Chapter 4

Correct and Symmetric Beliefs

In this chapter we will focus on two special kinds of belief hierarchies: *simple* belief hierarchies and *symmetric* belief hierarchies. Simple belief hierarchies have the property that a player believes that his opponents are *correct* about the beliefs he holds. We will show that simple belief hierarchies, together with common belief in rationality, lead to the concept of *Nash equilibrium.* Symmetric belief hierarchies may be viewed as a more permissive variant of simple belief hierarchies, where a player need not believe that his opponents are correct about the beliefs he holds. However, they display some form of symmetry between the beliefs a player holds himself, and the beliefs he thinks that other players hold. It is shown that symmetric belief hierarchies in combination with common belief in rationality yield the concept of *correlated equilibrium.* If, in addition, we impose *one theory per choice,* we arrive at the more restrictive concept of *canonical* correlated equilibrium.

4.1 Correct Beliefs

In the previous chapter we have investigated the reasoning concept of *common belief in rationality.* For every player it selects those belief hierarchies where the player believes that his opponents choose rationally, believes that every opponent believes that all other players choose rationally, and so on. In a sense, according to common belief in rationality, all belief hierarchies with these properties are regarded as equal.

However, we will see in this section that two belief hierarchies which both express common belief in rationality may display completely different properties. For instance, in some belief hierarchy you may believe that the opponents are *correct* about the actual beliefs that you hold, whereas this may not be true in other belief hierarchies. In this section we will formalize the idea of correct beliefs by means of *simple* belief hierarchies, and show that common belief in rationality in combination with simple belief hierarchies leads to the concept of *Nash equilibrium.* Subsequently we prove that in every game,

You	*Cinema Palace*	*Movie Corner*	*home*
Cinema Palace	4	0	0
Movie Corner	0	4	0
home	3	3	3

Barbara	*Cinema Palace*	*Movie Corner*	*home*
Cinema Palace	0	4	0
Movie Corner	4	0	0
home	3	3	3

Table 4.1.1 Decision problems in "Movie for two"

every player will have at least one simple belief hierarchy that expresses common belief in rationality. We finally discuss how reasonable the concept of Nash equilibrium really is.

4.1.1 Simple Belief Hierarchies

To explain what we mean by saying that "you believe that the opponent is correct about your beliefs", consider the following example.

Example 4.1: Movie for two.

This evening, both you and Barbara want to go to the movies. In town there are two movie theaters, Cinema Palace and The Movie Corner. The question for Barbara and you is: To which movie theater do you go? What complicates the matter is that last night you had a fight with Barbara. Now she is so upset that she would rather not talk to you this evening, and hence she prefers to go to a different theater than you. You, on the other hand, would really like to make up with her, and therefore you would like to go to the same theater as Barbara. The third option for Barbara and you is simply to stay at home, avoid any possible disappointment and watch your all-time favorite movie on Netflix.

The utilities are as follows: If you go to the same theater as Barbara, you will be happy to make up with her and your utility will be 4. If, on the other hand, you go out and Barbara goes to the other theater or stays at home, you will be very disappointed and your utility will be 0. If you stay at home you feel okay, but still regret not to have the chance to make up with Barbara, and your utility will be 3. For Barbara the utilities are similar: If she goes out and you go to the other cinema, she will be relieved and enjoy a utility of 4. If you go to the same cinema as she does, Barbara will not enjoy the movie at all and her utility will be 0. If she goes out and discover that you have stayed at home, she will feel sorry about this and her utility will be 0. If she stays at home, she regrets not to be watching a movie on the big screen and her utility will be 3. These utilities give rise to the decision problems for you and Barbara in Table 4.1.1.

In this game, it turns out that both you and Barbara can rationally make *any* choice under common belief in rationality. This is most easily seen from the beliefs diagram in Figure 4.1.1. Since all arrows in this beliefs diagram are solid, it follows from Theorem 3.3.2 that all belief hierarchies generated by this beliefs diagram express common belief in rationality.

