first stage, would play $\left(\frac{1}{5}\delta_L + \frac{4}{5}\delta_R\right)$ after the counterfactual choice N, and plays r after B. ▲

The OD Principle allows a relatively simple proof of the existence of randomized equilibria in finite games:

Theorem 49. *(Kuhn) Every finite game has at least one subgame perfect equilibrium in behavior strategies.*

Proof. We provide a recursive construction of a behavior strategy profile β^*. We start from histories of height 1, i.e., those immediately preceding the last stage, and go backward until the first stage.

Basis step. For every history h with height 1, the corresponding last-stage simultaneous game has at least one mixed equilibrium. Let $\beta^*(\cdot|h) \in \times_{i\in I} \Delta(\mathcal{A}_i(h))$ be one of the mixed equilibria, and consider

[4] If there is a selection β^2 that makes Bob (the only first-mover) indifferent between B and N, then there is a corresponding continuum of randomized SPE's parameterized by the probability of burning $\beta_b(B|\varnothing)$.

the corresponding expected payoff profile $(v_i^*(h))_{i\in I}$, where $v_i^*(h) := \sum_{a\in\mathcal{A}(h)}\left(\prod_{j\in I}\beta_j^*(a_j|h)\right)u_i(h,a)$.

Inductive step. Suppose that a mixed action profile $\beta^*(\cdot|h)$ and an expected payoff profile $(v_i^*(h))_{i\in I}$ have been assigned to each history with height k or less. Then one can assign a mixed action profile $\beta^*(\cdot|h)$ and a payoff profile $(v_i^*(h))_{i\in I}$ to each history h with height $k+1$: just pick a mixed equilibrium of the game $\left\langle I, \left(\mathcal{A}_i(h), v_i^h\right)_{i\in I}\right\rangle$ such that, for every $i\in I$ and for every $a\in\mathcal{A}(h)$, $v_i^h(a) = v_i^*(h,a)$ (note that (h,a) has height k or less, so $v_i^*(h,a)$ is well defined).

By construction, the behavior strategy profile β^* satisfies one-step optimality with correct conjectures. Thus, by the OD Principle, β^* is a subgame perfect equilibrium. ■

12.2 Subgame Perfection with Chance Moves

The definition of multistage game can be extended to allow for the possibility that some events in the game depend on chance. Chance is modeled as a fictitious player, denoted by 0, that chooses actions with exogenously given conditional probabilities. This means that the probability that a particular chance action a_0 is selected immediately after history h is a number $\beta_0(a_0|h)$ which is part of the description of the game.

We maintain the convention that at each stage there are simultaneous moves by all players, including chance. Since the set of active players (those with more than one feasible action) may depend on time and on the previous history, this assumption does not entail any loss of generality. For example, one may model games where chance moves occur before or after the moves of real players. To minimize on notational changes, we also ascribe a utility function u_0 to the chance player 0, but we assume that it is constant. This is just another notational trick.

A **multistage game with observed actions and chance moves** is a structure

$$\Gamma = \left\langle (A_i, \mathcal{A}_i(\cdot), u_i)_{i=0}^n, \beta_0 \right\rangle.$$

The symbols A_i, $\mathcal{A}_i(\cdot)$, u_i have the same meaning as before. From the feasibility correspondences $\mathcal{A}_i(\cdot)$ $(i = 0,...,n)$ we can derive the set of feasible histories $\bar{H}$, the set of terminal histories Z and the set of nonterminal histories $H = \bar{H}\backslash Z$ (see section 9.1). Functions u_i are real-valued and have domain Z. The only novelty is player 0, chance: for every $z \in Z$, $u_0(z) = 0$, and β_0 describes how the probabilities of chance moves depend on the previous history, that is,

$$\beta_0 = (\beta_0(\cdot|h))_{h \in H} \in \bigtimes_{h \in H} \Delta^\circ(\mathcal{A}_0(h)),$$

where $\Delta^\circ(\mathcal{A}_0(h))$ is the set of full-support probability measures on $\mathcal{A}_0(h)$.[5]

Next we define the strategic form. To avoid measure-theoretic technicalities we will henceforth assume that all the sets $\mathcal{A}_0(h)$ ($h \in H$) are *finite*. Let I be the set of the "real" players, i.e., $I = \{1, ..., n\}$. The set S_i of strategies of $i \in I$ is defined as before. Although "strategies of chance" are well-defined mathematical objects, we disregard them, as chance is not a strategic player. Thus, the set of strategy profiles is $S = \bigtimes_{i \in I} S_i$ and similarly $S_{-i} = \bigtimes_{j \in I\backslash\{i\}} S_j$. The presence of chance moves implies that more than one terminal history may be possible when the players follow a particular strategy profile s; let $Z(s)$ denote the set of such terminal histories. The probability of each $z \in Z(s)$ depends on the probability of the chance moves contained in z and is denoted $\hat{\zeta}(z|s)$ (if a player, or an external observer were uncertain about the true strategy profiles, $\hat{\zeta}(z|s)$ would represent the probability of z conditional on the players following the strategy profile s, which explains the notation). Therefore the path function when there are chance moves has the form $\hat{\zeta} : S \to \Delta(Z)$. The formal definition of $Z(s)$ and $\hat{\zeta}(\cdot|s)$ is as follows. Let $\mathbf{a}_i^t(z)$ denote the action of player i at stage t in history z, let $\mathbf{h}^{t-1}(z)$ denote the prefix (initial sub-history) of z of length t, and recall that $\ell(z)$ is the length of z. Then

$$\forall s \in S,\ Z(s) = \{z \in Z : \forall t \in \{1, ..., \ell(z)\}, \forall i \in I,\ \mathbf{a}_i^t(z) = s_i(\mathbf{h}^{t-1}(z))\},$$

$$\forall z \in Z,\ \hat{\zeta}(z|s) = \begin{cases} \prod_{t=1}^{\ell(z)} \beta_0(\mathbf{a}_0^t(z)|\mathbf{h}^{t-1}(z)), & \text{if } z \in Z(s), \\ 0, & \text{if } z \notin Z(s). \end{cases}$$

In words, $\hat{\zeta}(z|s)$ is the product of the probabilities of the actions taken by the chance player in history z.

[5] Recall that $\Delta^o(X) = \{\mu \in \Delta(X) : \text{supp}\mu = X\}$, where the superscript o stands for (relatively) "open" in the hyperplane of vectors whose elements sum to 1.

A **Nash equilibrium** of Γ is a strategy profile s^* such that for every $i \in I$ and for every $s_i \in S_i$,

$$\sum_z \hat{\zeta}(z|s^*)u_i(z) \geq \sum_z \hat{\zeta}(z|s_i, s^*_{-i})u_i(z).$$

In other words, a Nash equilibrium of Γ is an equilibrium of the static game where players simultaneously choose strategies and have payoff functions $U_i(s) = \sum_z \hat{\zeta}(z|s)u_i(z)$ $(i \in I,\ s \in S)$.

In order to define subgame perfect equilibria, we first define the conditional outcome function $\hat{\zeta}(\cdot|h, s)$: for each $h \in H$, the set of terminal histories that can be reached when s is followed starting from h (where h may be *in*consistent with s) is

$$Z(h, s) = \{z \in Z : h \prec z, \forall t \in \{\ell(h)+1, ..., \ell(z)\}, \forall i \in I, \mathbf{a}_i^t(z) = s_i(\mathbf{h}^{t-1}(z))\};$$

then

$$\forall z \in Z,\ \hat{\zeta}(z|h, s) = \begin{cases} \prod_{t=\ell(h)+1}^{\ell(z)} \beta_0(\mathbf{a}_0^t(z)|\mathbf{h}^{t-1}(z)), & \text{if } z \in Z(h, s), \\ 0, & \text{if } z \notin Z(h, s). \end{cases}$$

A strategy profile s^* is a **subgame perfect equilibrium** if for all $h \in H$, $i \in I$, and $s_i \in S_i$,

$$\sum_z \hat{\zeta}(z|h, s^*)u_i(z) \geq \sum_z \hat{\zeta}(z|h, s_i, s^*_{-i})u_i(z).$$

Let $\hat{\zeta}(\cdot|h, a_i, s) \in \Delta(Z)$ denote the probability measure on Z that results if s is followed starting from h, except that player i chooses $a_i \in \mathcal{A}_i(h)$ at h. A strategy profile s satisfies **one-step optimality with correct conjectures** if for all $h \in H$, $i \in I$, and $a_i \in \mathcal{A}_i(h)$,

$$\sum_z \hat{\zeta}(z|h, s)u_i(z) \geq \sum_z \hat{\zeta}(z|h, a_i, s)u_i(z).$$

It can be shown that in every game with finite horizon or with continuity at infinity (hence in every game with discounting) a strategy profile is a subgame perfect equilibrium if and only if it satisfies one-step optimality with correct conjectures.

13

Repeated Games

A subset of multistage games that has received particular attention in the literature is the one of *repeated games*. Loosely speaking, a repeated game is a strategic environment in which the same players play multiple times the same static game (or one-period game)[1]—the **stage game**. More generally, we could think of the repetition of the same static or sequential one-period game; but our analysis, like most of the literature on repeated games, will be restricted to the repetition of the same static (i.e., simultaneous-move) games, so that stages will coincide with periods. See our comments in Section 9.2.1.

Many real-life situations can be modelled as repeated games. Think of a couple which has to decide every weekend what to do. Or consider the daily relationship between two colleagues, who must independently decide whether to exert effort on a common project or to shirk. Or again, suppose that firms are involved in a cartel and must decide whether to set their individual price as suggested by the cartel or to undercut such level in order to reap a larger market share.

In addition to their practical relevance, repeated games are interesting also from a theoretical point of view. Indeed, given the dynamic nature of the strategic interaction, players can adjust their future behavior in order to incentive other players to behave in a certain way. Although this phenomenon may arise in any multistage game, the specific structure of repeated games (namely, the fact that the same stage game is repeated over and over again) makes them the ideal environment to study these

[1] In this chapter we assume that stages and periods coincide.

reward and punishment strategies.

Before we delve into the formal analysis, it is useful to introduce some taxonomy. First of all, the class of repeated games can be divided in *finitely repeated games* and *infinitely repeated games* depending on whether the stage game is repeated a finite number of times or infinitely many times. Furthermore, repeated games can be classified according to the information feedback that players receive about the behavior of their opponents at previous stages. *Repeated games with perfect monitoring* are games in which each player perfectly observes (and recalls) the actions of all players in previous rounds. Thus, this class of games is a special case of multistage games with observed actions. *Repeated games with imperfect public monitoring* are games in which players do not directly observe what their opponents chose in the past. Instead, in each round, they observe (and recall) a common public stochastic signal, whose realization is affected by the action profile chosen by players in the previous round. *Repeated games with imperfect private monitoring* are games in which players do not directly observe what their opponents chose in the past, nor do they observe a common public signal. Instead, each of them observes (and recalls) a private stochastic signal, whose realization depends on the action taken by players in the previous round.

One of the main questions addressed by the literature on repeated games is the following: "To what extent can the repeated strategic interaction enlarge the set of equilibrium payoffs with respect to those of the stage game?" An answer to this question is provided by so called "folk theorems," namely results with the following flavor.

Informal Folk Theorem. Consider an infinitely repeated game. If players are very patient and they can statistically detect the behavior of their opponents, then any payoff profile compatible with individual optimality (namely, each payoff profile in which players get at least as much as the worst possible payoff consistent with utility-maximizing behavior) is achievable in equilibrium.

In this chapter we follow a different approach and we stress the strategic nature of the repeated interaction. As such, we highlight how dynamic incentives come at play in repeated games and may enable players to punish or reward their opponents. As a by-product of this approach, we introduce the tools needed to prove folk theorems.

Specifically, we will only analyze repeated games with *perfect* *monitoring.* First, we will introduce some basic notation concerning the one-period game and the associated repeated games. Then, we will prove some simple results relating the Nash equilibria of the one-period game to the subgame perfect equilibria of the corresponding repeated game for a fixed "degree of patience" (discount factor). Finally, we will provide a characterization of the set of (pure) SPE payoff profiles in infinitely repeated games, again, for a fixed "degree of patience." Since it is implausible to assume that agents are arbitrarily patient, the latter kind of results (rather than the "folk theorems") are those most used in the applications of the theory of repeated games.

13.1 Repeated Games with Perfect Monitoring

For any static *compact-continuous* game $G = \langle I, (A_i, v_i)_{i\in I}\rangle$, we let $\Gamma^{\delta,T}(G)$ denote the T-stage (or T-period) game with observed actions obtained by repeating G for $T \in \mathbb{N} \cup \{\infty\}$ times and computing payoffs $u_i : Z \to \mathbb{R}$ (where $Z = A^T$) as discounted time averages with common discount factor $\delta \in (0, 1)$;[2] that is, for every terminal history $(a^t)_{t=1}^T \in Z$ and for every player i,

$$u_i(a^1, a^2, ...) = \sum_{t=1}^{T} \delta^{t-1} v_i(a^t). \qquad (13.1.1)$$

Note that u_i is well defined even if $T = \infty$, because G is compact-continuous, and so v_i is bounded; therefore, the series of discounted payoffs has a finite sum. In the analysis of finitely repeated games ($T \in \mathbb{N}$) one can also allow for $\delta = 1$, i.e., no discounting. Also note that the most important insights of the analysis can be extended to the more general case where G has *bounded* payoff functions, e.g., when G is an oligopoly game with price setting, homogeneous products, and competition *à la* Bertrand.

In the analysis of infinitely repeated games, it may be convenient to use an equivalent representation of the payoff associated with an infinite history:

$$\bar{u}_i(a^1, a^2, ...) = (1 - \delta) \sum_{t=1}^{T} \delta^{t-1} v_i(a^t), \qquad (13.1.2)$$

[2] The analysis can be easily extended to allow for heterogeneous discount factors.

which is a **discounted time average** of the stream of payoffs $\left(v_i(a^t)\right)_{t=1}^{\infty}$. When the stream of payoffs is constant, e.g., $v_i(a^t) = v_i(\bar{a})$ for each t, then the average (13.1.2) is precisely this constant:

$$\bar{u}_i(\bar{a}, \bar{a}, ...) = (1-\delta)\sum_{t=1}^{T}\delta^{t-1}v_i(\bar{a}) = v_i(\bar{a}).$$

In this strategic environment, every history h of length $t \in \mathbb{N}$ is an element of A^t. By convention, $A^0 = \{\varnothing\}$ is the set containing the empty sequence of action profiles.

Example 65. Consider the Prisoners' Dilemma in which the set of players is given by $I = \{1, 2\}$ and the set of actions available to each player i is given by "*C*ooperation" and "*D*efection," that is, $A_i = \{C, D\}$. Utilities are represented in the following table.

	C	D
C	$2,2$	$0,3$
D	$3,0$	$1,1$

The repeated game with perfect monitoring $\Gamma^{\delta,2}(G)$ is the multistage game obtained by repeating the Prisoners' Dilemma twice. In this case, the set of nonterminal histories is

$$H = \{\varnothing\} \cup A$$

and the payoff of player $i \in I$ at terminal history $\left(a^1, a^2\right)$ is given by

$$u_i\left(a^1, a^2\right) = v_i\left(a^1\right) + \delta v_i\left(a^2\right).$$

▲

Example 66. There are n firms which produce and sell a homogenous good: I is the set of firms, $q_i \in [0, \bar{q}_i]$ is the output of firm i and $\bar{q}_i$ is its capacity. The market (inverse) demand function $P : [0, \sum_i \bar{q}_i] \to \mathbb{R}_+$ is given by

$$P\left(\sum_{j\in I} q_j\right) = \max\{0, \bar{p} - \beta\sum_{j\in I} q_j\}.$$

Firms compete *à la* Cournot choosing the quantity to produce in order to maximize their total profits and keeping into account that the production of the good entails a unit cost equal to c. Thus, each firm i maximizes $v_i(q_1, ...q_n) = q_i P\left(q_i + \sum_{j \neq i} q_j\right) - cq_i$. Game $\Gamma^{\delta,\infty}(G)$ is the infinitely repeated game in which the set of histories is given by

$$\bar{H} = A^{\leq \mathbb{N}_0} = \{\varnothing\} \cup \left(\bigcup_{t \in \mathbb{N}} \left(\bigtimes_{i \in I} [0, \bar{q}_i] \right)^t \right) \cup \left(\bigtimes_{i \in I} [0, \bar{q}_i] \right)^{\mathbb{N}},$$

and the payoff function of firm i is given by

$$\bar{u}_i(\left(q_1^1, ..., q_n^1\right), \left(q_1^2, ..., q_n^2\right), ...) = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} \left(q_i^t \max\{0, \bar{p} - \beta \sum_{j \in I} q_j^t\} - cq_i^t \right).$$

▲

In the analysis of repeated games with perfect monitoring, we use subgame perfect equilibrium (SPE) as a solution concept. Notice that since G is compact-continuous and $\delta < 1$, the repeated game is continuous at infinity (see Section 10.5) and, consequently, we can use the OD Principle to characterize SPEs.

13.1.1 Automaton Representation of Strategy Profiles

The equilibrium analysis of repeated games often requires the construction of strategy profiles that attain a certain payoff. Since the set of histories in a repeated game may include many different elements, this task can lead to significant notational complexity. To simplify the analysis, it is often useful to represent strategy profiles by collecting histories in equivalence classes with respect to the strategic behavior of players. This approach underlies the so-called *automaton* representation of a strategy profile.

Definition 73. *An automaton is a profile* $(\Psi, \psi_0, \gamma, \varphi)$ *where:*

- Ψ *is a set of states;*
- $\psi_0 \in \Psi$ *is the initial state;*
- $\gamma : \Psi \to A$ *is a behavioral rule that specifies an action profile for each state;*

- $\varphi : \Psi \times A \to \Psi$ *is a transition rule that specifies how the automaton moves across states based on the current state and by the action profile chosen by players.*

The automaton representation of a strategy profile in a repeated game with perfect monitoring requires to specify a set of states. Each state corresponds to a set of histories in the repeated game at which the behavior of players (under the strategy profile) is constant. Put differently, the behavior of players under a certain strategy profile is measurable with respect to the partition of nonterminal histories represented by the states.

Every strategy profile $s \in S$ can be represented as an automaton. To see this, let $\Psi = H$, $\psi_0 = \varnothing$ and $\gamma(h) = (s_i(h))_{i \in I}$ for every $h \in H$. Finally, let $\varphi(h, a) = (h, a) \in \Psi$. In general, the automaton representation we just provided is not the only possible one; indeed it is often possible to reduce the set of states by grouping histories together—namely, by coarsening the equivalence classes.

Conversely, given an automaton $(\Psi, \psi_0, \gamma, \varphi)$, we can derive a profile of strategies as follows. For every $\ell \geq 1$ and for every $h^\ell \in A^\ell$, define iteratively function $\tilde{\varphi} : \Psi \times H \backslash \{\varnothing\} \to \Psi$ as follows. For $\ell = 1$, let $\tilde{\varphi}\left(\psi_0, h^1\right) = \varphi\left(\psi_0, a^1\right)$. Then, suppose that the map $\tilde{\varphi}\left(\psi_0, h^{\ell-1}\right)$ has been defined for every $h^{\ell-1} \in A^{\ell-1}$; hence let $\tilde{\varphi}\left(\psi_0, h^\ell\right) = \varphi\left(\tilde{\varphi}\left(\psi_0, h^{\ell-1}\right), a^\ell\right)$. So, for every player $i \in I$, we can specify the following strategy: $s_i(\varnothing) = \gamma(\psi_0)$ and, for every $\ell \geq 1$ and for every $h^\ell \in A^\ell$, let $s_i\left(h^\ell\right) = \gamma\left(\tilde{\varphi}\left(\psi_0, h^\ell\right)\right)$.

13.2 Multiplicity and Uniqueness of Equilibria

We prove three simple results relating the Nash equilibria of G to the subgame perfect equilibria of $\Gamma^{\delta,T}(G)$. We first apply the OD principle to show a rather obvious, but important result: playing repeatedly a Nash equilibrium of G is subgame perfect. Actually, playing any fixed sequence of Nash equilibria of G is subgame perfect (for example, alternating between the equilibrium preferred by Rowena and the equilibrium preferred by Colin in the repeated Battle of the Sexes is subgame perfect).

Theorem 50. *Let $\bar{z} = (\bar{a}^1, \bar{a}^2, ...) \in A^T$ be a sequence of Nash equilibria of G; then the strategy profile $\overline{s}$ that prescribes to play $\bar{a}^t$ in each stage* *t (irrespective of previous actions) is a subgame perfect equilibrium of* $\Gamma^{\delta,T}(G)$.

Proof. Let us show that there are no incentives to make one-shot deviations from $\bar{s}$. By the OD principle, this implies the thesis.

Fix an arbitrary stage t and a player i. The continuation-payoff of i if he chooses a_i in stage t and sticks to $\bar{s}_i$ afterward (assuming that everybody else sticks to $\bar{s}_{-i}$) is

$$v_i(a_i, \bar{a}^t_{-i}) + \sum_{\tau=t+1}^{T} \delta^{\tau-t} v_i(\bar{a}^\tau).$$

Note that the second term in this expression *does not depend on* a_i, because the equilibria of G played in the future may depend on calendar time, but do not depend on past actions. Since $\bar{a}^t$ is an equilibrium of G, for every a_i, $v_i(a_i, \bar{a}^t_{-i}) \leq v_i(\bar{a}^t_i, \bar{a}^t_{-i})$. Therefore i has no incentive to deviate from $\bar{a}^t_i$ in stage t. ■

By Theorem 50, any set of assumptions that guarantees the existence of a pure equilibrium of G also guarantees the existence of a subgame perfect equilibrium (in pure strategies) of $\Gamma^{\delta,T}(G)$.

The next result identifies situations in which a subgame perfect equilibrium *must* be the repetition of the Nash equilibrium of G.

Theorem 51. *Suppose that* G *is* finitely *repeated* ($T < \infty$) *and has a* unique *Nash equilibrium* a°, *then* $\Gamma^{\delta,T}(G)$ *has a* unique *subgame perfect equilibrium* s° *that prescribes to play* a° *always (for every* $h \in H$ *and for every* $i \in I$, ${s_i}^\circ(h) = {a_i}^\circ$*).*

For example, the finitely repeated Prisoners' Dilemma has a unique subgame perfect equilibrium: always defect.[3]

Proof. By Theorem 50, s° is a subgame perfect equilibrium. Let s be any subgame perfect equilibrium. We will show that $s = s^\circ$, that is,

[3] The finitely repeated Prisoners' Dilemma is a very simple game. Since the Prisoners' Dilemma has a dominant action equilibrium, the SPE can be obtained by backward induction. Another result that holds for the finitely repeated Prisoners' Dilemma is that, although it has many Nash equilibrium strategy profiles, they are all equivalent, and hence they all induce the permanent defection path.

for every $t \in \{0, ..., T-1\}$, for every $h \in A^t$, $s(h) = a^\circ$ (by convention, $A^0 = \{\varnothing\}$ is the set containing the empty sequence of action profiles). The result will be proved by induction on the number k of stages that still have to be played.

Basis step. Suppose that a history $h \in A^{T-1}$ has just occurred, i.e., the game reached the last stage. Since s is subgame perfect it must prescribe an equilibrium of G at the last stage. But there is only one equilibrium of G, a°; thus $s(h) = a^\circ$.

Inductive step. Suppose, by way of induction, that s prescribes a° in the last k periods ($k \geq 1$), that is, for every period $t \in \{T-k+1, ..., T\}$ and every history h' of length $t-1$ (that is, $h' \in A^{t-1}$), $s(h') = a^\circ$. We show that s prescribes a° in period $T-k$ for every history h of length $T-k-1$ ($h \in A^{T-k-1}$), $s(h) = a^\circ$. Since s is subgame perfect, it must be immune to deviations, in particular to one-shot deviations, in period $T-k$. Therefore, by the inductive hypothesis, for every $h \in A^{T-k-1}$, for every $i \in I$ and for every $a_i \in A_i$,

$$v_i(s(h)) + \sum_{\tau=1}^{k} \delta^\tau v_i(a^\circ) \geq v_i(a_i, s_{-i}(h)) + \sum_{\tau=1}^{k} \delta^\tau v_i(a^\circ).$$

Note that the summation term is the same on both sides of the inequality because the inductive hypothesis implies that if s is followed in the last k periods, future payoffs are independent of current actions. Therefore the action profile $s(h)$ is such that for every $i \in I$, for every $a_i \in A_i$, $v_i(s(h)) \geq v_i(a_i, s_{-i}(h))$, implying that $s(h)$ is a Nash equilibrium of G. But the only Nash equilibrium of G is a°; thus, $s(h) = a^\circ$. ■

Comments. First, there is no need to invoke the OD principle to prove the result above. The principle states that immunity to one-shot deviations implies immunity to all deviations, which holds under some (fairly general) assumptions. The proof above used the converse, immunity to all deviations implies immunity to one-shot deviations, which is trivially true by definition. Second, when a° is a profile of strictly *dominant* actions *in* G (or, more generally, when a° is the *unique rationalizable* profile *of* G), then the uniqueness result holds for continuation-rationalizability: $\hat{S}^\infty = \{s^\circ\}$. Furthermore, Theorem 46 implies that every profile of strongly rationalizable strategies of $\Gamma^{\delta,T}(G)$ induces path $z^\circ = (a^\circ, ..., a^\circ)$: $\zeta(S^\infty) = \{z^\circ\}$.

If one of the hypotheses of the previous theorem is removed, i.e., if either $T = \infty$ or G has multiple equilibria, then the thesis may fail. We illustrate this with an example and a general result.

Suppose that $T < \infty$ but G has multiple equilibria. Then the Basis Step of the proof above does not work anymore, because *the last stage equilibrium may depend on past actions*. Consider, for example, the following variation of the Prisoners' Dilemma game:

$1\backslash 2$	C	D	P
C	$4,4$	$0,5$	$-1,0$
D	$5,0$	$2,2$	$-1,0$
P	$0,-1$	$0,-1$	$0,0$

Figure 13.1: A modified Prisoners' Dilemma.

It can be checked that, for δ sufficiently large, the following strategies support the $(4,4)$ outcome in all periods but the last one:

- start playing C;
- if you are in stage $t < T$ and no deviation from (C, C) has occurred, play C;
- if you are in stage $t = T$ and no deviation from (C, C) has occurred, play D;
- if a deviation from (C, C) has occurred, play P in every stage until the last one.

The key observation is that in the second-to-last stage players do not defect because the immediate gain from defection is more than compensated by the future loss caused by a switch to a worse last-stage equilibrium.

Next suppose that $T = \infty$ and δ is "sufficiently high," then the repeated game may have many other equilibria beside the repetition of the stage-game equilibria. We show this for the case in which a stage-game equilibrium a° is strictly Pareto dominated by some other profile a^*.

Theorem 52. (Nash reversion, trigger strategy equilibria) *Let* $G = \langle I, (A_i, v_i)_{i \in I} \rangle$ *be a static game such that there are an equilibrium* a° *and an action profile* a^* *that strictly Pareto-dominates* a°*. Consider the repeated game* $\Gamma^{\delta,\infty}(G)$ *and the strategy profile* $s^* = (s_i^*)_{i \in I}$ *defined as follows:*

$$s_i^*(h) = \begin{cases} a_i^*, & \text{if } h = \varnothing \text{ or } h = (a^*, ..., a^*), \\ a_i{}^\circ, & \text{otherwise.} \end{cases}$$

Then s^* *is a subgame perfect equilibrium if and only if*

$$v_i(a^*) \geq (1 - \delta) \sup_{a_i \in A_i} v_i(a_i, a_{-i}^*) + \delta v_i(a^\circ), \tag{IC}$$

that is,

$$\delta \geq \frac{\sup_{a_i \in A_i} v_i(a_i, a_{-i}^*) - v_i(a^*)}{\sup_{a_i \in A_i} v_i(a_i, a_{-i}^*) - v_i(a^\circ)}.$$

Note that by definition $\sup_{a_i \in A_i} v_i(a_i, a_{-i}^*) \geq v_i(a^*)$ and by assumption $v_i(a^*) > v_i(a^\circ)$. Therefore

$$0 \leq \frac{\sup_{a_i \in A_i} v_i(a_i, a_{-i}^*) - v_i(a^*)}{\sup_{a_i \in A_i} v_i(a_i, a_{-i}^*) - v_i(a^\circ)} < 1.$$

Strategies such as those defined in the theorem are called **trigger strategies** because a deviation from the equilibrium path "triggers" some sort of punishment.

Proof of Theorem 52. We show that there are no incentives to make one-shot deviations if and only if condition (IC) holds. By the OD principle, this implies the result.

There are two types of histories: those for which a deviation from a^* has not yet occurred and the others. For the latter histories s^* prescribes to play a° forever after. Since a° is an equilibrium of G, Theorem 50 implies that there cannot be incentives to make one-shot deviations. Suppose now that no deviation from a^* has yet occurred. Then s^* prescribes to play a^*. Consider the incentives of player i. His continuation-payoff if he sticks to s_i^* is $v_i(a_i^*)$, because i expects to get $v_i(a_i^*)$ forever (recall that the continuation-payoff is a discounted time average). His continuation-payoff if he deviates and plays $a_i \neq a_i^*$ is $(1 - \delta)v_i(a_i, a_{-i}^*) + \delta v_i(a^\circ)$, because i expects that his opponents stick to a_{-i}^* in the current stage, and that his deviation triggers a switch to a° from the following stage (as prescribed by s^*). Therefore i does not have an incentive to make a one-shot deviation if and only if condition (IC) holds. ■

As an application of Theorem 52 consider, for example, an infinitely repeated Cournot oligopoly game: a is an output profile, $v_i(a)$ is the one-period profit of firm i, a° is the Cournot equilibrium profile, a^* is the joint-profit maximizing profile, i.e., the profile that would be chosen if the firms were allowed to merge and if the merged firm were not regulated. The trigger-strategy profile s^* may be interpreted as a collusive agreement to keep the price at the monopoly level. If δ is above the threshold the collusive agreement is self-enforcing. "Self-enforcing" in this context does not only mean that there are no incentive to deviate from the equilibrium path (all Nash equilibria satisfy this property), but also that the "punishment" or "war" that is supposed to take place after a deviation is itself self-enforcing, hence "credible," and this is indeed the case because playing the Cournot profile in each stage is an equilibrium (Theorem 50).

13.3 Characterization of the Equilibrium Set

In the modified Prisoners' Dilemma of Figure 13.1, players follow a quite intuitive punishment (or rewarding) code: if both players coordinate on (C, C) in the first $T-1$ stages, such behavior is rewarded by having players coordinating on the Pareto superior Nash equilibrium. Instead, if one of the players (or both) plays something other than C in some stage $t \leq T-1$, then players coordinate on the Pareto inferior Nash equilibrium (P, P) in every subsequent stage. Notice that the punishment code specified in the example is self-enforcing as it is based on players playing in the last stage a Nash equilibrium of the stage game. However, more generally, rewards and punishments can be supported as, in turn, players adjust their strategic behavior in subsequent rounds. We now provide a formal discussion of this issue in the context of infinitely repeated games.

We start making a simple observation that turns out to be extremely powerful: an infinitely repeated game exhibits a recursive structure. Indeed, once we rescale payoffs dividing by δ^ℓ, the subgame that starts after any history $h^\ell \in A^\ell$ is identical to the original repeated game. Therefore, a strategy profile of the whole game induces an SPE in the subgame with root h^ℓ if and only if it is an SPE in the original game. As a result, the set of SPE payoffs in the subgame with root h^ℓ is identical to the one of the original game.

Definition 74. *Let $X \subseteq \mathbb{R}^I$. Action profile $a^* = (a_i^*)_{i \in I}$ is **enforceable** on X, if there exists a function $\eta : A \to X$ such that for every $i \in I$ and every $a_i \in A_i$,*

$$(1-\delta) v_i(a^*) + \delta \eta_i(a^*) \geq (1-\delta) v_i\left(a_i, a_{-i}^*\right) + \delta \eta_i\left(a_i, a_{-i}^*\right),$$

where $\eta_i(\cdot)$ is the projection of function $\eta(\cdot)$ on the i-th dimension. With this, we say that function $\eta(\cdot)$ enforces a^* (on X) and we refer to $\eta(\cdot)$ as to the enforcing function.

In words, action profile a^* is enforceable on X if we can define "continuation-payoff" within subset X in a way that makes action a_i^* incentive compatible for every player i. Furthermore, we say that a payoff profile is pure-action decomposable on X if the payoff of each player can be decomposed in a (weighted) sum of the payoff associated with a profile a^* enforceable on X and of the associated enforcing function.

Definition 75. *Payoff profile $v^* \in \mathbb{R}^I$ is **pure-action decomposable** on $X \subseteq \mathbb{R}^I$ if there exist a profile a^* enforceable on X such that for every $i \in I$,*

$$v_i^* = (1-\delta) v_i(a^*) + \delta \eta_i(a^*),$$

where $\eta : A \to X$ enforces a^* on X.

If payoff profile v^* is pure-action decomposable, then it can be regarded as the equilibrium payoff profile of the static game $\hat{G} = \left\langle I, (A_i, \hat{v}_i)_{i \in I}\right\rangle$ obtained by modifying the payoff function of $G = \left\langle I, (A_i, v_i)_{i \in I}\right\rangle$ as follows:

$$\forall i,\ \forall a_i,\ \forall a_{-i},\ \hat{v}_i(a_i, a_{-i}) = (1-\delta) v_i(a_i, a_{-i}) + \delta \eta_i(a_i, a_{-i}).$$

The following definition introduces the important concept of self-generating set.

Definition 76. *A set $X \subseteq \mathbb{R}^I$ is **pure-action self-generating** if every payoff profile in X is pure-action decomposable on X.*

Thus, a set of payoff profiles X is pure-action self-generating if each payoff profile in such a set can be achieved by playing an action profile in the stage game and specifying continuation-payoffs that (i) enforce the action profile, and (ii) belong to the set X itself. Note that this last requirement exploits the recursive structure of the repeated game with perfect monitoring. Indeed, since continuation-payoffs belong to X, they can be pure-action decomposed as well with continuation-values in X. In turn, these second-degree continuation-payoffs can be further decomposed and so on.

Theorem 53. *Let $\Gamma^{\delta,\infty}(G)$ be a repeated game with perfect monitoring. Then the set W^* of pure SPE payoffs is the maximal (under set inclusion) pure-action self-generating set.*

Proof. The set of pure SPE payoff profiles is self-generating. Indeed, by definition, each payoff profile in such a set can be decomposed in a payoff profile $v(a^*)$ of the stage game and in a payoff profile $\eta(a^*)$ that emerges in the subgame that starts after a^*. By definition of subgame perfection, this payoff profile is also a pure SPE payoff profile.

