10. Rationality

10.1 Models of strategic-form games

In Chapter 2 (Section 2.5) and Chapter 6 (Section 6.4) we discussed the procedure of iterative elimination of strictly dominated strategies and claimed that it captures the notion of common knowledge of rationality in strategic-form games. We now have the tools to state this precisely.

The *epistemic foundation program* in game theory aims to identify, for every game, the strategies that might be chosen by rational and intelligent players who know the structure of the game and the preferences of their opponents and who recognize each other's rationality. The two central questions are thus:

(1) Under what circumstances is a player rational?

(2) What does 'mutual recognition of rationality' mean?

A natural interpretation of the latter notion is in terms of 'common knowledge of rationality'.[1] We already know how to model the notion of common knowledge (Chapter 8); thus, we only need to define what it means for a player to be rational. Intuitively, a player is rational if she chooses an action which is "best" given what she believes. In order to make this more precise we need to introduce the notion of a model of a game. We shall focus on strategic-form games with cardinal payoffs (Definition 6.1.2, Chapter 6).

[1] As pointed out in Chapter 8, a defining characteristic of knowledge is truth: if, at a state, a player knows event E then, at that state, E must be true. A more general notion is that of belief, which allows for the possibility of mistakes: believing E is compatible with E being false. Thus a more appealing notion is that of *common belief* of rationality; however, in order to simplify the exposition, we shall restrict attention to the notions of knowledge and common knowledge developed in Chapter 8. For the analysis of common belief and an overview of the epistemic foundation program in game theory the reader is referred to Battigalli and Bonanno (1999).

The definition of a strategic-form game specifies the choices available to the players and what motivates those choices (their preferences over the possible outcomes); however, it leaves out an important factor in the determination of players' choices, namely what they believe about the other players. Adding a specification of the players' knowledge and beliefs determines the context in which a particular game is played; this can be done with the help of an interactive knowledge-belief structure.

Recall that an interactive knowledge-belief structure consists of

- ◇ a set of states W,
- ◇ n partitions $\mathscr{I}_1, \mathscr{I}_2, \ldots, \mathscr{I}_n$ of W,
- ◇ a collection of probability distributions on W, one for each information set of each partition, whose support is a subset of that information set.

The probability distributions encode the beliefs of the players in each possible state of knowledge.
Recall also that if w is a state we denote by $I_i(w)$ the element of the partition $\mathscr{I}_i$ that contains w, that is, the information set of Player i at state w and by $P_{i,w}$ the probability distribution on $I_i(w)$, representing the beliefs of Player i at state w.

R In Chapter 6 we used the symbol σ_i to denote a mixed strategy of Player i. In this chapter the symbol σ_i will have a different meaning (see the following definition). Since mixed strategies play no role in this chapter, there is be no possibility of confusion. Only pure strategies are considered throughout this chapter.

Definition 10.1.1 Let G be a n-player strategic-form game with cardinal payoffs. A *model* of G is an interactive knowledge-belief structure together with a function

$$\sigma_i : W \to S_i$$

(for every Player i) that associates with every state a pure strategy of Player i (recall that S_i denotes the set of pure strategies of Player i). The interpretation of $s_i = \sigma_i(w)$ is that, at state w, Player i plays (or has chosen) strategy s_i. We impose the restriction that a player always knows what strategy she has chosen, that is, the function σ_i is constant on each information set of Player i:

$$\text{if } \ w' \in I_i(w) \ \text{ then } \ \sigma_i(w') = \sigma_i(w).$$

The addition of the functions σ_i to an interactive knowledge-belief structure yields an interpretation of events in terms of propositions about what strategies the players choose, thereby giving content to players' knowledge and beliefs.

Figure 10.1 shows a strategic-form game with cardinal payoffs and Figure 10.2 a model of it.

		Player 2		
		L	C	R
Player 1	T	4 , 6	3 , 2	8 , 0
	M	0 , 9	0 , 0	4 , 12
	B	8 , 3	2 , 4	0 , 0

Figure 10.1: A strategic-form game with cardinal payoffs.

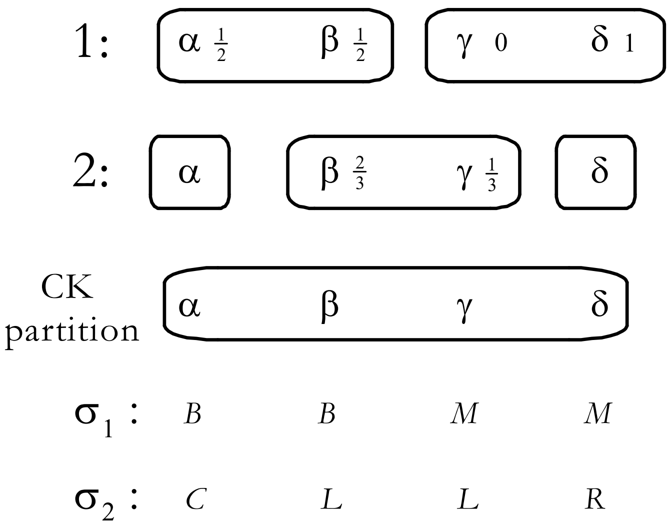


Figure 10.2: A model of the game of Figure 10.1.

State β in Figure 10.2 describes the following situation:

- Player 1 plays B ($\sigma_1(\beta) = B$) and Player 2 plays L ($\sigma_2(\beta) = L$).
- Player 1 (whose information set is $\{\alpha, \beta\}$) is uncertain as to whether Player 2 has chosen to play C ($\sigma_2(\alpha) = C$) or L ($\sigma_2(\beta) = L$); furthermore, he attaches probability $\frac{1}{2}$ to each of these two possibilities.
- Player 2 (whose information set is $\{\beta, \gamma\}$) is uncertain as to whether Player 1 has chosen to play B ($\sigma_1(\beta) = B$) or M ($\sigma_1(\gamma) = M$); furthermore, she attaches probability $\frac{2}{3}$ to Player 1 playing B and $\frac{1}{3}$ to Player 1 playing M.

Are the two players' choices rational at state β? Definition 10.1.2 provides an answer to this question.

To summarize: one can view a strategic-form game as only a partial description of an interactive situation: it specifies who the players are, what actions they can take and how they rank the possible outcomes; a model of the game completes this description by also specifying what each player actually does and what she believes about what the other players are going to do. Once we know what a player does and what she believes, then we are in a position to judge her choice to be either rational or irrational.