Note that going to *Cinema Palace* is optimal for the belief hierarchy that is obtained if we start at your choice Palace and follow the arrows. Indeed, the induced first-order belief is that you believe that Barbara goes to *Cinema Palace*, and going to *Cinema Palace* yourself is optimal for that belief. As this belief hierarchy expresses common belief in rationality, we conclude that you can rationally go to

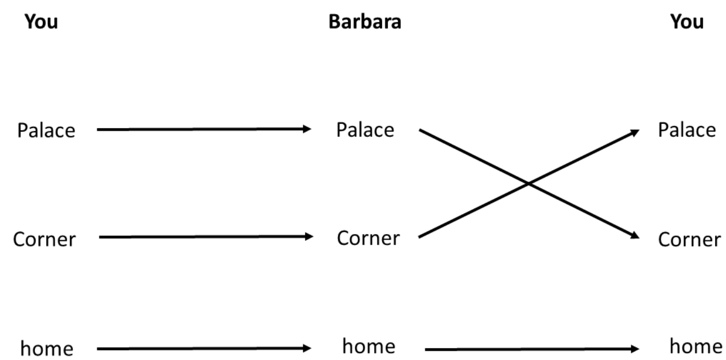

Figure 4.1.1 Beliefs diagram for "Movie for two"

Cinema Palace under common belief in rationality. In a similar fashion, we can conclude that under common belief in rationality you can also rationally go to *The Movie Corner* or stay at *home.* And also Barbara can rationally go to *Cinema Palace*, *The Movie Corner* or stay at *home* under common belief in rationality.

Consider now the belief hierarchy that starts at your choice to stay at *home.* In that belief hierarchy, you believe that Barbara stays at *home*. That is your first-order belief. At the same time, you believe that Barbara believes that you *indeed* believe that Barbara stays at *home*. In other words, you believe that Barbara is *correct* about your first-order belief.

Question 4.1.1 *In that same belief hierarchy, do you also believe that Barbara is correct about your second-order belief?*

Compare this to the belief hierarchy that starts at your choice *Palace*. There, your first-order belief is that you believe that Barbara goes to *Cinema Palace*. At the same time, however, you believe that Barbara believes that you believe that Barbara goes to *The Movie Corner* (and not *Cinema Palace*). Hence, you believe that Barbara is *not correct* about your first-order belief. The same can be said about the belief hierarchy that starts at your choice *Corner*. Please check this.

Question 4.1.2 *Consider the beliefs diagram from Figure 3.2.1 for "Going to a party". In which of your belief hierarchies do you believe that Barbara is correct about your first-order belief?*

Hence, at the belief hierarchies that start at your choices *Palace* and *Corner*, you believe that Barbara is *incorrect* about your first-order belief, while at the belief hierarchy that starts at your choice *home* you believe that Barbara is *correct* about your first-order belief. In fact, in the latter belief hierarchy you do not only believe that Barbara is correct about your first-order belief, you also believe that Barbara is correct about your second-order belief, your third-order belief, and so on. Indeed, in the belief hierarchy that starts at your choice *home* you believe that Barbara is correct about your *entire belief hierarchy.*

Note also that this belief hierarchy is completely generated by only two beliefs: the belief σ_2 that Barbara chooses to stay at *home,* and the belief σ_1 that you choose to stay at *home.* To see this, observe that your first-order belief about Barbara's choice is σ_2. Moreover, your second-order belief is that you believe that Barbara believes that you choose to stay at *home,* that is, you believe that Barbara has belief σ_1 about your choice. Your third-order belief is that you believe that Barbara believes that you believe that Barbara chooses to stay at *home,* which means that you believe that Barbara believes that you have belief σ_2 about Barbara's choice. And so on. We call this a *simple* belief hierarchy that is generated by the belief σ_1 about your choice and the belief σ_2 about Barbara's choice. Here, the subindex 1 in σ_1 indicates that this is a belief about player 1's (hence, your) choice, whereas the subindex 2 in σ_2 stresses that this belief is about player 2's (hence, Barbara's) choice.

More generally, a simple belief hierarchy generated by beliefs about choices can be defined as follows.

Definition 4.1.1 (Simple belief hierarchy) *Let σ_1 be a probabilistic belief about player 1's choice, σ_2 a probabilistic belief about player 2's choice, and so on, until σ_n being a probabilistic belief about player n's choice. The belief hierarchy for player i generated by the beliefs $(\sigma_1, \sigma_2, ..., \sigma_n)$ is defined as follows:*

(1) in the first-order belief, player i assigns to every opponents' choice combination $(c_j)_{j \neq i}$ the probability $\prod_{j \neq i} \sigma_j(c_j)$,

(2) in the second-order belief, player i believes with probability 1 that every opponent j assigns to every opponents' choice combination $(c_k)_{k \neq j}$ the probability $\prod_{k \neq j} \sigma_k(c_k)$,

(3) in the third-order belief, player i believes with probability 1 that every opponent j believes with probability 1 that every opponent k assigns to every opponents' choice combination $(c_l)_{l \neq k}$ the probability $\prod_{l \neq k} \sigma_l(c_l)$, and so on.