Thus, we need to show that any pure-action self-generating set is a subset of W^*. Consider an arbitrary pure-action self-generating set X, and pick any $\hat{v} \in X$. We show that $\hat{v}$ is a pure SPE payoff profile, that is, $\hat{v} \in W^*$. Since $\hat{v}$ is pure-action decomposable on X, we can find an action profile $a^{\hat{v}}$ and a function $\eta^{\hat{v}} : A \to X$ such that $\hat{v} = (1-\delta)\, v\left(a^{\hat{v}}\right) + \delta \eta^{\hat{v}}\left(a^{\hat{v}}\right)$. Since $\eta^{\hat{v}}(a) \in X$ for every $a \in A$, we can further conclude that there exists an action profile $a^{\eta^{\hat{v}}(a)} \in A$ and a continuation-payoff function $\eta^{\eta^{\hat{v}}(a)} : A \to X$ such that

$$\eta^{\hat{v}}(a) = (1-\delta)\, v\left(a^{\eta^{\hat{v}}(a)}\right) + \delta \eta^{\eta^{\hat{v}}(a)}\left(a^{\eta^{\hat{v}}(a)}\right).$$

Iterating this procedure, we can define for each $t \in \mathbb{N}$ the following maps $\bar{\eta}^{\hat{v},t} : A^t \to X$. For every history $h^t = \left(a^1, a^2, ..., a^t\right) = \left(h^{t-1}, a^t\right) \in A^t$, define $\bar{\eta}^{\hat{v},1}\left(a^1\right) = \eta^{\hat{v}}\left(a^1\right)$. Next suppose that the map $\bar{\eta}^{\hat{v},t-1} : A^{t-1} \to X$ has been defined for $t \geq 2$; then define $\bar{\eta}^{\hat{v},t}\left(h^t\right) = \eta^{\bar{\eta}^{\hat{v},t-1}\left(h^{t-1}\right)}\left(a^t\right)$. Since the set X is self-generating, we have $\bar{\eta}^{\hat{v},t}\left(h^t\right) \in X$. Hence, for every $i \in I$, we can construct a strategy $s_i^{\hat{v}}$ for player i as follows: define $s_i^{\hat{v}}(\varnothing) = a_i^{\hat{v}}$ and $s_i^{\hat{v}}\left(h^t\right) = a_i^{\bar{\eta}^{\hat{v},t}\left(h^t\right)}$ for every $h^t \in A^t$ and $t \in \mathbb{N}$. By construction, strategy profile $\left(s_1^{\hat{v}}, ..., s_n^{\hat{v}}\right)$ induces payoff $\hat{v}$. Furthermore, after each history $h \in H$, the action profile specified by the strategy profile $\left(s_i^{\hat{v}}\right)_{i\in I}$ is enforceable on X. Thus the behavior of each player is incentive compatible given the enforcing function. We conclude that $\left(s_i^{\hat{v}}\right)_{i\in I}$ is an SPE profile. Thus, $\hat{v}$ is a pure SPE payoff profile. ■

Theorem 53, due to Abreu, Pearce and Stacchetti [2], provides a characterization of the set of pure SPE payoff profiles in terms of self-generating sets. Although such characterization sheds light on some key properties of SPE payoffs, it does not provide any guidance on how to build a strategy profile that induces a certain equilibrium-payoff profile.

An important result due to Abreu [1] helps in this respect. Loosely speaking, this result says that to check whether a certain payoff profile is compatible with equilibrium, we can restrict attention to strategy profiles that specify an "optimal penal code" for each player. This penal code is implemented whenever a player unilaterally deviates from the equilibrium behavior *independently of the specific history at which this deviation occurs* (joint deviations are instead ignored). Put differently, the scheme used to incentivize player i to abide equilibrium behavior is specified independently of the circumstances in which such deviation occurred. In particular, the same punishment is implemented regardless on whether the player's deviation takes place on or off the equilibrium path.

We start by defining simple strategy profiles.

Definition 77. *Fix a path* $\mathbf{a}(0) \in A^{\mathbb{N}}$ *and a profile of paths* $\mathbf{a}_I = (\mathbf{a}(i))_{i\in I} \in \left(A^{\mathbb{N}}\right)^I$. *The* ***simple strategy profile*** $s^{\mathbf{a}(0),\mathbf{a}_I}$ *is the strategy profile represented by the following automaton* $(\Psi, \psi_0, \gamma, \varphi)$*:*

$$\begin{aligned}
\Psi &= (\{0\} \cup I) \times \mathbb{N}_0, \\
\psi_0 &= (0,0), \\
\gamma(i,t) &= a^t(i), \\
\varphi((j,t),a) &= \begin{cases} (i,0), & \text{if } a_i \neq a_i^t(j) \text{ and } a_{-i} = a_{-i}^t(j), \\ (j,t+1), & \text{otherwise,} \end{cases}
\end{aligned}$$

where $a^t(i)$ *denotes the* t*-th element of* $\mathbf{a}(i)$.

A simple strategy profile can be understood as follows. Players start playing the action profiles according to $\mathbf{a}(0)$ and keep doing so as long as no unilateral deviation is observed. In particular, joint deviations by more than one player are ignored. If player i unilaterally deviates in some round t, namely if $a_i^t \neq a_i^t(0)$ whereas $a_{-i}^t = a_{-i}^t(0)$, then players "punish" player i by beginning to play the profiles in $\mathbf{a}(i)$, the "punishment protocol" for player i. Once players start playing a punishment protocol against player i, namely once they are playing according to $\mathbf{a}(i)$, they keep doing so unless a unilateral deviation from $\mathbf{a}(i)$ is observed (once more, joint deviations are ignored). If for some t', player $j \neq i$ unilaterally deviates from the protocol $\mathbf{a}(i)$—namely, if $a_j^{t'} \neq a_j^{t'}(i)$ whereas $a_{-j}^{t'} = a_{-j}^{t'}(i)$—then players exit punishment protocol $\mathbf{a}(i)$ and start playing the action profiles in protocol $\mathbf{a}(j)$. If the deviator is agent i, then the punishment protocol $\mathbf{a}(i)$ starts again from the beginning.

Profile $\mathbf{a}_I$ can thus be regarded as a *penal code*: it specifies, for each player $i \in I$, a punishment protocol $\mathbf{a}(i)$ to be implemented each time player i unilaterally deviates. In particular, the punishment protocol against a player is independent of the time at which the deviation occurs and of its nature.

Two important examples of simple strategy profiles are the *Nash-reversion trigger strategy* and the *Nash-reversion grim trigger strategy.*

Definition 78. *Fix an infinitely repeated game $\Gamma^{\delta,\infty}(G)$ and a path $\mathbf{a}(0) \in A^{\mathbb{N}}$. For every $i \in I$, let $\mathbf{a}^*(i)$ denote the constant path playing $a^*(i)$ in every period, where $a^*(i)$ is a Nash equilibrium of G, and consider the profile of paths $\mathbf{a}_I^* = (\mathbf{a}^*(i))_{i\in I}$.The simple strategy profile $s^{\mathbf{a}(0),\mathbf{a}_I^*}$ is called the* ***Nash-reversion trigger strategy*** *profile. Furthermore, if, for every $i \in I$, the Nash equilibrium $a^*(i)$ satisfies $v_i(a^*(i)) = \min_{a_{-i}\in A_{-i}} \max_{a_i \in A_i} v_i(a_i, a_{-i})$, then the simple strategy profile $s^{\mathbf{a}(0),\mathbf{a}_I^*}$ is called the* ***Nash-reversion grim trigger strategy*** *profile.*

In words, a Nash-reversion trigger strategy profile is a simple strategy profile in which if player $i \in I$ deviates, the punishment protocol recommends to play a Nash equilibrium of the stage game from there onwards. Moreover, if the Nash equilibrium also "minmaxes" player i (namely, it minimizes the utility of the deviator once we take into account that he is always able to react to the punishment by maximizing his own payoff), then we have a Nash-reversion grim trigger strategy profile.

Example 67. Let $\Gamma^{\delta,\infty}(G)$ be the infinitely repeated game with stage game G given by the Prisoners' Dilemma of Example 65. Consider strategy profile (s_1, s_2) such that, for every $i \in \{1, 2\}$,

$$s_i(h) = \begin{cases} C, & \text{if } h \in \{\varnothing\} \cup \left\{(a_1, a_2, ...) : \forall t \geq 1, a^t \in \{(C,C), (D,D)\}\right\}, \\ D, & \text{otherwise.} \end{cases}$$

This strategy profile is Nash-reversion grim trigger strategy (hence a trigger strategy) profile. ▲

Although trigger strategy profiles are easy to describe, they are not, in general, the best way to incentivize players. Indeed, harder punishments are possible if we give up the requirement that (i) the punishment must specify a constant action profile, and (ii) such action profile must be a Nash equilibrium of the stage game. Once we give up these two requirements, we can obtain other relevant concepts, such as *optimal penal codes*. Their importance relies in the fact that they represent the optimal way to incentivize players in a repeated game. As such, whenever a payoff profile is attainable in a pure SPE, it is also attainable with a simple strategy profile in which deviators are punished according to the optimal penal code.

Before providing a formal definition of optimal penal codes, it is useful to prove a preliminary result that has its own relevance.

Lemma 36. *Let $\Gamma^{\delta,\infty}(G)$ be a repeated game with perfect monitoring. Then the set W^* of pure SPE payoffs is compact.*

Proof. We provide the proof for the case in which the stage game G is finite. In this proof we use the concepts of enforceability and self-generation introduced above. The proof for the case where G is compact-continuous can be found in Osborne and Rubinstein [65, Lemma 153.3]. Since the utility of the players in the stage game is bounded (by finiteness of G) and the discount factor δ is strictly smaller than 1, the set W^* is bounded. So we have to show that W^* is closed. To this end, let cl(W^*) denote the closure of W^*. By definition of closure of a set, we have $W^* \subseteq$cl(W^*). We will show that cl(W^*) is pure-action self-generating; by Theorem 53, this will imply cl(W^*) $\subseteq W^*$, and so W^* =cl(W^*), as required. Pick any $\hat{v} \in$cl(W^*). We prove that $\hat{v}$ is pure-action decomposable with action profile a^* and enforcing function η^*. First note that, since cl(W^*) is a compact subset of a metric space, there exists a sequence $(\hat{v}_n)_{n\in\mathbb{N}} \in (W^*)^{\mathbb{N}}$ of payoff profiles converging to $\hat{v}$. Theorem 53 implies that each $\hat{v}_n$ is a pure-strategy SPE profile and it is pure-action decomposable on W^*. Thus, we can define a sequence of action profiles $(a_n)_{n\in\mathbb{N}}$ and a sequence of (continuous) functions $(\eta_n : A \to W^*)_{n\in\mathbb{N}}$ such that, for every $n \in \mathbb{N}$,

$$\hat{v}_n = (1-\delta)\, v(a_n) + \delta \eta_n(a_n).$$

Next note that $A = \times_{i\in I} A_i$ is finite, hence compact. Moreover $(\mathrm{cl}(W^*))^A$ is a compact metrizable space, and $\eta_n \in (\mathrm{cl}(W^*))^A$ for every $n \in \mathbb{N}$. It follows that $A \times (\mathrm{cl}(W^*))^A$ is compact, and $(a_n, \eta_n) \in A \times (\mathrm{cl}(W^*))^A$ for every $n \in \mathbb{N}$. Therefore we can find a subsequence $(a_{n_k}, \eta_{n_k})_{k\in\mathbb{N}}$ of $(a_n, \eta_n)_{n\in\mathbb{N}}$ such that, as $k \to \infty$, a_{n_k} converges to $a^* \in A$ and η_{n_k} converges to a (continuous) function $\eta^* \in (\mathrm{cl}(W^*))^A$. Furthermore, $v(\cdot)$ is a continuous function, and so $\lim_{k\to\infty} v(a_{n_k}) = v(a^*)$. Hence

$$\begin{aligned}\hat{v} &= \lim_{n\to\infty} \hat{v}_n = \lim_{k\to\infty} \hat{v}_{n_k} = \lim_{k\to\infty} \left((1-\delta)\, v(a_{n_k}) + \delta \eta_{n_k}(a_{n_k})\right) \\ &= (1-\delta)\, v(a^*) + \delta \eta^*(a^*).\end{aligned}$$

Therefore, $\hat{v}$ is pure-action decomposable with action profile a^* and enforcing function η^*. ■

We are now ready to define optimal penal codes and to prove that they exist.

Definition 79. *An **optimal penal code** is a profile of strategy profiles $\mathbf{s} \in S^I$ such that $\mathbf{s} = \left(s^{\mathbf{a}(i),\mathbf{a}_I}\right)_{i\in I}$ for some profile of paths $\mathbf{a}_I = (\mathbf{a}(i))_{i\in I}$, where each simple strategy profile $s^{\mathbf{a}(i),\mathbf{a}_I}$ ($i \in I$) is an SPE.*

To understand this definition, interpret each $\mathbf{a}(i)$ as paths "punishing" player i. Then, the i-th element of an optimal penal code is the strategy profile that starts "punishing" player i and, if any player j unilaterally deviates from such "punishment path," then "punishes" j with path $\mathbf{a}(j)$.

Proposition 7. *If repeated game $\Gamma^{\delta,\infty}(G)$ has a (pure) subgame perfect equilibrium, then it has an optimal penal code.*

Proof. If there is an SPE, then the set W^* of SPE payoff profiles is nonempty. It follows from Lemma 36 that, for every $i \in I$, the set W_i^* is nonempty and compact. Define $\underline{\mathbf{v}}_i^* = \min W_i^*$ and let $\mathbf{a}_I = (\mathbf{a}(i))_{i\in I}$ be a profile of infinite sequences of actions such that $u_i(\mathbf{a}(i)) = \underline{\mathbf{v}}_i^*$ for every

$i \in I$. Lemma 36 implies that for every $i \in I$ there exists an SPE $s[i]$ that induces paths $\mathbf{a}(i)$, with $u_i(\mathbf{a}(i)) = \underline{\mathbf{v}}_i^*$. We now show that, for every $i \in I$, the simple strategy $s^{\mathbf{a}(i),\mathbf{a}_I}$ is an SPE. To this end, recall that for all $i \in I$ and $t \in \mathbb{N}$,

$$u_i^t\left(a^1, a^2, ...,\right) = (1-\delta)\left(\sum_{m=t}^{\infty} \delta^{m-t} v_i\left(a^m\right)\right)$$

is the payoff of player i if path $\left(a^1, a^2, ...,\right)$ is played. So we have to show that for all $i, j \in I$, $t \in \mathbb{N}$ and $a_j \in A_j$,

$$u_j^t(\mathbf{a}(i)) \geq (1-\delta)\, v_j\left(a_j, a_{-j}^t(i)\right) + \delta u_j^0(\mathbf{a}(j)), \qquad (13.3.1)$$

where $a_{-j}^t(i)$ is the profile of actions of j's co-players at time t in sequence $\mathbf{a}(i)$. The SPE conditions imply that for all $j \in I$, $t \in \mathbb{N}$ and $a_j \in A_j$,

$$u_j^t(\mathbf{a}(i)) \geq (1-\delta)\, v_j\left(a_j, a_{-j}^t(i)\right) + \delta u_j^{d,t}\left(a_j, a_{-j}^t(i)\right),$$

where $u_j^{d,t}\left(a_j, a_{-j}^t(i)\right)$ is the continuation-payoff of player j if he unilaterally deviates from $\mathbf{a}(i)$ at stage t and plays action a_j. Since $s[i]$ is an SPE, we have that $u_j^{d,t}\left(a_j, a_{-j}^t(i)\right) \in W^*$ and consequently $u_j^{d,t}\left(a_j, a_{-j}^t(i)\right) \geq u_j^0(\mathbf{a}(j)) = \underline{\mathbf{v}}_j^*$. We conclude that (13.3.1) holds. ■

An immediate corollary of Proposition 7 is the following: Consider a simple strategy profile such that the punishment protocols are given by an optimal penal code and the initial path is given by $\mathbf{a}(0)$; then such simple strategy profile is self-supporting as long as $\mathbf{a}(0)$ can be supported when the punishment protocols are prescribed by the optimal penal code. The following result states that a sequence of action profiles in the repeated game can be supported in a pure SPE if and only if it can be supported through simple strategies in which an optimal penal code is used.

Proposition 8. *Fix a path* $\mathbf{a}(0) \in A^{\mathbb{N}}$ *and a profile of paths* $\mathbf{a}_I \in \left(A^{\mathbb{N}}\right)^I$ *such that* $\mathbf{s} = \left(s^{\mathbf{a}(i),\mathbf{a}_I}\right)_{i \in I}$ *is an optimal penal code. There exists an SPE* $s[0]$ *that induces* $\mathbf{a}(0)$ *if and only if the simple strategy profile* $s^{\mathbf{a}(0),\mathbf{a}_I^*}$ *is an SPE.*

Proof. The "if" direction is immediate. To prove the "only if" direction, observe that, by assumption, there exist an SPE $s[0]$ that induces $\mathbf{a}(0)$. Then, we can replicate the proof of Proposition 7 with $s[0]$ replacing $s[i]$ and prove the result. ∎

14

Bargaining

Bargaining is a pervasive economic phenomenon. Many relevant economic settings involve the confrontation between two or more agents with conflicting interests. Examples abound. A seller and a buyer may bargain over the terms of trade; unions can bargain with entrepreneurs to attain a more favorable split of the total surplus; different political parties may bargain on the allocation of the national budget.

The goal of this Chapter is to introduce the workhorse model used in economics to analyze bargaining environments, namely *Rubinstein's Bargaining Game* (Rubinstein [72]).[1] In line with this goal, we first analyze a simpler setting that, strictly speaking, does not involve any actual bargaining, the *Ultimatum Game*. The analysis of the Ultimatum Game will highlight some insights which hold also in more complicated settings. Then, we will introduce some flavor of bargaining by considering a twice repeated Ultimatum Game. Finally, we will proceed with the analysis of the Rubinstein's game and we will characterize the set of its equilibria.

14.1 The Ultimatum Game

Two agents, Ann (A) and Bob (B), have to agree on the split of \$1, or a "pie" (surplus) of size 1. Ann moves first and proposes a feasible payoff

[1] See also chapter 7 in Osborne and Rubinstein [65].

split, namely a pair of non-negative numbers summing up to 1.[2] Formally, define

$$X = \left\{ (x_{\mathrm{A}}, x_{\mathrm{B}}) \in [0,1]^{\{\mathrm{A},\mathrm{B}\}} : x_{\mathrm{A}} + x_{\mathrm{B}} = 1 \right\}.$$

Ann proposes $x = (x_{\mathrm{A}}, x_{\mathrm{B}}) \in X$, where x_{B} is the share offered to Bob and $x_{\mathrm{A}} = 1 - x_{\mathrm{B}}$ is the share that Ann demands for herself. Upon observing the offer, Bob can either accept (action y) or reject (action n). If Bob accepts, the proposal is implemented and the game ends. If instead Bob rejects, a default exogenous split $(\bar{x}_{\mathrm{A}}, \bar{x}_{\mathrm{B}}) \in X$ is implemented with a *delay* of one period and the game ends.[3] Delay captures the idea that disagreement is costly and Pareto inefficient. We also assume that each default share is strictly positive, that is, $0 < \bar{x}_{\mathrm{A}} < 1$. This describes the rules of the game, also called *bargaining protocol*. As regards immediate agreements, players are expected utility maximizers with linear vNM utility indexes over the total amount of money they receive. But players are impatient in the following sense: When comparing x in the current period to $\bar{x}$ in the following period, each player i weakly (respectively, strictly) prefers x immediately if and only if $x_i \geq \delta \bar{x}_i$ (respectively, $x_i > \delta \bar{x}_i$), where $\delta \in (0,1)$ is the discount factor.[4] The bargaining protocol and players' preferences over (lotteries of) consequences are common knowledge.

The environment we just described is a two-stage game with complete and perfect information. Using our general notation for multistage games, we have $\Gamma = \langle I, Y, g, (A_i, \mathcal{A}_i(\cdot), v_i)_{i \in I} \rangle$ with the following elements:

- $A_{\mathrm{A}} = X \cup \{\mathrm{wait}\}$, $A_{\mathrm{B}} = \{\mathrm{y}, \mathrm{n}, \mathrm{wait}\}$,
- $\mathcal{A}_{\mathrm{A}}(\varnothing) = X$, $\mathcal{A}_{\mathrm{B}}(\varnothing) = \{\mathrm{wait}\}$, $\mathcal{A}_{\mathrm{A}}((x, \mathrm{wait})) = \{\mathrm{wait}\}$, $\mathcal{A}_{\mathrm{B}}((x, \mathrm{wait})) = \{\mathrm{y}, \mathrm{n}\}$,[5]

[2] More generally, the sum has to be at most 1. Here we apply the resource constraint with equality to simplify the analysis, as this does not affect the equilibria.

[3] More generally, the default shares satisfy $\bar{x}_{\mathrm{A}} \geq 0$, $\bar{x}_{\mathrm{B}} \geq 0$, and $\bar{x}_{\mathrm{A}} + \bar{x}_{\mathrm{B}} \leq 1$.

[4] Under risk neutrality, one can equivalently assume that there is no pure time-discounting, but in case of rejection of the proposal the pie is destroyed with probability $1-\delta$. Furthermore, all the statements and arguments given here extend seamlessly to the more general case of players with different discount factors. We omit this generalization for notational simplicity.

[5] Recall that $\varnothing$ denotes the sequence of length zero, i.e., the root of the game, or empty history.

- $Y = X \times \{1, 2\}$ is the set of dated splits, where (x, t) represents a split x implemented at date t,[6]
- $g\left(\left(x, \text{wait}\right), \left(\text{wait}, \text{y}\right)\right) = (x, 1)$, $g\left(\left(x, \text{wait}\right), \left(\text{wait}, \text{n}\right)\right) = (\bar{x}, 2)$ for all $x \in X$,
- $v_i(x, t) = \delta^{t-1} x_i$ for all $(x, t) \in Y$.

The game features complete information because all of the above is common knowledge, and it also features perfect information because players perfectly observe past histories and there is only one active player at each nonterminal history. We maintain such assumptions about the information structure hold for all the bargaining games analyzed here. Henceforth, we ease notation and identify histories with the sequence of actions taken by the active players, e.g., we write (x,n) instead of $\left(\left(x, \text{wait}\right), \left(\text{wait}, \text{n}\right)\right)$.

Before we analyze the Ultimatum Game as defined above, it is useful to consider an *approximation with finite action spaces.* For any $k \in \mathbb{N}$, let

$$X_k = \left\{ x \in X : x_{\text{A}} \in \left\{ 0, \frac{1}{k}, ..., \frac{k-1}{k}, 1 \right\} \right\}.$$

For example, if $k = 100$, the proposer can only offer to the respondent 0%, 1%, 2%, ..., 99%, or 100% of the pie. Let the corresponding game be denoted by $\Gamma^{(k)}$. In what follows, we assume k to be "large enough," that is, we assume the finite *grid of possible offers* to be *sufficiently fine* to obtain equilibria that approximate the equilibrium of the game with a continuum of offers.

To analyze the game, we have to identify the **minimally acceptable offer** for Bob, considering that a rejection is worth $\delta \bar{x}_{\text{B}}$ to him:

$$x_{\text{B}}^{(k)} := \min_{x \in X_k} \left\{ x_{\text{B}} : x_{\text{B}} \geq \delta \bar{x}_{\text{B}} \right\}.$$

By definition, $\delta \bar{x}_{\text{B}} \leq x_{\text{B}}^{(k)} < \delta \bar{x}_{\text{B}} + 1/k$, with $\delta \bar{x}_{\text{B}} = x_{\text{B}}^{(k)}$ if $\delta \bar{x}_{\text{B}} \in \left\{ 0, \frac{1}{k}, ..., \frac{k-1}{k}, 1 \right\}$. Thus, $\lim_{k \to \infty} x_{\text{B}}^{(k)} = \delta \bar{x}_{\text{B}}$, which implies that $x_{\text{B}}^{(k)} - \delta \bar{x}_{\text{B}} < 1 - \delta$ for k large enough. There are two possibilities, either

[6] Recall that Y denotes the space of possible outcomes, or consequences. Do not confuse Y with "yes."

$\delta\bar{x}_{\mathrm{B}} \in \left\{0, \frac{1}{k}, ..., \frac{k-1}{k}, 1\right\}$, or not; in the latter case, $x_{\mathrm{B}}^{(k)} > \delta\bar{x}_{\mathrm{B}}$. If $x_{\mathrm{B}}^{(k)} > \delta\bar{x}_{\mathrm{B}}$, Bob is strictly better off accepting offer $x^{(k)} = \left(1 - x_{\mathrm{B}}^{(k)}, x_{\mathrm{B}}^{(k)}\right)$, and we can use backward induction to find the unique subgame perfect equilibrium (SPE hereafter): If $x_{\mathrm{A}}^{(k)} > \delta\bar{x}_{\mathrm{A}}$ (i.e., $x_{\mathrm{B}}^{(k)} - \delta\bar{x}_{\mathrm{B}} < 1 - \delta$, which holds if $k > 1/(1-\delta)$), Ann offers $x^{(k)}$ and Bob accepts an offer $x \in X_k$ if and only if $x_{\mathrm{B}} \geq x_{\mathrm{B}}^{(k)}$ (Ann is better off making Bob accept this offer if k is sufficiently large). If instead $x_{\mathrm{B}}^{(k)} = \delta\bar{x}_{\mathrm{B}}$, Bob is indifferent between accepting and rejecting the minimally acceptable offer, because the cost of delay exactly offsets the immediate consumption of a smaller share. Therefore, we have to use "case-by-case" backward induction to find the SPEs:

Case 1 When indifferent, Bob accepts; anticipating this,[7] Ann offers $x^{(k)} = \left(1 - x_{\mathrm{B}}^{(k)}, x_{\mathrm{B}}^{(k)}\right)$; Bob accepts offer $x = (x_{\mathrm{A}}, x_{\mathrm{B}})$ if and only if $x_{\mathrm{B}} \geq x_{\mathrm{B}}^{(k)}$.

Case 2 When indifferent, Bob rejects; anticipating this, Ann has to offer slightly more to make him accept. Assuming that $k > 1/(1-\delta)$ (k large) Ann is better off by offering $\left(1 - x_{\mathrm{B}}^{(k)} - \frac{1}{k}, x_{\mathrm{B}}^{(k)} + \frac{1}{k}\right)$ and make Bob accept. Indeed, $x_{\mathrm{B}}^{(k)} = \delta\bar{x}_{\mathrm{B}} = \delta(1 - \bar{x}_{\mathrm{A}})$ and $1 - \delta(1 - \bar{x}_{\mathrm{A}}) - \frac{1}{k} > \delta\bar{x}_{\mathrm{A}}$ if $k > 1/(1-\delta)$; thus, the minimal offer that Bob is willing to accept gives Ann more than the present values of her default share. To summarize, we have:

Proposition 9. *Suppose that k is large enough, that is, $k > 1/(1-\delta)$. Then, the following is a pure strategy SPE of $\Gamma^{(k)}$: $s_{\mathrm{A}}(\varnothing) = \left(1 - x_{\mathrm{B}}^{(k)}, x_{\mathrm{B}}^{(k)}\right)$ and $s_{\mathrm{B}}(x) = y$ if and only if $x_{\mathrm{B}} \geq x_{\mathrm{B}}^{(k)}$; this is the* unique *SPE if $x_{\mathrm{B}}^{(k)} > \delta\bar{x}_{\mathrm{B}}$. If $x_{\mathrm{B}}^{(k)} = \delta\bar{x}_{\mathrm{B}}$, there is another pure strategy SPE of $\Gamma^{(k)}$: $s_{\mathrm{A}}(\varnothing) = \left(1 - x_{\mathrm{B}}^{(k)} - \frac{1}{k}, x_{\mathrm{B}}^{(k)} + \frac{1}{k}\right)$ and $s_{\mathrm{B}}(x) = y$ if and only if $x_{\mathrm{B}} > x_{\mathrm{B}}^{(k)}$.*[8]

[7] Recall that, in a subgame perfect equilibrium, Ann's conjecture about Bob is correct.

[8] If the equilibrium is unique, it coincides with the unique rationalizable strategy pair, for any multistage version of the rationalizability solution concept. The same holds for backward rationalizability in finite-horizon bargaining games with a finite grid of offers. Furthermore, by Theorem 46, the induced path is the same for every strongly rationalizable strategy pair.

If $x_{\mathrm{B}}^{(k)} = \delta\bar{x}_{\mathrm{B}}$, there is also a continuum of SPEs in partially randomized behavior strategies, we omit the details. As $k \to \infty$ the grid of possible offers approximates the continuum ever more finely, $x_{\mathrm{B}}^{(k)} \to \delta\bar{x}_{\mathrm{B}}$ and all the strategy pairs in the SPE set converge to the pure equilibrium whereby Ann makes the least acceptable offer and Bob says Yes. It turns out that this is the unique subgame perfect equilibrium of Γ, the Ultimatum Game with a continuum of offers, even if we allow for randomized strategies.[9]

Proposition 10. *The Ultimatum Game has a unique SPE* $(s_{\mathrm{A}}, s_{\mathrm{B}})$, *specifically:*

$$s_{\mathrm{A}}(\varnothing) = (1 - \delta\bar{x}_{\mathrm{B}}, \delta\bar{x}_{\mathrm{B}}), \; s_{\mathrm{B}}(x_{\mathrm{A}}, x_{\mathrm{B}}) = \begin{cases} \mathrm{y}, & \textit{if } x_{\mathrm{B}} \geq \delta\bar{x}_{\mathrm{B}}, \\ \mathrm{n}, & \textit{if } x_{\mathrm{B}} < \delta\bar{x}_{\mathrm{B}}. \end{cases}$$

Proof. Let $(\beta_{\mathrm{A}}, \beta_{\mathrm{B}})$ be an SPE in behavior strategies. We will proceed backwards. Faced with an offer $(x_{\mathrm{A}}, x_{\mathrm{B}})$, if Bob accepts this offer, he gets x_{B} immediately. If he rejects, he gets $\bar{x}_{\mathrm{B}}$ one period later, which is worth $\delta\bar{x}_{\mathrm{B}}$. Thus, rationality implies that he replies y if $x_{\mathrm{B}} > \delta\bar{x}_{\mathrm{B}}$, n if $x_{\mathrm{B}} < \delta\bar{x}_{\mathrm{B}}$ and either y or n if $x_{\mathrm{B}} = \delta\bar{x}_{\mathrm{B}}$. As a result, the rejection probability $\beta_{\mathrm{B}}(\mathrm{n}|x)$ may be different from 0 and 1 only if $x_{\mathrm{B}} = \delta\bar{x}_{\mathrm{B}}$. To ease notation, let $\rho = \beta_{\mathrm{B}}(\mathrm{n}|(1 - \delta\bar{x}_{\mathrm{B}}, \delta\bar{x}_{\mathrm{B}}))$ denote this rejection probability. With this, the expected payoff of Ann when she proposes $(x_{\mathrm{A}}, x_{\mathrm{B}})$ is

$$\mathbb{E}_{\beta_{\mathrm{B}}}(v_{\mathrm{A}}|(1 - x_{\mathrm{B}}, x_{\mathrm{B}})) = \begin{cases} \delta\bar{x}_{\mathrm{A}}, & \text{if } x_{\mathrm{B}} < \delta\bar{x}_{\mathrm{B}}, \\ \rho\delta\bar{x}_{\mathrm{A}} + (1-\rho)(1 - \delta\bar{x}_{\mathrm{B}}), & \text{if } x_{\mathrm{B}} = \delta\bar{x}_{\mathrm{B}}, \\ 1 - x_{\mathrm{B}}, & \text{if } x_{\mathrm{B}} > \delta\bar{x}_{\mathrm{B}}. \end{cases}$$

Now note that

$$1 - \delta\bar{x}_{\mathrm{B}} > \delta(1 - \bar{x}_{\mathrm{B}}) = \delta\bar{x}_{\mathrm{A}},$$

which implies that $\mathbb{E}_{\beta_{\mathrm{B}}}(v_{\mathrm{A}}|(1 - x_{\mathrm{B}}, x_{\mathrm{B}}))$ is discontinuous at $x_{\mathrm{B}} = \delta\bar{x}_{\mathrm{B}}$, with

$$\lim_{x_{\mathrm{B}} \searrow \delta\bar{x}_{\mathrm{B}}} \mathbb{E}_{\beta_{\mathrm{B}}}[v_{\mathrm{A}}|(1 - x_{\mathrm{B}}, x_{\mathrm{B}})] = 1 - \delta\bar{x}_{\mathrm{B}} > \delta\bar{x}_{\mathrm{A}} = \lim_{x_{\mathrm{B}} \nearrow \delta\bar{x}_{\mathrm{B}}} \mathbb{E}_{\beta_{\mathrm{B}}}(v_{\mathrm{A}}|(1 - x_{\mathrm{B}}, x_{\mathrm{B}})).$$

If $\rho > 0$ then

$$1 - \delta\bar{x}_{\mathrm{B}} > \rho\delta\bar{x}_{\mathrm{A}} + (1-\rho)(1 - \delta\bar{x}_{\mathrm{B}}).$$

[9] From now, when we claim that a bargaining game has a unique equilibrium, we mean unique among all the SPEs in behavior strategies, which include those in pure strategies. It will turn out that the unique equilibrium is always in pure strategies.

If $\rho = 0$, then $\mathbb{E}_{\beta_{\mathrm{B}}}(v_{\mathrm{A}}|\,(1 - x_{\mathrm{B}}, x_{\mathrm{B}}))$ attains the supremum $1 - \delta\bar{x}_{\mathrm{B}}$, which is therefore a maximum, at offer $(1 - \delta\bar{x}_{\mathrm{B}}, \delta\bar{x}_{\mathrm{B}})$. Hence, the candidate equilibrium of the statement is indeed a (pure) SPE. If instead $\rho > 0$, then payoff $1 - \delta\bar{x}_{\mathrm{B}}$ can be approached arbitrarily closely, but not attained by $\mathbb{E}_{\beta_{\mathrm{B}}}(v_{\mathrm{A}}|\,(1 - x_{\mathrm{B}}, x_{\mathrm{B}}))$, and Ann does not have a best reply. This shows that the candidate equilibrium is the unique SPE. ■

For future reference, we want to stress a few key implications of Proposition 10. First, the equilibrium is *unique*. In particular, in such equilibrium the respondent is kept at the present value of his disagreement outcome, $\delta\bar{x}_{\mathrm{B}}$, while the proposer extracts all the remaining surplus, $1-\delta\bar{x}_{\mathrm{B}}$. Second, the equilibrium payoff is determined by the disagreement share of the (last-stage) respondent or equivalently by the value he can obtain by rejecting the offer coming from the proposer. Third, the equilibrium outcome is Pareto efficient, namely the disagreement outcome $((1 - \delta\bar{x}_{\mathrm{B}}, \delta\bar{x}_{\mathrm{B}}), 2)$, which is inefficient because consumption is delayed, never arises. In the rest of the chapter we show that these insights hold true even if Bob can make counteroffers and bargaining lasts for (infinitely) many periods.

14.2 2-Period Alternating Offer Game

In the Ultimatum Game, Bob can only accept or reject the offer coming from Ann and he is prevented from proposing counteroffers. To allow for the latter, we extend the Ultimatum Game by adding a subgame following the rejection of Bob. In this subgame, the interaction between the two agents is a new Ultimatum Game in which the roles of the two agents are switched. Thus, Bob becomes the proposer and Ann the respondent. We call this game **2-Period Alternating Offer Game**.

As before, suppose that Ann and Bob have to split \$1. The bargaining protocol is described as follows:[10]

1.P at the beginning of the game Ann makes a proposal $x^1 = \left(x^1_{\mathrm{A}}, x^1_{\mathrm{B}}\right) \in X$;

[10] We leave to the reader the exercise of representing this bargaining protocol with the general mathematical notation for dynamic games.

1.R Bob can either accept (y) or reject (n) offer x^1; if Bob accepts, the proposed split x^1 is implemented in period 1 and the game ends; if Bob rejects, the game moves to period 2;

2.P if the play moves to period 2, Bob makes a proposal $x^2 = \left(x^2_{\mathrm{A}}, x^2_{\mathrm{B}}\right) \in X$;

2.R Ann can either accept (y) or reject (n) offer x^2; if Ann accepts, the proposed split x^2 is implemented in period 2; if Ann rejects the disagreement split $(\bar{x}_{\mathrm{A}}, \bar{x}_{\mathrm{B}})$ is implemented in period 3 and the game ends.

The next proposition states that also the 2-Period Alternating Offer Game has a unique subgame perfect equilibrium and provides its characterization.