The following definition says that at a state (of a model) a player is rational if her choice at that state is optimal, given her beliefs, that is, if there is no other pure strategy that would give her a higher expected payoff, given what she believes about the choices of her opponents.

Given a state w and a Player i, we denote by $\sigma_{-i}(w)$ the profile of strategies chosen at w by the players other than i : $\sigma_{-i}(w) = (\sigma_1(w), ..., \sigma_{i-1}(w), \sigma_{i+1}(w), ..., \sigma_n(w))$. Furthermore, as in Chapter 8, we denote by $P_{i,w}$ the beliefs of Player i at state w, that is, $P_{i,w}$ is a probability distribution over $I_i(w)$ (although we can think of $P_{i,w}$ as a probability distribution over the entire set of states W satisfying the property that if $w' \notin I_i(w)$ then $P_{i,w}(w') = 0$). Note that, for every Player i, there is just one probability distribution over $I_i(w)$ that is, if $w' \in I_i(w)$ then $P_{i,w'} = P_{i,w}$.

Definition 10.1.2 Player i is *rational* at state w if, given her beliefs $P_{i,w}$, there is no pure strategy $s_i' \in S_i$ of hers that yields a higher expected payoff than $\sigma_i(w)$ (recall that $\sigma_i(w)$ is the strategy chosen by Player i at state w): for all $s_i' \in S_i$,

$$\sum_{w' \in I_i(w)} P_{i,w}(w') \times \pi_i\left(\sigma_i(w), \sigma_{-i}(w')\right) \geq \sum_{w' \in I_i(w)} P_{i,w}(w') \times \pi_i\left(s'_i, \sigma_{-i}(w')\right).$$

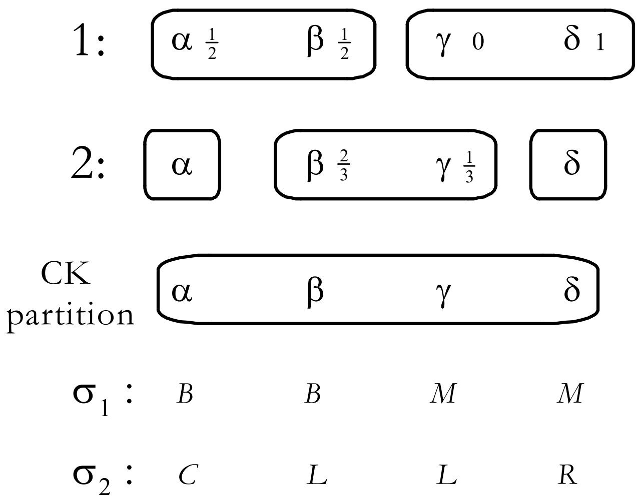
To illustrate Definition 10.1.2 let us re-consider the example of Figures 10.1 and 10.2, which are reproduced on the following page.

Consider first Player 1. Given his beliefs and his choice of B, his expected payoff is: $\frac{1}{2}\,\pi_1(B,C) + \frac{1}{2}\,\pi_1(B,L) = \frac{1}{2}(2) + \frac{1}{2}(8) = 5$. Given his beliefs about Player 2, could he obtain a higher expected payoff with a different choice?
With M his expected payoff would be $\frac{1}{2}\,\pi_1(M,C) + \frac{1}{2}\,\pi_1(M,L) = \frac{1}{2}(0) + \frac{1}{2}(0) = 0$, while with T his expected payoff would be $\frac{1}{2}\,\pi_1(T,C) + \frac{1}{2}\,\pi_1(T,L) = \frac{1}{2}(3) + \frac{1}{2}(4) = 3.5$. Thus, given his beliefs, Player 1's choice of B is optimal and we can say that Player 1 is rational at state β.

Consider now Player 2. Given her beliefs and her choice of L, her expected payoff is: $\frac{2}{3}\,\pi_2(B,L) + \frac{1}{3}\,\pi_2(M,L) = \frac{2}{3}(3) + \frac{1}{3}(9) = 5$. Given her beliefs about Player 1, could she obtain a higher expected payoff with a different choice?
With C her expected payoff would be $\frac{2}{3}\,\pi_2(B,C) + \frac{1}{3}\,\pi_2(M,C) = \frac{2}{3}(4) + \frac{1}{3}(0) = \frac{8}{3}$, while with R her expected payoff would be $\frac{2}{3}\,\pi_2(B,R) + \frac{1}{3}\,\pi_2(M,R) = \frac{2}{3}(0) + \frac{1}{3}(12) = 4$.

Thus, given her beliefs, Player 2's choice of L is optimal and we can say that also Player 2 is rational at state β.

		Player 2		
		L	C	R
Player 1	T	4 , 6	3 , 2	8 , 0
	M	0 , 9	0 , 0	4 , 12
	B	8 , 3	2 , 4	0 , 0



Given a model of a game, using Definition 10.2 we can determine, for every player, the set of states where that player is rational. Let $\mathbf{R}_i$ be the event that (that is, the set of states at which) Player i is rational and let $\mathbf{R} = \mathbf{R}_1 \cap \ldots \cap \mathbf{R}_n$ be the event that all players are rational.

In the example of Figures 10.1 and 10.2 we have that $\mathbf{R}_1 = \{\alpha, \beta\}$.

Indeed, it was shown above that $\beta \in \mathbf{R}_1$; since Player 1's choice and beliefs at α are the same as at β, it follows that also $\alpha \in \mathbf{R}_1$.

On the other hand, Player 1 is not rational at state γ because he is certain (that is, attaches probability 1 to the fact) that Player 2 plays R and his choice of M is not optimal against R (the unique best response to R is T). The same is true of state δ.

For Player 2 we have that $\mathbf{R}_2 = \{\alpha, \beta, \gamma, \delta\}$, that is, Player 2 is rational at every state (the reader should convince himself/herself of this).

Thus $\mathbf{R} = \mathbf{R}_1 \cap \mathbf{R}_2 = \{\alpha, \beta\} \cap \{\alpha, \beta, \gamma, \delta\} = \{\alpha, \beta\}$.

Test your understanding of the concepts introduced in this section, by going through the exercises in Section 10.5.1 at the end of this chapter.