*A belief hierarchy is called **simple** if it is generated by a combinations of such beliefs $(\sigma_1, \sigma_2, ..., \sigma_n)$.*

Here, we denote by $\prod_{j \neq i} \sigma_j(c_j)$ the product of the probabilities $\sigma_j(c_j)$ for every opponent $j \neq i$. For instance, if there are two players, then $\prod_{j \neq 1} \sigma_j(c_j) = \sigma_2(c_2)$. If there are three players, then $\prod_{j \neq 1} \sigma_j(c_j) = \sigma_2(c_2) \cdot \sigma_3(c_3)$. To illustrate the notion of a simple belief hierarchy for games with more than two players, consider the game "When Chris joins the party", with the decision problems as depicted in Table 3.2.1. Let

$$\sigma_1 = green$$

be the belief about player 1's (your) choice that assigns probability 1 to your choice *green.* Let

$$\sigma_2 = (0.3) \cdot blue + (0.7) \cdot red$$

be the probabilistic belief about player 2's (Barbara's) choice that assigns probability 0.3 to her choice *blue* and probability 0.7 to her choice *red.* Finally, let

$$\sigma_3 = (0.6) \cdot blue + (0.4) \cdot yellow$$

be the belief about player 3's (Chris') choice that assigns probability 0.6 to Chris choosing *blue* and probability 0.4 to Chris choosing *yellow.*

Consider your simple belief hierarchy generated by these beliefs σ_1, σ_2 and σ_3. How does this belief hierarchy look like? Let us start with your first-order belief about the choice combinations by Barbara and Chris. By definition, your belief about Barbara's choice is σ_2 and your belief about Chris' choice is σ_3. Hence, the probability that you assign to Barbara wearing *blue* and Chris wearing *blue* is

$$\sigma_2(blue) \cdot \sigma_3(blue) = (0.3) \cdot (0.6) = 0.18.$$

Similarly, the probability that you assign to Barbara wearing *blue* and Chris wearing *yellow* is

$$(0.3) \cdot (0.4) = 0.12,$$

the probability you assign to Barbara wearing *red* and Chris wearing *blue* is

$$(0.7) \cdot (0.6) = 0.42,$$

and the probability you assign to Barbara wearing *red* and Chris wearing *yellow* is

$$(0.7) \cdot (0.4) = 0.28.$$

The probability you assign to every other choice combination by Barbara and Chris is zero. This constitutes your first-order belief.

What is your second-order belief? As a part of your second-order belief, let us describe what you believe that Barbara believes about your choice and Chris' choice? You believe that Barbara has belief σ_1 about your choice and belief σ_3 about Chris' choice. Hence, the probability you believe that Barbara assigns to you wearing *green* and Chris wearing *blue* is

$$\sigma_1(green) \cdot \sigma_3(blue) = 1 \cdot (0.6) = 0.6.$$

Similarly, the probability you believe that Barbara assigns to you wearing *green* and Chris wearing *yellow* is

$$1 \cdot (0.4) = 0.4,$$

whereas you believe that Barbara assigns probability zero to all other choice combinations by you and Chris. This is your belief about Barbara's first-order belief.

Question 4.1.3 *As another part of your second-order belief, describe your belief about Chris' first-order belief. As a part of your third-order belief, describe what you believe that Barbara believes about Chris' first-order belief.*

In this way, we can derive the full simple belief hierarchy for you by only using the three beliefs σ_1, σ_2 and σ_3.

Throughout this book, we will use simple belief hierarchies as a way to express *correct beliefs.* To see why, consider a simple belief hierarchy for player i generated by the beliefs $\sigma_1, \sigma_2, ..., \sigma_n$. Then, by construction, player i does not only have belief σ_j about j's choice, but i also believes that every opponent k believes that i indeed has this particular belief σ_j about j's choice. That is, in this simple belief hierarchy player i believes that each of his opponents is *correct* about his first-order belief.