Proposition 11. *The 2-Period Alternating Offer Game has a unique SPE* $(s_{\mathrm{A}}, s_{\mathrm{B}})$ *determined as follows: for every* $(x^1, x^2) \in X \times X$

$$s_{\mathrm{A}}(\varnothing) = (1 - \delta(1 - \delta\bar{x}_{\mathrm{A}}), \delta(1 - \delta\bar{x}_{\mathrm{A}})), \ s_{\mathrm{B}}\left(x^1\right) = \begin{cases} \mathrm{y}, & \text{if } x^1_{\mathrm{B}} \geq \delta(1 - \delta\bar{x}_{\mathrm{A}}), \\ \mathrm{n}, & \text{if } x^1_{\mathrm{B}} < \delta(1 - \delta\bar{x}_{\mathrm{A}}), \end{cases}$$

and

$$s_{\mathrm{B}}\left(x^1, \mathrm{n}\right) = (\delta\bar{x}_{\mathrm{A}}, 1 - \delta\bar{x}_{\mathrm{A}}), \ s_{\mathrm{A}}\left(x^1, \mathrm{n}, x^2\right) = \begin{cases} \mathrm{y}, & \text{if } x^2_{\mathrm{A}} \geq \delta\bar{x}_{\mathrm{A}}, \\ \mathrm{n}, & \text{if } x^2_{\mathrm{A}} < \delta\bar{x}_{\mathrm{A}}. \end{cases}$$

Proof. As in the Ultimatum Game with a continuum of offers, if the proposer anticipates that—when indifferent—the respondent says No (with positive probability), then the reduced-form payoff function of the proposer cannot be maximized. Thus, the proposer has a best response to the rational strategy of the respondent only if the respondent accepts when indifferent. This allows to solve for the unique SPE by a kind of modified backward induction procedure, whereby the respondent breaks ties in favour of the proposer.

In particular, the analysis of the second-period game, is the same as for the Ultimatum Game with reversed roles for Bob (proposer) and Ann (respondent). This explains the second-period strategies. With this, we obtain a reduced-form Ultimatum Game where the disagreement split implemented with one-period delay in case of rejection is the equilibrium offer of the second period $(\hat{x}_{\mathrm{A}}, \hat{x}_{\mathrm{B}}) = (\delta \bar{x}_{\mathrm{A}}, 1 - \delta \bar{x}_{\mathrm{A}})$. The solution of the Ultimatum Game with (delayed) disagreement split $(\hat{x}_{\mathrm{A}}, \hat{x}_{\mathrm{B}})$ is that Bob's response strategy is

$$\begin{aligned} s_{\mathrm{B}}\left(x^{1}\right) &= \begin{cases} \mathrm{y}, & \text{if } x_{\mathrm{B}}^{1} \geq \delta \hat{x}_{\mathrm{B}}, \\ \mathrm{n}, & \text{if } x_{\mathrm{B}}^{1} < \delta \hat{x}_{\mathrm{B}}, \end{cases} \\ &= \begin{cases} \mathrm{y}, & \text{if } x_{\mathrm{B}}^{1} \geq \delta\left(1 - \delta \bar{x}_{\mathrm{A}}\right), \\ \mathrm{n}, & \text{if } x_{\mathrm{B}}^{1} < \delta\left(1 - \delta \bar{x}_{\mathrm{A}}\right), \end{cases} \end{aligned}$$

and Ann offers

$$s_{\mathrm{A}}(\varnothing) = (1 - \delta \hat{x}_{\mathrm{B}}, \delta \hat{x}_{\mathrm{B}}) = \left(1 - \delta\left(1 - \delta \bar{x}_{\mathrm{A}}\right), \delta\left(1 - \delta \bar{x}_{\mathrm{A}}\right)\right).$$

■

Proposition 11 and its proof show that the equilibrium implications of the 2-period model are similar to those of the 1-period model (Ultimatum Game) stated in Proposition 10. First, the equilibrium is unique. Second, the equilibrium outcome[11] is determined by the disagreement share of the first-round respondent, $\hat{x}_{\mathrm{B}}$. However, differently from the Ultimatum Game, the disagreement share $\hat{x}_{\mathrm{B}}$ is endogenous, as it takes into account that if Bob rejects the offer, he becomes the new proposer and, in equilibrium, he "forces" Ann to accept share $\delta \bar{x}_{\mathrm{A}}$. Third, the equilibrium outcome is Pareto efficient, because costly disagreement (rejection and delay) does not occur. Finally, we observe that the strategy of the second-period respondent depends only on the last offer, i.e., it is independent of what happened in the first round.

14.3 Bargaining with Infinite Horizon

The analysis developed in the previous Sections can be easily extended to n-Periods Alternating Offer Games. For this reason, in this section we will jump to the case in which the bargaining game has an infinite horizon and agents keep alternating in the role of proposer and respondent until an agreement is reached. This setting has been studied in an extremely influential paper by Ariel Rubinstein (Rubinstein [72]) and it has become a building block of several applied works.

This is the bargaining protocol:

[11] That is, split $\left(1 - \delta\left(1 - \delta \bar{x}_{\mathrm{A}}\right), \delta\left(1 - \delta \bar{x}_{\mathrm{A}}\right)\right)$ in period 1.

- (*First-Step Rule*) In period 1 Ann can offer any split $x = (x_{\mathrm{A}}, x_{\mathrm{B}}) \in X$. Bob can accept (y) or reject (n). If Bob accepts offer x, the agreement is immediately implemented, i.e., the outcome (consequence) is $(x, 1)$; if Bob rejects, the play moves to the following period.

- (*Inductive-Step Rule*) If no agreement was reached before period $t > 1$, the player who rejected the offer in period $t-1$ becomes the proposer in period t. The proposer can offer any split $x \in X$. If offer x is accepted, the agreement is immediately implemented and the outcome is (x, t); if it is rejected the play moves to period $t+1$.

Thus, in the Rubinstein's bargaining game, Ann and Bob keep switching roles until an agreement is reached. In particular, Ann plays in the role of the proposer in period $t \in \mathbb{N}$ if and only if t is odd, and Bob plays in the role of proposer in period $t \in \mathbb{N}$ if and only if t is even.

Formally, this strategic environment can be represented as an infinite-horizon game with perfect information. In particular,[12]

- The set of **nonterminal histories** is the union $H = H_P \cup H_R$ of the set of histories where a proposer is active,

$$H_P := \bigcup_{t \in \mathbb{N}} (X \times \{\mathrm{n}\})^{t-1},$$

and the set of histories where a respondent is active

$$H_R := \bigcup_{t \in \mathbb{N}} \left((X \times \{\mathrm{n}\})^{t-1} \times X \right).$$

- For all $t \in \mathbb{N}$, $h \in (X \times \{\mathrm{n}\})^{t-1}$ and $x \in X$, $\mathcal{A}_{\mathrm{A}}(h) = X$, $\mathcal{A}_{\mathrm{B}}(h) = \{\mathrm{wait}\} = \mathcal{A}_{\mathrm{A}}(h, x)$ and $\mathcal{A}_{\mathrm{B}}(h, x) = \{\mathrm{y,n}\}$ if t is odd, whereas $\mathcal{A}_{\mathrm{B}}(h) = X$, $\mathcal{A}_{\mathrm{A}}(h) = \{\mathrm{wait}\} = \mathcal{A}_{\mathrm{B}}(h, x)$ and $\mathcal{A}_{\mathrm{A}}(h, x) = \{\mathrm{y,n}\}$ if t is even.

- The set of **finite terminal histories**—or agreement histories—is

$$Z_{\mathrm{y}} := \bigcup_{t \in \mathbb{N}} \left((X \times \{\mathrm{n}\})^{t-1} \times (X \times \{\mathrm{y}\}) \right).$$

[12] Recall that $\mathbb{N} = \{1, 2, ...\}$ and that we adopt the convention $(X \times \{\mathrm{n}\})^0 = \{\varnothing\}$, where $\varnothing$ is the empty history.

- The set of **infinite terminal histories**—or permanent disagreement histories—is

$$Z_{\mathrm{n}} := (X \times \{\mathrm{n}\})^{\infty},$$

the set of terminal histories is therefore $Z := Z_{\mathrm{y}} \cup Z_{\mathrm{n}}$.

- The **outcome** (or consequence) **space** is $Y = (X \times \mathbb{N}) \cup \{D\}$, where (x,t) means that agreement on x is reached in period t, and D conveniently denotes the **permanent disagreement** outcome; the **outcome** (or consequence) **function** $g : Z \to Y$ is

$$g(z) = \begin{cases} (x,t), & \text{if } z = (h, x, \mathrm{y}) \in (X \times \{\mathrm{n}\})^{t-1} \times (X \times \{\mathrm{y}\}), \\ D, & \text{if } z \in Z_{\mathrm{n}}. \end{cases}$$

- The **utility function** $v_i : Y \to \mathbb{R}$ of each player i is

$$\forall (x,t) \in (X \times \mathbb{N}), \quad v_i(x,t) = \delta^{t-1} x_i,$$
$$v_i(D) = 0.$$

The implied **payoff function** $u_i = v_i \circ g : Z \to \mathbb{R}$ of each player $i \in \{\mathrm{A}, \mathrm{B}\}$ is given by

$$u_i(z) = \begin{cases} \delta^{t-1} x_i, & \text{if } z = (h, x, \mathrm{y}) \in (X \times \{\mathrm{n}\})^{t-1} \times (X \times \{\mathrm{y}\}), \\ 0, & \text{if } z \in Z_{\mathrm{n}}. \end{cases}$$

Note that u_i satisfies the property of **continuity at infinity**: for every $\varepsilon > 0$ there exists a sufficiently large t, such that if two terminal histories z' and z'' have a t-period common prefix, then $|u_i(z') - u_i(z'')| \leq \varepsilon$. Indeed, for every $t \in \mathbb{N}$, every nonterminal history $h \in (X \times \{\mathrm{n}\})^t$, and every pair of terminal continuations of h, viz. $z' = (h, ...) \in Z$ and $z'' = (h, ...) \in Z$, we have $|u_i(z') - u_i(z'')| \leq \delta^t$. This implies that the *OD principle* holds: a strategy s_i is sequentially optimal given conjecture β^i (possibly a deterministic conjecture s_{-i}) if and only if s_i is one-step optimal given β^i.[13]

In the following analysis, it is convenient to partition H into four subsets according to two binary criteria: whether A or B is the active player, and whether the active player is a proposer or a respondent. We

[13]See Section 10.5.

let $H_{i,P}$ (respectively $H_{i,R}$) denote the set of histories where $i \in \{\text{A}, \text{B}\}$ is a proposer (respectively, a respondent). Thus

$$H = H_{\text{A},P} \cup H_{\text{B},R} \cup H_{\text{B},P} \cup H_{\text{A},R},$$

with

$$\begin{aligned}
H_{\text{A},P} &= \bigcup_{t \in \{1,3,5,...\}} (X \times \{\text{n}\})^{t-1}, \\
H_{\text{B},R} &= \bigcup_{t \in \{1,3,5,...\}} (X \times \{\text{n}\})^{t-1} \times X, \\
H_{\text{B},P} &= \bigcup_{t \in \{2,4,6,...\}} (X \times \{\text{n}\})^{t-1}, \\
H_{\text{A},R} &= \bigcup_{t \in \{2,4,6,...\}} (X \times \{\text{n}\})^{t-1} \times X.
\end{aligned}$$

The set of strategies in the Rubinstein's bargaining game allows rather complex behavior and history-dependence. For instance, an agent may react to a low offer by rejecting it and by offering an even lower share to his opponent once he becomes proposer in the following period. However, we can define a simple subset of strategies, namely stationary strategies. A strategy for player i is stationary if:

(i) the proposal made by i when he plays in the role of the proposer is always the same, hence it is independent of the previous history;

(ii) the reply of agent i when he plays in the role of the respondent depends only on the last offer, hence it is independent of the history that precedes such an offer.

Formally, a strategy s_i ($i \in \{\text{A,B}\}$) is **stationary** if $s_i(h) = s_i(h')$ for all $h, h' \in H_{i,P}$, and $s_i(h,x) = s_i(h',x)$ for all $h, h' \in H_{-i,P}$, and $x \in X$.[14] An SPE is stationary if both players use stationary strategies. Notice that if each player follows a stationary strategy, each bargaining round is independent of the past: the proposer does not condition his behavior on the past history, while the reply of the respondent depends only on the offer that is currently on the table. Moreover, by their very nature,

[14]Note that, if $h \in H_{-i,P}$, then $(h,x) \in H_{i,R}$.

stationary strategies can be easily described by the proposal that an agent makes when he plays in the role of the proposer and by a reply function that specifies how the respondent reacts to any offer $x \in X$.

Since Rubinstein's game has infinite horizon, its SPE cannot be computed by backward induction. Nevertheless, we can use a "conjecture-and-verify" approach that exploits our analysis of the 2-Period Alternating Offer Game to characterize an SPE.[15]

In the finite-horizon equilibrium, offers are accepted and reply functions are "monotone," that is, every offer that gives the respondent less than the equilibrium offer is rejected and every offer that gives more is accepted. However, the sequence of equilibrium offers is not stationary, because as the deadline comes closer equilibrium offers are more influenced by the default split.[16]

In the infinite horizon game there is no deadline and this implies that the game has a **stationary structure**. Specifically, for every $i \in \{\mathrm{A}, \mathrm{B}\}$ and every pair of histories $h, h' \in H_{i,P}$, the subgames starting after h and h' look exactly the same: i is the first proposer of an infinite-horizon alternating-offer game; furthermore, for every offer $x \in X$, the subgames starting after (h, x) and (h', x) again look exactly the same: $-i$ has to respond and—in case of rejection—he will be the first proposer in an infinite-horizon alternating-offer game. Therefore we conjecture that there is a stationary SPE whereby equilibrium offers are always accepted and reply functions are "monotone" in the sense explained above, with symmetry when the roles of the players are switched, that is, the proposer always demand the same for herself and the acceptance threshold of the respondent is what is left for himself by such proposal. Let x_P^* denote the **share that the proposer demands for herself**. Then we conjecture that there is a stationary SPE $(s_\mathrm{A}, s_\mathrm{B})$ of the following form: for all $h \in H_{\mathrm{A},P}$, $h' \in H_{\mathrm{B},P}$, $x = (x_\mathrm{A}, x_\mathrm{B}) \in X$:

$$
\begin{aligned}
s_\mathrm{A}(h) &= (x_P^*, 1 - x_P^*), \; s_\mathrm{A}(h', x) = \begin{cases} \mathrm{y}, & \text{if } x_\mathrm{A} \geq 1 - x_P^*, \\ \mathrm{n}, & \text{if } x_\mathrm{A} < 1 - x_P^*, \end{cases} \\
s_\mathrm{B}(h') &= (1 - x_P^*, x_P^*), \; s_\mathrm{B}(h, x) = \begin{cases} \mathrm{y}, & \text{if } x_\mathrm{B} \geq 1 - x_P^*, \\ \mathrm{n}, & \text{if } x_\mathrm{B} < 1 - x_P^*. \end{cases}
\end{aligned}
$$

[15] This approach is borrowed from Gibbons [48].

[16] We proved this for two periods of barganing, but it is intuitive that the same holds for finitely many periods.

With this, we can find the implied value of x_P^* relying on the analysis of the 2-Period Alternating Offer Game: according to the candidate equilibrium, if the play moves to period $t = 3$ the split $(x_P^*, 1 - x_P^*)$ is implemented in that period. Therefore, the first two periods strategies must be determined by backward induction (breaking ties in favor of the proposer) as in the 2-period model with default split $(\bar{x}_A, \bar{x}_B) = (x_P^*, 1 - x_P^*)$. We have seen that Ann's equilibrium offer in such game is

$$(1 - \delta (1 - \delta \bar{x}_A), \delta (1 - \delta \bar{x}_A)) = (1 - \delta (1 - \delta x_P^*), \delta (1 - \delta x_P^*)).$$

Therefore, to have a stationary equilibrium of the infinite-horizon game we must have

$$x_P^* = 1 - \delta (1 - \delta x_P^*).$$

Solving this equation we find

$$x_P^* = \frac{1-\delta}{1-\delta^2} = \frac{1-\delta}{(1-\delta)(1+\delta)} = \frac{1}{1+\delta}.$$

To check that the stationary strategies specified above with $x_P^* = 1/(1+\delta)$ indeed form an SPE it is enough to verify that each one is one-step optimal given the conjecture that the opponent follows his strategy in the candidate equilibrium. We leave this as an exercise.

Proposition 12. *Rubinstein's bargaining game has an SPE in stationary strategies* (s_A, s_B) *determined as follows: for all* $h \in H_{A,P}$*,* $h' \in H_{B,P}$*, and* $x = (x_A, x_B) \in X$*:*

$$\begin{aligned} s_A(h) &= \left(\frac{1}{1+\delta}, \frac{\delta}{1+\delta}\right),\ s_A(h', x) = \begin{cases} \text{y}, & \textit{if } x_A \geq \frac{\delta}{1+\delta}, \\ \text{n}, & \textit{if } x_A < \frac{\delta}{1+\delta}, \end{cases} \\ s_B(h') &= \left(\frac{\delta}{1+\delta}, \frac{1}{1+\delta}\right),\ s_B(h, x) = \begin{cases} \text{y}, & \textit{if } x_B \geq \frac{\delta}{1+\delta}, \\ \text{n}, & \textit{if } x_B < \frac{\delta}{1+\delta}. \end{cases} \end{aligned}$$

The next result establishes that this is the unique SPE of the Rubinstein's game.

Theorem 54. *The strategy pair of Proposition 12 is the unique SPE of Rubinstein's bargaining game.*

Proof. Consider any subgame that starts with player i being active as the proposer. All these subgames have the same structure up to a normalization of utilities. In particular, for every $h, h' \in H_{i,P}$, the set of continuation-strategies available to agent j ($j \in \{\text{A,B}\}$) in the subgame with initial node h is isomorphic to the set of continuation-strategies available to agent j in the subgame with initial node h'. The only difference among these subgames is that the respective payoff functions can be obtained one from the other by a positive linear transformation, which does not affect incentives.

To prove our result, it is useful to introduce some additional notation. Take any $h \in H_{i,P}$. Let Γ_i^h be the subgame with initial node h in which we normalize payoffs so that immediate agreement on split $x = (x_\text{A}, x_\text{B})$ yields payoff x_A to Ann (x_B to Bob).[17] Given this normalization, the sum of the players' payoffs at any terminal history of the subgame Γ_i^h is at most 1. Denote with $E\left(\Gamma_i^h\right)$ the set of SPE equilibria in subgame Γ_i^h and with $V_j\left(\Gamma_i^h\right)$ ($j \in \{\text{A,B}\}$) the set of SPE payoffs of agent j. Since all subgames where i is the proposer are isomorphic, $E\left(\Gamma_i^h\right)$ and $V_j\left(\Gamma_i^h\right)$ are independent of $h \in H_{i,P}$, and it makes sense to write $E\left(\Gamma_i\right)$ and $V_j\left(\Gamma_i\right)$. Finally, denote with $\underline{v}_j\left(\Gamma_i\right)$ and $\overline{v}_j\left(\Gamma_i\right)$ the infimum and the supremum of $V_j\left(\Gamma_i\right)$. By Proposition 12, we know that $E\left(\Gamma_i\right)$ and $V_j\left(\Gamma_i\right)$ are nonempty and that $\overline{v}_j\left(\Gamma_i\right) > 0$.

Consider any subgame that starts at period-t history $h \in (X \times \{\text{n}\})^{t-1} \subseteq H_{A,P}$ (t odd). At such history Ann is the proposer. Rationality and the expectation of an equilibrium continuation imply that Bob accepts any offer x such that $x_\text{B} > \delta\overline{v}_\text{B}\left(\Gamma_\text{B}\right)$ and rejects any offer x such that $x_\text{B} < \delta\underline{v}_\text{B}\left(\Gamma_\text{B}\right)$. In words, Bob accepts (respectively, rejects) any offer that gives him more than the maximum (respectively, less than the minimum) equilibrium payoff he can get in the subgame that would start after his rejection. Thus if Ann offers $x = (x_\text{A}, x_\text{B})$ with $x_\text{B} > \delta\overline{v}_\text{B}\left(\Gamma_\text{B}\right)$, the behavior of players would yield consequence (x, t) and Ann would get a payoff equal to $1 - x_\text{B}$. As a result, $\lim_{x_\text{B} \searrow \delta\overline{v}_\text{B}(\Gamma_\text{B})} (1 - x) = 1 - \delta\overline{v}_\text{B}\left(\Gamma_\text{B}\right)$ is a lower bound on the equilibrium payoff that Ann can get in subgame Γ_A. Formally:

$$\underline{v}_\text{A}\left(\Gamma_\text{A}\right) \geq 1 - \delta\overline{v}_\text{B}\left(\Gamma_\text{B}\right). \tag{14.3.1}$$

Moreover, we have already argued that in equilibrium Bob rejects any offer

[17] That is, if $h \in (X \times \{\text{n}\})^{t-1}$, the payoff function of i in Γ_i^h satisfies $u_i^h(h, x, \text{y}) = x_i = \delta^{1-t} u_i(h, x, \text{y})$.

$x = (x_A, x_B)$ with $x_B < \delta \underline{v}_B(\Gamma_B)$. Thus, if Ann offers $x = (x_A, x_B)$ with $x_B < \delta \underline{v}_B(\Gamma_B)$, the maximum she can get is smaller than the maximum sum of payoffs attainable in the subgame starting after the rejection (which, from the perspective of period t, is equal to $\delta \cdot 1 = \delta$) minus the minimum Bob can get (which, from the perspective of period t, is equal to $\delta \underline{v}_B(\Gamma_B)$). Instead, if Ann offers $x = (x_A, x_B)$ with $x_B \geq \delta \underline{v}_B(\Gamma_B)$, the maximum she can obtain is at most $1 - \delta \underline{v}_B(\Gamma_B)$. Indeed, if Bob accepts the offer, Ann would get $1 - x_B$, while if he rejects, the maximum Ann can get is once more lower than $\delta(1 - \underline{v}_B(\Gamma_B))$. Since $\delta < 1$ and $\lim_{x_B \searrow \delta \underline{v}_B(\Gamma_B)} (1 - x) = 1 - \underline{v}_B(\Gamma_B)$, we conclude that :

$$\overline{v}_A(\Gamma_A) \leq 1 - \delta \underline{v}_B(\Gamma_B) \tag{14.3.2}$$

By replicating the previous analysis for subgame Γ_B, we get:

$$\underline{v}_B(\Gamma_B) \geq 1 - \delta \overline{v}_A(\Gamma_A) \tag{14.3.3}$$

$$\overline{v}_B(\Gamma_B) \leq 1 - \delta \underline{v}_A(\Gamma_A) \tag{14.3.4}$$

Combining (14.3.1) with (14.3.4) and (14.3.2) with (14.3.3), we get:

$$\begin{aligned} \underline{v}_A(\Gamma_A) &\geq 1 - \delta(1 - \delta \underline{v}_A(\Gamma_A)) \Rightarrow \underline{v}_A(\Gamma_A) \geq \frac{1}{1+\delta}, \\ \overline{v}_A(\Gamma_A) &\leq 1 - \delta(1 - \delta \overline{v}_A(\Gamma_A)) \Rightarrow \overline{v}_A(\Gamma_A) \leq \frac{1}{1+\delta}. \end{aligned}$$

By definition, $\underline{v}_A(\Gamma_A) \leq \overline{v}_A(\Gamma_A)$. Thus, $\underline{v}_A(\Gamma_A) = \overline{v}_A(\Gamma_A) = v_A(\Gamma_A) = \frac{1}{1+\delta}$. Therefore, $V_A(\Gamma_A) = \left\{\frac{1}{1+\delta}\right\}$. Following similar steps, we can get $\underline{v}_B(\Gamma_B) = \overline{v}_B(\Gamma_B) = v_B(\Gamma_B) = \frac{1}{1+\delta}$, so that $V_B(\Gamma_B) = \left\{\frac{1}{1+\delta}\right\}$. Given the normalization in the payoffs, we conclude that in every subgame of the Rubinstein's game that starts at period-t history $h \in (X \times \{\text{n}\})^{t-1} \subseteq H_{i,P}$ (t odd for $i = \text{A}$, even for $i = \text{B}$) player i (the proposer) gets a payoff equal to $\delta^{t-1}\frac{1}{1+\delta}$.

Now, consider subgame Γ_A from the perspective of Bob. We have just showed that by rejecting the offer, Bob will obtain $\delta v_B(\Gamma_B) = \frac{\delta}{1+\delta}$. Obviously, Bob can follow a strategy that prescribes to reject all offers made in period t. Thus, $\underline{v}_B(\Gamma_A) \geq \delta v_B(\Gamma_B) = \frac{\delta}{1+\delta}$. Moreover, the maximum payoff Bob can get in subgame Γ_A is smaller or equal than the maximum sum of payoffs available (namely, 1) minus the minimum payoff

Ann can get (by our previous reasoning, we know that $\underline{v}_{\mathrm{A}}(\Gamma_{\mathrm{A}}) = v_{\mathrm{A}}(\Gamma_{\mathrm{A}}) = \frac{1}{1+\delta}$). Thus, $\overline{v}_{\mathrm{B}}(\Gamma_{\mathrm{A}}) \leq 1 - v_{\mathrm{A}}(\Gamma_{\mathrm{A}}) = \frac{\delta}{1+\delta}$. By definition $\underline{v}_{\mathrm{B}}(\Gamma_{\mathrm{A}}) \leq \overline{v}_{\mathrm{B}}(\Gamma_{\mathrm{A}})$. We conclude that $\overline{v}_{\mathrm{B}}(\Gamma_{\mathrm{A}}) = \underline{v}_{\mathrm{B}}(\Gamma_{\mathrm{A}}) = v_{\mathrm{B}}(\Gamma_{\mathrm{A}}) = \frac{\delta}{1+\delta}$ and $V_{\mathrm{B}}(\Gamma_{\mathrm{A}}) = \left\{\frac{\delta}{1+\delta}\right\}$. A similar reasoning implies that $\overline{v}_{\mathrm{A}}(\Gamma_{\mathrm{B}}) = \underline{v}_{\mathrm{A}}(\Gamma_{\mathrm{B}}) = v_{\mathrm{A}}(\Gamma_{\mathrm{B}}) = \frac{\delta}{1+\delta}$ and $V_{\mathrm{A}}(\Gamma_{\mathrm{B}}) = \left\{\frac{\delta}{1+\delta}\right\}$. Thus, in every subgame of the Rubinstein's game that starts at period-t history $h \in (X \times \{\mathrm{n}\})^{t-1} \subseteq H_{i,P}$ (t even for $i = \mathrm{A}$, odd for $i = \mathrm{B}$) player j (the receiver) gets a payoff equal to $\delta^{t-1}\frac{\delta}{1+\delta}$.

Now, pick $i \in \{\mathrm{A,B}\}$ and consider any subgame with initial node $h \in (X \times \{\mathrm{n}\})^{t-1} \subseteq H_{i,P}$. By the previous argument, we know that the sum of players' equilibrium payoffs in this subgame must be equal to $\delta^t \frac{1}{1+\delta} + \delta^t \frac{\delta}{1+\delta} = \delta^t$. Therefore, the only consequence compatible with equilibrium payoffs is (x^*, t), where $x^* = (x_P^*, x_R^*) = \left(\frac{1}{1+\delta}, \frac{\delta}{1+\delta}\right)$ and x_P^* (x_R^*) is the share the proposer demands for himself (offers to the respondent). Indeed, any consequence (x, t) with $x \neq x^*$ would yield payoffs different from the unique equilibrium payoffs we identified above. Similarly, any consequence (x, t') with $t' > t$ would yield payoffs whose sum could be at most $\delta^{t+1} < \delta^t$. Obviously, in the subgame that starts with initial node $h \in (X \times \{\mathrm{n}\})^{t-1} \subseteq H_{i,P}$, consequence (x, t) is reached if and only if i proposes x^* and j accepts at history (h, x^*). Finally, consider any history (h, x) with $x = (x_P, x_R) \neq x^*$. Rationality and the expectation of an equilibrium continuation on the respondent's side imply that in equilibrium Bob accepts at any history $h \in (X \times \{\mathrm{n}\})^{t-1} \times \{x\} \subseteq H_{i,R}$ with $x_R > \delta v_{\mathrm{B}}(\Gamma_{\mathrm{B}}) = \frac{\delta}{1+\delta}$, and rejects at any history $h \in (X \times \{\mathrm{n}\})^{t-1} \times \{x\} \subseteq H_{i,R}$ with $x_R < \delta v_{\mathrm{B}}(\Gamma_{\mathrm{B}}) = \frac{\delta}{1+\delta}$. ■

Theorem 54 shows that infinite-horizon alternating-offers game shares the same features that we highlighted in a finitely repeated Ultimatum Game. In particular, although this game has infinite horizon and we cannot apply backward induction, the game has a unique subgame perfect equilibrium, which is stationary and rather simple to describe.[18] In this equilibrium, the proposer always asks for himself a fraction $\frac{1}{1+\delta}$ and offers

[18] This is in stark contrast with other infinite-horizon games that admit multiple equilibria and equilibrium outcomes. Such uniqueness is one of the reasons for the success of Rubinstein's model for applied work.

to the respondent $\frac{\delta}{1+\delta}$. The respondent, in turn, accepts any offer that gives him at least $\frac{\delta}{1+\delta}$ and rejects everything else.

The difference $\frac{1}{1+\delta} - \frac{\delta}{1+\delta} = \frac{1-\delta}{1+\delta}$ can be interpreted as the proposer's advantage and it is decreasing in δ. This is not surprising. When players are impatient, i.e., when δ is low, disagreement is costly and the respondent will be reluctant to reject the offer. As a result, the proposer will be able to extract a higher share of the total surplus.

Furthermore, the equilibrium characterization implies that an agreement is reached in the first period. To put it differently, costly disagreement never happens. This is a consequence of the fact that disagreement is costly for both players and that such cost is common knowledge between them.

15

Multistage Games with Incomplete Information

In this chapter we extend the analysis of static games with incomplete information to game forms with a multistage structure. As in Chapter 8, we first analyze rationalizability in games with payoff uncertainty and then we move on to the analysis of equilibria in Bayesian games. We observed in Chapters 9 and 12 that some Nash equilibria of multistage games are un-intuitive because the strategies of one or more players are "irrational" in some subgame. This led most game theorists to adopt subgame perfect equilibrium, a refinement of Nash equilibrium, as the standard solution concept. Similarly, there are Bayesian equilibria of multistage Bayesian games that prescribe "irrational" behavior at histories that cannot be reached in equilibrium. The reaction of most game theorists was to extend the subgame perfect equilibrium concept to multistage Bayesian games to refine Bayesian equilibrium and obtain a notion of "perfect Bayesian equilibrium." Unlike the complete-information case, however, they could not agree on a canonical definition, except for simple cases like leader-follower games. Here we provide the most general definition that satisfies some basic consistency requirements. Then we focus on a special class of leader-follower games, the signaling games, where all the definitions in the literature coincide with ours.

15.1 Multistage Games with Payoff Uncertainty

Consider a **multistage game tree** with observed actions (see Section 9.2 of Chapter 9), i.e., a structure

$$\left\langle I, (A_i, \mathcal{A}_i(\cdot))_{i\in I}\right\rangle .$$

The sets of histories H, Z and $\bar{H} = H \cup Z$, where H (respectively, Z) are nonterminal (respectively, terminal), are defined in the usual way.

A multistage environment with incomplete information is obtained by adding to the previous structure a set of states of nature and payoff functions that depend on the state of nature and the terminal history (cf. Section 8.1 of Chapter 8). A **multistage game with payoff uncertainty** and observed actions is given by the following mathematical structure

$$\hat{\Gamma} = \left\langle I, \Theta_0, (\Theta_i, A_i, \mathcal{A}_i(\cdot), u_i)_{i\in I}\right\rangle ,$$

where $\left\langle I, (A_i, \mathcal{A}_i(\cdot))_{i\in I}\right\rangle$ is a multistage game tree, all the sets Θ_j ($j \in I \cup \{0\}$) are nonempty, Θ_0 is the space of residual uncertainty, and Θ_i the set of information-types of player i. As usual, we let $\Theta = \times_{i\in I\cup\{0\}} \Theta_i$. The outcome function $g : \Theta \times Z \to Y$ and the utility functions $v_i : \Theta_i \times Y \to \mathbb{R}$ depend on the vector of parameters $\theta \in \Theta$ which is not commonly known. For each $i \in I$, payoff function u_i is written in a parameterized form

$$\begin{array}{llll} u_i : & \Theta \times Z & \to & \mathbb{R}, \\ & (\theta, z) & \mapsto & v_i\left(\theta_i, g\left(\theta, z\right)\right). \end{array}$$

If the set Θ is a singleton (or, more generally, if each payoff function u_i is independent of $\theta \in \Theta$), then the game is said to have **complete information**.

Note that the definition of $\hat{\Gamma}$ requires that each feasible set $\mathcal{A}_i(h)$ be independent of the vector of parameters $\theta \in \Theta$. In some applications it makes sense to allow the set of feasible actions of some player i at some history h to depend on θ_i. The above definition of $\hat{\Gamma}$ can be generalized as follows: for each $i \in I$, the feasibility correspondence $\mathcal{A}_i(\cdot,\cdot) : \Theta_i \times A^{<\mathbb{N}_0} \rightrightarrows A_i$ associates with each information-type θ_i and finite sequence of action profiles $h = (a^k)_{k=1}^{\ell}$ a set of actions $\mathcal{A}_i(\theta_i, h)$ that are feasible for θ_i immediately after h. Note, since player i knows θ_i he always knows his feasible actions. Such generalization would not change the substance of the results. Furthermore, to simplify the probabilistic analysis to come, we assume that the game is finite, i.e., $\bar{H}$ *and* Θ *are finite*.

In line with the genuine incomplete-information interpretation, we assume that there is *no ex ante stage*: players are "born" with their private information. Thus, we do not define strategies as choice rules that depend on private information and history, but rather as choice rules that only depend on history. As usual, we let $S_i = \times_{h \in H} \mathcal{A}_i(h)$ denote the set of pure strategies of player $i \in I$, and we let $S = \times_{i \in I} S_i$ denote the set of strategy profiles.

We obtained multistage games with payoff uncertainty by putting together ideas and concepts related to static games with incomplete information and multistage games with (observed actions and) complete information. Both are special cases of this larger class of games. This transition may appear so seamless that one risks neglecting an important new issue: if there is incomplete information, some player i does not know (θ_0 or) the information-types θ_{-i} of his co-players. If there were only one stage, it would be enough to posit *exogenous* (type-dependent) beliefs about such exogenous unknowns. But, if there are at least two stages, one has to consider i's beliefs about co-players' information-types conditional on their observed actions. The key issue is that such beliefs are necessarily *endogenous*, because—besides i's initial exogenous beliefs—they depend on what i thinks about the relationship between co-players' information-types and their behavior. In other words, player i has to update his beliefs about the information-types of others. To understand this issue, one first has to go back to probability theory and recall how it models beliefs about unknown parameters as they are updated upon observing new evidence.

15.2 Intermezzo: Bayes Rule

Consider an agent i who is initially uncertain about two things: a parameter $\theta \in \Theta$ and a variable (or a vector of variables) with realizations $x \in X$, where Θ and X are assumed to be *finite*. In a later stage i observes x. Even though i cannot observe θ, he is able to assess the probability of each realization $x \in X$ conditional on each $\theta \in \Theta$, which is denoted by $\mathrm{P}(x|\theta)$.

For example, consider an urn of unknown composition. It is only known that the urn contains between 1 and 10 balls, which may differ only in the color, *Black* or *White*. A first ball is drawn, its color is observed and then it is put back into the urn. Then a second ball is drawn (possibly the same as before) and its color is observed. Let θ denote the proportion of *White* balls in the urn. The set Θ of possible values of this parameter is finite:

$$\Theta = \bigcup_{n=1}^{10} \left\{ \theta : \exists k \in \{0, ..., n\}, \theta = \frac{k}{n} \right\}.$$

Let *Black* draws correspond to number 0 (since black is the absence of light) and *White* draws correspond to number 1. Then $X = \{0,1\}^2$ and the probability of $x = (x_1, x_2)$ conditional on the proportion of white balls being θ is

$$\mathrm{P}(x|\theta) := \theta^{x_1}(1-\theta)^{1-x_1}\theta^{x_2}(1-\theta)^{1-x_2}.$$

For example, $\mathrm{P}((1,0)|\theta) = \theta(1-\theta)$.