10.2 Common knowledge of rationality in strategic-form games

Since we have defined the rationality of a player and the rationality of all the players as events, we can apply the knowledge operator and the common knowledge operator to these events; thus we can determine, for example, if there are any states where the rationality of all the players is common knowledge and see what strategy profiles are compatible with common knowledge of rationality.

R Note that, for every player i, $K_i\mathbf{R}_i = \mathbf{R}_i$, that is, every player is rational if and only if she knows it (this is a consequence of the assumption that a player always knows what strategy she has chosen and knows her own beliefs:
if $w' \in I_i(w)$ then $\sigma_i(w') = \sigma_i(w)$ and $P_{i,w'} = P_{i,w}$).

Note also that $CK\mathbf{R} \subseteq \mathbf{R}$, that is, if it is common knowledge that all the players are rational, then they are indeed rational (while, in general, it is **not** true that $\mathbf{R}$ is a subset of $CK\mathbf{R}$, that is, it is possible for all the players to be rational without it being common knowledge).

In Chapters 2 (for ordinal strategic-form games) and 6 (for cardinal strategic-form games) we claimed that the iterated deletion of strictly dominated strategies corresponds to the notion of common knowledge of rationality. We are now able to state this precisely. The following two theorems establish the correspondence between the two notions for the case of strategic-form games with cardinal payoffs. A corresponding characterization holds for strategic-form games with ordinal payoffs.[2] The proofs of Theorems 10.2.1 and 10.2.2 are given in Section 10.4.

Theorem 10.2.1 Given a finite strategic-form game with cardinal payoffs G, the following is true: for any model of G and any state w in that model, if $w \in CK\mathbf{R}$ (that is, at w there is common knowledge of rationality) then the pure-strategy profile associated with w must be one that survives the Cardinal Iterated Deletion of Strictly Dominated Strategies (Definition 6.4.1, Chapter 6).

Theorem 10.2.2 Given a finite strategic-form game with cardinal payoffs G, the following is true: if $s \in S$ is a pure-strategy profile that survives the Cardinal Iterated Deletion of Strictly Dominated Strategies, then there is a model of G and a state w in that model, such that $w \in CK\mathbf{R}$ and the strategy profile associated with w is s.

[2] See Bonanno (2015).

As an application of Theorem 10.2.1 consider the game of Figure 10.3 (which reproduces Figure 10.2).

		Player 2		
		L	C	R
Player 1	T	4 , 6	3 , 2	8 , 0
	M	0 , 9	0 , 0	4 , 12
	B	8 , 3	2 , 4	0 , 0

Figure 10.3: Copy of Figure 10.2.

In this game the iterated deletion of strictly dominated strategies leads to the following strategy profiles: $(T,L),(T,C),(B,L)$ and (B,C).[3] Thus these are the only strategy profiles that are compatible with common knowledge of rationality: by Theorem 10.1, at a state in a model of this game where there is common knowledge of rationality the players can play only one of these strategy profiles. Furthermore, by Theorem 10.2.2, any of these four strategy profiles can in fact be played in a situation where there is common knowledge of rationality.

It is worth stressing that common knowledge of rationality does *not* imply that the players play a Nash equilibrium: indeed, none of the above four strategy profiles is a Nash equilibrium. However, since a pure-strategy Nash equilibrium always survives the iterated deletion of strictly dominated strategies, the pure-strategy Nash equilibria (if any) are always compatible with common knowledge of rationality.

Test your understanding of the concepts introduced in this section, by going through the exercises in Section 10.5.2 at the end of this chapter.

[3]For Player 1, M is strictly dominated by any mixed strategy $\begin{pmatrix} T & B \\ p & 1-p \end{pmatrix}$ with $p > \frac{1}{2}$. After deletion of M, for Player 2 R becomes strictly dominated by either of the other two strategies.

10.3 Common knowledge of rationality in extensive-form games

So far we have only discussed the implications of common knowledge of rationality in strategic-form games, where the players make their choices simultaneously (or in ignorance of the other players' choices). We now turn to a brief discussion of the issues that arise when one attempts to determine the implications of common knowledge of rationality in dynamic (or extensive-form) games.

How should a model of a dynamic game be constructed? One approach in the literature has been to consider models of the corresponding strategic-form. However, there are several conceptual issues that arise in this context. In the models considered in Section 10.1 the interpretation of $s_i = \sigma_i(w)$ is that at state w Player i "plays" or "chooses" strategy s_i.

Consider the perfect-information game shown in Figure 10.4 and any model of the strategic-form game associated with it.
Let w be a state in that model where

$$\sigma_1(w) = (d_1, a_3).$$

What does it mean to say that Player 1 "chooses" strategy (d_1, a_3)? The first part of the strategy, namely d_1, can be interpreted as a description of Player 1's actual behavior (what he actually does: he plays d_1), but the second part of the strategy, namely a_3, has no such interpretation: if Player 1 in fact plays d_1 then he knows that he will not have to make any further choices and thus it is not clear what it means for him to "choose" to play a_3 in a situation that is made impossible by his decision to play d_1.

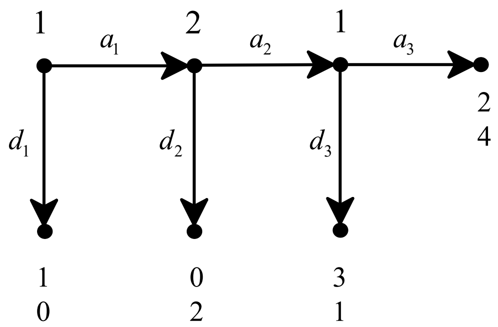


Figure 10.4: A prefect-information game.

Thus it does not seem to make sense to interpret $\sigma_1(w) = (d_1, a_3)$ as 'at state w Player 1 chooses (d_1, a_3)'. Perhaps the correct interpretation is in terms of a more complex sentence such as 'Player 1 chooses to play d_1 and if – contrary to this plan – he were to play a_1 and Player 2 were to follow with a_2, then Player 1 would play a_3.'

While in a simultaneous game the association of a strategy of Player i to a state can be interpreted as a description of Player i's actual behavior at that state, in the case of dynamic games this interpretation is no longer valid, since one would end up describing not only the actual behavior of Player i at that state but also his counterfactual behavior.

Methodologically, this is not satisfactory: if it is considered to be necessary to specify what a player would do in situations that do not occur at the state under consideration, then one should model the counterfactual explicitly.