We can say even more: In the simple belief hierarchy above, player i does not only believe that player j has belief σ_k about k's choice, but i also believes that every opponent l believes that i indeed believes that j has belief σ_k about k's choice. That is, player i believes that every opponent is correct about his second-order beliefs also. In the same fashion we can conclude that player i also believes that his opponents are correct about his third-order beliefs and higher. In other words, in a simple belief hierarchy player i believes that his opponents are correct about his *entire belief hierarchy.*

Question 4.1.4 *Consider the simple belief hierarchy for player i generated by the beliefs $\sigma_1, \sigma_2, ..., \sigma_n$. Explain why player i believes that every opponent j believes that every other player is correct about j's first-order belief.*

			Chris	
Probability			0.6	0.4
			blue	*yellow*
Barbara	0.3	*blue*	0.18	0.12
	0.7	*red*	0.42	0.28

Table 4.1.2 Simple belief hierarchy induces independent beliefs about opponents' choices

In the same way as in Question 4.1.4 it can be shown that in a simple belief hierarchy, player i also believes that every opponent j believes that every other player is correct about j's entire belief hierarchy. And so on.

For games with more than two players, a simple belief hierarchy also displays other properties that go beyond the idea of correct beliefs. Consider a game with at least three players, and the simple belief hierarchy for player i generated by the beliefs $\sigma_1, \sigma_2, ..., \sigma_n$. Choose two different opponents j and k of player i. Then, player i does not only have the belief σ_k about k's choice, but also believes that player j has the *same* belief σ_k about k's choice. In other words, player i believes that every opponent j *shares* his belief about a third player k's choice. In the example for "When Chris joins the party" above, for instance, we see that you do not only assign probability 0.6 to Chris wearing *blue,* but you additionally believe that Barbara *also* assigns probability 0.6 to Chris wearing *blue.*

Another property, beyond correct beliefs, that follows for games with more than two players is that player i's belief about j's choice must be *independent* from i's belief about k's choice. To see what this means, consider again the example above for "When Chris joins the party". Your belief about the choice combinations by Barbara and Chris can be summarized by Table 4.1.2. Note that the probability you assign to the choice combination (*blue, blue*) by Barbara and Chris is obtained by taking the product of the probability $\sigma_2(blue) = 0.3$ about Barbara's choice and the probability $\sigma_3(blue) = 0.6$ about Chris' choice. The same holds for the three other choice combinations in the table. We thus see that your first-order belief about Barbara's and Chris' choice combinations can be written as the product of your belief σ_2 about Barbara's choice and the belief σ_3 about Chris' choice. In this case, we say that your belief about Barbara's choice is *independent* from your belief about Chris' choice.

Generally, the first-order belief generated by a simple belief hierarchy is always independent in this sense. Consider the simple belief hierarchy for player i generated by the beliefs $\sigma_1, \sigma_2, ..., \sigma_n$. Then, the probability that player i assigns to the choice combination where j chooses c_j and k chooses c_k is given by the product

$$\sigma_j(c_j) \cdot \sigma_k(c_k),$$

and hence i's belief about j's choice is independent from i's belief about k's choice.

A belief hierarchy that is not simple may easily contain first-order beliefs that are *not independent.* In the example "When Chris joins the party", for instance, consider the first-order belief

$$(0.5) \cdot (blue,\ blue) + (0.5) \cdot (red,\ yellow),$$

in which you assign probability 0.5 to the event that Barbara and Chris both wear *blue,* and assign probability 0.5 to the event that Barbara wears *red* and Chris wears *yellow.* Intuitively, your belief about Barbara's choice is heavily dependent on your belief about Chris' choice. Indeed, you only consider Barbara wearing *blue* if at the same time Chris wears *blue* also, and you only consider

Barbara wearing *red* if at the same time Chris wears *yellow.* One can also formally show that this first-order belief cannot be written as the product of a belief about Barbara's choice and a belief about Chris' choice.

Question 4.1.5 *Explain why the first-order belief above cannot be written as the product of a probabilistic belief σ_2 about Barbara's choice and a probabilistic belief σ_3 about Chris' choice.*

Such non-independent (or *correlated*) first-order beliefs are thus excluded if we concentrate on simple belief hierarchies. Summarizing we thus see that player i, in a simple belief hierarchy, (a) believes that every opponent is correct about his belief hierarchy, (b) believes that every opponent j believes that every other player is correct about j's belief hierarchy, (c) believes that every opponent j has the same belief as i himself about a third player k's choice, and (d) has independent beliefs about the choice combinations of two different opponents j and k. Of course, the properties (c) and (d) are only relevant if there are three players or more.