Agent i also assigns a *subjective probability* $\mathrm{P}(\theta)$ to all the possible values of θ. Note that $\mathrm{P}(x|\theta)$ is well defined even if i assigns probability 0 to θ. For example, for some reason i may be certain that the urn does not contain more than 5 balls, and hence $\mathrm{P}\big(\theta = \frac{1}{10}\big) = 0$. Yet i thinks that if θ *were* $\frac{1}{10}$ then the probability of $x = (0,0)$ *would be* $\frac{9}{10} \cdot \frac{9}{10}$, that is, $\mathrm{P}\big((0,0)|\frac{1}{10}\big) = \frac{81}{100}$.

Fix any finite uncertainty space Ω (later, we will relate Ω to Θ and X). The law of conditional probabilities says that for every pair of events $E, F \subseteq \Omega$,

$$\mathrm{P}(F) > 0 \Rightarrow \mathrm{P}(E|F) = \frac{\mathrm{P}(E \cap F)}{\mathrm{P}(F)}.$$

Note that the same condition can be written more compactly as

$$\mathrm{P}(E \cap F) = \mathrm{P}(E|F)\mathrm{P}(F).$$

Here we may assume that the space of uncertainty is $\Omega = \Theta \times X$. Each realization x corresponds to event $E = \Theta \times \{x\}$, each parameter value θ corresponds to event $F = \{\theta\} \times X$, each pair (θ, x) is the singleton $\{(\theta, x)\}$. Then the joint probability of the pair (θ, x) is

$$\mathrm{P}(\theta, x) = \mathrm{P}(x|\theta)\mathrm{P}(\theta).$$

Note that the collection of events $\{F \subseteq \Omega : \exists \theta \in \Theta, F = \{\theta\} \times X\}$ forms a partition of $\Omega = \Theta \times X$. This partition of Ω induces a corresponding partition of every event $E \subseteq \Omega$, and in particular of events of the form $E = \Theta \times \{x\}$. Then, we can compute the probability of every x (that is, event $\Theta \times \{x\}$) starting from the probabilities of the "cells" $\{(\theta, x)\} = (\Theta \times \{x\}) \cap (\{\theta\} \times X)$, where $\theta \in \Theta$. In turn, these probabilities can be obtained from the conditional and prior probabilities $\mathrm{P}(x|\theta)$ and $\mathrm{P}(\theta)$, $\theta \in \Theta$. Thus, we obtain the formula expressing the marginal, or **predictive probability** of x as

$$\mathrm{P}(x) = \sum_{\theta' \in \Theta} \mathrm{P}(x, \theta') = \sum_{\theta' \in \Theta} \mathrm{P}(x|\theta')\mathrm{P}(\theta'). \tag{15.2.1}$$

The problem is to derive from these elements $\mathrm{P}(\theta|x)$, the probability that i would assign to each $\theta \in \Theta$ upon observing evidence $x \in X$. There are two possibilities, either $\mathrm{P}(x) = 0$ or $\mathrm{P}(x) > 0$.

If $\mathrm{P}(x) = 0$, then $\mathrm{P}(\theta|x)$ cannot be derived from the previous data. *This does not mean that i is unable to assess the conditional probability* $\mathrm{P}(\theta|x)$, it only means that $\mathrm{P}(\theta|x)$ is not determined by the other probabilistic assessments expressed above.

If $\mathrm{P}(x) > 0$, then $\mathrm{P}(\theta|x) = \frac{\mathrm{P}((\theta,x))}{\mathrm{P}(x)}$. Substituting $\mathrm{P}(x)$ with the expression given by (15.2.1) we obtain

$$\mathrm{P}(\theta|x) = \frac{\mathrm{P}(x|\theta)\mathrm{P}(\theta)}{\sum_{\theta' \in \Theta} \mathrm{P}(x|\theta')\mathrm{P}(\theta')}. \tag{15.2.2}$$

Eq. (15.2.2) is known as **Bayes Formula,** that is, an equation expressing $\mathrm{P}(\theta|x)$ as a function of the conditional probabilities $\mathrm{P}(x|\theta')$ and the prior probabilities $\mathrm{P}(\theta')$, with $\theta' \in \Theta$.

We call **Bayes Rule** the rule that says that whenever (15.2.2) *can* be applied, then it *must* be applied. Since (15.2.2) can be applied if and only if $\mathrm{P}(x) > 0$, we may write Bayes Rule in the following compact form:

$$\forall x \in X, \forall \theta \in \Theta,\ \mathrm{P}(\theta|x) \left(\sum_{\theta' \in \Theta} \mathrm{P}(x|\theta')\mathrm{P}(\theta') \right) = \mathrm{P}(x|\theta)\mathrm{P}(\theta). \tag{15.2.3}$$

Bayes Rule is *not violated* if either $\mathrm{P}(x) = 0$ (both sides of (15.2.3) are zero), or $\mathrm{P}(x) > 0$ and $\mathrm{P}(\theta|x)$ is computed with (15.2.2). Therefore, whenever an assessment of prior and conditional probabilities satisfies (15.2.3), we say that it is **consistent with Bayes rule**. In particular, Bayes rule "holds" if $\mathrm{P}(x) = 0$: another way to express this point is that if the antecedent in the material implication

$$\mathrm{P}(x) > 0 \Rightarrow \mathrm{P}(\theta|x) = \frac{\mathrm{P}((\theta, x))}{\sum_{\theta' \in \Theta} \mathrm{P}(x|\theta')\mathrm{P}(\theta')}$$

is false, then the material implication holds.[1]

15.3 Rational Planning

Before we move on to the analysis of solution concepts for multistage games with incomplete information, we have to extend the analysis of rational planning of Chapter 10. Specifically, we need to take into account that player i, given his (information-)type θ_i, is uncertain about the types θ_{-i} of the co-players. To simplify the notation, we assume that *there is no residual uncertainty*, that is, Θ_0 is a *singleton*, and we remove θ_0 from formulas.[2]

Definition 80. *Fix player i and type θ_i in a multistage game with payoff uncertainty $\hat{\Gamma}$. A* ***conjecture*** *is an array of probability measures*

$$\beta^i = (\beta^i(\cdot|\theta_{-i}, h))_{\theta_{-i} \in \Theta_{-i}, h \in H} \in \left(\bigtimes_{h \in H} \Delta(\mathcal{A}_{-i}(h)) \right)^{\Theta_{-i}}.$$

A ***personal system of beliefs*** *of type θ_i is an array of probability measures*

$$\mu_i\left(\cdot|\theta_i, \cdot\right) = (\mu_i(\cdot|\theta_i, h))_{h \in H} \in \Delta(\Theta_{-i})^H.$$

A pair $\left(\beta^i, \mu_i\left(\cdot|\theta_i, \cdot\right)\right)$ is called ***personal assessment*** *of type θ_i.*

As with complete-information games, to ease intuition, it is better to start thinking about two-person games, so that $-i$ is the co-player other than i. In this case, $\beta^i(\cdot|\theta_{-i}, \cdot) \in \bigtimes_{h \in H} \Delta(\mathcal{A}_{-i}(h))$ is like a behavior strategy of $-i$. Thus, this is a generalization of the notion of conjecture in two-person static games with complete information, where we represented conjectures about $-i$ and mixed actions of $-i$ with the same mathematical

[1] The material implication $p \Rightarrow q$ is verified if either p is false, or both p and q are true.

[2] Alternatively, we could interpret θ_{-i} as including θ_0.

object, a probability measure over A_{-i}. With more than two players, for some h, $\beta^i(\cdot|\theta_{-i},h) \in \Delta(\mathcal{A}_{-i}(h))$ may be a correlated probability measure over $\times_{j\neq i}\mathcal{A}_j(h)$. In the analysis of perfect Bayesian equilibrium of Section 15.7, we will assume that each $\beta^i(\cdot|\theta_{-i},h)$ is the product of the marginal measures (mixed actions) $\beta_j(\cdot|\theta_j,h) \in \Delta(\mathcal{A}_j(h))$, where $\beta_j(\cdot|\theta_j,\cdot)$ is the equilibrium behavior strategy of j. Thus, it will make sense to write $\beta^i = \beta_{-i}$.

As a matter of interpretation, conjecture β^i is like a family of statistical models that type θ_i deems possible. This family is parameterized by θ_{-i}, the unknown type profile of the co-players (cf. Section 15.2). Thus, $\beta^i(a_{-i}|\theta_{-i},h)$ is the probability of a_{-i} given θ_{-i} and conditional on "evidence" h, $\mu_i(\theta_{-i}|\theta_i,h)$ is the probability assigned by type θ_i of i to θ_{-i} conditional on observing h, and $\mu_i(\cdot|\theta_i,\varnothing)$ is the exogenous initial belief of type θ_i.[3]

Given $\hat{\Gamma}$ and type θ_i (hence, payoff function $u_{i,\theta_i}: \Theta_{-i}\times Z \to \mathbb{R}$), personal assessment $\left(\beta^i, \mu_i(\cdot|\theta_i,\cdot)\right)$ yields a **subjective decision tree**. The relevant conditional probabilities are derived as follows.

Suppose that the plan of type θ_i is described by behavior strategy $\beta_i(\cdot|\theta_i,\cdot)$ and fix a personal assessment $\left(\beta^i, \mu_i(\cdot|\theta_i,\cdot)\right)$. Let

$$\begin{aligned}
\mathbb{P}_\beta(a|\theta_i,\theta_{-i},h) &= \beta_i(a_i|\theta_i,h)\beta^i(a_{-i}|\theta_{-i},h), \qquad (15.3.1)\\
\mathbb{P}_{\beta^i,\mu_i}(a_{-i}|\theta_i,h) &= \sum_{\theta_{-i}\in\Theta_{-i}} \beta^i(a_{-i}|\theta_{-i},h)\mu_i(\theta_{-i}|\theta_i,h),\\
\mathbb{P}_{\beta^i,\mu_i}(a|\theta_i,h) &= \sum_{\theta_{-i}\in\Theta_{-i}} \mathbb{P}_\beta(a|\theta_i,\theta_{-i},h)\mu_i(\theta_{-i}|\theta_i,h)
\end{aligned}$$

for all $h\in H$, $a=(a_i,a_{-i})\in\mathcal{A}(h)$, and $\theta_{-i}\in\Theta_{-i}$. We can use the chain rule to obtain the probability of history $h'=\left(a^1,...,a^{\ell(h')}\right)$ conditional on prefix $h=\left(a^1,...,a^{\ell(h)}\right)$ (with $\ell(h)<\ell(h')$) given type profile $\theta=(\theta_i,\theta_{-i})$:

$$\mathbb{P}_\beta\left(h'|\theta,h\right) = \prod_{t=\ell(h)+1}^{\ell(h')} \mathbb{P}_\beta\left(a^t|\theta,\left(h,...,a^{t-1}\right)\right), \qquad (15.3.2)$$

[3] According to the interpretation given above, $\mu_i(\cdot|\theta_i,\varnothing)$ is the "prior" of type θ_i in the sense of Bayesian statistics, and each $\mu_i(\cdot|\theta_i,h)$ (with $h\neq\varnothing$) is the "posterior" of θ_i given evidence h.

with the convention that $\mathbb{P}_\beta\left(a^t|\theta,\left(h,...,a^{t-1}\right)\right) = \mathbb{P}_\beta\left(a^t|\theta,h\right)$ if $t-1 = \ell(h)$. Consistency with Bayes rule requires that, for all type profiles $\theta_{-i} \in \Theta_{-i}$ and histories $(h,a) \in H$,

$$\mathbb{P}_{\beta^i,\mu_i}(a|\theta_i,h) > 0 \Rightarrow \mu_i(\theta_{-i}|\theta_i,(h,a)) = \frac{\mathbb{P}_\beta(a|\theta,h)\mu_i(\theta_{-i}|\theta_i,h)}{\mathbb{P}_{\beta^i,\mu_i}(a|\theta_i,h)}. \quad (15.3.3)$$

(Eq. (15.3.3) is a "one-step" Bayes rule; given eq. (15.3.1) and the chain rule (15.3.2), eq. (15.3.3) implies the "multi-step" version of Bayes rule.) Eq. (15.3.1) yields

$$\mathbb{P}_{\beta^i,\mu_i}(a|\theta_i,h) = \beta_i(a_i|\theta_i,h)\mathbb{P}_{\beta^i,\mu_i}(a_{-i}|\theta_i,h).$$

Hence, for every $a_i \in$supp$\beta_i(\cdot|\theta_i,h)$, eq. (15.3.3) yields a version of Bayes rule whereby the belief about the co-players' types depends on what co-players just did:

$$\mathbb{P}_{\beta^i,\mu_i}(a_{-i}|\theta_i,h) > 0 \Rightarrow \mu_i(\theta_{-i}|\theta_i,(h,(a_i,a_{-i}))) = \frac{\beta^i(a_{-i}|\theta_{-i},h)\mu_i(\theta_{-i}|\theta_i,h)}{\mathbb{P}_{\beta^i,\mu_i}(a_{-i}|\theta_i,h)},$$

or, equivalently,

$$\mu_i(\theta_{-i}|\theta_i,(h,(a_i,a_{-i})))\mathbb{P}_{\beta^i,\mu_i}(a_{-i}|\theta_i,h) = \beta^i(a_{-i}|\theta_{-i},h)\mu_i(\theta_{-i}|\theta_i,h). \quad (15.3.4)$$

Thus, $\mu_i(\theta_{-i}|\theta_i,(h,(a_i,a_{-i})))$ is *independent* of own action a_i within subset supp$\beta_i(\cdot|\theta_i,h) \subseteq \mathcal{A}_i(h)$. The following definition of Bayes consistency requires that eq. (15.3.4) holds within the whole set $\mathcal{A}_i(h)$, not just the subset supp$\beta_i(\cdot|\theta_i,h)$.

Definition 81. *Personal assessment* $\left(\beta^i,\mu_i\left(\cdot|\theta_i,\cdot\right)\right)$ *is* ***Bayes consistent*** *if (15.3.4) holds for all* $\theta_{-i} \in \Theta_{-i}$ *and* $(h,(a_i,a_{-i})) \in H$.

Under personal Bayes consistency, *known results about rational planning—such as the OD principle—extend to the incomplete information case.* A detailed and formal analysis is provided in the appendix of this chapter. Here we only specify the elements to be used in the equilibrium analysis below.

To better connect with the analysis of perfect Bayesian equilibria of Section 15.7, we consider the possibility that i's plan, which depends on his type θ_i, is a behavior strategy $\beta_i\left(\cdot|\theta_i,\cdot\right) \in B_i = \bigtimes_{h\in H}\Delta(\mathcal{A}_i(h))$. To ease notation, in the remainder of this section we let $\beta_i = \beta_i(\cdot|\theta_i, \cdot)$ and $\mu_i = \mu_i(\cdot|\theta_i, \cdot) \in \Delta(\Theta_{-i})^H$ respectively denote the behavior strategy and personal system of beliefs *of a given type* θ_i of player i.[4] Thus, we fix a type θ_i who plans according to behavior strategy β_i and has personal assessment (β^i, μ_i), and we first define the value for θ_i of using β_i starting from history h *given* a type profile θ_{-i}:

$$V_{\theta_i}^{\beta_i,\beta^i}(\theta_{-i}, h) = \sum_{z \in Z(h)} \mathbb{P}_{\beta_i,\beta^i}(z|\theta_i, \theta_{-i}, h)\, u_i(\theta_i, \theta_{-i}, z).$$

Since type θ_i does not know θ_{-i}, to obtain the value of using β_i from h he must use the system of beliefs μ_i:

$$V_{\theta_i}^{\beta_i,\beta^i,\mu_i}(h) = \sum_{\theta_{-i} \in \Theta_{-i}} \mu_i(\theta_{-i}|\theta_i, h)\, V_{\theta_i}^{\beta_i,\beta^i}(\theta_{-i}, h).$$

Similarly, the value of choosing action a_i at h with future behavior given by β_i is

$$\begin{aligned} & V_{\theta_i}^{\beta_i,\beta^i,\mu_i}(h, a_i) \\ = & \sum_{\theta_{-i} \in \Theta_{-i}} \mu_i(\theta_{-i}|\theta_i, h) \sum_{a_{-i} \in \mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i}, h)\, V_{\theta_i}^{\beta_i,\beta^i}(\theta_{-i}, (h, (a_i, a_{-i}))). \end{aligned}$$

Note that these values depend only on the specification of β_i and β^i at histories that weakly follow h.

Definition 82. *Fix a type θ_i with personal assessment (β^i, μ_i) and a behavior strategy $\bar{\beta}_i \in B_i$. We say that $\bar{\beta}_i$ is*

- ***one-step optimal given** (β^i, μ_i) if*

$$\forall h \in H,\ V_{\theta_i}^{\bar{\beta}_i,\beta^i,\mu_i}(h) = \max_{a_i \in \mathcal{A}_i(h)} V_{\theta_i}^{\bar{\beta}_i,\beta^i,\mu_i}(h, a_i); \qquad (15.3.5)$$

- ***sequentially optimal given** (β^i, μ_i) if*

$$\forall h \in H,\ V_{\theta_i}^{\bar{\beta}_i,\beta^i,\mu_i}(h) = \max_{\beta_i \in B_i} V_{\theta_i}^{\beta_i,\beta^i,\mu_i}(h). \qquad (15.3.6)$$

[4] This makes sense because feasibility constraints and hard information about co-players are type-independent. Hence, each type of i has the same set of behavior strategies B_i and the same exogenous uncertainty space Θ_{-i}.

The following extension of the OD principle is proved in the appendix.

Theorem 55. (OD principle with incomplete information) *Fix a type* θ_i *with personal assessment* (β^i, μ_i)*. If* (β^i, μ_i) *is Bayes consistent, then one-step optimality given* (β^i, μ_i) *is equivalent to sequential optimality given* (β^i, μ_i)*.*

The following example illustrates how the equivalence between one-step and sequential optimality depends on Bayes consistency.

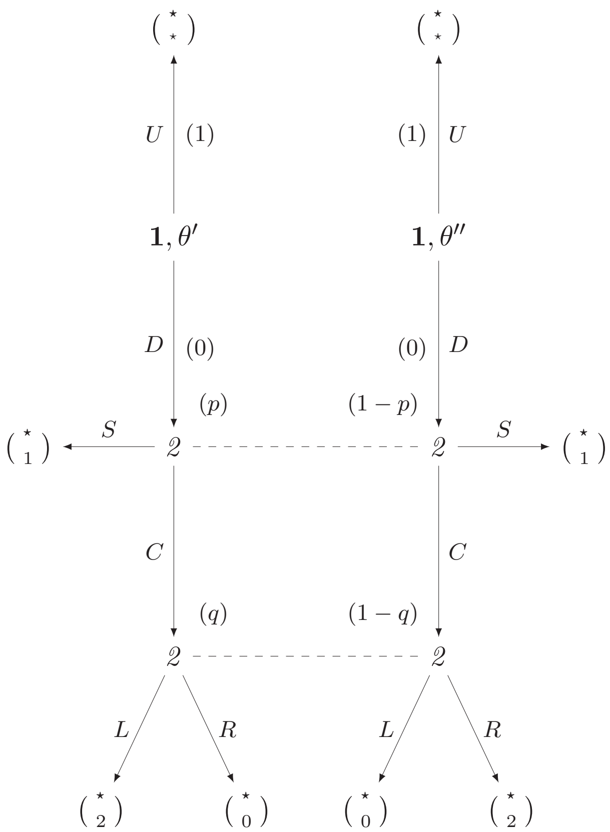


Figure 15.1: A subjective decision tree for player 2.

Example 68. Consider the subjective decision tree for player 2 depicted in Figure 15.1. Player 1 knows the true state of nature (his type) and moves first, choosing between U (up) and D (down). Player 2 has no private information, and his assessment is represented by the numbers in parentheses. The conjecture of 2 is given by $\beta^2(U|\theta') = 1 = \beta^2(U|\theta'')$. Since he initially expects player 1 to go up (U) with probability 1, his initial belief about θ cannot affect his conditional beliefs about θ given (D) and (D, C), because they cannot be pinned down by Bayes rule. Hence, the initial belief is not shown. The system of beliefs is given by $\mu_2(\theta'|D) = p$ and $\mu_2(\theta'|D, C) = q$. This assessment of player 2 is Bayes consistent if and only if $p = q$, because eq. (15.3.4) implies that his belief conditional on (D, C) cannot depend on the fact the he chose C instead of S. Note that the expected value of action L (respectively R) given (D, C) is $2q$ (respectively $2(1-q)$). Given Bayes consistency, if $p > 1/2$ (respectively, $p < 1/2$) the unique sequentially optimal strategy of 2 is also the only one that satisfies one-step optimality, $C.L$ (respectively, $C.R$). Now suppose that Bayes consistency is violated because $p < 1/2$ and $q > 1/2$. Then the unique strategy that satisfies one-step optimality is $S.L$: indeed, $2q > 2(1-q)$ which justifies going left (L) after (D, C); furthermore, player 2 *predicts* that he would go left after (D, C), thus the expected value of taking action C given D is $2p < 1$, which justifies S after D. With this inconsistent assessment player 2 changes his beliefs about θ and therefore his preferences over continuation-strategies. This makes it impossible to satisfy sequential optimality, which requires to continue and go right ($C.R$) after D and to go left (L) after (D, C). ▲

15.3.1 Justifiability and Dominance

In Chapter 8 we modeled conjectures as probability measures $\mu^i \in \Delta(\Theta_{-i} \times A_{-i})$, we considered the sets $r_i(\mu^i, \theta_i)$ of best replies to conjectures μ^i for types θ_i, we called an action a_i "justifiable" for a type θ_i if $a_i \in r_i(\mu^i, \theta_i)$ for some $\mu^i \in \Delta(\Theta_{-i} \times A_{-i})$, and we related justifiability and dominance. In Chapter 10 we defined conditional probability systems (CPSs) $\mu^i \in \Delta^H(S_{-i})$, we considered the sets of weakly sequential best replies $r_i(\mu^i)$ to such CPSs, we called a strategy s_i "justifiable" if $s_i \in r_i(\mu^i)$ for some $\mu^i \in \Delta^H(S_{-i})$, and we related justifiability to conditional dominance. The reason to consider *weak* sequential optimality in the definition of justifiability was that it is invariant to behavioral equivalence, and we took the perspective of an external observer, or a co-player, who wonders whether i may exhibit some pattern of behavior s_i

and who cannot distinguish, nor cares to distinguish between behaviorally equivalent strategies $s_i' \approx_i s_i''$, because they are necessarily realization-equivalent (Lemma 26). Merging these ideas, here we represent how player i would update or revise his beliefs about the co-players as the play unfolds with CPSs on $\Theta_{-i} \times S_{-i}$, we define justifiability for a type by means of such CPSs, and we relate it to conditional dominance for a type.

Let

$$\mathcal{H}_{-i} = \{C_{-i} \subseteq \Theta_{-i} \times S_{-i} : \exists h \in H, C_{-i} = \Theta_{-i} \times S_{-i}(h)\}.$$

A CPS on $\Theta_{-i} \times S_{-i}$ is an array of conditional probability measures

$$\bar{\mu}^i = \left(\bar{\mu}^i(\cdot|C_{-i})\right)_{C_{-i} \in \mathcal{H}_{-i}} \in \bigtimes_{C_{-i} \in \mathcal{H}_{-i}} \Delta(C_{-i})$$

such that the chain rule holds, that is, for all $C_{-i}, D_{-i} \in \mathcal{H}_{-i}$, and $E_{-i} \subseteq \Theta_{-i} \times S_{-i}$,

$$E_{-i} \subseteq D_{-i} \subseteq C_{-i} \Rightarrow \bar{\mu}^i(E_{-i}|C_{-i}) = \bar{\mu}^i(E_{-i}|D_{-i})\,\bar{\mu}^i(D_{-i}|C_{-i}).$$

As in Section 10.4 of Chapter 10, whenever convenient we ease notation by writing $\bar{\mu}^i(\cdot|h) = \bar{\mu}^i(\cdot|\Theta_{-i} \times S_{-i}(h))$, and we let $\Delta^H(\Theta_{-i} \times S_{-i})$ denote the set of CPSs on $\Theta_{-i} \times S_{-i}$. With this, it should be always kept in mind that if h' and h'' differ only because of actions taken by player i, so that $S_{-i}(h') = S_{-i}(h'')$, then $\bar{\mu}^i(\cdot|h') = \bar{\mu}^i(\cdot|h'')$.

Note that CPSs can be related to Bayes consistent personal assessments. Recall that

$$S_{-i}(h, a_{-i}) = \{s_{-i} \in S_{-i}(h) : s_{-i}(h) = a_{-i}\}$$

is the set of strategies of others allowing h and selecting a_{-i} given h.

Definition 83. *Personal assessment* (β^i, μ_i) *is* ***consistent with*** *CPS* $\bar{\mu}^i$ *if, for all* $h \in H$, $\theta_{-i} \in \Theta_{-i}$, *and* $a_{-i} \in \mathcal{A}_{-i}(h)$,

$$\beta^i(a_{-i}|\theta_{-i}, h)\,\mu_i(\theta_{-i}|h) = \bar{\mu}^i(\{\theta_{-i}\} \times S_{-i}(h, a_{-i})\,|h). \tag{15.3.7}$$

Note that, taking the summation with respect to co-players' actions in eq. (15.3.7), we obtain

$$\begin{aligned}\mu_i(\theta_{-i}|h) &= \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i},h)\,\mu_i(\theta_{-i}|h)\\ &= \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \bar{\mu}^i(\{\theta_{-i}\}\times S_{-i}(h,a_{-i})\,|h)\\ &= \bar{\mu}^i\left(\{\theta_{-i}\}\times\left(\cup_{a_{-i}\in\mathcal{A}_{-i}(h)}S_{-i}(h,a_{-i})\right)|h\right)\\ &= \bar{\mu}^i(\{\theta_{-i}\}\times S_{-i}(h)\,|h),\end{aligned}$$

because $\sum_{a_{-i}\in\mathcal{A}_{-i}(h)}\beta^i(a_{-i}|\theta_{-i},h) = 1$ and $\cup_{a_{-i}\in\mathcal{A}_{-i}(h)}S_{-i}(h,a_{-i}) = S_{-i}(h)$. Furthermore, this implies that if (β^i,μ_i) is consistent with $\bar{\mu}^i$, then

$$\bar{\mu}^i(\{\theta_{-i}\}\times S_{-i}(h)\,|h) > 0 \Rightarrow \beta^i(a_{-i}|\theta_{-i},h) = \frac{\bar{\mu}^i(\{\theta_{-i}\}\times S_{-i}(h,a_{-i})\,|h)}{\bar{\mu}^i(\{\theta_{-i}\}\times S_{-i}(h)\,|h)}.$$

By inspection of these definitions and formulas, one can show the following:

Remark 49. *For each CPS $\bar{\mu}^i \in \Delta^H(\Theta_{-i}\times S_{-i})$, there is a* Bayes consistent *personal assessment that is consistent with $\bar{\mu}^i$.*

This observation is important because it shows that the only parts of a personal assessment (β^i,μ_i) consistent with a CPS $\bar{\mu}^i$ that are not determined by eq. (15.3.7) are the conditional probabilities of actions $\beta^i(a_{-i}|\theta_{-i},h)$ given types deemed negligible conditional on h, and such probabilities do not determine the conditional values of carrying out strategies. Thus, we may as well define such values by means of CPSs: for all $\theta_i\in\Theta_i$, $s_i\in S_i$, $\bar{\mu}^i\in\Delta^H(\Theta_{-i}\times S_{-i})$, $h\in H_i(s_i)$, and $z\in Z(h)$, let

$$\mathbb{P}_{s_i,\bar{\mu}^i}(\theta_{-i},z|h) = \begin{cases} 0, & \text{if } s_i\notin S_i(z),\\ \bar{\mu}^i(\{\theta_{-i}\}\times S_{-i}(z)\,|h), & \text{if } s_i\in S_i(z),\end{cases}$$

$$V_{\theta_i}^{s_i,\bar{\mu}^i}(h) = \sum_{(\theta_{-i},z')\in\Theta_{-i}\times Z(h)} u_i(\theta_i,\theta_{-i},z')\,\mathbb{P}_{s_i,\bar{\mu}^i}(\theta_{-i},z'|h).$$

Definition 84. *Strategy s_i^* is* ***weakly sequentially optimal for type*** *θ_i given CPS $\bar{\mu}^i$, written $s_i^*\in r_i(\bar{\mu}^i,\theta_i)$, if*

$$\forall h\in H_i(s_i^*),\ V_{\theta_i}^{s_i^*,\bar{\mu}^i}(h) = \max_{s_i\in S_i(h)} V_{\theta_i}^{s_i,\bar{\mu}^i}(h).$$

Strategy s_i^* *is* ***justifiable for type*** θ_i *if there exists some CPS* $\bar{\mu}^i \in \Delta^H(\Theta_{-i} \times S_{-i})$ *such that* $s_i^* \in r_i(\bar{\mu}^i, \theta_i)$.

As in Chapter 10, we can give an analogous definition of weak sequential optimality given a Bayes consistent personal assessment and use Remark 49 to show that it is equivalent to the foregoing definition with CPSs (cf. Proposition 4).

For any fixed type θ_i, it is convenient to consider the section $r_{i,\theta_i} : \Delta^H(\Theta_{-i} \times S_{-i}) \rightrightarrows S_i$ at θ_i of the weak sequential optimality correspondence $r_i : \Delta^H(\Theta_{-i} \times S_{-i}) \times \Theta_i \rightrightarrows S_i$, that is, $r_{i,\theta_i}(\bar{\mu}^i) = r_i(\bar{\mu}^i, \theta_i)$ for all $\bar{\mu}^i$. With this, the set of weakly sequentially optimal strategies for type θ_i is $r_{i,\theta_i}(\Delta^H(\Theta_{-i} \times S_{-i}))$.

Using essentially the same arguments as in Chapter 10, we can relate justifiability and conditional dominance. To ease notation, let us define the parameterized normal-form (or strategic-form) payoff function

$$\begin{array}{rccl} \hat{U}_i: & \Theta \times S & \to & \mathbb{R} \\ & (\theta, s) & \mapsto & u_i(\theta, \zeta(s)). \end{array}$$

For any mixed strategy $\sigma_i \in \Delta(S_i)$, let

$$\hat{U}_i(\theta, \sigma_i, s_{-i}) = \mathbb{E}_{\sigma_i}\left(\hat{U}_{i,\theta,s_{-i}}\right) = \sum_{s_i \in S_i} \hat{U}_i(\theta, s_i, s_{-i})\,\sigma_i(s_i)\,.$$

Definition 85. *Strategy* $\bar{s}_i$ *is* ***conditionally dominated for type*** θ_i *if there are* $h \in H_i(\bar{s}_i)$ *and* $\sigma_i \in \Delta(S_i(h))$ *such that*

$$\forall(\theta_{-i}, s_{-i}) \in \Theta_{-i} \times S_{-i}(h)\,,\ \hat{U}_i(\theta_i, \theta_{-i}, \bar{s}_i, s_{-i}) < \hat{U}_i(\theta_i, \theta_{-i}, \sigma_i, s_{-i})\,.$$

The set of pairs (s_i, θ_i) *such that* s_i *is not conditionally dominated for* θ_i *is denoted* NCD_i*, and we let* $NCD = \times_{i \in I} NCD_i$.

Lemma 37. *A strategy is not conditionally dominated for a type if and only if it is justifiable for that type, that is, for each* $i \in I$,

$$NCD_i = \bigcup_{\theta_i \in \Theta_i} \{\theta_i\} \times r_{i,\theta_i}\left(\Delta^H(\Theta_{-i} \times S_{-i})\right).$$

15.4 Rationalizability

In Section 8.2 of Chapter 8 we analyzed rationalizability in static games with payoff uncertainty. In Chapter 11 we analyzed versions of the rationalizability solution concept for multistage games with complete information, characterizing the behavioral implications of rationality and different ways to extend the idea of common belief in rationality. Here we can take stock of those analyses and merge them to obtain an analysis of rationalizability in multistage games with payoff uncertainty, focusing on initial and strong rationalizability.[5] Similarly to Chapter 11, the behavioral implication of mere rationality is justifiability. As in Section 15.3, to ease notation we assume that there is no residual uncertainty, that is, Θ_0 is a singleton and can be neglected.[6]

15.4.1 Initial Rationalizability

Recall that initial rationalizability considers the implications of strategic reasoning on players *initial* beliefs and the related behavioral implications (see Section 11.2 of Chapter 11). Of course, beliefs have to be updated according to the rules of conditional probability when possible, but no assumption is made on how players *revise* their beliefs after they observe actions to which they had previously assigned probability 0. Thus, initial rationalizability characterizes the behavioral implications of Rationality and Common Initial Belief in Rationality. Following the same template of Chapters 8 and 11, we first extend the justification operator to multistage games with payoff uncertainty.

Let $\mathcal{C}$ denote the collection of Cartesian subsets $C = \times_{i\in I} C_i$ with $C_i \subseteq \Theta_i \times S_i$ for each $i \in I$. We interpret a set $C \in \mathcal{C}$ as type-dependent restrictions on behavior implied by strategic reasoning and we assume that each player *initially believes* that such restrictions hold for the co-players.[7]

---

[5] One can also put forth a meaningful definition and an interesting analysis of continuation rationalizability for multistage games with payoff uncertainty. See Penta [67] and Catonini and Penta [34].

[6] Analyzing the more general case does not create any conceptual difficulty.

[7] The restrictions given by C_i may be type-dependent because the sections of C_i at different types θ_i' and θ_i'' ($C_{i,\theta_i'}$ and $C_{i,\theta_i''}$) may be different. Furthermore, it makes sense to require that $\mathrm{proj}_{\Theta_i} C_i = \Theta_i$, because types are exogenous and cannot be deleted as such. But this is not strictly necessary for our analysis.

Let

$$\Delta^H_\varnothing(E_{-i}) = \left\{\bar{\mu}^i \in \Delta^H(\Theta_{-i} \times S_{-i}) : \bar{\mu}^i(E_{-i}|\varnothing) = 1\right\}$$

denote the set of CPSs that initially assign probability 1 to event $E_{-i} \subseteq \Theta_{-i} \times S_{-i}$ concerning the types and behavior of co-players. The initial justification operator $\rho : \mathcal{C} \to \mathcal{C}$ is the self-map defined as follows: for every $C \in \mathcal{C}$,

$$\begin{aligned} \rho(C) &= \bigtimes_{i \in I} \left\{(\theta_i, s_i) : \exists \bar{\mu}^i \in \Delta^H_\varnothing(C_{-i}), s_i \in r_i\left(\bar{\mu}^i, \theta_i\right)\right\} \\ &= \bigtimes_{i \in I} \left(\bigcup_{\theta_i \in \Theta_i} \{\theta_i\} \times r_{i,\theta_i}\left(\Delta^H_\varnothing(C_{-i})\right) \right). \end{aligned}$$

It can be verified that ρ is *monotone*, because $E_{-i} \subseteq F_{-i}$ implies $\Delta^H_\varnothing(E_{-i}) \subseteq \Delta^H_\varnothing(F_{-i})$. Thus, iterating ρ starting from the whole set of profiles of types and strategies we obtain a (weakly) decreasing sequence of subsets. To ease notation, we write $\Theta \times S$ instead of $\bigtimes_{i \in I}(\Theta_i \times S_i)$, given the obvious isomorphism between the two sets. Similarly, for games with more than two players, we identify $\Theta_{-i} \times S_{-i}$ with $\bigtimes_{j \neq i}(\Theta_j \times S_j)$. With this, we consider the sequence $(\rho^n(\Theta \times S))_{n \in \mathbb{N}}$, where

$$\rho^1(\Theta \times S) = \bigtimes_{i \in I} \left(\bigcup_{\theta_i \in \Theta_i} \{\theta_i\} \times r_{i,\theta_i}\left(\Delta^H(\Theta_{-i} \times S_{-i})\right) \right)$$

contains all the profiles $(\theta_i, s_i)_{i \in I}$ such that s_i is justifiable for θ_i, thus representing the type-dependent behavioral implications of rationality. Since $(\rho^n(\Theta \times S))_{n \in \mathbb{N}}$ is decreasing, we can define

$$\rho^\infty(\Theta \times S) = \bigcap_{n \in \mathbb{N}} \rho^n(\Theta \times S).$$

By finiteness of $\Theta \times S$, the limit $\rho^\infty(\Theta \times S)$ is attained after finitely many iterations.