But why should it be necessary to specify at state w (where Player 1 is playing d_1) what he would do at the counterfactual node following actions a_1 and a_2? Perhaps what matters is not so much what Player 1 would actually do in that situation but what Player 2 believes that Player 1 would do: after all, Player 2 might not know that Player 1 has decided to play d_1 and she needs to consider what to do in the eventuality that Player 1 actually ends up playing a_1.

So, perhaps, the strategy of Player 1 is to be interpreted as having two components: (1) a description of Player 1's behavior and (2) a conjecture in the mind of Player 2 about what Player 1 would do. If this is the correct interpretation, then one could object – again from a methodological point of view – that it would be preferable to disentangle the two components and model them explicitly.[4]

An alternative – although less common – approach in the literature dispenses with strategies and considers models of games where

1. states are described in terms of players' *actual behavior* and
2. players' conjectures concerning the actions of their opponents in various hypothetical situations are modeled by means of a generalization of the knowledge-belief structures considered in Section 10.1. The generalization is obtained by encoding not only the initial beliefs of the players (at each state) but also their *dispositions to revise those beliefs* under various hypotheses.[5]

A third approach has been to move away from static models (like the models considered in Section 10.1) and consider dynamic models where time is introduced explicitly into the analysis. These are behavioral models where strategies play no role and the only beliefs that are specified are the actual beliefs of the players at the time of choice. Thus players' beliefs are modeled as temporal, rather than conditional, beliefs and rationality is defined in terms of actual choices, rather than hypothetical plans.[6]

A discussion of the implications of common knowledge (or common belief) of rationality in dynamic games would require the introduction of several new concepts and definitions. For the sake of brevity, we shall not pursue this topic in this book.

Test your understanding of the concepts introduced in this section, by going through the exercises in Section 10.5.3 at the end of this chapter.

[4]For an extensive discussion of these issues see Bonanno (2015).

[5]The interested reader is referred to Perea (2012).

[6]The interested reader is referred to Bonanno (2014).

10.4 Proofs of Theorems

In order to prove Theorem 10.2.1 we will need the following extension of Theorem 6.4.1 (Chapter 6) to the case of an arbitrary number of players. A proof of Theorem 10.4.1 can be found in Osborne and Rubinstein (1994, Lemma 60.1, p. 60).

Theorem 10.4.1 Consider a finite n-player game in strategic form with cardinal payoffs $\langle I, (S_i)_{i\in I}, (\pi_i)_{i\in I}\rangle$.
Select an arbitrary Player i and a pure strategy s_i of Player i.
For every Player $j \neq i$, let $S'_j \subseteq S_j$ be a subset of j's set of pure strategies and let $S'_{-i} = S'_1 \times \cdots \times S'_{i-1} \times S'_{i+1} \times \cdots \times S'_n$ be the Cartesian product of the sets S'_j $(j \neq i)$.
Then the following are equivalent:
(1) There is a belief of Player i on S'_{-i} (that is, a probability distribution $P : S'_{-i} \to [0,1]$) that makes s_i a best response, that is, for every $x \in S_i$,

$$\sum_{s'_{-i}\in S_{-i}} \pi_i(s_i, s'_{-i})\,P(s'_{-i}) \;\geq\; \sum_{s'_{-i}\in S'_{-i}} \pi_i(x, s'_{-i})\,P(s'_{-i}),$$

(2) s_i is not strictly dominated by a mixed strategy of Player i in the restriction of the game to the sets of pure strategies $S'_1, \ldots, S'_{i-1}, S_i, S'_{i+1}, \ldots, S'_n$.

Given a finite strategic-form game with cardinal payoffs $G = \langle I, (S_i)_{i\in I}, (\pi_i)_{i\in I}\rangle$ we shall denote by S_i^∞ the set of pure strategies of Player i that survive the Cardinal Iterated Deletion of Strictly Dominated Strategies (Definition 6.4.1, Chapter 6) and by S^∞ the corresponding set of strategy profiles (that is, $S^\infty = S_1^\infty \times \cdots \times S_n^\infty$).

Proof of Theorem 10.2.1. Consider a finite strategic-form game G with cardinal payoffs and an epistemic model of G.
Let w be a state in that model such that $w \in CK\mathbf{R}$.
We want to show that $\sigma(w) = (\sigma_1(w), \ldots, \sigma_n(w)) \in S^\infty$.
We shall prove the stronger claim that, for every Player i and for every $w' \in I_{CK}(w)$ (recall that $I_{CK}(w)$ is the cell of the common knowledge partition that contains w) and for every $m \geq 0$, $\sigma_i(w') \notin D_i^m$, where $D_i^m \subseteq S_i$ is the set of pure strategies of Player i that are strictly dominated in game G^m (G^m is the subgame of G obtained at step m of the iterated deletion procedure; we define G^0 to be G itself). Note that this is equivalent to stating that, for every $w' \in I_{CK}(w)$, $\sigma(w') \in S^\infty$. Since $w \in I_{CK}(w)$, it follows from this that $\sigma(w) \in S^\infty$.
The proof is by induction.

1. BASE STEP (m = 0). Select an arbitrary $w' \in I_{CK}(w)$ and an arbitrary Player i.
   Since $w \in CK\mathbf{R}$, $I_{CK}(w) \subseteq \mathbf{R}$ and thus $w' \in \mathbf{R}$;
   furthermore, since $\mathbf{R} \subseteq \mathbf{R}_i$, $w' \in \mathbf{R}_i$.
   Thus $\sigma_i(w')$ is a best response to the beliefs of Player i at state w'.
   Hence, by Theorem 10.4.1, $\sigma_i(w') \notin D_i^0$.

2. INDUCTIVE STEP: assuming that the statement is true for all $k \leq m$ (for some $m \geq 0$), we prove that it is true for $k+1$.
   The hypothesis is that, for every Player i and for every $w' \in I_{CK}(w)$, $\sigma_i(w') \notin D_i^k$; that is, for every $w' \in I_{CK}(w)$, $\sigma_i(w') \in S^{k+1}$.
   Select an arbitrary $w' \in I_{CK}(w)$ and an arbitrary Player i.
   Since $w \in CK\mathbf{R}$, $w' \in \mathbf{R}_i$
   and thus $\sigma_i(w')$ is a best reply to Player i's beliefs at w' which, by hypothesis, attach positive probability only to strategy profiles of the other players that belong to S_{-i}^{k+1} (note that, since $w' \in I_{CK}(w)$, $I_{CK}(w') = I_{CK}(w)$ and, by definition of common knowledge partition, $I_i(w') \subseteq I_{CK}(w')$).
   By Theorem 10.4.1, it follows that $\sigma_i(w') \notin D_i^{k+1}$.