Definition 86. *A profile of types and strategies $(\theta_i, s_i)_{i \in I}$ is **initially rationalizable** if $(\theta_i, s_i)_{i \in I} \in \rho^\infty(\Theta \times S)$.*

Leveraging on the monotonicity of ρ, one can prove an analog of Theorem 28. We postpone a discussion of this to Section 15.4.4, where we present Directed Rationalizability. As in Theorem 42 of Chapter 11, one can characterize initial rationalizability by means of iterated dominance. For each Cartesian set $C \in \mathcal{C}$, let $\mathrm{ND}(C) \subseteq C$ denote the set of profiles $(\theta_i, s_i)_{i \in I} \in C$ such that s_i is not dominated within C for θ_i by any mixed strategy $\sigma_i \in \Delta(C_{i,\theta_i})$, where $C_{i,\theta_i} = \{s_i' \in S_i : (\theta_i, s_i') \in C_i\}$ is the section at θ_i of C_i. In other words, $\mathrm{ND}(C)$ is the set of non-dominated profiles in the restricted strategic form $\langle I, (C_i, U_i|_C)_{i \in I} \rangle$. The restriction operator ND: $\mathcal{C} \to \mathcal{C}$ is a self-map that can be iterated starting from any $C \in \mathcal{C}$. Note that, by Lemma 37, $\rho(\Theta \times S) = NCD$. Thus, it makes sense to start the iteration from the set NCD of profiles of types and strategies such that the latter are not conditionally dominated for the given types. The proof of the following result is similar to the proof of Theorem 42.

Theorem 56. $\rho^\infty(\Theta \times S) = \mathrm{ND}^\infty(NCD)$.

In words, initial rationalizability can be computed by first eliminating (for every player i) each (θ_i, s_i) such that s_i is conditionally dominated for θ_i, and then iteratively eliminating pairs (θ_i, s_i) such that s_i is dominated for θ_i in the restricted strategic form obtained from the previous eliminations. The reason why conditional dominance is used only in the first step is the same as in Chapter 11: initial rationalizability allows "surprised" players to abandon their initial belief in the rationality of the co-players even if the co-players' observed behavior is consistent with rationality.

Example 69. In the multistage game with payoff uncertainty depicted in Figure 15.2, D is dominated by U for type θ' of player 1, but not for type θ''. The reduced strategy S is dominated for (the unique type of) player 2 by mixed strategy $\frac{1}{2}\delta_{\mathrm{C.L}} + \frac{1}{2}\delta_{\mathrm{C.R}}$ conditional on D. The remaining two strategies C.L and C.R are justifiable. For example, any $\bar{\mu}^2$ such that $\bar{\mu}^2(\theta', \mathrm{D}|\mathrm{D}) = \frac{1}{2}$ justifies both. Thus,

$$\rho^1(\Theta \times S) = NCD = \{(\theta', \mathrm{U}), (\theta'', \mathrm{U}), (\theta'', \mathrm{D})\} \times \{\mathrm{C.L}, \mathrm{C.R}\}.$$

Initial rationalizability stops here. If player 2 initially believes that both types of player 1 would choose U, which is consistent with initial belief in rationality, then no revised belief about player 1's type conditional upon observing D violates the chain rule (or Bayes consistency). In particular,

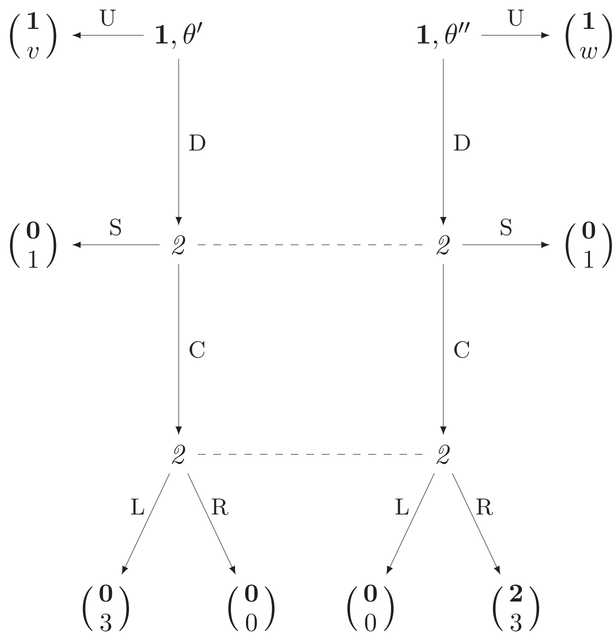


Figure 15.2: A 3-stage game with payoff uncertainty.

player 2 may deem θ' at least as likely as θ'' conditional on D despite the fact that D is dominated for θ' but not for θ'', thus justifying C.L. If, instead, player 2 strongly believed in the rationality of 1, then he should assign probability 0 to θ' conditional on D. Only C.R is optimal given such beliefs. Yet, initial rationalizability does not rely on strong belief in rationality. ▲

15.4.2 Strong Rationalizability

Unlike initial rationalizability, strong rationalizability relies on specific assumptions on how players *revise* their beliefs when surprised by the co-players' behavior (see Section 11.3 of Chapter 11). These assumptions use the concept of "strong belief": player i **strongly believes** an event E_{-i} concerning the co-players if he *initially believes* E_{-i} *and* assigns probability 1 to E_{-i} whenever he observes evidence consistent with E_{-i}. With this, strong rationalizability characterizes the behavioral implications of Rationality and Common Strong Belief in Rationality, which in turn captures the "**best rationalization principle**": rational players strongly believe in the rationality of their co-players, and more generally ascribe to them the highest degree of strategic sophistication consistent with their observed behavior.

For any nonempty event $E_{-i} \subseteq \Theta_{-i} \times S_{-i}$, let $\Delta_{\text{sb}}^{H}(E_{-i})$ denote the set of CPSs of player i that **strongly believe** $E_{-i} \subseteq \Theta_{-i} \times S_{-i}$, that is,

$$\Delta_{\text{sb}}^{H}(E_{-i}) = \left\{ \bar{\mu}^{i} \in \Delta_{\varnothing}^{H}(E_{-i}) : \begin{array}{l} \forall h \in H \backslash \{\varnothing\}, \\ (\Theta_{-i} \times S_{-i}(h)) \cap E_{-i} \neq \emptyset \Rightarrow \bar{\mu}^{i}(E_{-i}|h) = 1 \end{array} \right\}.$$

With this, we define a (weakly) decreasing sequence of Cartesian subsets $(\Sigma^{n})_{n\in\mathbb{N}}$ representing the type-dependent behavioral implications of rationality and the strategic reasoning steps reflecting the best rationalization principle.

Definition 87. *Consider the following elimination procedure:*
*(**Step** $n = 0$) For each $i \in I$, let $\Sigma_{i}^{0} = \Theta_{i} \times S_{i}$, $\Sigma_{-i}^{0} = \Theta_{-i} \times S_{-i}$, and $\Sigma_{\text{sb}}^{0} = \Theta \times S$.*
*(**Step** $n > 0$) For each $i \in I$, let*

$$\begin{aligned} \Delta_{i}^{n} &= \bigcap_{m=0}^{n-1} \Delta_{\text{sb}}^{H}\left(\Sigma_{-i}^{m}\right), \\ \Sigma_{i}^{n} &= \left\{(\theta_{i}, s_{i}) \in \Theta_{i} \times S_{i} : \exists \bar{\mu}^{i} \in \Delta_{i}^{n}, s_{i} \in r_{i}\left(\bar{\mu}^{i}, \theta_{i}\right)\right\} \\ &= \bigcup_{\theta_{i}\in\Theta_{i}} \{\theta_{i}\} \times r_{i,\theta_{i}}\left(\Delta_{i}^{n}\right), \end{aligned}$$

$\Sigma_{-i}^{n} = \times_{j\neq i} \Sigma_{j}^{n}$, $\Sigma^{n} = \times_{i\in I} \Sigma_{i}^{n}$.
*A profile of types and strategies $(\theta_{i}, s_{i})_{i\in I}$ is **strongly rationalizable** if $(\theta_{i}, s_{i})_{i\in I} \in \Sigma^{\infty} = \bigcap_{n>0} \Sigma^{n}$, that is, if $(\theta_{i}, s_{i}) \in \Sigma_{i}^{\infty} = \bigcap_{n>0} \Sigma_{i}^{n}$ for each player $i \in I$.*

Note that the sequence $(\Sigma^{n})_{n\in\mathbb{N}}$ is indeed (weakly) decreasing because, by definition, $\Delta_{i}^{n} \subseteq \Delta_{i}^{n-1}$ and thus $r_{i,\theta_{i}}(\Delta_{i}^{n}) \subseteq r_{i,\theta_{i}}\left(\Delta_{i}^{n-1}\right)$ for every $i \in I$ and $\theta_{i} \in \Theta_{i}$.

Let us compare strong and initial rationalizability. To ease the comparison, write the sequence of subsets obtained with the initial rationalizability procedure as $\left(\Sigma_{\varnothing}^{n}\right)_{n\in\mathbb{N}} = (\rho^{n}(\Theta \times S))_{n\in\mathbb{N}}$ with $\Sigma_{\varnothing}^{n} =$

$\times_{i\in I}\Sigma^n_{i,\varnothing}$. Since $\Sigma^0_{-i}=\Theta_{-i}\times S_{-i}$, $\Delta^1_i=\Delta^H(\Theta_{-i}\times S_{-i})$ and Σ^1_i is the set of pairs (θ_i,s_i) such that s_i is justifiable for θ_i. Thus, $\Sigma^1=\rho(\Theta\times S)=\Sigma^1_{\varnothing}$. For $n=2$,

$$\Delta^2_i=\Delta^H_{\text{sb}}\left(\Sigma^1_{-i}\right)\subseteq\Delta^H_{\varnothing}\left(\Sigma^1_{-i}\right)=\Delta^H_{\varnothing}\left(\Sigma^1_{-i,\varnothing}\right)$$

$(i\in I)$, where the inclusion follows from the fact that strong belief implies initial belief; thus, $\Sigma^2\subseteq\Sigma^2_{\varnothing}$. In the game of Example 69 (which will be analyzed in detail below), $\Delta^2_i=\left\{\bar{\mu}^2:\bar{\mu}^2(\theta',\text{D}|\text{D})=0\right\}$ and $\Sigma^2_2=\{\text{C.R}\}$, while $\Sigma^2_{2,\varnothing}=\Sigma^1_{2,\varnothing}=\{\text{C.L},\text{C.R}\}$. The inclusion indeed holds for all steps $n>1$. The proof of the following result is a rather straightforward extension of the proof of Remark 42 in Chapter 11.

Remark 50. *Initial rationalizability is weaker than strong rationalizability, that is,*

$$\Sigma^n\subseteq\Sigma^n_{\varnothing}:=\rho^n(\Theta\times S)$$

for all $n\in\mathbb{N}\cup\{\infty\}$.

Extending the analysis of Chapter 11 to allow for incomplete information, we could define a restriction operator $\bar{\rho}_{\text{sb}}:\mathcal{C}\to\mathcal{C}$ based on constrained optimization so that $\Sigma^n=\bar{\rho}^n_{\text{sb}}(\Theta\times S)$ for every $n\in\mathbb{N}\cup\{\infty\}$; we skip the details. This in turn yields a characterization of strong rationalizability by means of iterated conditional dominance, which we report below without proof.

Recall that, for any subset $C_i\subseteq\Theta_i\times S_i$ and type θ_i, we let $C_{i,\theta_i}\subseteq S_i$ denote the section of $C_i\subseteq\Theta_i\times S_i$ at θ_i. Similarly, we let $C_{i,\theta_i}(h)\subseteq S_i(h)$ denote the section of $C_i\cap(\Theta_i\times S_i(h))$ at θ_i, that is,

$$C_{i,\theta_i}(h)=\{s_i\in S_i(h):(\theta_i,s_i)\in C_i\}.$$

Also, for any nonempty subset $C\subseteq\Theta\times S$, let

$$H(C)=\{h\in H:\exists(\theta,s)\in C,h\prec\zeta(s)\}$$

denote the set of nonterminal histories consistent with C.

Definition 88. *Fix a nonempty Cartesian subset $C\in\mathcal{C}$, a player $i\in I$, and a pair $(\theta_i,\bar{s}_i)\in C_i$. Strategy $\bar{s}_i$ is **conditionally dominated in** C **for type** θ_i if there are $h\in H_i(\bar{s}_i)\cap H(C)$ and $\sigma_i\in\Delta(C_{i,\theta_i}(h))$, such that*

$$\forall(\theta_{-i},s_{-i})\in C_{-i}\cap(\Theta_{-i}\times S_{-i}(h)),\ \hat{U}_i(\theta_i,\theta_{-i},\bar{s}_i,s_{-i})<\hat{U}_i(\theta_i,\theta_{-i},\sigma_i,s_{-i}).$$

We say that $\bar{s}_i \in C_{i.\theta_i}$ *is* ***conditionally undominated in*** C ***for type*** θ_i *if it is* not *conditionally dominated in* C *for type* θ_i*. With this, we let* $\mathrm{NCD}(C)$ *denote the set of profiles* $(\theta_i, s_i)_{i\in I}$ *such that* s_i *is* conditionally undominated in C *for* θ_i*, for every* $i \in I$*.*

Theorem 57. $\Sigma^n = \mathrm{NCD}^n(\Theta \times S)$ *for all* $n \in \mathbb{N} \cup \{\infty\}$*.*

Example 70. Go back to the multistage game with payoff uncertainty of Example 69. Strong rationalizability yields a unique solution. We already know that

$$\Sigma^1 = \mathrm{NCD}(\Theta \times S) = \Sigma^1_{\varnothing} = \left\{(\theta', \mathrm{U}), (\theta'', \mathrm{U}), (\theta'', \mathrm{D})\right\} \times \{\mathrm{C.L}, \mathrm{C.R}\}.$$

Next, note that $\Sigma^1_1 \cap (\Theta_1 \times \{\mathrm{D}\}) = \{(\theta'', \mathrm{D})\}$. Thus, C.L is conditionally dominated in Σ^1. No other elimination is possible in Step 2, because U is a best reply for θ'' to C.L, D is a best reply for θ'' to C.R, and C.R is sequentially optimal given any $\bar{\mu}^2$ such that $\bar{\mu}_2(\theta', \mathrm{D}|\mathrm{D}) = 0$. Therefore,

$$\Sigma^2 = \mathrm{NCD}^2(\Theta \times S) = \left\{(\theta', \mathrm{U}), (\theta'', \mathrm{U}), (\theta'', \mathrm{D})\right\} \times \{\mathrm{C.R}\}.$$

Finally, in Step 3 we eliminate (θ'', U), because D is the unique best reply to C.R for type θ'':

$$\Sigma^3 = \mathrm{NCD}^3(\Theta \times S) = \left\{(\theta', \mathrm{U}), (\theta'', \mathrm{D})\right\} \times \{\mathrm{C.R}\}.$$

This leaves only one strategy for each type. Therefore, the elimination procedure stops at Step 3: $\Sigma^3 = \Sigma^\infty$. ▲

15.4.3 Rationalizability and Iterated Admissibility

As in Chapter 11, one can establish a tight connection (a generic equivalence) between conditional dominance and weak dominance in the strategic form, and move from there to establish a connection between strong rationalizability and a form of iterated admissibility. Recall that, for any subset $C_i \subseteq \Theta_i \times S_i$, we let $C_{i,\theta_i} \subseteq S_i$ denote its section at θ_i.

Definition 89. *Fix a nonempty Cartesian subset* $C \in \mathcal{C}$*, a player* $i \in I$*, and a pair* $(\theta_i, \bar{s}_i) \in C_i$*. Strategy* $\bar{s}_i$ *is* ***weakly dominated for type*** θ_i ***in*** C *if there is a mixed strategy* $\sigma_i \in \Delta(C_{i,\theta_i})$ *such that*

$$\begin{aligned} \forall (\theta_{-i}, s_{-i}) &\in C_{-i}, U_i(\theta_i, \theta_{-i}, \bar{s}_i, s_{-i}) \leq U_i(\theta_i, \theta_{-i}, \sigma_i, s_{-i}) \text{ and} \\ \exists \left(\bar{\theta}_{-i}, \bar{s}_{-i}\right) &\in C_{-i}, U_i\left(\theta_i, \bar{\theta}_{-i}, \bar{s}_i, \bar{s}_{-i}\right) < U_i\left(\theta_i, \bar{\theta}_{-i}, \sigma_i, \bar{s}_{-i}\right). \end{aligned}$$

Strategy $\bar{s}_i$ *is* ***admissible for*** θ_i ***in*** C *if it is* not *weakly dominated for* θ_i *in* C*. We let* $\mathrm{NWD}(C)$ *denote the set of profiles* $(\theta_i, s_i)_{i\in I} \in C$ *such that* s_i *is admissible in* C *for* θ_i *for every player* $i \in I$*.*

It is relatively straightforward to show that if a strategy $\bar{s}_i$ is conditionally dominated in C for a type θ_i then it is weakly dominated for θ_i in C. By essentially the same arguments used in Chapter 11, one can show that the converse is true for almost all parameterized payoff functions. To make this precise, note that appending to structure $\left\langle I, (\Theta_i, A_i, \mathcal{A}_i(\cdot))_{i\in I}\right\rangle$ a profile of parameterized payoff functions $(u_i)_{i\in I} \in \mathbb{R}^{\Theta\times Z\times I}$, we obtain a game with payoff uncertainty. Recall that a subset of a Euclidean space is deemed **negligible** if its closure has 0 Lebesgue measure.

Lemma 38. *Fix a finite structure* $\left\langle I, (\Theta_i, A_i, \mathcal{A}_i(\cdot))_{i\in I}\right\rangle$ *and a nonempty* $C \in \mathcal{C}$*. For all profiles of parameterized payoff functions* $(u_i)_{i\in I} \in \mathbb{R}^{\Theta\times Z\times I}$ *except at most a negligible set,* $\mathrm{NCD}(C) = \mathrm{NWD}(C)$*.*

Using Lemma 38, one can prove that strong rationalizability is generically equivalent to iterated admissibility, that is, the maximal iterated removal of pairs (θ_i, s_i) such that s_i is weakly dominated for θ_i in the restricted set of profiles not eliminated in previous steps. Furthermore, initial rationalizability is generically equivalent to the removal of all pairs (θ_i, s_i) such that s_i is weakly dominated for θ_i (in $\Theta \times S$) followed by the iterated elimination of pairs (θ_i, s_i) such that s_i is dominated for θ_i.

Theorem 58. *Fix a finite structure* $\left\langle I, (\Theta_i, A_i, \mathcal{A}_i(\cdot))_{i\in I}\right\rangle$*. For all profiles* $u \in \mathbb{R}^{\Theta\times Z\times I}$ *except at most a negligible set,*

$$\Sigma^n = \mathrm{NWD}^n(\Theta\times S)\,,\ \rho^n(\Theta\times S) = \left(\mathrm{ND}^{n-1}\circ\mathrm{NWD}\right)(\Theta\times S)$$

for all $n \in \mathbb{N}\cup\{\infty\}$*.*

Example 71. Go back again to the multistage game with payoff uncertainty of Examples 69 and 70. We can verify that its parameterized payoffs are "generic" in the sense of Theorem 58, that is, they do not belong to the negligible set for which strong rationalizability differs from iterated admissibility. Its (reduced) strategic form can be represented by a pair of matrices, one for each θ, where player 2 chooses the rows and only player 1 knows the true matrix (payoffs of player 1 are in **bold**).

θ'	U	D
S	**1**, v	**0**, 1
C.L	**1**, v	**0**, 3
C.R	**1**, v	**0**, 0

θ''	U	D
S	**1**, w	**0**, 1
C.L	**1**, w	**0**, 0
C.R	**1**, w	**2**, 3

Note, in particular, that S is weakly dominated by $\frac{1}{2}\delta_{\text{C.L}} + \frac{1}{2}\delta_{\text{C.R}}$. Then we obtain

$$\begin{aligned}
\text{NWD}^1(\Theta \times S) &= \{(\theta', \text{U}), (\theta'', \text{U}), (\theta'', \text{D})\} \times \{\text{C.L}, \text{C.R}\} = \Sigma^1, \\
\text{NWD}^2(\Theta \times S) &= \{(\theta', \text{U}), (\theta'', \text{U}), (\theta'', \text{D})\} \times \{\text{C.R}\} = \Sigma^2, \\
\text{NWD}^3(\Theta \times S) &= \{(\theta', \text{U}), (\theta'', \text{D})\} \times \{\text{C.R}\} = \Sigma^3, \\
\text{NWD}^\infty(\Theta \times S) &= \text{NWD}^3(\Theta \times S) = \Sigma^3 = \Sigma^\infty.
\end{aligned}$$

▲

15.4.4 Directed Rationalizability

Following up on the analysis of Section 8.2 of Chapter 8, we provide notions of "directed rationalizability" for multistage games with payoff uncertainty. That is, we provide extensions of the notions of initial and strong rationalizability to situations where some contextual restrictions on players' beliefs are transparent. Let us first clarify the restrictions we consider and what we mean by "transparent" in the analysis of multistage games, where conditional beliefs as the game unfolds play a crucial role.

It is natural to consider restrictions concerning *initial exogenous beliefs*, that is, assumptions about what each player i would believe about θ_{-i} at the beginning of the game if he were of type θ_i (cf. Section 8.3 of Chapter 8). Such assumptions can be represented by restricted belief sets $\bar{\Delta}_{\theta_i} \subseteq \Delta(\Theta_{-i})$. However, it may be plausible and interesting to consider also restrictions on initial beliefs about both types and behavior, and/or restrictions on beliefs conditional on observing some histories of moves. Therefore, we will be more flexible and posit, for each type θ_i of each player i, a (possibly) restricted set of CPSs $\Delta_{i,\theta_i} \subseteq \Delta^H(\Theta_{-i} \times S_{-i})$.

Given this, how should we extend the meaning of "transparency" to the present environment? One candidate is a notion of "initial transparency," according to which an event is initially transparent if it is true and *commonly believed at the beginning of the game.* This is the simplest extension of the transparency idea from static games to multistage games, and it is sufficient to motivate and interpret the notion of initial directed rationalizability presented below. However, initial transparency is not germane to the idea of strong rationalizability, which captures forward-induction reasoning. Thus, we consider a strengthening based on the notion of "common full belief." We say that a player **fully believes** an event E if he assigns probability 1 to E conditional on *every* history $h \in H$. We say that event E is **transparent** if E is true and there is common full belief of E.

Note that some events cannot be fully believed, hence they cannot be transparent, because they are contradicted by some nonterminal history. Here, however, we only consider the transparency of events concerning what players believe about the state of nature θ and (possibly) behavior. Events that exclusively concern what players believe cannot be contradicted by the observation of players' behavior. Therefore, it is always possible to assume that some features of players' beliefs are transparent.[8]

Let us start with initial directed rationalizability. This solution concept posits restrictions on players' beliefs and characterizes the information-dependent behavioral implications of (a) Rationality, (b) Common Initial Belief in Rationality, and (c) (initial) transparency of the posited restrictions. For each $i \in I$ and $\theta_i \in \Theta_i$, let $\Delta_{i,\theta_i} \subseteq \Delta^H(\Theta_{-i} \times S_{-i})$ denote the restricted, nonempty set of CPSs for information-type θ_i of player i. For a given profile $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$, define the **initial Δ-justification operator** $\rho_\Delta : \mathcal{C} \to \mathcal{C}$ as follows: for every $C \in \mathcal{C}$,

$$\begin{aligned}\rho_\Delta(C) &= \bigtimes_{i\in I} \left\{(\theta_i, s_i) : \exists \bar{\mu}^i \in \Delta^H_\varnothing(C_{-i}) \cap \Delta_{i,\theta_i}, s_i \in r_i\left(\bar{\mu}^i, \theta_i\right)\right\} \\ &= \bigtimes_{i\in I} \left(\bigcup_{\theta_i\in\Theta_i} \{\theta_i\} \times r_{i,\theta_i}\left(\Delta^H_\varnothing(C_{-i}) \cap \Delta_{i,\theta_i}\right)\right).\end{aligned}$$

(Recall that $\Delta^H_\varnothing(C_{-i})$ denotes the set of CPSs that initially assign

[8]Someone might object that a player's behavior may reveal something about his beliefs. But this objection would be based on a misunderstanding. Any inference about beliefs based on observed behavior relies on an hypothesized link between beliefs and behavior, such as that chosen actions must maximize subjective expected utility. Thus, observations about behavior can only contradict the *joint hypothesis* that a player's beliefs have some features *and* such link holds.

probability 1 to C_{-i}.)

For a fixed Δ, self-map ρ_Δ is *monotone*, because $E_{-i} \subseteq F_{-i}$ implies $\Delta^H_\varnothing(E_{-i}) \subseteq \Delta^H_\varnothing(F_{-i})$. Thus, as we iterate ρ_Δ starting from the whole set of profiles of types and strategies, we obtain the (weakly) decreasing sequence of subsets $(\rho^n_\Delta(\Theta \times S))_{n\in\mathbb{N}}$. In particular, $\rho^1_\Delta(\Theta \times S) = \rho_\Delta(\Theta \times S)$ contains all the profiles $(\theta_i, s_i)_{i\in I}$ such that, for each i, s_i is justified for θ_i by some CPS $\bar{\mu}^i \in \Delta_{i,\theta_i}$. Let

$$\rho^\infty_\Delta(\Theta \times S) = \bigcap_{n\in\mathbb{N}} \rho^n_\Delta(\Theta \times S).$$

Definition 90. *Fix a profile $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$ of restricted (nonempty) sets of CPSs. A profile of types and strategies $(\theta_i, s_i)_{i\in I}$ is **initially Δ-rationalizable** if $(\theta_i, s_i)_{i\in I} \in \rho^\infty_\Delta(\Theta \times S)$.*

As a matter of fact, whether the belief restrictions are assumed to be transparent or only *initially* transparent does not make any difference here, because initial rationalizability is silent on what players would believe if they were surprised by the realization of an unexpected history.

Similarly to Section 8.2 of Chapter 8, for each $C \in \mathcal{C}$, map $\Delta \mapsto \rho_\Delta(C)$ is monotone in the following sense: if $\Delta_{i,\theta_i} \subseteq \Delta'_{i,\theta_i}$ for every i and θ_i, written $\Delta \subseteq \Delta'$, then $\rho_\Delta(C) \subseteq \rho_{\Delta'}(C)$. By monotonicity of ρ_Δ, this implies that the initial directed rationalizability map $\Delta \mapsto \rho^\infty_\Delta(\Theta \times A)$ is monotone.

Remark 51. *Fix two profiles of restrictions, Δ and Δ'. If $\Delta \subseteq \Delta'$, then $\rho^n_\Delta(\Theta \times S) \subseteq \rho^n_{\Delta'}(\Theta \times S)$ for every $n \in \mathbb{N} \cup \{\infty\}$.*

As argued in Section 8.2 of Chapter 8, the special case of restrictions that only concern exogenous beliefs deserves attention. We say that $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$ is a profile of **restrictions on exogenous beliefs** if, for each $i \in I$ and each $\theta_i \in \Theta_i$, there is a nonempty set $\bar{\Delta}_{i,\theta_i} \subseteq \Delta(\Theta_{-i})$ such that

$$\Delta_{i,\theta_i} = \left\{\bar{\mu}^i \in \Delta^H(\Theta_{-i} \times S_{-i}) : \mathrm{marg}_{\Theta_{-i}}\bar{\mu}^i(\cdot|\varnothing) \in \bar{\Delta}_{i,\theta_i}\right\},$$

where $\mathrm{marg}_{\Theta_{-i}}\bar{\mu}^i(\cdot|\varnothing)$ denotes the initial marginal belief about θ_{-i} derived from $\bar{\mu}^i$, that is, $\mathrm{marg}_{\Theta_{-i}}\bar{\mu}^i(\theta_{-i}|\varnothing) = \bar{\mu}^i(\{\theta_{-i}\} \times S_{-i}|\varnothing)$ for each $\theta_{-i} \in \Theta_{-i}$.

Given the monotonicity of ρ_Δ, one can prove the following extension of Theorem 31 in Chapter 8 to the present multistage framework.

Theorem 59. *Consider a profile* $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$ *of restrictions on* exogenous *beliefs. Then*

$$\mathrm{proj}_{\Theta}\rho_{\Delta}^{n}(\Theta\times S)=\Theta$$

for all $n\in\mathbb{N}\cup\{\infty\}$*, and*

$$\rho_{\Delta}\left(\rho_{\Delta}^{\infty}(\Theta\times S)\right)=\rho_{\Delta}^{\infty}(\Theta\times S).$$

Furthermore, for every $C\in\mathcal{C}$*,*

$$C\subseteq\rho_{\Delta}(C)\Rightarrow C\subseteq\rho_{\Delta}^{\infty}(\Theta\times S).$$

As noticed in Chapter 8, $\mathrm{proj}_{\Theta}\rho_{\Delta}^{\infty}(\Theta\times S)=\Theta$ implies that, for every type θ_i of each player i, the set of initially Δ-rationalizable strategies for type θ_i is nonempty.

The following example illustrates initial directed rationalizability. Although we neglected the residual uncertainty Θ_0 to simplify our abstract notation, we reintroduce it in this particular case.

Example 72. Consider a leader-follower game, where player 1 (she) is the leader and player 2 (he) is the follower. Each player has to choose between two restaurants, one on the left-hand side of the main street[9] and one on the right-hand side. Thus, $A_i=\{\ell_i,r_i\}$, $S_1=A_1$, and $S_2=\{\ell_2,r_2\}^{\{\ell_1,r_1\}}$. The unknown state $\theta_0\in\Theta_0=\{\theta_0^{\ell},\theta_0^{r}\}$ determines whether the best restaurant is on the left or on the right. Each player i has private information $\theta_i\in\Theta_i=\{\theta_i^{\ell},\theta_i^{r}\}$, which is believed to be correlated to the quality of restaurants, but the payoff depends only on the chosen restaurant and on θ_0:

$$u_i(\theta,(a_1,a_2))=\begin{cases}1, & \text{if } a_i=\ell_i \text{ and } \theta_0=\theta_0^{\ell}\text{, or } a_i=r_i \text{ and } \theta_0=\theta_0^{r},\\ 0, & \text{otherwise.}\end{cases}$$

It is transparent that player 1 thinks that her private information signals quality, and that player 2 thinks that player 1 is an expert who has more accurate information in the following sense: if he knew θ_1 he would deem the state θ_0 "signalled" by θ_1 more likely than the other *independently of his private information* θ_2. Formally, we consider the following sets

[9] Given a "canonical direction."

of CPSs featuring restrictions on exogenous initial beliefs: using obvious abbreviations for initial, marginal, and conditional beliefs,[10] for each "street side" $x \in \{\ell, r\}$,

$$\Delta_{1,\theta_1^x} = \left\{ \bar{\mu}^1 : \bar{\mu}^1 (\theta_0^x) > \frac{1}{2} \right\},$$

and for each $\theta_2 \in \left\{\theta_2^\ell, \theta_2^r\right\}$,

$$\Delta_{2,\theta_2} = \left\{ \bar{\mu}^2 : \forall x \in \{\ell, r\}, \bar{\mu}^2 (\theta_0^x, \theta_1^x) > 0, \bar{\mu}^2 (\theta_0^x | \theta_1^x) > \frac{1}{2} \right\}.$$

Then, initial Δ-rationalizability is computed in two steps: it implies that player 1 follows her private information and player 2 follows the choice of player 1 independently of his own private information:

$$\rho_\Delta^2 (\Theta \times S) = \Theta_0 \times \left\{ \left(\theta_1^\ell, \ell_1\right), (\theta_1^r, r_1) \right\} \times \{(\theta_2, s_2) : s_2 (\ell_1) = \ell_2, s_2 (r_1) = r_2\}.$$

We leave the formal proof as an exercise. The key insight is that, given the belief restrictions, if the follower is initially certain of the rationality of the leader, then he initially assigns strictly positive probability to each action of the leader; hence, he cannot be "surprised" and, for each action of the leader, he believes that this action reveals her private information. Since he "trusts" the private information of the leader more than his own, he follows suit. ▲

We now consider strong directed rationalizability. This solution concept posits restrictions on players' beliefs and characterizes the behavioral implications of (a) Rationality, (b) transparency of the posited restrictions, and (c) Common Strong Belief in (a)-(b). Fix a profile $\Delta = (\Delta_{i,\theta_i})_{i \in I, \theta_i \in \Theta_i}$ of (nonempty) restricted belief sets. For each $i \in I$ and $\theta_i \in \Theta_i$, and for any event $E_{-i} \subseteq \Theta_{-i} \times S_{-i}$, we let $\Delta_{\mathrm{sb},\theta_i}^H (E_{-i})$ denote the **set of CPSs in Δ_{i,θ_i} that strongly believe** $E_{-i} \subseteq \Theta_{-i} \times S_{-i}$, that is,

$$\Delta_{\mathrm{sb},\theta_i}^H (E_{-i}) = \Delta_{\mathrm{sb}}^H (E_{-i}) \cap \Delta_{i,\theta_i}.$$

Note that $\Delta_{\mathrm{sb},\theta_i}^H (E_{-i}) = \Delta_{i,\theta_i}$ if $E_{-i} = \Theta_{-i} \times S_{-i}$.

[10]For example, $\bar{\mu}^1 (\theta_0) = \bar{\mu}^1 (\{\theta_0\} \times \Theta_2 \times S_2 | \varnothing)$ and $\bar{\mu}^2 (\theta_0^x | \theta_1^x) = \frac{\bar{\mu}^2 \left(\left\{\theta_0^x, \theta_1^x\right\} \times S_1 | \varnothing\right)}{\bar{\mu}^2 \left(\Theta_0 \times \left\{\theta_1^x\right\} \times S_1 | \varnothing\right)}$.

Definition 91. *Fix a profile* $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$ *of restricted sets of CPSs. Consider the following elimination procedure:*
*(**Step** $n = 0$) For each $i \in I$, let $\Sigma_i^{\Delta,0} = \Theta_i \times S_i$, $\Sigma_{-i}^{\Delta,0} = \Theta_{-i} \times S_{-i}$, and $\Sigma^{\Delta,0} = \Theta \times S$.*
*(**Step** $n > 0$) For each $i \in I$ and $\theta_i \in \Theta_i$, let*

$$\Delta_{i,\theta_i}^n = \bigcap_{m=0}^{n-1} \Delta_{\mathrm{sb},\theta_i}^H \left(\Sigma_{-i}^{\Delta,m}\right),$$

and for each $i \in I$ let

$$\begin{aligned}\Sigma_i^{\Delta,n} &= \left\{(\theta_i, s_i) \in \Theta_i \times S_i : \exists \bar{\mu}^i \in \Delta_{i,\theta_i}^n, s_i \in r_i\left(\bar{\mu}^i, \theta_i\right)\right\} \\ &= \bigcup_{\theta_i\in\Theta_i} \{\theta_i\} \times r_{i,\theta_i}\left(\Delta_{i,\theta_i}^n\right).\end{aligned}$$

Also, let $\Sigma_{-i}^{\Delta,n} = \times_{i\neq i} \Sigma_j^{\Delta,n}$ and $\Sigma^{\Delta,n} = \times_{i\in I} \Sigma_i^{\Delta,n}$.
*A profile of types and strategies $(\theta_i, s_i)_{i\in I}$ is **strongly Δ-rationalizable** if $(\theta_i, s_i)_{i\in I} \in \Sigma^{\Delta,\infty} = \bigcap_{n>0} \Sigma^{\Delta,n}$, that is, if $(\theta_i, s_i) \in \Sigma_i^{\Delta,\infty} = \bigcap_{n>0} \Sigma_i^{\Delta,n}$ for each player $i \in I$.*

It is also convenient to keep track of the (n-) strongly Δ-rationalizable strategies of type θ_i letting

$$S_i^{\Delta,n}(\theta_i) = \left(\Sigma_i^{\Delta,n}\right)_{\theta_i} = \left\{s_i \in S_i : \exists \bar{\mu}^i \in \Delta_{i,\theta_i}^n, s_i \in r_i\left(\bar{\mu}^i, \theta_i\right)\right\}$$

(the section at θ_i of $\Sigma_i^{\Delta,n} \subseteq \Theta_i \times S_i$). With this,

$$S^{\Delta,n}(\theta) = \bigtimes_{i\in I} S_i^{\Delta,n}(\theta_i)$$

denotes the set of (n-) strongly rationalizable strategy profiles at state of nature θ. Note that the sequence $\left(\Sigma^{\Delta,n}\right)_{n\in\mathbb{N}}$ is (weakly) decreasing because, by definition, $\Delta_{i,\theta_i}^n \subseteq \Delta_{i,\theta_i}^{n-1}$ and thus $r_{i,\theta_i}\left(\Delta_{i,\theta_i}^n\right) \subseteq r_{i,\theta_i}\left(\Delta_{i,\theta_i}^{n-1}\right)$ for every $i \in I$ and $\theta_i \in \Theta_i$. When $\Delta_{i,\theta_i} = \Delta^H(\Theta_{-i} \times S_{-i})$ for all i and θ_i, that is, without belief restrictions, we get "plain" strong rationalizability and therefore we suppress the Δ superscript from out notation.

The proof of the following result is left to the reader as an exercise.

Remark 52. *The first steps of the strong and initial directed rationalizability procedures coincide, but strong directed rationalizability refines initial directed rationalizability:* $\Sigma^{\Delta,1} = \rho^1_\Delta(\Theta \times S)$ *and* $\Sigma^{\Delta,n} \subseteq \rho^n_\Delta(\Theta \times S)$ *for every* $n \geq 2$. *If* $\Delta = (\Delta_{i,\theta_i})_{i\in I,\theta_i\in\Theta_i}$ *is a profile of restrictions on* exogenous *beliefs, then*

$$\mathrm{proj}_\Theta \Sigma^{\Delta,n} = \Theta$$

for all $n \in \mathbb{N} \cup \{\infty\}$*, which implies that, for every type* θ_i *of each player* i*, the set of* n*-strongly* Δ*-rationalizable strategies* $S_i^{\Delta,n}(\theta_i)$ *is nonempty, that is,* $S^{\Delta,n}(\theta) \neq \emptyset$.

The following example illustrates strong directed rationalizability.

Example 73. Player 1 is a young unemployed person who has to decide whether to get an education (e) or not (n). There is only one kind of job, with a fixed wage W, and education is *not* necessary to apply for the job. Furthermore, education does *not* enhance productivity. Player 2 is an employer who has to decide whether to hire player 1 or not, after observing his decision.

Player 1 can be a high type θ^H or a low type θ^L. Higher types are more productive, yield more additional revenue if hired ($R^H > R^L$), and incur a lower subjective cost to get an education ($c^H < c^L$). Player 1 knows his type. We assume that $0 < c^H < W$ and $R^H = R > W > R^L = 0$. This game can be represented as in Figure 15.3.

We let $x.y$ denote the strategy of the employer who selects action $x \in \{j, n\}$ at history (n) and action $y \in \{J, N\}$ at history (e). Let us first consider strong rationalizability. Note that, for type θ^{L}, action n is dominant. Strong rationalizability yields

$$\left\{\left(\theta^{\mathrm{H}}, \mathrm{e}\right), \left(\theta^{\mathrm{H}}, \mathrm{n}\right), \left(\theta^{\mathrm{L}}, \mathrm{n}\right)\right\} \times \{j.J, n.J\}.$$

Indeed, at the second step strategies $j.N$ and $n.N$ are eliminated: strong belief in rationality implies that, if player 1 gets an education, then player 2 must conclude that his type is high; this entails that J is optimal.

Assume now that (it is transparent that) at the beginning of the game player 2 assigns a low probability to the high type. Specifically, every CPS $\bar{\mu}^2 \in \Delta_2$ is such that

$$\bar{\mu}^2\left(\left\{\theta^{\mathrm{H}}\right\} \times \{\mathrm{n}, \mathrm{e}\} | \varnothing\right) < \frac{W}{R}.$$

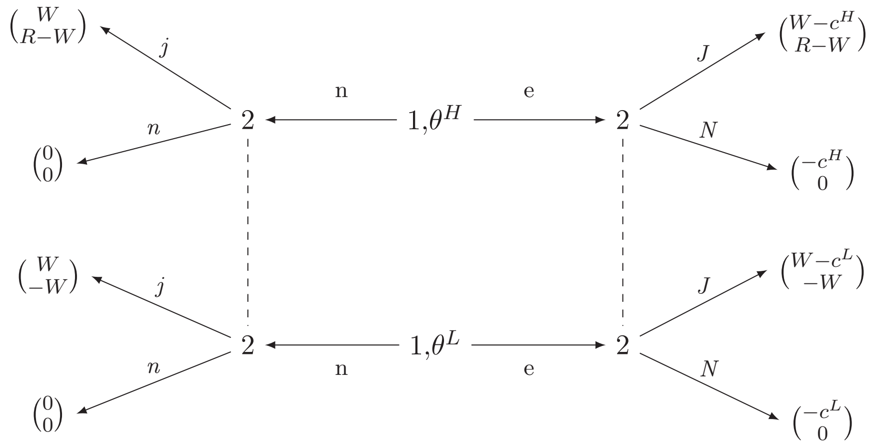


Figure 15.3: A "job market" game.

(Note that $\frac{W}{R} \in (0,1)$.) With this restriction on the beliefs of player 2 and no restriction on the beliefs of any type of player 1, directed rationalizability is computed in three steps:

$$\Sigma^{\Delta,3} = \left\{ \left(\theta^{\mathrm{H}}, \mathrm{e}\right), \left(\theta^{\mathrm{L}}, \mathrm{n}\right) \right\} \times \{n.J\}.$$

At the first step, $\left(\theta^{\mathrm{L}}, \mathrm{e}\right)$ is eliminated because, as noted above, action n is dominant for type θ^{L}. Strategy $j.J$ is also eliminated at the first step, because at least one of the two beliefs after observing e and n must be derived from $\bar{\mu}^2(\cdot|\varnothing)$ by conditioning; hence at least one of the two conditional beliefs cannot assign probability higher than $\frac{W}{R}$ to θ^{H}. The same argument as the one used for strong rationalizability shows that strategies $j.N$ and $n.N$ are eliminated at the second step. Since both $j.J$ and $j.N$ were eliminated, at the third step type θ^{H} chooses action e. We leave the formal proof as an exercise. ▲

We also note a rather subtle point: unlike for initial directed rationalizability, the set of strong directed rationalizable strategy profiles for any given state θ and even the set of paths of play induced by such profiles *are not monotone in the beliefs restrictions.* In particular, strong Δ-rationalizability may not refine strong rationalizability. This follows from the fact that the strong belief set $\Delta^{H}_{\mathrm{sb}}(E_{-i})$ is not monotone in E_{-i}

(see the discussion in Section 11.3 of Chapter 11). The following example illustrates this point.

Example 74. Consider the following game. Player 1 moves first, choosing between In and Out. After Out, the game ends. After In, player 2 moves, choosing between $left$, $center$, and $right$, then the game ends. Player 1 is privately informed of a payoff-relevant parameter $\theta_1 \in \{0,1\}$. Players' payoffs depend on the terminal history and on θ_1 as follows (see Figure 15.4).

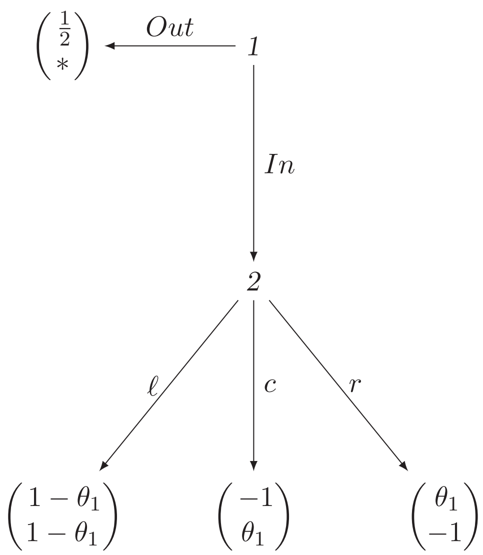


Figure 15.4: A leader-follower game with asymmetric information.

We first analyze the game with strong rationalizability (that is, without belief restrictions), which can be computed by iterated conditional dominance. Since r is conditionally dominated, it is eliminated in Step 1. Given this, in Step 2, In is eliminated for type $\theta_1 = 1$. In Step 3, player 2 rationalizes In assuming that it was chosen by type $\theta_1 = 0$ (forward induction), therefore c is eliminated. Finally, in Step 4, Out is eliminated for type $\theta_1 = 0$. To conclude, Out is the only strongly rationalizable strategy for type $\theta_1 = 1$, In is the only strongly rationalizable strategy for type $\theta_1 = 0$, and ℓ is the only strongly rationalizable strategy for player 2:

$$\Sigma^\infty = \Sigma^4 = \{(0, In), (1, Out)\} \times \{\ell\}.$$

Next we consider directed rationalizability assuming that (only) the following is transparent: player 2 becomes certain of type $\theta_1 = 1$ upon observing In, that is,

$$\Delta_2 = \left\{\bar{\mu}^2 \in \Delta^H(\Theta_1 \times S_1) : \mu_2((1, In) \,|(In)) = 1\right\}.$$

Given this, both ℓ and r are eliminated in Step 1 of directed rationalizability. Then, in Step 2 In is eliminated for *both* types of player 1. *This makes it impossible to rationalize* In in Step 3. Hence, the only strongly Δ-rationalizable strategy of both types of player 1 is Out, and the only strongly Δ-rationalizable strategy of player 2 is c:

$$\Sigma^{\Delta,\infty} = \Sigma^{\Delta,3} = \{(0, Out), (1, Out)\} \times \{c\}.$$

▲

Note that in Example 74 we considered a restriction on beliefs conditional on observed behavior, which are *endogenous*. Indeed, one can show that, *when we consider only restrictions on exogenous beliefs*, *the set of strongly* Δ*-rationalizable paths is monotone in* Δ, despite the non-monotonicity of strong belief.

Here, as in the example above, we just focus on the comparison between some profile Δ of subsets of CPSs and the case of no restrictions, that is, strong rationalizability. The proof of the result, which is beyond the scope of this textbook, can be easily adapted to obtain the more general path-monotonicity claim.[11]

Theorem 60. *Fix a profile* $\Delta = (\Delta_{i,\theta_i})_{i,\in I,\theta_i\in\Theta_i}$ *of restrictions on* exogenous *beliefs. Then, for all steps* $n > 0$ *and states* $\theta \in \Theta$, $\emptyset \neq \zeta\left(S^{\Delta,n}(\theta)\right) \subseteq \zeta\left(S^n(\theta)\right)$; *therefore, for each* $(\theta, s) \in \Sigma^{\Delta,\infty} \neq \emptyset$, *there exists* $s' \in S$ *such that* $(\theta, s') \in \Sigma^\infty$ *and* $\zeta(s) = \zeta(s')$.

The assumption that the belief restrictions only apply to exogenous beliefs is tight, as the example above shows. In the literature, there are many examples of strong directed rationalizability with restrictions on the initial beliefs about the co-player's *strategy* (also in games with complete information) yielding non-strongly-rationalizable outcomes.[12]

[11] See Battigalli and Catonini [10] for the complex proof of this result.
[12] See, e.g., Battigalli and Friedenberg [15] and Catonini [31].

The following example illustrates how the monotonicity result holds only for the possible paths of play, whereas the addition of transparent restrictions on exogenous beliefs may significantly change the implications about the reactions to moves that deviate from the paths allowed by the solution.

Example 75. Consider again a signaling game: Γ is a leader-follower game where only the leader (pl. 1) has private information. In particular:

$$\begin{aligned} \Theta_1 &= \{x, y, z\}, \Theta_2 = \{\bar{\theta}_2\}, \text{ (a singleton)}, \\ H &= \{\varnothing, (\ell), (r)\}, Z = \{(\ell, a), (\ell, b), (r, c), (r, d), (r, e)\}, \bar{H} = H \cup Z, \\ \mathcal{A}_1(\varnothing) &= \{\ell, r\}, \mathcal{A}_2(\ell) = \{a, b\}, \mathcal{A}_2(r) = \{c, d, e\}, \end{aligned}$$

and the type-dependent payoff functions $u_i : \Theta \times Z \to \mathbb{R}$ are described by the following tables:

$u_1(\cdot, \ell, \cdot), u_2(\cdot, \ell, \cdot)$	a	b
$\theta_1 = x$	3 1	1 0
$\theta_1 = y$	1 0	1 1
$\theta_1 = z$	3 1	1 0

$u_1(\cdot, r, \cdot), u_2(\cdot, r, \cdot)$	c	d	e
$\theta_1 = x$	0 0	0 0	0 1
$\theta_1 = y$	0 0	0 1	3 0
$\theta_1 = z$	0 1	2 0	2 0

Since player 2 is uninformed and inactive in the first stage, $\Theta_2 \times S_2$ is isomorphic to S_2 and $\Delta^H(\Theta_2 \times S_2)$ is isomorphic to $\Delta(S_2)$ (by the chain rule and $S_2(\ell) = S_2(r) = S_2$). Without belief restrictions, "plain" strong rationalizability yields:

$$\begin{aligned} \Sigma_1^1 &= \{(x, \ell), (y, \ell), (y, r), (z, \ell), (z, r)\} \text{ (thus, } S_1^1(x) = \{\ell\}\text{)}, \Sigma_2^1 = S_2; \\ \Sigma_1^2 &= \Sigma_1^1, \Sigma_2^2 = \{a, b\} \times \{c, d\}; \\ \Sigma_1^3 &= \{(x, \ell), (y, \ell), (z, \ell), (z, r)\}, \text{ (thus, } S_1^3(y) = \{\ell\}\text{)}, \Sigma_2^3 = \Sigma_2^2; \\ \Sigma_1^4 &= \Sigma_1^3, \Sigma_2^4 = \{a.c, b.c\}; \\ \Sigma_1^5 &= \Theta_1 \times \{\ell\} \ (\forall \theta_1 \in \Theta_1, S_1^5(\theta_1) = \{\ell\}), \Sigma_2^5 = \Sigma_2^4; \\ \Sigma_1^\infty &= \Theta_1 \times \{\ell\}, \Sigma_2^\infty = \{a.c, b.c\}. \end{aligned}$$

Thus, for all $\theta \in \Theta$,

$$\zeta(S^\infty(\theta)) = \{(\ell, a), (\ell, b)\}.$$

Now we *add restrictions on exogenous beliefs*. Since Θ_2 is a singleton, we can only have restrictions on the exogenous beliefs of player 2. In particular, we assume that she is *a priori certain that* $\theta_1 = z$. Thus the restricted belief set is

$$\Delta_2 = \left\{\mu_2 \in \Delta^H(\Theta_1 \times S_1) : \left(\text{marg}_{\Theta_1}\mu_2(\cdot|\varnothing)\right)(z) = 1\right\}.$$

With this,

$$\begin{aligned}
\Sigma_1^{\Delta,1} &= \{(x,\ell),(y,\ell),(y,r),(z,\ell),(z,r)\}, \ \Sigma_2^{\Delta,1} = \{a.c, b.c, a.d, a.e\};\\
\Sigma_1^{\Delta,2} &= \{(x,\ell),(y,\ell),(y,r),(z,\ell)\}, \ \Sigma_2^{\Delta,2} = \{a.c, b.c, a.d\};\\
\Sigma_1^{\Delta,3} &= \Theta_1 \times \{\ell\}, \ \Sigma_2^{\Delta,3} = \{a.d\};\\
\Sigma_1^{\Delta,\infty} &= \Theta_1 \times \{\ell\}, \ \Sigma_2^{\Delta,\infty} = \{a.d\}.
\end{aligned}$$

Thus $\zeta\left(S^{\Delta,\infty}(\theta)\right) = \{(\ell,a)\}$ for all $\theta \in \Theta$, but $\Sigma_2^{\Delta,\infty} \not\subseteq \Sigma_2^{\infty}$, actually $\Sigma_2^{\Delta,\infty} \cap \Sigma_2^{\infty} = \emptyset$. ▲

15.5 Multistage Bayesian Games

Structure $\left\langle I, \Theta_0, (\Theta_i, A_i, \mathcal{A}_i(\cdot), u_i)_{i\in I}\right\rangle$ does not specify the exogenous interactive beliefs of the players, i.e., their beliefs about each other's types. Therefore, it is not rich enough to define traditional notions of equilibrium, e.g. Bayesian equilibrium and refinements thereof.

As in the case of static games, we should add exogenous interactive beliefs structures *à la* Harsanyi to carry out a traditional analysis of games with incomplete information. Since the dynamics involve additional complications of strategic analysis, we *simplify* the *interactive beliefs* aspect: we assume that the set of types *à la* Harsanyi T_i coincides with (more precisely, is isomorphic to) Θ_i. To emphasize this assumption we use the phrase "simple Bayesian game" (cf. Section 8.4 of Chapter 8). Compared to the previous analysis of rationalizability, we allow for the presence of residual uncertainty, that is, we allow for the possibility that $|\Theta_0| > 1$.[13]

A multistage Bayesian game with information-types, or **simple multistage Bayesian game** (with observed actions), is a (*finite*) structure

$$\Gamma = \left\langle I, \Theta_0, \left(\Theta_i, A_i, \mathcal{A}_i(\cdot), u_i, (p_i(\cdot|\theta_i))_{\theta_i\in\Theta_i}\right)_{i\in I}\right\rangle$$

[13] In the analysis of rationalizability, we assumed distributed knowledge of the θ parameter (that is, $|\Theta_0| = 1$) and neglected Θ_0 merely for notational simplicity.

where

- $\langle I, \Theta_0, (\Theta_i, A_i, \mathcal{A}_i(\cdot), u_i)_{i\in I}\rangle$ is a (*finite*) game with payoff uncertainty
- for every $i \in I$ and $\theta_i \in \Theta_i$, $p_i(\cdot|\theta_i) \in \Delta(\Theta_0 \times \Theta_{-i})$ is the *initial* exogenous belief of information-type θ_i.

Alternatively, we could specify an ex ante belief, or "prior," $P_i \in \Delta(\Theta)$ for every player i such that $P_i(\theta_i) = P_i(\Theta_0 \times \{\theta_i\} \times \Theta_{-i}) > 0$ for each $\theta_i \in \Theta_i$, and derive the initial belief of each type θ_i as a conditional probability, that is,

$$p_i(\theta_0, \theta_{-i}|\theta_i) = \frac{P_i(\theta_0, \theta_i, \theta_{-i})}{P_i(\theta_i)} \tag{15.5.1}$$

for all $(\theta_0, \theta_{-i}) \in \Theta_0 \times \Theta_{-i}$.[14]

Note also that for every profile of initial beliefs of types $(p_i(\cdot|\theta_i))_{\theta_i\in\Theta_i}$ we can find a class of ex ante beliefs that generate them according to (15.5.1): just pick any strictly positive measure $\lambda_i \in \Delta^\circ(\Theta_i)$ and compute the ex ante probability of each $(\theta_0,\theta_i, \theta_{-i})$ as follows:

$$P_{\lambda_i}(\theta_0,\theta_i, \theta_{-i}) = p_i(\theta_0, \theta_{-i}|\theta_i)\lambda_i(\theta_i).$$

Since λ_i is strictly positive,

$$\begin{aligned} P_{\lambda_i}(\theta_i) &= \sum_{(\theta_0,\theta_{-i})\in\Theta_0\times\Theta_{-i}} p_i(\theta_0, \theta_{-i}|\theta_i)\lambda_i(\theta_i) \\ &= \lambda_i(\theta_i) \sum_{(\theta_0,\theta_{-i})\in\Theta_0\times\Theta_{-i}} p_i(\theta_0, \theta_{-i}|\theta_i) = \lambda_i(\theta_i) > 0. \end{aligned}$$

Therefore P_{λ_i} is an ex ante belief that generates $(p_i(\cdot|\theta_i))_{\theta_i\in\Theta_i}$, that is,

$$\frac{P_{\lambda_i}(\theta_0, \theta_i, \theta_{-i})}{P_{\lambda_i}(\theta_i)} = \frac{p_i(\theta_0, \theta_{-i}|\theta_i)\lambda_i(\theta_i)}{\lambda_i(\theta_i)} = p_i(\theta_0, \theta_{-i}|\theta_i)$$

[14] Game theorists use the term "prior" somewhat liberally to mean any kind of subjective probability measure assigned before receiving any kind of information. In Bayesian inferential statistics, instead, a **prior** is a subjective probability measure over statistical models, e.g., urn models. In the present context, the priors of Bayesian statistics are the initial beliefs of types $p_i(\cdot|\theta_i)$, because they are the starting points of the Bayesian updating process.

for all $(\theta_0, \theta_i, \theta_{-i})$. If there are at least two information-types for player i, then there is a continuum of ex ante beliefs that generate $(p_i(\cdot|\theta_i))_{\theta_i \in \Theta_i}$. If instead there is only one type of player i, say $\bar{\theta}_i$, then there is no essential difference between i's ex ante belief and i's belief given his (unique) type $\bar{\theta}_i$, because $\Theta = \Theta_0 \times \{\bar{\theta}_i\} \times \Theta_{-i}$ is isomorphic to $\Theta_0 \times \Theta_{-i}$.

We keep writing $p_i(\cdot|\theta_i)$ for the initial exogenous belief of type θ_i—as if there were an ex ante stage—for several reasons. First, it is traditional in much of the applied-theory literature. Second, in some applications, it does make sense to think of an ex ante stage in which all player have the same information (i.e., they just know that the set of possible states of nature is Θ). Third, many formulas involving conditional probabilities are more easily readable under the "ex-ante interpretation" that θ is determined at random and each i observes the realization θ_i (cf. Section 8.6 in Chapter 8).

Simple multistage Bayesian games are traditionally analyzed by means of a refined notion of Bayesian equilibrium, just like multistage games with observed actions and complete information are traditionally analyzed by computing subgame perfect Nash equilibria. The following sections are devoted to such traditional analysis. However, we point out that traditional equilibrium analysis is not a must. For example, we can use directed rationalizability assuming that there are transparent restrictions on exogenous beliefs determined by the belief maps $\theta_i \mapsto p_i(\cdot|\theta_i)$. Alternatively, one can determine the ex ante strategic form of the Bayesian game and apply iterated admissibility. Both approaches can be very insightful.

15.6 Bayesian Equilibrium

We now show how the notion of Bayesian equilibrium given for the static case (see Section 8.4 of Chapter 8) can be naturally extended to simple multistage Bayesian games.

Fix a simple multistage Bayesian game Γ. A **choice rule**, or (extended) decision function of player i is an element of the set $S_i^{\Theta_i}$ that can be interpreted as the deterministic conjecture of player $j \neq i$ about i's (history-dependent) behavior as a function of his information-type θ_i. For each $i \in I$, we let

$$\Sigma_i = S_i^{\Theta_i} = \left(\bigtimes_{h \in H} \mathcal{A}_i(h) \right)^{\Theta_i}$$

denote the set of all choice rules of player i.[15] A generic element of Σ_i is denoted by σ_i (not to be confused with a mixed strategy). Going back to the definition of Bayesian equilibrium of Chapter 8 and replacing actions a_i with strategies s_i, we obtain the following:

Definition 92. *A (pure)* ***Bayesian equilibrium*** *of* Γ *is a profile of choice rules* $(\sigma_i)_{i \in I} \in \bigtimes_{i \in I} \Sigma_i$ *such that, for all* $i \in I$, $\theta_i \in \Theta_i$, *and* $s_i \in S_i$,

$$\begin{aligned} & \sum_{(\theta_0, \theta_{-i}) \in \Theta_0 \times \Theta_{-i}} p_i(\theta_0, \theta_{-i} | \theta_i) \, u_i(\theta_0, \theta_i, \theta_{-i}, \zeta(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i}))) \\ \geq & \sum_{(\theta_0, \theta_{-i}) \in \Theta_0 \times \Theta_{-i}} p_i(\theta_0, \theta_{-i} | \theta_i) \, u_i(\theta_0, \theta_i, \theta_{-i}, \zeta(s_i, \sigma_{-i}(\theta_{-i}))) , \end{aligned}$$

where $\sigma_{-i}(\theta_{-i}) = (\sigma_j(\theta_j))_{j \neq i}$.

The definition can be extended to allow for randomization, which requires a suitable notion of randomized choice rules. As in multistage games with complete information, we can think of players choosing pure strategies at random (given their type), or players who randomize over actions as the play unfolds. The latter notion of randomization is more convenient. A behavioral (randomized) choice rule, or **extended behavior strategy** of player i is an element of the set $B_i^{\Theta_i}$, where $B_i = \bigtimes_{h \in H} \Delta(\mathcal{A}_i(h))$ denotes the set of behavior strategies. A generic element of the set $B_i^{\Theta_i}$ is denoted by

$$(\beta_i(\cdot|\theta_i, \cdot))_{\theta_i \in \Theta_i} = (\beta_i(\cdot|\theta_i, h))_{\theta_i \in \Theta_i, h \in H} ,$$

where $\beta_i(a_i|\theta_i, h)$ is the probability of choosing a_i given θ_i and h.

If we define Γ using "priors" $P_i \in \Delta(\Theta)$, we can characterize the set of Bayesian equilibria of Γ as the Nash equilibria of the *ex ante* strategic form. For each $i \in I$, define $U_i : \bigtimes_{j \in I} \Sigma_j \to \mathbb{R}$ by

$$U_i(\sigma_i, \sigma_{-i}) = \sum_{\theta \in \Theta} P_i(\theta) u_i(\theta, \zeta(\sigma_i(\theta_i), \sigma_{-i}(\theta_{-i}))) ,$$

[15] Note that the sets Σ_i^k of Section 15.4 are subsets of $\Theta \times S_i$, not subsets of $S_i^{\Theta_i} = \Sigma_i$. Given the different context, the similarity of symbols should cause no confusion.

where $P_i \in \Delta(\Theta)$ is any ex ante belief generating $(p_i(\cdot|\theta_i))_{\theta_i \in \Theta_i}$. The **ex ante strategic form** of Γ is the static game

$$\mathcal{AS}(\Gamma) = \langle I, (\Sigma_i, U_i)_{i \in I} \rangle .$$

It can be checked that *a profile* σ *is a Bayesian equilibrium of* Γ *if and only if it is a Nash equilibrium of* $\mathcal{AS}(\Gamma)$. An analogous result holds for randomized Bayesian equilibria.

As anticipated above, it is convenient to represent randomized Bayesian equilibria of multistage Bayesian games as profiles of extended behavior strategies $\beta \in \times_{i \in I} B_i^{\Theta_i}$ rather than profiles of (type-dependent) mixed strategies.[16] We can interpret extended behavior strategies as conditional distributions in heterogeneous populations of agents. To see this, suppose that $\pi_i(\theta_i, s_i)$ is the fraction of agents in population i of type θ_i who play s_i.[17] For every $X_i \subseteq S_i$, set $\pi_i(\theta_i, X_i) = \sum_{s_i \in X_i} \pi_i(\theta_i, s_i)$. Next recall that $S_i(h)$ denotes the set of strategies of (population) i that allow (do not prevent) h, and $S_i(h, a_i)$ denotes the set of strategies of i that allow h *and* select a_i given h. Then, let

$$\beta_i(a_i|\theta_i, h) = \frac{\pi_i(\theta_i, S_i(h, a_i))}{\pi_i(\theta_i, S_i(h))} \text{ if } \pi_i(\theta_i, S_i(h)) > 0.$$

With this, $\beta_i(a_i|\theta_i, h)$ is the fraction of agents in population i taking action a_i at h, among those whose type is θ_i and whose strategy does not prevent h.

We can also characterize the Bayesian equilibria of Γ as the Nash equilibria of the *interim* strategic form. Assume for notational simplicity that $\Theta_i \cap \Theta_j = \emptyset$ for each $i, j \in I$, $i \neq j$ (this is just a matter of labelling). The **interim strategic form** of Γ is the static $\left(\sum_{i \in I} |\Theta_i|\right)$-player game

$$\mathcal{IS}(\Gamma) = \left\langle \bigcup_{i \in I} \Theta_i, (S_{\theta_i}, U_{\theta_i})_{i \in I, \theta_i \in \Theta_i} \right\rangle$$

where

- $S_{\theta_i} = S_i$ for each θ_i,

[16] By an incomplete-information version of Theorem 37, the two approaches are equivalent.

[17] Milgrom and Weber [59] call π_i a "distributional strategy."

- the expected payoff function $U_{\theta_i} : \times_{j\in I} S_j^{\Theta_j} \to \mathbb{R}$ is given by

$$U_{\theta_i}((s_{\theta_j})_{j\in I,\theta_j\in\Theta_j}) = \sum_{(\theta_0,\theta_{-i})\in\Theta_0\times\Theta_{-i}} p_i(\theta_0,\theta_{-i}|\theta_i)u_i\left(\theta_0,\theta_i,\theta_{-i},\zeta(s_{\theta_i},s_{\theta_{-i}})\right),$$

  with $s_{\theta_{-i}} = (s_{\theta_j})_{j\in I\setminus\{i\}} \in S_{-i}$.

It can be checked that *a profile σ is a Bayesian equilibrium of Γ if and only if it is a Nash equilibrium of $\mathcal{IS}(\Gamma)$*. An analogous result holds for randomized Bayesian equilibria.

15.7 Perfect Bayesian Equilibrium

Like Nash equilibrium, Bayesian equilibrium allows for non maximizing choices at histories that are not supposed to occur in equilibrium. For multistage games with complete information, this problem was addressed using the notion of subgame perfect equilibrium. Thus, traditional game theorists find it natural to try to extend this notion to multistage Bayesian games and define some kind of "perfect Bayesian equilibrium" (PBE) concept. Here we define the weakest (i.e., the most general) among the meaningful versions of the PBE concept.[18]

As we did in Section 15.3, to ease notation we neglect residual uncertainty, that is, we consider a simple multistage Bayesian game Γ where Θ_0 is a *singleton* and remove θ_0 from formulas. It is convenient to look at *randomized* equilibria as this makes it more transparent how the analysis relies on Bayes rule. To connect to the analysis of rational planning, a candidate equilibrium profile $\beta = (\beta_i)_{i\in I}$ of extended behavior strategies yields a "correct conjecture" for (each type of) each player i, with

$$\beta^i\left(a_{-i}|\theta_{-i},h\right) = \prod_{j\neq i} \beta_j\left(a_j|\theta_j,h\right) \tag{15.7.1}$$

for all $h \in H$, $\theta_{-i} \in \Theta_{-i}$, and $a_{-i} \in \mathcal{A}_{-i}(h)$. Furthermore, we need to specify a profile $(\mu_i(\cdot|\theta_i,h))_{i\in I,\theta_i\in\Theta_i}$ of personal systems of beliefs, one for each type of each player, otherwise we cannot compute conditional

[18] Some versions of PBE found in textbooks are either ambiguously defined, or are serioulsy flawed because they do not assume that Bayes consistency holds starting from every history (cf. Definition 81).

expected payoffs. Of course, for each player $i \in I$, the initial belief of each type $\theta_i \in \Theta_i$ about the types of the other players must be the belief $p_i(\cdot|\theta_i)$ exogenously specified in Bayesian game Γ.

Definition 93. *A* ***system of beliefs*** *in* Γ *is a profile* $\mu = (\mu_i)_{i \in I} \in \left(\Delta(\Theta_{-i})^{\Theta_i \times H}\right)^I$ *where each* $\mu_i(\cdot|\theta_i, \cdot) \in \Delta(\Theta_{-i})^H$ *(*$i \in I$*,* $\theta_i \in \Theta_i$*) is a personal system of beliefs such that* $\mu_i(\cdot|\theta_i, \varnothing) = p_i(\cdot|\theta_i)$*. A pair* (β, μ) *given by a profile of extended behavior strategies and a system of beliefs is called* ***assessment****.*

Clearly, beliefs μ must be related to β and are therefore *endogenous*. Hence, a candidate equilibrium cannot be just an extended behavior strategy profile, it has to be an *assessment* (β, μ). With this, fix a candidate equilibrium assessment (β, μ); when can we say that (β, μ) is a PBE? First note that, for each $h \in H$, beliefs $(\mu_i(\cdot|\theta_i, h))_{i \in I, \theta_i \in \Theta_i}$ define a (h, μ)-**continuation** (Bayesian) **game** $\Gamma(h, \mu)$: consider the set of feasible continuations of h, $\{h' \in A^{\leq \mathbb{N}} : (h, h') \in \bar{H}\}$, the resulting θ-dependent payoffs and the interactive beliefs $(\mu_i(\cdot|\theta_i, h))_{i \in I, \theta_i \in \Theta_i}$. (Of course, the $(\varnothing, \mu)$-"continuation-game" is the Bayesian game itself, $\Gamma(\varnothing, \mu) = \Gamma$, because $\mu_i(\cdot|\theta_i, \varnothing) = p_i(\cdot|\theta_i)$ for all $i \in I$, and $\theta_i \in \Theta_i$.) Intuitively, (β, μ) is a PBE if it satisfies two conditions:

- (**Interpersonal**) **Bayes consistency:** the system of beliefs μ and the behavior strategies β must be related to each other *via Bayes rule* (thus, the players use the candidate equilibrium strategies of the co-players as conjectures in order to update beliefs *via* Bayes rule).

- **Continuation equilibrium** (often called "sequential rationality"): for each $h \in H$, β induces a Bayesian equilibrium of the (h, μ)-continuation-game $\Gamma(h, \mu)$.

However, *it is not obvious how to specify Bayes consistency.* Game theorists have proposed different definitions. The reason is that, on top of mere consistency with Bayes rule (see Definition 81), they wanted to incorporate additional assumptions in the spirit of traditional equilibrium analysis, such as that players "update in the same way" (differences in beliefs are only due to differences in information), and that beliefs satisfy independence across opponents. Unfortunately, it was not very clear which additional assumptions one was trying to incorporate in the PBE concept, nor how to exactly express those assumptions in a mathematical language. Appeals to intuition and references to particular examples dominated the analysis. Hence the mess: *there is no universally accepted, or "canonical," notion of PBE* despite the fact that many applied papers refer to the PBE concept as if such a universally accepted notion existed. We consider below a minimalistic (i.e., general) notion of PBE based *only* on Bayes consistency as explained in Section 15.3.

Each player i, given his type θ_i, is uncertain about the types θ_{-i} of the co-players. As in the SPE analysis of complete information games, the candidate equilibrium behavior strategies β are commonly believed to be complied with, starting from any history h and whatever the number of deviations from β implied by h.[19] Thus, the probabilities of simultaneous and future actions are always believed to be determined by β given the players' types. A new action profile a taken at h is the evidence upon which player i updates his subjective belief about θ_{-i}.

Definition 94. *Assessment (β, μ) is a **perfect Bayesian equilibrium (PBE)** of Γ if, for every player $i \in I$ and type $\theta_i \in \Theta_i$, (1) personal assessment $\left(\beta^i, \mu_i(\cdot|\theta_i, \cdot)\right)$ (with conjecture β^i given by eq. (15.7.1)) is Bayes consistent and (2) behavior strategy $\beta_i(\cdot|\theta_i, \cdot)$ is sequentially optimal given $\left(\beta^i, \mu_i(\cdot|\theta_i, \cdot)\right)$.*

Note that assessment (β, μ) satisfies condition (2) of Definition 94 if and only if the continuation of β from each $h \in H$ induces a Bayesian equilibrium in the (h, μ)-continuation-game $\Gamma(h, \mu)$. Letting $h = \varnothing$ (empty history), we obtain the intuitive observation that, *if (β, μ) is a PBE of Γ, then β is a Bayesian equilibrium of Γ.*[20]

Let $\mathbb{E}_{\beta,\mu_i}(u_i|\theta_i, h, a_i) = V_{\theta_i}^{\beta,\mu_i}(h, a_i)$ denote the conditional expected payoff for type θ_i given h and a_i, under assessment (β, μ). By an application of the OD Principle (Theorem 55), we obtain the following result:

---

[19] A deviation from extended behavior strategy β_j is *detected* at history h if h contains an action a_j taken at some earlier history $h' \prec h$ such that $\beta_j(a_j|\theta_j, h') = 0$ *for every* θ_j.