Proof of Theorem 10.2.1. Given a strategic-form game $G = \langle I, (S_i)_{i\in I}, (\pi_i)_{i\in I}\rangle$ with cardinal payoffs, construct the following epistemic model of G:

- $W = S^\infty$,
- for every Player i and every $s \in S^\infty$, $I_i(s) = \{s' \in S^\infty : s'_i = s_i\}$ (that is, the information set of Player i that contains s consists of all the strategy profiles in S^∞ that have the same strategy of Player i as in s);
- finally, let $\sigma_i(s) = s_i$ (s_i is the i^{th} component of s).

Select an arbitrary $s \in S^\infty$ and an arbitrary Player i.
By definition of S^∞, it is not the case that s_i is strictly dominated in the restriction of G to the sets of pure strategies $S_1^\infty, \ldots, S_{i-1}^\infty, S_i, S_{i+1}^\infty, \ldots, S_n^\infty$.
Thus, by Theorem 10.4.1, there is a probability distribution over S_{-i}^∞ that makes $s_i = \sigma_i(s)$ a best reply.
Choose one such probability distribution and let that probability distribution give Player i's beliefs at s.
Then $s \in \mathbf{R}_i$.
Since i was chosen arbitrarily, $s \in \mathbf{R}$;
hence, since $s \in S^\infty$ was chosen arbitrarily, $\mathbf{R} = S^\infty$.
It follows that $s \in CK\mathbf{R}$ for every $s \in S^\infty$.

10.5 Exercises

10.5.1 Exercises for Section 10.1: Model of strategic-form games

The answers to the following exercises are in Section 10.6 at the end of this chapter.

Exercise 10.1

Consider the game of Figure 10.5 (where the payoffs are von Neumann-Morgenstern payoffs) and the model of it shown in Figure 10.6.

(a) Find the event $\mathbf{R}_1$ (that is, the set of states where Player 1 is rational).

(b) Find the event $\mathbf{R}_2$ (the set of states where Player 2 is rational).

(c) Find the event $\mathbf{R}$ (the set of states where both players are rational).

		Player 2		
		L	*M*	*R*
Player 1	*A*	3 , 5	2 , 0	2 , 2
	B	5 , 2	1 , 2	2 , 1
	C	9 , 0	1 , 5	3 , 2

Figure 10.5: A strategic-form game with cardinal payoffs.

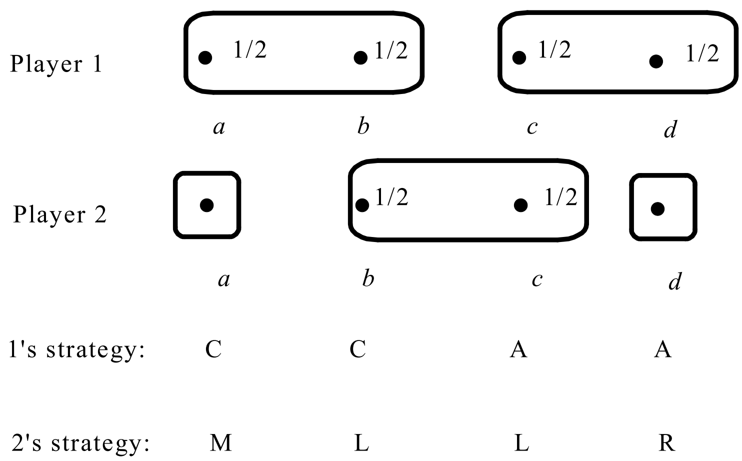


Figure 10.6: A model of the game of Figure 10.5.

Exercise 10.2

Consider the three-player game and model shown in Figures 10.7 and 10.8.

(a) Find the event $\mathbf{R}_1$ (that is, the set of states where Player 1 is rational).

(b) Find the event $\mathbf{R}_2$ (that is, the set of states where Player 2 is rational).

(c) Find the event $\mathbf{R}_3$ (that is, the set of states where Player 3 is rational).

(d) Find the event $\mathbf{R}$ (that is, the set of states where all players are rational). ■

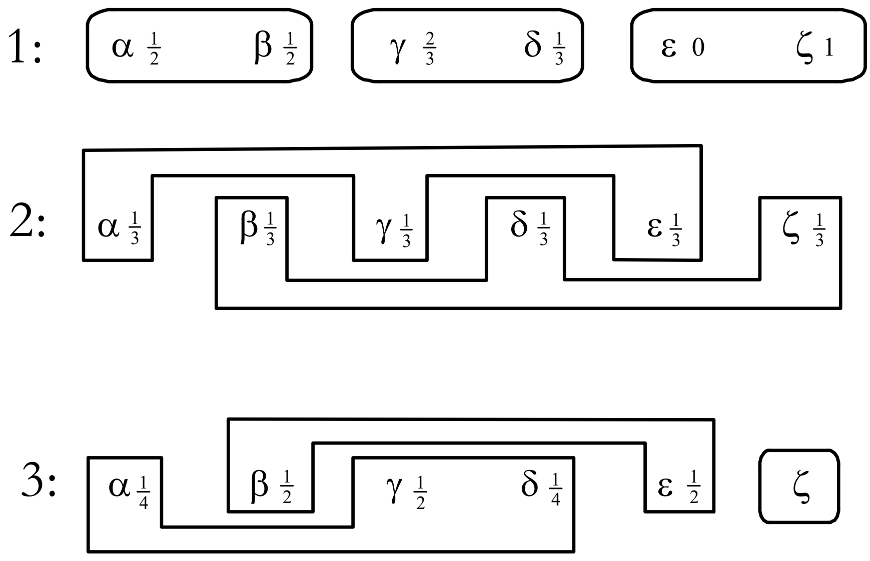
Player 3 chooses *E*

Player 1 \ Player 2	*C*	*D*
A	2 , 3 , 2	1 , 0 , 4
B	3 , 6 , 4	0 , 8 , 0

Player 3 chooses *F*

Player 1 \ Player 2	*C*	*D*
A	0 , 0 , 3	2 , 9 , 7
B	2 , 1 , 3	0 , 3 , 1

Figure 10.7: A three-player strategic-form game with cardinal payoffs.



σ_1 :	*A*	*A*	*B*	*B*	*B*	*B*
σ_2 :	*C*	*D*	*C*	*D*	*C*	*D*
σ_3 :	*E*	*F*	*E*	*E*	*F*	*F*

Figure 10.8: A model of the game of Figure 10.7.

10.5.2 Exercises for Section 10.2: Common knowledge of rationality in strategic-form games

The answers to the following exercises are in Section 10.6 at the end of this chapter.

Exercise 10.3

(a) For the game and model of Exercise 10.1 find the following events:

(i) $K_1\mathbf{R}_2$,

(ii) $K_2\mathbf{R}_1$,

(iii) $K_2K_1\mathbf{R}_2$,

(iv) $K_1K_2\mathbf{R}_1$,

(v) $CK\mathbf{R}_1$,

(vi) $CK\mathbf{R}_2$,

(vii) $CK\mathbf{R}$.

(b) Suppose that you found a model of the game of Exercise 10.1 (Figure 10.5) and a state w in that model such that $w \in CK\mathbf{R}$.
What strategy profile could you find at w?

Exercise 10.4

(a) For the game and model of Exercise Exercise 10.2 find the following events:

(i) $K_1\mathbf{R}$,

(ii) $K_2\mathbf{R}$,

(iii) $K_3\mathbf{R}$,

(iv) $CK\mathbf{R}_3$,

(v) $CK\mathbf{R}$.

(b) Suppose that you found a model of the game of Exercise 10.2 (Figure 10.7) and a state w in that model such that $w \in CK\mathbf{R}$.
What strategy profile could you find at w?

Exercise 10.5

For the game of Exercise 10.1 construct a model where there is a state at which there is common knowledge of rationality and the strategy profile played there is (C, L).
[Hints: (1) four states are sufficient, (2) it is easiest to postulate degenerated beliefs where a player assigns probability 1 to a particular state in his information set.]

Exercise 10.6 For the game of Exercise 10.2 construct a model where there is a state at which there is common knowledge of rationality. [Hint: think carefully, you don't need many states!]

■

10.5.3 Exercises for Section 10.3: Common knowledge of rationality in extensive-form games

Exercise 10.7 — ⋆⋆⋆Challenging Question⋆⋆⋆. Consider the extensive-form game shown in Figure 10.9 (where o_1, o_2, o_3 and o_4 are the possible outcomes).

(a) What is the backward-induction outcome?

(b) Write the strategic-form game associated with the extensive-form game and find all the outcomes that are supported by a pure-strategy Nash equilibrium.

(c) Considering models of the strategic form, what strategy profiles are compatible with common knowledge of rationality?

(d) Choose a strategy profile which is not a Nash equilibrium and construct a model of the strategic form where at a state there is common knowledge of rationality and the associated strategy profile is the one you selected.

■

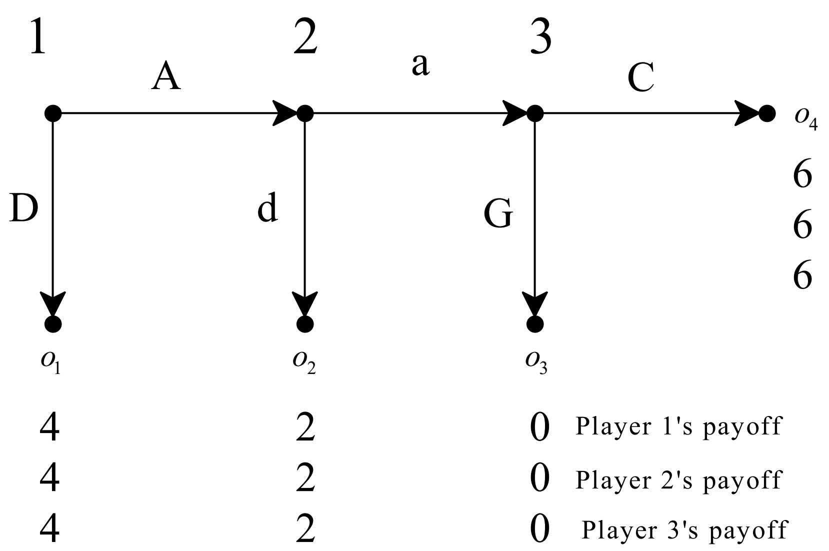


Figure 10.9: A perfect-information game with cardinal payoffs.

10.6 Solutions to Exercises

Solutions to Exercise 10.1

(a) $\mathbf{R}_1 = \{a, b\}$. At states a and b Player 1 believes that Player 2 is playing either L or M with equal probability.
Thus his expected payoff from playing A is $\frac{1}{2}(3) + \frac{1}{2}(2) = 2.5$,
from playing B is $\frac{1}{2}(5) + \frac{1}{2}(1) = 3$
and from playing C is $\frac{1}{2}(9) + \frac{1}{2}(1) = 5$.
Hence the best choice is C and this is indeed his choice at those two states.
Thus $a, b \in \mathbf{R}_1$.
At states c and d Player 1 plays A with an expected payoff of $\frac{1}{2}(3) + \frac{1}{2}(2) = 2.5$
but he could get a higher expected payoff with C (namely $\frac{1}{2}(9) + \frac{1}{2}(3) = 6$);
thus $c, d \notin \mathbf{R}_1$. Hence $\mathbf{R}_1 = \{a, b\}$.

(b) $\mathbf{R}_2 = \{a, b, c\}$.

(c) $\mathbf{R} = \{a, b\}$. □

Solutions to Exercise 10.2

(a) For each state let us calculate Player 1's expected payoff (denoted by $\mathbb{E}\pi_1$) from playing A and from playing B.

At states α and β,
$\mathbb{E}\pi_1(A) = \frac{1}{2}\pi_1(A,C,E) + \frac{1}{2}\pi_1(A,D,F) = \frac{1}{2}(2) + \frac{1}{2}(2) = 2$,
$\mathbb{E}\pi_1(B) = \frac{1}{2}\pi_1(B,C,E) + \frac{1}{2}\pi_1(B,D,F) = \frac{1}{2}(3) + \frac{1}{2}(0) = 1.5$.
Thus A is optimal and hence, since $\sigma_1(\alpha) = \sigma_1(\beta) = A$, $\alpha, \beta \in \mathbf{R}_1$.