[20] We will consider the relationship between PBE and Bayesian equilibrium more in detail in the analysis of a special case; see Section 15.8.

Corollary 8. *An assessment* (β, μ) *such that each type of each player satisfies Bayes consistency is a PBE if and only if, for all* $i \in I$, $\theta_i \in \Theta_i$, *and* $h \in H$,

$$\mathrm{supp}\beta_i(\cdot|\theta_i, h) \subseteq \arg\max_{a_i \in \mathcal{A}_i(h)} \mathbb{E}_{\beta,\mu}\left(u_i|\theta_i, h, a_i\right).$$

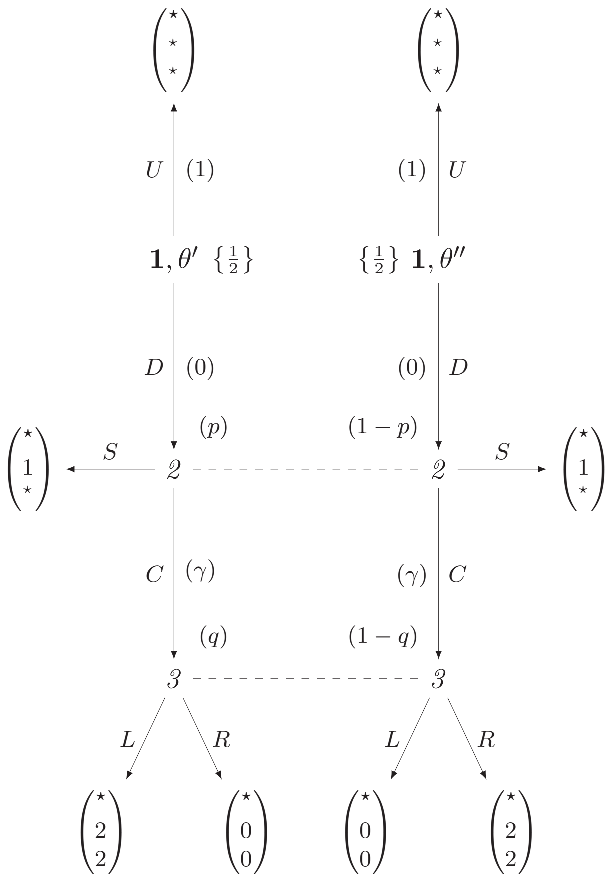


Figure 15.5: Assessment in a three-person game.

The PBE concept of Definition 94 is quite weak. For example, it allows players who start with the same information and beliefs in a given stage to update in different ways, although they observe exactly the same evidence, that is, the actions played at that stage. Also, it allows player i to change his beliefs about the type of j upon observing the action of some other player k, as if k had some private signal about the type of j. The reason is that the *interpersonal* Bayes consistency requirements bite at histories with positive probability, but not at histories with 0 probability.

Example 76. Consider the assessment partially shown in Figure 15.5, where p denotes the belief of player 2 conditional on D and q denotes the belief of player 3 conditional on (D, C). First of all, since D has zero probability, even if players 2 and 3 have the same (uniform) prior and observe the same evidence (action D is taken by player 1), they may have different beliefs conditional on D, i.e., it is possible that $\mu_3(\theta'|D) \neq p =: \mu_2(\theta'|D)$. In a sense, they can come up with different explanations of the unexpected move by player 1. Now suppose for the sake of the discussion that instead they have the same belief given D: $\mu_3(\theta'|D) = p$. If the candidate equilibrium behavior strategy of 2 assigns strictly positive probability to C—that is, $\gamma =: \beta_2(C|D) > 0$—then Bayes consistency implies that 2 and 3 hold the same belief conditional on (D, C), i.e., $p = q$. If instead $\gamma = 0$ then Bayes consistency only implies $\mu_2(\theta'|D, C) = p$, but it does *not* imply $p = q$; in other words, 3 may change his belief about 1's type just because he is surprised by 2. Stronger notions of PBE require, in this example, that $\mu_3(\theta'|D) = p = q$. More generally, they require that players with the same initial beliefs and information update in the same way and that updated beliefs about a co-player's type depend only on that co-player's actions. ▲

15.8 Equilibrium in Signaling Games

We now study an easy case widely considered in applications, for which everybody agrees on the meaning of PBE:

Definition 95. *A* ***leader-follower game*** *is a two-person, two-stage game with observed actions where only one player is active at each stage and the second mover is different from the first mover, for each non-terminating action of the first mover. A* ***signaling game*** *(or* ***sender-receiver game****) is a leader-follower game with payoff uncertainty $\hat{\Gamma}$ where the second mover has no private information.*

So, a **signaling game** is a two-stage game with payoff uncertainty and perfect information (only one active player at stage 1, at most one active player at stage 2) where only the first mover has private information.[21] Thus, we assume that θ is known to player 1, whose action[22] $a_1 \in A_1$ is observed by player 2 before he chooses $a_2 \in \mathcal{A}_2(a_1)$.[23] The possible parameter values $\theta \in \Theta$ are the **types** of player 1, the informed player. The actions of the informed player are also called **signals** or **messages** because they may be thought to reveal her private information. Thus, the first player is also called "**sender"** and the second player is also called **"receiver."** Let $A_2 = \bigcup_{a_1 \in A_1} \mathcal{A}_2(a_1)$. The payoffs of sender and receiver are given by the functions

$$\begin{aligned} u_1 &: \Theta \times A_1 \times A_2 \to \mathbb{R}, \\ u_2 &: \Theta \times A_1 \times A_2 \to \mathbb{R}. \end{aligned}$$

As we notice more generally for multistage games with incomplete information, a signaling game can be analyzed by some versions of rationalizability. We have already done this in previous examples of signaling games, such as those in Examples 73 and 74. Yet, in order to define the traditional notion of equilibrium we need to add a description of players' exogenous beliefs. In particular, to obtain a **Bayesian signaling game** Γ, we just have to append to signaling game (with payoff uncertainty) $\hat{\Gamma}$ an initial belief $p \in \Delta(\Theta)$ of the uninformed player, i.e., of the receiver. We assume for simplicity that $p(\theta) > 0$ for every $\theta \in \Theta$. This *prior* belief is an *exogenously given* element of the model, whereas (behavior) strategies have to be determined through equilibrium analysis. Furthermore, consistently with our interpretation of Bayesian games, it should be informally assumed that the prior p of the receiver is "transparent," that is, the sender is certain that p is the prior of the receiver, the receiver is certain that the sender is certain of this, etc.

A(n extended) **behavior strategy for the sender** is an array of probability measures $\beta_1 = (\beta_1(\cdot|\theta))_{\theta\in\Theta} \in (\Delta(A_1))^{\Theta}$: under the

[21] This is not a crucial assumption, it just allows to simplify the notation. The extension of the analysis to signaling games with two-sided incomplete information is conceptually straightforward.

[22] In some models the set of feasible actions of player 1 depends on θ and is denoted by $\mathcal{A}_1(\theta)$. The set of potentially feasible actions of player 1 is $A_1 = \bigcup_{\theta\in\Theta} \mathcal{A}_1(\theta)$.

[23] If player 2 is not active after action a_1, then $\mathcal{A}_2(a_1)$ is a singleton.

interpretation that the sender is "born" of a given type θ, β_1 is not a strategy in the intuitive sense, but rather a randomized decision rule expressing how the behavior of the sender depends on his type. A **behavior strategy for the receiver** is an array of probability measures $\beta_2 = (\beta_2(\cdot|a_1))_{a_1 \in A_1} \in \times_{a_1 \in A_1} \Delta(\mathcal{A}_2(a_1))$.

Here the receiver has the role of agent i in Section 15.2 on Bayes rule, with $X = A_1$. The receiver is initially uncertain about (θ, a_1) and has a prior probability measure on $\Theta \times A$ with marginal $p \in \Delta(\Theta)$. *In an equilibrium assessment* (β, μ), β_1 represents the conjecture of the receiver about the probability of each action of the sender conditional on each possible type (parameter value) θ. Thus, the probabilistic assessment of the receiver is such that

$$\begin{aligned}
\forall \theta &\in \Theta,\, \mathbb{P}_{\beta,\mu}(\theta) = \mu(\theta|\varnothing) = p(\theta), \\
\forall \theta &\in \Theta, \forall a_1 \in A_1,\, \mathbb{P}_{\beta,\mu}(a_1|\theta) = \beta_1(a_1|\theta),\, \mathbb{P}_{\beta,\mu}(\theta, a_1) = \beta_1(a_1|\theta)p(\theta), \\
\forall a_1 &\in A_1,\, \mathbb{P}_{\beta,\mu}(a_1) = \sum_{\theta' \in \Theta} \beta_1(a_1|\theta')p(\theta').
\end{aligned}$$

Here $\mu(\theta|a_1)$ denotes the probability that the receiver would assign to θ upon observing action a_1 of the sender. Since the receiver chooses a_2 after he has observed a_1 so as to maximize the expectation of $u_2(\theta, a_1, a_2)$, the system of conditional probabilities $\mu = (\mu(\cdot|a_1))_{a_1 \in A_1}$ is an essential ingredient of equilibrium analysis. Applying the terminology of Section 15.7 to the special case of signaling games, μ is called "system of beliefs."

If the predictive probability of action a_1 is positive, $\mathbb{P}_{\beta,\mu}(a_1) = \sum_{\theta' \in \Theta} \beta_1(a_1|\theta')p(\theta') > 0$, then Bayes formula applies and we must have

$$\mu(\theta|a_1) = \frac{\mathbb{P}_{\beta,\mu}(\theta, a_1)}{\mathbb{P}_{\beta,\mu}(a_1)} = \frac{\beta_1(a_1|\theta)p(\theta)}{\sum_{\theta' \in \Theta} \beta_1(a_1|\theta')p(\theta')}.$$

Since β_1 is endogenous, also the system of beliefs μ *is endogenous*. Thus, we have to determine through equilibrium analysis the triple (β_1, β_2, μ). As we have explained in Section 15.7, (β_1, β_2, μ) (a profile of behavior strategies plus a system of beliefs) is called "assessment."

For any given assessment (β_1, β_2, μ) we use the following notation to abbreviate conditional expected payoff formulas:

$$\begin{aligned}\mathbb{E}_{\beta_2}\left(u_1\left(\theta, a_1, \cdot\right)\right) &= \sum_{a_2 \in \mathcal{A}_2(a_1)} \beta_2(a_2|a_1)u_1(\theta, a_1, a_2), \\ \mathbb{E}_{\mu}\left(u_2\left(\cdot, a_1, a_2\right)\right) &= \sum_{\theta \in \Theta} \mu(\theta|a_1)u_2(\theta, a_1, a_2).\end{aligned}$$

Thus, $\mathbb{E}_{\beta_2}\left(u_1\left(\theta, a_1, \cdot\right)\right)$ is the expected payoff for the sender of choosing action a_1 given that his type is θ and assuming that his conjecture about the behavior of the receiver is given by β_2.[24] Similarly, $\mathbb{E}_{\mu}\left(u_2\left(\cdot, a_1, a_2\right)\right)$ is the expected payoff for the receiver of choosing action a_2 given that he has observed a_1 and assuming that his conditional beliefs about θ are determined by μ .

In equilibrium, an action of player i can have positive conditional probability only if it maximizes the conditional expected payoff of i. Furthermore, the equilibrium assessment must be consistent with Bayes rule. Therefore we obtain three equilibrium conditions for the three components of (β_1, β_2, μ):

Definition 96. *Assessment* (β_1, β_2, μ) *is a* ***perfect Bayesian equilibrium*** (PBE) *of the Bayesian signaling game* Γ *if it satisfies the following conditions:*

$$\forall \theta \in \Theta, \text{supp}\beta_1(\cdot|\theta) \subseteq \arg\max_{a_1 \in A_1} \mathbb{E}_{\beta_2}\left(u_1\left(\theta, a_1, \cdot\right)\right) \qquad (\text{BR}_1)$$

$$\forall a_1 \in A_1, \text{supp}\beta_2(\cdot|a_1) \subseteq \arg\max_{a_2 \in \mathcal{A}_2(a_1)} \mathbb{E}_{\mu}\left(u_2\left(\cdot, a_1, a_2\right)\right) \qquad (\text{BR}_2)$$

$$\forall a_1 \in A_1, \forall \theta \in \Theta, \mu(\theta|a_1) \sum_{\theta' \in \Theta} \beta_1(a_1|\theta')p(\theta') = \beta_1(a_1|\theta)p(\theta). \qquad (\text{CONS})$$

Thus, from the perspective of an analyst who wants to compute the set of equilibrium assessments (β_1, β_2, μ), there are three systems of weak inequalities and equalities, one for each group of "unknowns" (endogenous

[24] There is no loss of generality in representing a conjecture of player 1 as a behavior strategy of player 2. If player 1 had a conjecture of the form $\sigma_2 \in \Delta(S_2)$, where S_2 is the set of pure strategies of player 2, then we could derive from σ_2 a realization-equivalent behavior strategy. A similar argument holds for the conjecture of player 2 about player 1.

variables) β_1, β_2, and μ. Note that (CONS)—consistency with Bayes rule—can also be expressed as

$$\begin{aligned}\forall a_1 &\in A_1, \forall\theta\in\Theta,\\ \sum_{\theta'\in\Theta}\beta_1(a_1|\theta')p(\theta') &> 0 \Rightarrow \mu(\theta|a_1) = \frac{\beta_1(a_1|\theta)p(\theta)}{\sum_{\theta'\in\Theta}\beta_1(a_1|\theta')p(\theta')}.\end{aligned}$$

Each equilibrium condition involves two out of the three groups of endogenous variables β_1, β_2 and μ: (BR_1) says that each mixed action $\beta_1(\cdot|\theta) \in \Delta(A_1)$ is a best reply to β_2 for type θ of the sender (for each $\theta \in \Theta$), (BR_2) says that each mixed action $\beta_2(\cdot|a_1) \in \Delta(\mathcal{A}_2(a_1))$ is a best reply for the receiver to the conditional belief $\mu(\cdot|a_1) \in \Delta(\Theta)$ (for each $a_1 \in A_1$), and (CONS) says that the triple (p, β_1, μ) is consistent with Bayes rule.

As we noticed more generally for multistage Bayesian games, PBE refines Bayesian equilibrium. Here we provide an explicit proof of this observation for the special case of signaling games.

Proposition 13. *Each PBE (β_1, β_2, μ) of a Bayesian signaling game Γ is such that the (extended) behavior strategy pair (β_1, β_2) is a Bayesian equilibrium of Γ, that is, a (randomized) Nash equilibrium of the ex ante strategic form of Γ.*

Proof. Fix a PBE (β_1, β_2, μ). By inspection of the definition of PBE, the best reply condition for each type of the sender is satisfied. To see that also the strategic-form best reply condition of the receiver is satisfied note that the ex ante expected utility of the receiver can be expressed as

$$\begin{aligned}\mathbb{E}_{\beta_1,\beta_2}(u_2) &= \sum_{\theta\in\Theta} p(\theta) \sum_{a_1\in A_1} \beta_1(a_1|\theta) \sum_{a_2\in A_2} \beta_2(a_2|a_1)\, u_2(\theta, a_1, a_2)\\ &= \sum_{a_1\in A_1}\sum_{\theta\in\Theta} \beta_1(a_1|\theta)\, p(\theta) \sum_{a_2\in A_2} \beta_2(a_2|a_1)\, u_2(\theta, a_1, a_2)\\ &= \sum_{a_1\in A_1}\sum_{\theta\in\Theta} \mu(\theta|a_1)\mathbb{P}(a_1) \sum_{a_2\in A_2} \beta_2(a_2|a_1)\, u_2(\theta, a_1, a_2)\\ &= \sum_{a_1\in A_1} \mathbb{P}(a_1) \sum_{a_2\in A_2} \beta_2(a_2|a_1) \sum_{\theta\in\Theta} \mu(\theta|a_1) u_2(\theta, a_1, a_2)\\ &= \sum_{a_1\in A_1} \mathbb{P}(a_1) \sum_{a_2\in A_2} \beta_2(a_2|a_1)\, \mathbb{E}_\mu(u_2(\cdot, a_1, a_2))\end{aligned}$$

where the third equality follows from Bayes rule. Thus, β_2 maximizes $\mathbb{E}_{\beta_1,\beta_2}(u_2)$ if and only if each term $\sum_{a_2 \in A_2} \beta_2(a_2|a_1)\,\mathbb{E}_\mu(u_2(\cdot, a_1, a_2))$ with $\mathbb{P}(a_1) > 0$ is maximized. Since condition (BR$_2$) requires that *each* term $\sum_{a_2 \in A_2} \beta_2(a_2|a_1)\,\mathbb{E}_\mu(u_2(\cdot, a_1, a_2))$ is maximized (whether or not $\mathbb{P}(a_1) > 0$), it follows that also the best reply condition of the receiver holds, and (β_1, β_2) is a Bayesian equilibrium of the signaling game. ∎

It is worth noting that the equilibrium probability

$$\mathbb{P}_{\beta,\mu}(a_1) = \sum_{\theta' \in \Theta} \beta_1(a_1|\theta')p(\theta')$$

of some action a_1 *may be zero*. Fix a_1 and suppose, for example, that for each $\theta \in \Theta$, there is some action a_1^θ such that $\mathbb{E}_{\beta_2}\left(u_1\left(\theta, a_1^\theta, \cdot\right)\right) > \mathbb{E}_{\beta_2}(u_1(\theta, a_1, \cdot))$. Then action/signal a_1 must have zero probability because it is not a best reply to β_2 for any type θ of the sender. *Yet we assume that the belief* $\mu(\cdot|a_1)$ *is well defined* and the receiver takes a best reply to this belief. This is a *perfection* requirement analogous to the subgame perfection condition for games with observed actions and complete information. A perfect Bayesian equilibrium is a Bayesian equilibrium that satisfies perfection and consistency with Bayes rule.

Furthermore, even if $\mu(\cdot|a_1)$ cannot be computed with Bayes formula, it may still be the case that the equilibrium conditions put incentive constraints on the possible values of $\mu(\cdot|a_1)$, because mixed action $\beta_2(\cdot|a_1)$ has to be a best reply to $\mu(\cdot|a_1)$. The following example illustrates this point.

Example 77. Consider the signaling game depicted in Figure 15.6. The payoffs of the sender are in **bold.** The set $\mathcal{A}_2(l)$ is a singleton and therefore the action of player 2 after l is not shown. If the sender goes right (r) then the receiver can go up (u) or down (d), i.e., $\mathcal{A}_2(r) = \{u, d\}$.

Note that action r is dominated for type θ' of the sender. Therefore $\beta_1(r|\theta') = 0$ in every PBE. Now we show that in equilibrium we also have $\beta_1(r|\theta'') = 0$. Suppose, by way of contradiction, that $\beta_1(r|\theta'') > 0$ in a PBE. Then Bayes formula applies and $\mu(\theta''|r) = 1$. But then the best reply of the receiver is down, i.e., $\beta_2(d|r) = 1$, and the best reply of type θ'' is left, i.e., $\beta_1(l|\theta'') = 1 - \beta_1(r|\theta'') = 1$, contradicting our initial assumption. We conclude that, in every PBE, r is chosen with probability zero and $\mu(\cdot|r)$ cannot be determined with Bayes formula. Yet the equilibrium conditions

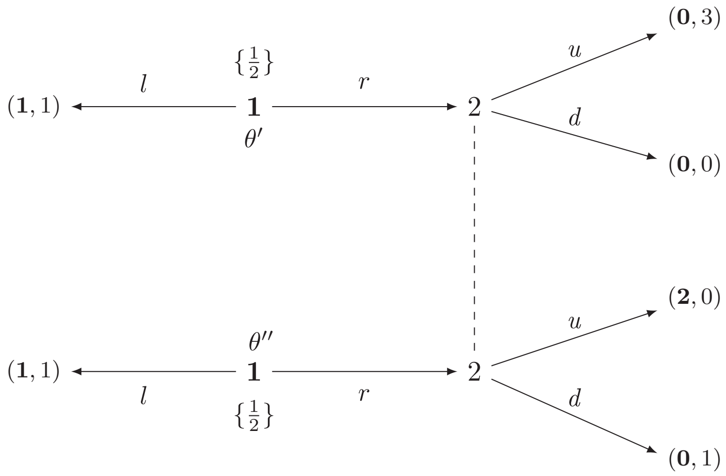


Figure 15.6: A signaling game

put a constraint on $\mu(\cdot|r)$: in equilibrium d must be (weakly) preferred to u (if the receiver chooses u after r then type θ'' chooses r and we have just shown that this cannot happen in equilibrium). Therefore

$$\begin{aligned} \mu(\theta''|r) &\geq 3\mu(\theta'|r) \\ \text{or } \mu(\theta''|r) &\geq \frac{3}{4}. \end{aligned}$$

The set of equilibrium assessments is

$$\left\{(\beta_1, \beta_2, \mu) : \beta_1(l|\theta') = \beta_1(l|\theta'') = 1, \beta_2(d|r) = 1, \mu(\theta''|r) > \frac{3}{4}\right\}$$
$$\cup \left\{(\beta_1, \beta_2, \mu) : \beta_1(l|\theta') = \beta_1(l|\theta'') = 1, \beta_2(d|r) \geq \frac{1}{2}, \mu(\theta''|r) = \frac{3}{4}\right\}.$$

▲

The PBE assessments in Example 77 are examples of "pooling" equilibria. A **pooling equilibrium** is a PBE assessment where all types of the sender choose the same pure action with probability one, i.e., there exists $a_1^* \in A_1$ such that for every $\theta \in \Theta$, $\beta_1(a_1^*|\theta) = 1$. In this case Bayes rule implies that the posterior on θ conditional on the equilibrium action a_1^* is the same as the prior: $\mu(\cdot|a_1^*) = p(\cdot)$.

The polar case is when different types choose different pure actions: a **separating equilibrium** is a PBE assessment such that each type θ of player 1 chooses some action $a_1(\theta)$ with probability one ($\beta_1(a_1(\theta)|\theta) = 1$) and $a_1(\theta') \neq a_1(\theta'')$ for all θ' and θ'' with $\theta' \neq \theta''$. A separating equilibrium may exist only if A_1 has at least as many elements as Θ. *If* A_1 *and* Θ *have the same number of elements* (cardinality) then in a separating equilibrium each action is chosen with ex ante positive probability (because $p(\theta) > 0$ for each $\theta \in \Theta$) and the action of player 1 *perfectly reveals* her private information (if A_1 has more elements than Θ then the actions that in equilibrium are chosen by some type are perfectly revealing, the others need not be revealing).

We provide an example of a signaling game with a separating equilibrium.

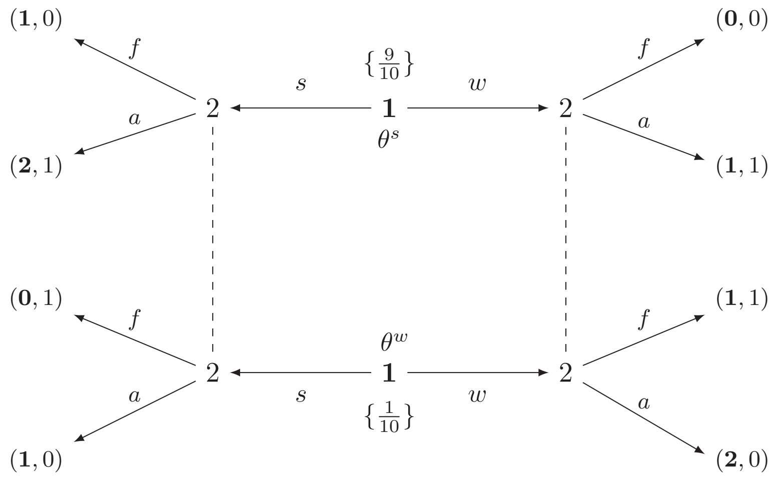


Figure 15.7: The "sausage-whipped cream" game

Example 78. Consider the signaling game depicted in Figure 15.7. The game can be interpreted as follows: a truck driver (sender) enters in a pub where an aggressive customer (receiver) has to decide whether to start a fight (*f*, upward) or acquiesce (a, downward). There are two types of truck drivers: 90% of them are surly (θ^s) and like to eat sausages (s) for breakfast; the remaining 10% are wimps (θ^w) and prefer a dessert with whipped cream (w). Each type of truck driver receives an incremental utility equal to 2 from his favorite breakfast and an incremental utility equal to 1 from the other breakfast. Furthermore both types incur a loss of 1 util if they have to fight. The receiver prefers to fight with a wimp and to avoid the fight with a surly driver.

This game has only one "reasonable" PBE and it is separating.[25] Indeed, the payoffs are such that for each type of driver it is weakly dominant to have his preferred breakfast; thus, iterated deletion of weakly dominated actions yields the equilibrium $\beta_1(s|\theta^s) = 1 = \beta_1(w|\theta^w)$, $\beta_2(a|s) = 1 = \beta_2(f|w)$, $\mu(\theta^s|s) = 1 = \mu(\theta^w|w)$.

The game has also two sets of pooling equilibria (meaning that one type of player 1 chooses a weakly dominated action). In the first set of assessments each type has sausages for breakfast and player 2 would fight if and only he observed a whipped-cream breakfast: $\beta_1(s|\theta^s) = 1 = \beta_1(s|\theta^w)$, $\beta_2(a|s) = 1 = \beta_2(f|w)$, $\mu(\theta^s|s) = \frac{9}{10}$, $\mu(\theta^w|w) \geq \frac{1}{2}$. In the second set of assessments each type has whipped cream for breakfast and player 2 would fight if and only if he observed a sausage breakfast: $\beta_1(w|\theta^s) = 1 = \beta_1(w|\theta^w)$, $\beta_2(a|w) = 1 = \beta_2(f|s)$, $\mu(\theta^s|w) = \frac{9}{10}$, $\mu(\theta^w|s) \geq \frac{1}{2}$. ▲

15.8.1 An Algorithm to Compute Pure PBE's.

As for other finite-horizon games, also the equilibria of signaling games can be computed with a "case-by-case backward induction" algorithm. For simplicity, we focus on pure equilibria, that is, any PBE (β_1, β_2, μ) such that β_1 and β_2 are pure ($|\text{supp}\beta_1(\cdot|\theta)| = 1$ for each θ, and $|\text{supp}\beta_2(\cdot|a_1)| = 1$ for each a_1). We denote by $\mathbf{a}_1 \in A_1^\Theta$ the decision function of the sender, whereas $s_2 \in \times_{a_1 \in A_1} \mathcal{A}_2(a_1)$ denotes the pure strategy of the receiver (i.e., $\beta_1(\mathbf{a}_1(\theta)|\theta) = 1$ for each θ, and $\beta_2(s_2(a_1)|a_1) = 1$ for each a_1). Working backward, we start from the analysis of stage 2 (receiver) and then we move back to the analysis of stage 1 (sender). The analysis of stage 2, in turn, is broken down into a preliminary step and a main step.

Stage 2, preliminary step: deletion of dominated actions. For any a_1, an action $a_2 \in \mathcal{A}_2(a_1)$ is **conditionally dominated** given a_1 if

[25] Actually, this example is a modification of a well-known game where the cost of a fight is *larger* than the marginal benefit from having the preferred breakfast and all equilibria are pooling.

there exists some $\alpha_2 \in \Delta(\mathcal{A}_2(a_1))$ such that

$$\forall \theta \in \Theta, \sum_{a_2' \in \mathcal{A}_2(a_1)} u_2(\theta, a_1, a_2') \alpha_2(a_2') > u_2(\theta, a_1, a_2);$$

otherwise, a_2 is **conditionally undominated** given a_1. Denote by $\mathcal{ND}(a_1) \subseteq \mathcal{A}_2(a_1)$ the set of conditionally undominated actions given a_1. An obvious adaptation of Lemma 2 (the *Wald-Pearce Lemma*) implies that, for every $a_1 \in A_1$ and $a_2 \in \mathcal{A}(a_1)$,

$$a_2 \in \mathcal{ND}(a_1) \Longleftrightarrow \left(\exists \nu \in \Delta(\Theta), a_2 \in \arg\max_{a_2' \in \mathcal{A}_2(a_1)} \mathbb{E}_\nu \left(u_2(\cdot, a_1, a_2')\right)\right)$$

(see Lemma 37). In other words, a second-mover action is conditionally undominated if and only if it is a best reply to some conditional belief about the types of the first mover given the observed action of the first mover. With this, it is obvious that the conditionally dominated actions of player 2 are unjustifiable and must have zero probability in equilibrium. Therefore, the preliminary step requires to delete all the conditionally dominated actions of the receiver.[26]

Stage 2, main step. Pick a strategy $s_2 \in \times_{a_1 \in A_1} \mathcal{ND}(a_1)$, that is, pick one action $s_2(a_1) \in \mathcal{ND}(a_1)$ among the undominated actions given each a_1. For each $a_2 \in \mathcal{ND}(a_1)$, let

$$J(a_2|a_1) = \left\{\nu \in \Delta(\Theta) : a_2 \in \arg\max_{a_2' \in \mathcal{A}_2(a_1)} \mathbb{E}_\nu \left(u_2(\cdot, a_1, a_2')\right)\right\}$$

denote the nonempty set of beliefs $\nu \in \Delta(\Theta)$ that justify a_2 as a best reply given a_1. A strategy $s_2 \in \times_{a_1 \in A_1} \mathcal{ND}(a_1)$ (i.e., a selection from the undominated action correspondence $a_1 \mapsto \mathcal{ND}(a_1)$) is a "**case.**" Clearly, there are $\prod_{a_1 \in A_1} |\mathcal{ND}(a_1)|$ cases to be considered. For every

[26] A similar preliminary step could be applied to player 1, but we do not explicitly include such step in the algorithm for the following reason: Suppose that a_1 is dominated for type θ of player 1. Clearly, $\beta_1(a_1|\theta) = 0$ in any equilibrium. But the "deletion" of such a_1 for type θ may induce a naive user of the algorithm to illegitimately infer that $\mu(\theta|a_1) = 0$. Such inference is correct and follows from Bayes rule if $\beta_1(a_1|\theta') > 0$ for some θ', but this cannot be presumed. It may be the case that, in equilibrium, $\beta_1(a_1|\theta') = 0$ for every θ'; hence, Bayes formula cannot determine the "off-path" conditional belief $\mu(\cdot|a_1)$ and there may be an equilibrium assessment such that $\mu(\theta|a_1) > 0$.

such case s_2 (by Lemma 2 and Lemma 37), there is a system of beliefs $\mu \in \bigtimes_{a_1 \in A_1} J(s_2(a_1)|a_1)$ that makes s_2 a kind of "stage-2 equilibrium."[27]

Stage 1. For each case $s_2 \in \bigtimes_{a_1 \in A_1} \mathcal{ND}(a_1)$, pick a decision function $\mathbf{a}_1 \in \bigtimes_{\theta \in \Theta} \arg\max_{a_1 \in A_1} u_1(\theta, a_1, s_2(a_1))$, that is, pick an action

$$\mathbf{a}_1(\theta) \in \arg\max_{a_1 \in A_1} u_1(\theta, a_1, s_2(a_1))$$

for each type $\theta \in \Theta$. Finally, verify whether there exists a system of beliefs $\mu \in \bigtimes_{a_1 \in A_1} J(s_2(a_1)|a_1)$ that is consistent with Bayes rule given decision function $\mathbf{a}_1$, i.e., such that, for each $a_1 \in A_1$,

$$\sum_{\theta' : \mathbf{a}_1(\theta') = a_1} p(\theta') > 0 \Rightarrow \left(\forall \theta \in \mathbf{a}_1^{-1}(a_1), \mu(\theta|a_1) = \frac{p(\theta)}{\sum_{\theta' : \mathbf{a}_1(\theta') = a_1} p(\theta')} \right),$$

where

$$\sum_{\theta' : \mathbf{a}_1(\theta') = a_1} p(\theta') = p\left(\mathbf{a}_1^{-1}(a_1)\right) = \mathbb{P}_{\mathbf{a}_1}(a_1)$$

is the probability of action/signal a_1 determined by decision function $\mathbf{a}_1$. If such μ exists, then assessment $(\mathbf{a}_1, s_2, \mu)$ is a pure PBE; otherwise, there is no μ such that $(\mathbf{a}_1, s_2, \mu)$ is a PBE. Note that here we are applying Bayes rule in the special case where the behavior strategy of the sender is deterministic, that is, $\beta_1(a_1|\theta) = 1$ if $\mathbf{a}_1(\theta) = a_1$, and $\beta_1(a_1|\theta) = 0$ if $\mathbf{a}_1(\theta) \neq a_1$.

We illustrate the algorithm computing the equilibria of the previous "sausage-whipped cream" example (Figure 15.7). First note the preliminary stage-2 step is vacuous because the receiver has no conditionally dominated action. Since the receiver weakly prefers to fight if and only if the conditional probability of the surly type is not larger

[27] An element of $\bigtimes_{a_1 \in A_1} J(s_2(a_1)|a_1) \subseteq \Delta(\Theta)^{A_1}$ is indeed a system of beliefs because it associates a probability measure on Θ with each action a_1 of the sender.

than 1/2, the sets of justifying beliefs are as follows:

$$\begin{aligned} J(f|s) &= \left\{\mu(\cdot|s) \in \Delta(\{\theta^s, \theta^w\}) : \mu(\theta^s|s) \leq \frac{1}{2}\right\}, \\ J(a|s) &= \left\{\mu(\cdot|s) \in \Delta(\{\theta^s, \theta^w\}) : \mu(\theta^s|s) \geq \frac{1}{2}\right\}, \\ J(f|w) &= \left\{\mu(\cdot|w) \in \Delta(\{\theta^s, \theta^w\}) : \mu(\theta^s|w) \leq \frac{1}{2}\right\}, \\ J(a|w) &= \left\{\mu(\cdot|w) \in \Delta(\{\theta^s, \theta^w\}) : \mu(\theta^s|w) \geq \frac{1}{2}\right\}. \end{aligned}$$

Also note that if the expected reply of the receiver does not depend on the signal—the breakfast of the truck driver—, then the sender eats his preferred breakfast. Thus, there are 2×2 possible cases, or undominated strategies of the receiver (we list first the reply to *s*ausage, second the reply to *w*hipped cream):

- **Case** $f.f$: θ^s chooses s, θ^w chooses w, thus $\mu(\theta^s|s) = 1$ and $\mu(\theta^s|w) = 0$. Since $1 = \mu(\theta^s|s) > \frac{1}{2}$, $\mu(\cdot|s) \notin J(f|s)$ and there is no PBE where the receiver plays $f.f$.
- **Case** $a.a$: θ^s chooses s, θ^w chooses w, thus $\mu(\theta^s|s) = 1$ and $\mu(\theta^w|w) = 1$. Since $0 = \mu(\theta^s|w) < \frac{1}{2}$, $\mu(\cdot|w) \notin J(a|w)$ and there is no PBE where the receiver plays $a.a$.
- **Case** $a.f$: θ^s chooses s, θ^w is indifferent because eating his preferred breakfast (whipped cream) would be compensated by the cost of a fight.
  - **Subcase** $w|\theta^w$**:** θ^w chooses w. Thus, $\mu(\theta^s|s) = 1 \geq \frac{1}{2}$ and $\mu(\theta^s|w) = 0 \leq \frac{1}{2}$, which implies that $\mu(\cdot|s) \in J(a|s)$ and $\mu(\cdot|w) \in J(f|w)$ as required. We thus obtain the separating PBE (β_1, β_2, μ) with $\beta_1(s|\theta^s) = \beta_1(w|\theta^w) = 1$, $\beta_2(a|s) = \beta_2(f|w) = 1$, $\mu(\theta^s|s) = \mu(\theta^w|w) = 1$.
  - **Subcase** $s|\theta^w$**:** also θ^w chooses s. Thus, $\mu(\theta^s|s) = p(\theta^s) = \frac{9}{10} \geq \frac{1}{2}$, which implies that $\mu(\cdot|s) \in J(a|s)$; on the other hand, $\mu(\cdot|w)$ is not determined by Bayes rule and we can pick any $\mu(\cdot|w) \in J(f|w)$, e.g., $\mu(\theta^s|w) = \frac{1}{3}$, to obtain a

set of pure PBE's (β_1, β_2, μ) with $\beta_1(s|\theta^s) = \beta_1(s|\theta^w) = 1$, $\beta_2(a|s) = \beta_2(f|w) = 1$, $\mu(\theta^s|s) = \frac{9}{10}$, $\mu(\theta^s|w) \leq \frac{1}{2}$.