At states γ and δ,
$\mathbb{E}\pi_1(A) = \frac{2}{3}\pi_1(A,C,E) + \frac{1}{3}\pi_1(A,D,E) = \frac{2}{3}(2) + \frac{1}{3}(1) = \frac{5}{3}$ and
$\mathbb{E}\pi_1(B) = \frac{2}{3}\pi_1(B,C,E) + \frac{1}{3}\pi_1(B,D,E) = \frac{2}{3}(3) + \frac{1}{3}(0) = \frac{6}{3}$.
Thus B is optimal and hence, since $\sigma_1(\gamma) = \sigma_1(\delta) = B$, $\gamma, \delta \in \mathbf{R}_1$.

At states ε and ζ,
Player 1 assigns probability 1 to (D,F) against which A is the unique best reply and yet $\sigma_1(\varepsilon) = \sigma_1(\zeta) = B$;
hence $\varepsilon, \zeta \notin \mathbf{R}_1$. Thus $\mathbf{R}_1 = \{\alpha, \beta, \gamma, \delta\}$.

(b) For each state let us calculate Player 2's expected payoff (denoted by $\mathbb{E}\pi_2$) from playing C and from playing D.

At states α, γ and ε,
$\mathbb{E}\pi_2(C) = \frac{1}{3}\pi_2(A,C,E) + \frac{1}{3}\pi_2(B,C,E) + \frac{1}{3}\pi_2(B,C,F) = \frac{1}{3}(3) + \frac{1}{3}(6) + \frac{1}{3}(1) = \frac{10}{3}$,
$\mathbb{E}\pi_2(D) = \frac{1}{3}\pi_2(A,D,E) + \frac{1}{3}\pi_2(B,D,E) + \frac{1}{3}\pi_2(B,D,F) = \frac{1}{3}(0) + \frac{1}{3}(8) + \frac{1}{3}(3) = \frac{11}{3}$.
Thus D is optimal and hence, since $\sigma_2(\alpha) = \sigma_2(\gamma) = \sigma_2(\varepsilon) = C$, $\alpha, \gamma, \varepsilon \notin \mathbf{R}_2$.

At states β, δ and ζ,
$\mathbb{E}\pi_2(C) = \frac{1}{3}\pi_2(A,C,F) + \frac{1}{3}\pi_2(B,C,E) + \frac{1}{3}\pi_2(B,C,F) = \frac{1}{3}(0) + \frac{1}{3}(6) + \frac{1}{3}(1) = \frac{7}{3}$,
$\mathbb{E}\pi_2(D) = \frac{1}{3}\pi_2(A,D,F) + \frac{1}{3}\pi_2(B,D,E) + \frac{1}{3}\pi_2(B,D,F) = \frac{1}{3}(9) + \frac{1}{3}(8) + \frac{1}{3}(3) = \frac{20}{3}$.
Thus D is optimal and hence, since $\sigma_2(\beta) = \sigma_2(\delta) = \sigma_2(\zeta) = D$, $\beta, \delta, \zeta \in \mathbf{R}_2$.
Thus $\mathbf{R}_2 = \{\beta, \delta, \zeta\}$.

(c) For each state let us calculate Player 3's expected payoff (denoted by $\mathbb{E}\pi_3$) from playing E and from playing F.

At states α, γ and δ
$\mathbb{E}\pi_3(E) = \frac{1}{4}\pi_3(A,C,E) + \frac{1}{2}\pi_3(B,C,E) + \frac{1}{4}\pi_3(B,D,E) = \frac{1}{4}(2) + \frac{1}{2}(4) + \frac{1}{4}(0) = 2.5,$
$\mathbb{E}\pi_3(F) = \frac{1}{4}\pi_3(A,C,F) + \frac{1}{2}\pi_3(B,C,F) + \frac{1}{4}\pi_3(B,D,F) = \frac{1}{4}(3) + \frac{1}{2}(3) + \frac{1}{4}(1) = 2.5.$
Thus both E and F are optimal and hence $\alpha, \gamma, \delta \in \mathbf{R}_3$.

At states β, and ε
$\mathbb{E}\pi_3(E) = \frac{1}{2}\pi_3(A,D,E) + \frac{1}{2}\pi_3(B,C,E) = \frac{1}{2}(4) + \frac{1}{2}(4) = 4,$
$\mathbb{E}\pi_3(F) = \frac{1}{2}\pi_3(A,D,F) + \frac{1}{2}\pi_3(B,C,F) = \frac{1}{2}(7) + \frac{1}{2}(3) = 5.$
Thus F is optimal and hence, since $\sigma_3(\beta) = \sigma_3(\varepsilon) = F,\ \beta, \varepsilon \in \mathbf{R}_3$.

At state ζ Player 3 knows that Players 1 and 2 play (B,D) and she is best replying with F. Thus $\zeta \in \mathbf{R}_3$.

Hence $\mathbf{R}_3 = \{\alpha, \beta, \gamma, \delta, \varepsilon, \zeta\}$.

(d) $\mathbf{R} = \mathbf{R}_1 \cap \mathbf{R}_2 \cap \mathbf{R}_3 = \{\beta, \delta\}$. □

Solutions to Exercise 10.3
In Exercise 10.1 we determined that $\mathbf{R}_1 = \{a,b\}$, $\mathbf{R}_2 = \{a,b,c\}$ and $\mathbf{R} = \{a,b\}$. Thus

(a) (i) $K_1\mathbf{R}_2 = \{a,b\}$, (ii) $K_2\mathbf{R}_1 = \{a\}$, (iii) $K_2K_1\mathbf{R}_2 = \{a\}$, (iv) $K_1K_2\mathbf{R}_1 = \emptyset$.
The common knowledge partition consists of a single information set containing all the states. Thus (v) $CK\mathbf{R}_1 = \emptyset$, (vi) $CK\mathbf{R}_2 = \emptyset$, (vii) $CK\mathbf{R} = \emptyset$.