- **Case** $f.a$: θ^w chooses w, θ^s is indifferent because eating his preferred breakfast (sausage) would be compensated by the cost of a fight.

  - **Subcase** $s|\theta^s$**:** θ^s chooses s. Thus, $\mu(\theta^s|w) = 0 < \frac{1}{2}$**,** $\mu(\cdot|w) \notin J(a|w)$ and there is no PBE where the receiver plays $f.a$ and the sender eats his preferred breakfast.
  - **Subcase** $w|\theta^s$**:** also θ^s chooses w. Thus, $\mu(\theta^s|w) = p(\theta^s) = \frac{9}{10} \geq \frac{1}{2}$, which implies that $\mu(\cdot|w) \in J(a|w)$; on the other hand, $\mu(\cdot|s)$ is not determined by Bayes rule and we can pick any $\mu(\cdot|s) \in J(f|s)$, e.g., $\mu(\theta^s|s) = \frac{1}{3}$, to obtain a set of pure PBE's (β_1, β_2, μ) with $\beta_1(w|\theta^s) = \beta_1(w|\theta^w) = 1$, $\beta_2(a|w) = \beta_2(f|s) = 1$, $\mu(\theta^s|w) = \frac{9}{10}$, $\mu(\theta^s|s) \leq \frac{1}{2}$.

15.9 Appendix

Here we study an extension of the analysis of rational planning of Chapter 10 to encompass incomplete information. Compared to Chapter 10, we focus on uncertainty about the true map from terminal histories to "utils" and on Bayesian updating, while we trim some lengthy arguments explained in detail there. In what follows, we keep the same notation as in Section 15.3 and let $\beta^i \in \left(\times_{h\in H} \Delta(\mathcal{A}_{-i}(h))\right)^{\Theta_{-i}}$ denote a conjecture of i about the type-dependent behavior of the co-players, keeping in mind that in the perfect Bayesian equilibrium analisys such conjectures are determined by profiles β_{-i} of type-dependent behavior strategies.

Fix once and for all a type θ_i of player i. To ease notation we let $\mu_{\theta_i} = \mu_{\theta_i}(\cdot|\cdot) \in \Delta(\Theta_{-i})^H$ denote the personal system of beliefs of type θ_i, that is, with reference to the notation used for Bayesian games we write $\mu_{\theta_i}(\theta_{-i}|h)$ instead of $\mu_i(\theta_{-i}|\theta_i, h)$ for the probability assigned by type θ_i to profile θ_{-i} conditional on h. Hence we write a personal assessment of θ_i as a pair $\left(\beta^i, \mu_{\theta_i}\right)$. Similarly, we write $u_{\theta_i}(\theta_{-i}, z) = u_i(\theta_i, \theta_{-i}, z)$ (that is, u_{θ_i} is the section of u_i at θ_i). With this, we *fix* a *Bayes consistent* personal assessment $\left(\beta^i, \mu_{\theta_i}\right)$ and analyze the resulting subjective decision problem for type θ_i with parameterized payoff function $u_{\theta_i} : \Theta_{-i} \times Z \to \mathbb{R}$.

For a *fixed* θ_{-i}, the probability of action profile $a = (a_i, a_{-i})$ conditional on h given strategy s_i and conjecture β^i is

$$\mathbb{P}_{s_i,\beta^i}(a|\theta_{-i}, h) = \begin{cases} 0, & \text{if } a_i \neq s_i(h), \\ \beta^i(a_{-i}|\theta_{-i}, h), & \text{if } a_i = s_i(h). \end{cases}$$

Using the chain rule of conditional probabilities, we can define the probability of any history $h' \in H$ conditional on a prefix h ($h \prec h'$), given that s_i is played from h onward: let $h = (a^1, ..., a^{\ell(h)})$ and $h' = (a^1, ..., a^{\ell(h)}, ..., a^{\ell(h')})$, then

$$\mathbb{P}_{s_i,\beta^i}(h'|\theta_{-i}, h) = \prod_{t=\ell(h)+1}^{\ell(h')} \mathbb{P}_{s_i,\beta^i}\left(a^t|\theta_{-i}, \left(h, ..., a^{t-1}\right)\right),$$

with the convention that $\mathbb{P}_{s_i,\beta^i}\left(a^t|\theta_{-i}, \left(h, ..., a^{t-1}\right)\right) = \mathbb{P}_{s_i,\beta^i}\left(a^t|\theta_{-i}, h\right)$ if $t - 1 = \ell(h)$. With this, we can determine for all $s_i \in S_i$ and $h \in H$ the value for type θ_i of reaching h *given* θ_{-i} and given that s_i is followed from h onward:

$$V_{\theta_i}^{s_i,\beta^i}(\theta_{-i}, h) = \sum_{z \in Z(h)} \mathbb{P}_{s_i,\beta^i}(z|\theta_{-i}, h)\, u_{\theta_i}(\theta_{-i}, z),$$

which, of course, *depends only on the behavior of* s_i *at histories that weakly follow* h. Next we define, for all $s_i \in S_i$ and $h \in H$, the **subjective value for** θ_i **conditional on reaching** h given that s_i is followed from h onward:[28]

$$V_{\theta_i}^{s_i,\beta^i,\mu_i}(h) = \sum_{\theta_{-i} \in \Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h)\, V_{\theta_i}^{s_i,\beta^i}(\theta_{-i}, h)$$

$$= \sum_{\theta_{-i} \in \Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h) \sum_{z \in Z(h)} \mathbb{P}_{s_i,\beta^i}(z|\theta_{-i}, h) u_{\theta_i}(\theta_{-i}, z),$$

which—again—depends only on the behavior of s_i at histories that weakly follow h. Similarly, for every $a_i \in \mathcal{A}_i(h)$ we define the **value of** taking

[28] To ease notation, here and in the following formulas we write function symbols like V and $\mathbb{P}$ with μ_i instead of μ_{θ_i} in the superscript, and with θ_i as a subscript of the symbol. Consider the following interpretation: $\mu_i : \Theta_i \times H \to \Delta(\Theta_{-i})$ is the system of beliefs of player i before he observes his private information θ_i. Writing θ_i in the subscript of V or $\mathbb{P}$ means that we select from μ_i its section at θ_i, that is, μ_{θ_i}.

action a_i at h given that s_i will be followed from the next stage:

$$\begin{aligned} & V_{\theta_i}^{s_i,\beta^i,\mu_i}(h,a_i) \\ = & \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h) \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i},h) V_{\theta_i}^{s_i,\beta^i}(\theta_{-i},(h,(a_i,a_{-i}))), \end{aligned}$$

which depends only on the behavior of s_i at histories that follow h after taking action a_i.

Lemma 39. *For every $s_i \in S_i$ and for every $h \in H$,*

$$V_{\theta_i}^{s_i,\beta^i,\mu_i}(h,s_i(h)) = V_{\theta_i}^{s_i,\beta^i,\mu_i}(h).$$

Proof. By inspection of the definitions we have

$$\begin{aligned} & V_{\theta_i}^{s_i,\beta^i,\mu_i}(h,s_i(h)) \\ = & \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h) \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i},h)\, V_{\theta_i}^{s_i,\beta^i}(\theta_{-i},(h,(s_i(h),a_{-i}))) \\ = & \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h) \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i},h) \\ & \cdot \left(\sum_{z\in Z(h,(s_i(h),a_{-i}))} u_{\theta_i}(\theta_{-i},z)\, \mathbb{P}_{s_i,\beta^i}(z|\theta_{-i},(h,(s_i(h),a_{-i}))) \right) \\ = & \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h) \sum_{z\in Z(h)} u_{\theta_i}(\theta_{-i},z)\, \mathbb{P}_{s_i,\beta^i}(z|\theta_{-i},h) \\ = & \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h)\, V_{\theta_i}^{s_i,\beta^i}(\theta_{-i},h) \\ = & V_i^{s_i,\beta^i}(h) \end{aligned}$$

where the third equality follows from the chain rule. ■

As we did in Chapter 10, we define recursively the subjective value of reaching a history h and of taking an action a_i at h, under the presumption that the behavior of player i will be subjectively rational in the following stages. We use the symbol $\hat{V}_{\theta_i}^{\beta^i,\mu_i}$ to denote such values to emphasize that they are optimal given personal assessment $(\beta^i, \mu_{\theta_i})$. Similarly, for all $h \in H$ and $a_{-i} \in \mathcal{A}_{-i}(h)$, we let

$$\mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) = \sum_{\theta_{-i}\in\Theta_{-i}} \beta^i(a_{-i}|\theta_{-i}, h)\mu_{\theta_i}(\theta_{-i}|h) \tag{15.9.1}$$

denote the probability assigned by type θ_i to a_{-i} conditional on h according to personal assessment $(\beta^i, \mu_{\theta_i})$.[29] With this, the recursion is based on the height of histories $h \in H$, here denoted by $L(h) = \max_{z\in Z(h)} \ell(z) - \ell(h)$.[30]

- If $L(h) = 1$, then $(h, a) \in Z$ for every $a \in \mathcal{A}(h)$. Taking this into account, for each $a_i \in \mathcal{A}_i(h)$ let

$$\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i) = \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h) \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i}, h)u_{\theta_i}(\theta_{-i}, (h, (a_i, a_{-i}))),$$

  and let

$$\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h) = \max_{a_i\in\mathcal{A}_i(h)} \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i).$$

- Suppose $\hat{V}_{\theta_i}^{\beta^i,\mu_i}$ has been defined for every h with $L(h) \le k$. Then if $L(h) = k + 1$ let

$$\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i) = \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h)\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, (a_i, a_{-i})),$$

$$\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h) = \max_{a_i\in\mathcal{A}_i(h)} \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i).$$

Definition 97. *A strategy* $\bar{s}_i \in S_i$ *is*

[29] Here, we specify a system of beliefs *for each type,* and we focus on a specific type θ_i. Thus, we have to make it clear in our notation that the resulting probabilities depend on θ_i. To ease notation, we write $\mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(\cdot)$, putting the assessment in superscript.

[30] In Chapter 10, each $h \in H$ determined a subgame $\Gamma(h)$ and the height of h was written as the maximum length of such subgame, $L(\Gamma(h))$. Here we do not have a subgame in the same sense as Chapter 10, because payoffs are unknown, but we still have a subtree, which is all that matters.

- ***folding-back optimal given*** $\left(\beta^i, \mu_{\theta_i}\right)$ *if*

$$\forall h \in H,\ \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, \bar{s}_i(h)) = \max_{a_i \in \mathcal{A}_i(h)} \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i) = \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h);$$

- ***one-step optimal given*** $\left(\beta^i, \mu_{\theta_i}\right)$ *if*

$$\forall h \in H,\ V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, \bar{s}_i(h)) = \max_{a_i \in \mathcal{A}_i(h)} V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, a_i);$$

- ***sequentially optimal given*** $\left(\beta^i, \mu_{\theta_i}\right)$ *if*

$$\forall h \in H,\ V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h) = \max_{s_i \in S_i} V_{\theta_i}^{s_i,\beta^i,\mu_i}(h).$$

The following result is immediate from Definition 97:

Remark 53. *If a strategy is sequentially optimal given* $\left(\beta^i, \mu_{\theta_i}\right)$*, then it is also one-step optimal given* $\left(\beta^i, \mu_{\theta_i}\right)$*.*

The following results show that, under personal Bayes consistency, folding-back optimality and sequential optimality are equivalent to one-step optimality.

Theorem 61. *(*Folding Back Principle) *Fix a Bayes consistent assessment* $\left(\beta^i, \mu_{\theta_i}\right)$*. Then*
(I) A strategy $\bar{s}_i$ *is one-step optimal given* $\left(\beta^i, \mu_{\theta_i}\right)$ *if and only if*

$$\begin{aligned} V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, a_i) &= \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i), \\ V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h) &= \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h) \end{aligned} \tag{15.9.2}$$

for every $h \in H$ *and* $a_i \in \mathcal{A}_i(h)$*.*
(II) A strategy $\bar{s}_i$ *is folding-back optimal given* $\left(\beta^i, \mu_{\theta_i}\right)$ *if and only if* $\bar{s}_i$ *is one-step optimal given* $\left(\beta^i, \mu_{\theta_i}\right)$*.*

Proof. We prove by induction on the height of nonterminal histories that *if* $\bar{s}_i$ is *one-step optimal* given $\left(\beta^i, \mu_{\theta_i}\right)$ then (15.9.2) holds for every $h \in H$ and $a_i \in \mathcal{A}_i(h)$. As in the proof of Theorem 39 in Chapter 10, we omit (and leave to the reader) the proofs that the converse of this statement holds, that (II) holds, and that (I) is equivalent to (II).

Suppose that $\bar{s}_i$ is one-step optimal given $\left(\beta^i, \mu_{\theta_i}\right)$.

Basis step. Fix any $h \in H$ such that $L(h) = 1$. Then, for each $a_i \in \mathcal{A}_i(h)$, by inspection of the definitions of $V_{\theta_i}^{s_i,\beta^i,\mu_i}(h, a_i)$ and $\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i)$ (see the basis step in the recursion) we obtain

$$\begin{aligned} & V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, a_i) \\ = & \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h) \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i}, h) u_i(\theta_i, \theta_{-i}, (h, (a_i, a_{-i}))) \\ = & \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i). \end{aligned}$$

Since $\bar{s}_i$ is one-step optimal,

$$\begin{aligned} V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h) &= V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, \bar{s}_i\,(h)) = \max_{a_i\in\mathcal{A}_i(h)} V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, a_i) \\ &= \max_{a_i\in\mathcal{A}_i(h)} \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i) = \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h). \end{aligned}$$

Inductive step. Suppose that (15.9.2) holds for each $h \in H$ with $L(h) \leq k$. Now, fix any h with $L(h) = k + 1$. Then $L(h, a) \leq k$ for each $a \in \mathcal{A}(h)$. Therefore, for every $a_i \in \mathcal{A}_i(h)$,

$$\begin{aligned} & V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, a_i) \\ \overset{\text{(def.)}}{=} & \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|h) \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i}, h) V_{\theta_i}^{\bar{s}_i,\beta^i}(\theta_{-i}, (h, (a_i, a_{-i}))) \\ = & \sum_{\theta_{-i}\in\Theta_{-i}} \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \beta^i(a_{-i}|\theta_{-i}, h) \mu_{\theta_i}(\theta_{-i}|h) V_{\theta_i}^{\bar{s}_i,\beta^i}(\theta_{-i}, (h, (a_i, a_{-i}))) \\ \overset{(15.3.4)}{=} & \sum_{\theta_{-i}\in\Theta_{-i}} \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \mu_{\theta_i}(\theta_{-i}|\,(h, (a_i, a_{-i}))) \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) V_{\theta_i}^{\bar{s}_i,\beta^i}(\theta_{-i}, (h, (a_i, a_{-i}))) \\ = & \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) \sum_{\theta_{-i}\in\Theta_{-i}} \mu_{\theta_i}(\theta_{-i}|\,(h, (a_i, a_{-i}))) V_{\theta_i}^{\bar{s}_i,\beta^i}(\theta_{-i}, (h, (a_i, a_{-i}))) \\ \overset{\text{(def.)}}{=} & \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h, (a_i, a_{-i})) \\ \overset{\text{(I.H.)}}{=} & \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, (a_i, a_{-i})) \overset{\text{(def.)}}{=} \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h, a_i), \end{aligned}$$

where the first, fifth, and seventh equalities hold by definition, the third equality follows from eq. (15.3.4) in the definition of Bayes consistency, the sixth equality follows from the inductive hypothesis, and the remaining equalities are immediate. To see that the third equality follows from Bayes consistency, note that eq. (15.3.4) yields

$$\mu_{\theta_i}(\theta_{-i}|\,(h,(a_i,a_{-i})))\mathbb{P}^{\beta^i,\mu_i}_{\theta_i}(a_{-i}|h) = \beta^i(a_{-i}|\theta_{-i},h)\mu_{\theta_i}(\theta_{-i}|h),$$

where both sides of the equality represent the probability of (θ_{-i}, a_{-i}) given h.

Thus we obtain

$$\begin{aligned} V^{\bar{s}_i,\beta^i,\mu_i}_{\theta_i}(h) &= V^{\bar{s}_i,\beta^i,\mu_i}_{\theta_i}(h,\bar{s}_i(h)) \overset{\text{(loc.opt.)}}{=} \max_{a_i\in\mathcal{A}_i(h)} V^{\bar{s}_i,\beta^i,\mu_i}_{\theta_i}(h,a_i) \\ &= \max_{a_i\in\mathcal{A}_i(h)} \hat{V}^{\beta^i,\mu_i}_{\theta_i}(h,a_i) = \hat{V}^{\beta^i,\mu_i}_{\theta_i}(h), \end{aligned}$$

where the first equality follows from Lemma 39 and the second equality holds because $\bar{s}_i$ is locally optimal (loc.opt.) at h. ■

Theorem 62. (Optimality principle) *Fix a Bayes consistent assessment* $\left(\beta^i, \mu_{\theta_i}\right)$. *A strategy of player i is sequentially optimal given* $\left(\beta^i, \mu_{\theta_i}\right)$ *if and only if it is folding-back optimal given* $\left(\beta^i, \mu_{\theta_i}\right)$.

Proof. (If) Let $\bar{s}_i$ be *folding-back optimal* given $\left(\beta^i, \mu_{\theta_i}\right)$. We will prove by induction on the height of nonterminal histories that

$$\forall h \in H,\ \hat{V}^{\beta^i,\mu_i}_{\theta_i}(h) \geq \max_{s_i\in S_i} V^{s_i,\beta^i,\mu_i}_{\theta_i}(h).$$

With this, since $\bar{s}_i$ is folding-back optimal, it is one-step optimal (Theorem 61 II), therefore $V^{\bar{s}_i,\beta^i,\mu_i}_{\theta_i}(h) = \hat{V}^{\beta^i,\mu_i}_{\theta_i}(h)$ for every $h \in H$ (Theorem 61 I); this implies that $\bar{s}_i$ is sequentially optimal given $\left(\beta^i, \mu_{\theta_i}\right)$.

Basis step. Fix any $h \in H$ such that $L(h) = 1$. By Theorem 61, $\hat{V}^{\beta^i,\mu_i}_{\theta_i}(h,a_i) = V^{\bar{s}_i,\beta^i,\mu_i}_{\theta_i}(h,a_i)$ for every $a_i \in \mathcal{A}_i(h)$. Since, starting from h, every action profile $(a_i, a_{-i}) \in \mathcal{A}(h)$ terminates the game,

$$\max_{a_i\in\mathcal{A}_i(h)} V^{\bar{s}_i,\beta^i,\mu_i}_{\theta_i}(h,a_i) = \max_{s_i\in S_i} V^{s_i,\beta^i,\mu_i}_{\theta_i}(h).$$

Therefore

$$\begin{aligned}\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h) &= \max_{a_i\in\mathcal{A}_i(h)} \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h,a_i)\\ &= \max_{a_i\in\mathcal{A}_i(h)} V_{\theta_i}^{\bar{s}_i,\beta^i,\mu_i}(h,a_i) = \max_{s_i\in S_i} V_{\theta_i}^{s_i,\beta^i,\mu_i}(h).\end{aligned}$$

Inductive step. Suppose by way of induction that $\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h') \geq \max_{s_i\in S_i} V_{\theta_i}^{s_i,\beta^i,\mu_i}(h')$ for every $h' \in H$ with $L(h') \leq k$. Now fix any h with $L(h) = k+1$. Then $L(h,a) \leq k$ for each $a \in \mathcal{A}(h)$, and the inductive assumption (I.H.) yields

$$\begin{aligned}&\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h)\\ &\overset{\text{(def.)}}{\geq} \hat{V}_{\theta_i}^{\beta^i,\mu_i}(h,a_i)\\ &\overset{\text{(def.)}}{=} \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h)\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h,(a_i,a_{-i}))\\ &\overset{\text{(I.H.)}}{\geq} \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) \max_{s_i\in S_i} V_{\theta_i}^{s_i,\beta^i,\mu_i}(h,(a_i,a_{-i}))\end{aligned}$$

for every $a_i \in \mathcal{A}_i(h)$. Therefore,

$$\begin{aligned}&\hat{V}_{\theta_i}^{\beta^i,\mu_i}(h)\\ \geq\;& \max_{a_i\in\mathcal{A}_i(h)} \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) \max_{s_i\in S_i} V_{\theta_i}^{s_i,\beta^i,\mu_i}(h,(a_i,a_{-i}))\\ =\;& \max_{a_i\in\mathcal{A}_i(h)} \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \max_{s_i\in S_i} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) V_{\theta_i}^{s_i,\beta^i,\mu_i}(h,(a_i,a_{-i}))\\ =\;& \max_{a_i\in\mathcal{A}_i(h)} \max_{s_i\in S_i} \sum_{a_{-i}\in\mathcal{A}_{-i}(h)} \mathbb{P}_{\theta_i}^{\beta^i,\mu_i}(a_{-i}|h) V_{\theta_i}^{s_i,\beta^i,\mu_i}(h,(a_i,a_{-i}))\\ =\;& \max_{s_i\in S_i} V_{\theta_i}^{s_i,\beta^i,\mu_i}(h)\end{aligned}$$

where the first equality is obvious, the second holds because each term in the summation can be maximized independently of the others, and the third equality holds because global maximization from h is equivalent to picking at h an action that yields the largest maximal continuation-value.

(Only if) If $\bar{s}_i$ is sequentially optimal, then it is one-step optimal (Remark 53), which in turn implies that it is folding-back optimal (Theorem 61 II). ■

Theorem 61 and Theorem 62 yield the OD principle:

Corollary 9. (One-deviation principle) *Fix a Bayes consistent assessment* $\left(\beta^i, \mu_{\theta_i}\right)$. *A strategy of player i is sequentially optimal given* $\left(\beta^i, \mu_{\theta_i}\right)$ *if and only if it is one-step optimal given* $\left(\beta^i, \mu_{\theta_i}\right)$.

16

Imperfectly Observed Actions

TBD

Bibliography

[1] Abreu, D. (1988): "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica,* 56, 383-396.

[2] Abreu, D., D. Pearce, and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica,* 58, 1041-1063.

[3] Aliprantis, C. R. and K. C. Border (2006 [1999]): *Infinite Dimensional Analysis*. Berlin: Springer-Verlag.

[4] Attanasi, G. and R. Nagel (2007): "A Survey of Psychological Games: Theoretical Findings and Experimental Evidence," in *Games, Rationality and Behaviour. Essays on Behavioural Game Theory and Experiments*, ed. by A. Innocenti, and P. Sbriglia. Houndmills: Palgrave McMillan, 204-232.

[5] Aumann, R. J. (1974): "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics,* 1, 67-96.

[6] Aumann, R. J. (1987): "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica,* 55, 1-18.

[7] Battigalli, P. (1996): "Strategic Rationality Orderings and the Best Rationalization Principle," *Games and Economic Behavior,* 13, 178-200.

[8] Battigalli P. (2003): "Rationalizability in Infinite, Dynamic Games of Incomplete Information," *Research in Economics,* 57, 1–38.

[9] Battigalli, P. and G. Bonanno (1999): "Recent Results on Belief, Knowledge and the Epistemic Foundations of Game Theory," *Research in Economics*, 53, 149-226.

[10] Battigalli, P. and E. Catonini (2025): "Monotonicity and Robust Implementation Under Forward-Induction Reasoning," *Econometrica*, forthcoming.

[11] Battigalli, P. and N. De Vito (2021): "Beliefs, Plans, and Perceived Intentions in Dynamic Games," *Journal of Economic Theory*, 195, 105283.

[12] Battigalli, P. and M. Dufwenberg (2009): "Dynamic Psychological Games," *Journal of Economic Theory*, 144, 1-35.

[13] Battigalli, P. and M. Dufwenberg (2022): "Belief-Dependent Motivations and Psychological Game Theory," *Journal of Economic Literature*, 60, 833-882.

[14] Battigalli, P., and N. Generoso (2024): "Information Flows and Memory in Games." *Games and Economic Behavior*, 145, 356-376.

[15] Battigalli, P. and A. Friedenberg (2012): "Forward Induction Reasoning Revisited," *Theoretical Economics*, 7, 57-98.

[16] Battigalli, P. and M. Siniscalchi (2002): "Strong Belief and Forward Induction Reasoning," *Journal of Economic Theory*, 106, 356-391.

[17] Battigalli, P. and M. Siniscalchi (2003): "Rationalization and Incomplete Information," *Advances in Theoretical Economics*, 3 (1), Article 3.

[18] Battigalli, P., R. Corrao, and M. Dufwenberg (2019): "Incorporating Belief-Dependent Motivation in Games," *Journal of Economic Behavior and Organization*, 167, 185-218.

[19] Battigalli, P., M. Gilli, and M. C. Molinari (1992): "Learning and Convergence to Equilibrium in Repeated Strategic Interaction," *Research in Economics (Ricerche Economiche),* 46, 335-378.

[20] Battigalli, P., S. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci (2015): "Selfconfirming Equilibrium and Model Uncertainty," *American Economic Review*, 105, 646-677.

[21] Battigalli, P., S. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci (2016): "A Note on Comparative Ambiguity Aversion and Justifiability," *Econometrica*, 84, 1903-1916.

[22] Battigalli, P., A. Francetich, G. Lanzani, and M. Marinacci (2019): "Learning and Self-Confirming Long-Run Biases," *Journal of Economic Theory,* 183, 740-785.

[23] Battigalli, P., S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and T. Sargent (2022): "A Framework for the Analysis of Self-Confirming Policies," *Theory and Decisions,* 92, 455-512.

[24] Battigalli, P., A. Di Tillio, E. Grillo, and A. Penta (2011): "Interactive Epistemology and Solution Concepts for Games with Asymmetric Information," *The B.E. Journal of Theoretical Economics*, 11 (1) (*Advances*), Article 6.

[25] Ben-Porath, E. and E. Dekel (1992): "Signaling Future Actions and the Potential for Sacrifice," *Journal of Economic Theory*, 57, 36-51.

[26] Bernheim, D. (1984): "Rationalizable Strategic Behavior," *Econometrica*, 52, 1007-1028.

[27] Binmore, K. G., (1982 [1977]): *Mathematical Analysis: A Straightforward Approach.* Cambridge UK: Cambridge University Press.

[28] Brandenburger, A. (2007): "The Power of Paradox," *International Journal of Game Theory,* 35, 465-492.

[29] Brandenburger, A. and E. Dekel (1987): "Rationalizability and Correlated Equilibria," *Econometrica*, 55**,** 1391-1402.

[30] Boergers, T. and M. Janssen (1995): "On the Dominance Solvability of Large Cournot Games," *Games and Economic Behavior*, 8, 297-321.

[31] CATONINI, E. (2019): "Rationalizability and Epistemic Priority Orderings," *Games and Economic Behavior*, 114, 101-117.

[32] CATONINI, E. (2020): "On Non-Monotonic Strategic Reasoning," *Games and Economic Behavior*, 120, 209-224.

[33] CATONINI, E. (2024): "Iterated Admissibility Does Not Refine Extensive-Form Rationalizability," *Economic Journal*, 134, 3017-3026.

[34] CATONINI, E., AND A. PENTA (2025): "Backward Induction Reasoning beyond Backward Induction," *American Economic Journal: Microeconomics*, forthcoming.

[35] CERREIA-VIOGLIO, S., F. MACCHERONI, M. MARINACCI AND L. MONTRUCCHIO (2013): "Ambiguity and Robust Statistics," *Journal of Economic Theory*, 148, 974-1049.

[36] CHEN, J., AND S. MICALI (2013): "The Order Independence of Iterated Dominance in Extensive Games," *Theretical Economics*, 8, 125-163.

[37] DAVEY, B.A., AND H.A. PRIESTLEY (2002): *Introduction to Lattices and Order*. Cambridge UK: Cambridge University Press.

[38] DEKEL, E. AND M. SINISCALCHI (2014): "Epistemic Game Theory," in P. Young and S. Zamir (Eds.) *Handbook of Game Theory*, Vol. 4, 619-702.

[39] DEKEL, E., D. FUDENBERG, AND D. LEVINE (2004): "Learning to Play Bayesian Games," *Games and Economic Behavior*, 46, 282-303.

[40] DEKEL, E., D. FUDENBERG, AND S. MORRIS (2007): "Interim Correlated Rationalizability," *Theoretical Economics*, 2, 15-40.

[41] DUFWENBERG, M. AND M. STEGEMAN (2002): "Existence and Uniqueness of Maximal Reductions Under Iterated Strict Dominance," *Econometrica*, 70, 2007-2023.

[42] ESPONDA, I. (2013): "Rationalizable Conjectural Equilibrium: A Framework for Robust Predictions," *Theoretical Economics*, 8, 467-501.

[43] Eyster, E. and M. Rabin (2005): "Cursed Equilibrium," *Econometrica*, 73, 1623-1672.

[44] Friedman, J.W. (1987): *Game Theory with Applications to Economics,* Oxford: Oxford University Press.

[45] Fudenberg, D. and D. Levine (1998): *The Theory of Learning in Games*, Cambridge MA: MIT Press.

[46] Fudenberg, D. and J. Tirole (1991): *Game Theory,* Cambridge MA: MIT Press.

[47] Geanakoplos, J., D. Pearce, and E. Stacchetti (1989): "Psychological Games and Sequential Rationality," *Games and Economic Behavior*, 1, 60-79.

[48] Gibbons, R. (1992): *Game Theory for Applied Economists.* Princeton: Princeton University Press.

[49] Gilboa, I. and M. Marinacci (2013): "Ambiguity and the Bayesian Paradigm," in D. Acemoglu, M. Arellano, and E. Dekel (Eds.) *Advances in Economics and Econometrics: Theory and Applications (Vol. 1)*, Cambridge UK: Cambridge University Press, 179-242.

[50] Gilli, M. (1999): "Adaptive Learning in Imperfect Monitoring Games," *Review of Economic Dynamics*, 2, 472-485.

[51] Guesnerie, R. (1992): "An Exploration on the Eductive Justifications of the Rational-Expectations Hypothesis," *American Economic Review*, 82, 1254-1278.

[52] Harsanyi, J. (1967-68): "Games of Incomplete Information Played by Bayesian Players. Parts I, II, III," *Management Science*, 14, 159-182, 320-334, 486-502.

[53] Harsanyi, J. (1995): "Games with Incomplete Information," *American Economic Review*, 85, 291-303.

[54] Hellwig, M., W. Leininger, P. Reny, and A. Robson (1990): "Subgame Perfect Equilibrium in Continuous Games of

Perfect Information: An Elementary Approach to Existence and Approximation by Discrete Games," *Journal of the Economic Theory,* 52, 406-422.

[55] Kuhn, H. W. (1953): "Extensive Games and the Problem of Information," in: H.W. Kuhn and A.W. Tucker (Eds.), *Contributions to the Theory of Games II*, Princeton NJ: Princeton University Press, 193-216.

[56] Marinacci, M. (2015): "Model Uncertainty," *Journal of the European Economic Association*, 13, 1022-1100.

[57] Maynard Smith, J. (1982): *Evolution and the Theory of Games,* Cambridge: Cambridge University Press.

[58] Milgrom, P. and J. Roberts (1991): "Adaptive and Sophisticated Learning in Normal-Form Games," *Games and Economic Behavior*, 3, 82-100.

[59] Milgrom, P., and R. Weber (1985): "Distributional Strategies for Games with Incomplete Information," *Mathematics of Operation Research,* 10, 619-632.

[60] Motta, M. (2004): *Competition Policy: Theory and Practice.* Cambridge University Press.

[61] Moulin, H. (1984): "Dominance-Solvability and Cournot Stability," *Mathematical Social Sciences*, 7, 83-102.

[62] Nash, J. (1950): "Equilibrium Points in N-Person Games," *Proceedings of the National Academy of Sciences*, 36**,** 48-49.

[63] Nash, J. (1950): *Non-cooperative Games.* Ph.D. thesis, Princeton University.

[64] Ok, E. A. (2007): *Real Analysis with Economic Applications*, Princeton NJ: Princeton University Press.

[65] Osborne, M. and A. Rubinstein (1994): *A Course in Game Theory*, Cambridge MA: MIT Press.

[66] Pearce, D. (1984): "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, 52, 1029-1050.

[67] Penta, A. (2015): "Robust Dynamic Implementation," *Journal of Economic Theory*, 160, 280–316.

[68] Perea, A. (2001): *Rationality in Extensive Form Games*, Boston, Dordrecht: Kluwer Academic Publishers.

[69] Perea, A. (2014): "Belief in the Opponents' Future Rationality," *Games and Economic Behavior,* 83, 231-254.

[70] Perea, A. (2018): "Why Forward Induction Leads to the Backward Induction Outcome: A New Proof for Battigalli's Theorem," *Games and Economic Behavior*, 110, 120-138.

[71] Reny, P. (1992): "Backward Induction, Normal Form Perfection and Explicable Equilibria," *Econometrica*, 60, 626-649.

[72] Rubinstein, A. (1982): "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50, 97-109.

[73] Rubinstein, A. (1989): "The Electronic Mail Game: Strategic Behavior Under Almost Common Knowledge," *American Economic Review*, 79, 385-391.

[74] Rubinstein, A. (1989): "Comments on the Interpretation of Game Theory," *Econometrica*, 59, 909-924.

[75] Sandholm, W. (2010): *Population Games and Evolutionary Dynamics,* Cambridge MA: MIT Press.

[76] Selten, R. (1975): "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory,* 4, 25-55.

[77] Sion, M. (1958): "On General Minimax Theorems," *Pacific Journal of Mathematics*, 171-176.

[78] Shimoji, M. (2004): "On the Equivalence of Weak Dominance and Sequential Best Response," *Games and Economic Behavior*, 48, 385-402.

[79] Shimoji, M. (2004): "On Forward Induction in Money-Burning Games," *Economic Theory*, 19, 637-648.

[80] Shimoji, M. and J. Watson (1998): "Conditional Dominance, Rationalizability, and Game Forms," *Journal of Economic Theory,* 83, 161-195.

[81] Topkis, D. M. (1998): *Supermodularity and Complementarity.* Princeton University Press.

[82] Vickrey, W. (1961). "Counterspeculation, Auctions, and Competitive Sealed Tenders," *Journal of Finance*, 16, 8-37.

[83] von Neumann, J. (1928): "Zur Theorie der Gesellschaftsspiele," *Math. Annalen*, 100, 295-320.

[84] von Neumann, J. and O. Morgenstern (2004 [1944]): *Theory of Games and Economic Behavior,* Princeton NJ: Princeton University Press.

[85] Wald, A. (1949): "Statistical Decision Functions," *Annals of Mathematical Statistics*, 20, 165-205.

[86] Wald, A. (1950): *Statistical Decision Functions*. New York: Wiley.

[87] Weibull, J. (1995): *Evolutionary Game Theory,* Cambridge MA: MIT Press.

[88] Weibull, J. W. (1996): "The Mass Action Interpretation" in H. Kuhn et al. "The work of John Nash in game theory,"*Journal of Economic Theory*, 69, 153-185.

[89] Weinstein, J. (2016): "The Effect of Changes in Risk Attitude on Strategic Behavior," *Econometrica*, 84, 1881-1902.