(b) By Theorem 10.2.1, at a state at which there is common knowledge of rationality one can only find a strategy profile that survives the cardinal iterated deletion of strictly dominates strategies.
In this game, for Player 1 strategy B is strictly dominated by the mixed strategy

$$\begin{pmatrix} A & C \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix};$$

after deleting B, for Player 2 strategy R is strictly dominated by the mixed strategy

$$\begin{pmatrix} L & M \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Thus the iterated deletion of strictly dominated strategies yields the set of strategy profiles $\{(A,L),(A,M),(C,L),(C,M)\}$. Hence at a state where there is common knowledge of rationality one could only find one of these four strategy profiles. □

Solutions to Exercise 10.4
In Exercise 10.2 we determined that $\mathbf{R} = \{\beta, \delta\}$.

(a) (i) $K_1\mathbf{R} = \emptyset$, (ii) $K_2\mathbf{R} = \emptyset$, (iii) $K_3\mathbf{R} = \emptyset$. The common knowledge partition consists of a single information set containing all the states.
Thus (since $\mathbf{R}_3 = \{\alpha, \beta, \gamma, \delta, \varepsilon, \zeta\}$)
(iv) $CK\mathbf{R}_3 = \{\alpha, \beta, \gamma, \delta, \varepsilon, \zeta\}$, (v) $CK\mathbf{R} = \emptyset$.

(b) By Theorem 10.2.1, at a state at which there is common knowledge of rationality one can only find a strategy profile that survives the cardinal iterated deletion of strictly dominated strategies. Since in this game there are no strictly dominated strategies, at such a state one could find *any* strategy profile. □

Solutions to Exercise 10.5 In the model shown in Figure 10.10 there is common knowledge of rationality at every state. At state α the strategy profile played is (C, L). □

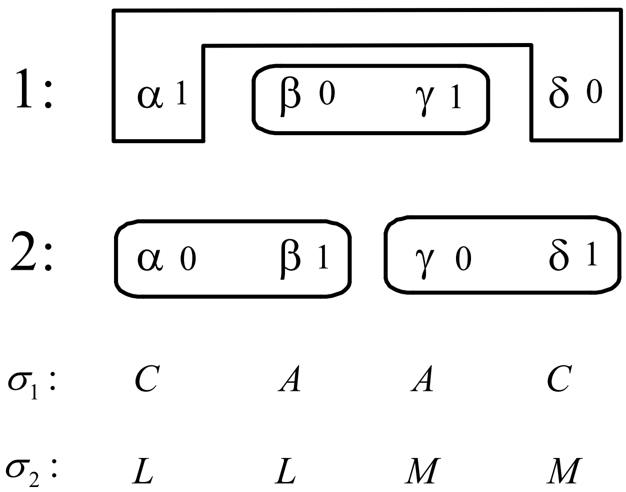


Figure 10.10: A model of the game of Figure 10.5 where there is common knowledge of rationality at every state.

Solutions to Exercise 10.6 Whenever a game has a Nash equilibrium in pure strategies, a one-state model where that Nash equilibrium is played is such that there is common knowledge of rationality at that state. Thus the model shown in Figure 10.11 provides a possible answer. □

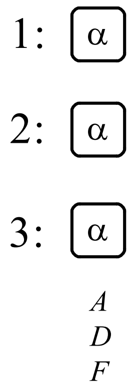


Figure 10.11: A one-state model of the game of Figure 10.5 where there is common knowledge of rationality and the strategy profile is (A, D, F).

Solutions to Exercise 10.7

(a) The backward induction outcome is o_4, with associated strategy profile (A, a, C) and payoffs (6,6,6).

(b) The strategic form is shown in Figure 10.12 (the Nash equilibria have been highlighted):
The Nash equilibria are: (D, d, G), (D, a, G), (D, d, C) and (A, a, C).
Thus the only two outcomes sustained by a Nash equilibrium are o_1, with payoffs (4,4,4), and o_4, with payoffs (6,6,6).

Player 3 chooses G

		Player 2 d	**Player 2** a
Player 1	D	4 , 4 , 4	4 , 4 , 4
	A	2 , 2 , 2	0 , 0 , 0

Player 3 chooses C

		Player 2 d	**Player 2** a
Player 1	D	4 , 4 , 4	4 , 4 , 4
	A	2 , 2 , 2	6 , 6 , 6

Figure 10.12: The strategic-form associated with the extensive-form game of Figure 10.9.

(c) By Theorem 10.2.1, the outcomes that are compatible with common knowledge of rationality are those associated with strategy profiles that survive the iterated deletion of strictly dominated strategies. Since no player has any strictly dominated strategies, all the outcomes are compatible with common knowledge of rationality. (Note that, for Player 3, C *weakly* dominates G, but not strictly.)

(d) In the model shown in Figure 10.13 at state α the players choose (A,d,G), which is not a Nash equilibrium; furthermore, there is no Nash equilibrium whose associated outcome is o_2 (with payoffs (2,2,2)).

In this model $\alpha \in CK\mathbf{R}$ (in particular, Player 1's choice of A is rational, given his belief that Player 2 plays d and a with equal probability). □

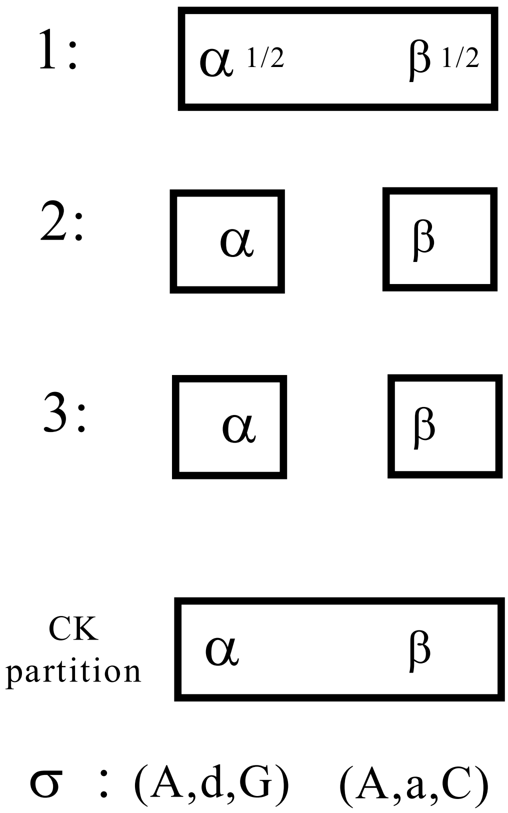


Figure 10.13: A model of the strategic-form of the game of Figure 10.9 where the strategy profile (A,d,G), which is not a Nash equilibrium, is played under common knowledge of rationality.