Question 4.1.6 *Consider the example "When Chris joins the party", and the beliefs diagram for this game in Figure 3.2.2. Look at your belief hierarchy that starts at your choice green. Is this belief hierarchy simple or not? Answer the same question for Barbara's belief hierarchy that starts at her choice yellow.*

4.1.2 Relation with Nash Equilibrium

In the previous subsection we have introduced the notion of a simple belief hierarchy, to express the belief that your opponents are correct about your beliefs. Suppose we now combine the conditions of common belief in rationality with that of a simple belief hierarchy. What kind of belief hierarchies do we get? And what choices can a player rationally make if he holds a simple belief hierarchy that expresses common belief in rationality? These are the questions we wish to answer in this subsection.

Consider a belief hierarchy for player i generated by a combination of beliefs $\sigma_1, \sigma_2, ..., \sigma_n$. Remember that σ_j is a probabilistic belief about player j's choice, for every j. As a short-hand notation, denote by σ_{-j} the belief that j holds about his opponents' choice combinations. Hence, the belief σ_{-j} assigns to every opponents' choice combination $(c_k)_{k \neq j}$ the probability

$$\prod\nolimits_{k \neq j} \sigma_k(c_k).$$

Then, in his simple belief hierarchy, player i has belief σ_{-i} about the opponents' choices, believes that every opponent j holds the belief σ_{-j} about his opponents' choices, and so on.

Suppose that, in addition, this simple belief hierarchy expresses common belief in rationality. Then, in particular, player i believes in every opponent j's rationality. That is, if player i assigns positive probability to j's choice c_j, then c_j must be optimal for j, given what i thinks that j believes about his opponents' choices. Now, if i assigns positive probability to j's choice c_j, then it must be that $\sigma_j(c_j) > 0$, because σ_j is i's belief about j's choice. On the other hand, i thinks that j holds the belief σ_{-j} about his opponents' choices. Therefore, this choice c_j must be optimal for j under the belief σ_{-j}. Overall, we thus see that if $\sigma_j(c_j) > 0$, then c_j must be optimal for j under the belief σ_{-j}. This must hold for every opponent j, and every choice c_j with $\sigma_j(c_j) > 0$.

Under common belief in rationality, i must also believe that each opponent j believes in i's rationality. That is, if i believes that j assigns positive probability to his own choice c_i, then i must believe that j believes that c_i is optimal for i, given what i thinks that j thinks is i's belief about his opponents' choices. Now, if i believes that j assigns positive probability to his choice c_i, then we must have that $\sigma_i(c_i) > 0$, since i believes that j holds the belief σ_i about i's choice. At the same time, i thinks that j thinks that i's belief about his opponents' choices is σ_{-i}, and hence this choice c_i must be optimal for i under the belief σ_{-i}. We thus see that if $\sigma_i(c_i) > 0$, then c_i must be optimal for i under the belief σ_{-i}.

Overall, we conclude that for every player j (including player i himself), the belief σ_j about j's choice can only assign positive probability to choices c_j that are optimal for j under the belief σ_{-j}. Belief combinations $(\sigma_1, \sigma_2, ..., \sigma_n)$ with this property are called *Nash equilibria.*

Definition 4.1.2 (Nash equilibrium) *Consider a combination of beliefs $(\sigma_1, \sigma_2, ..., \sigma_n)$, where σ_i is a probabilistic belief about i's choice for every player i. The belief combination $(\sigma_1, \sigma_2, ..., \sigma_n)$ is a **Nash equilibrium** if for every player i, the belief σ_i only assigns positive probability to choices c_i that are optimal for i under the belief σ_{-i}.*

From our arguments above we can thus conclude the following: If player i holds the simple belief hierarchy generated by the belief combination $(\sigma_1, ..., \sigma_n)$, and this belief hierarchy expresses common belief in rationality, then necessarily this belief combination $(\sigma_1, ..., \sigma_n)$ must be a Nash equilibrium.

In fact, we will show that the other direction is also true: If a simple belief hierarchy is generated by a Nash equilibrium, then it will express common belief in rationality. To see this, consider the simple belief hierarchy for player i generated by a Nash equilibrium $(\sigma_1, ..., \sigma_n)$. We first show that i believes in the opponents' rationality. Suppose that i assigns positive probability to j's choice c_j. Then, $\sigma_j(c_j) > 0$. Since $(\sigma_1, ..., \sigma_n)$ is a Nash equilibrium, c_j must be optimal for j under the belief σ_{-j}. Since i thinks that j has belief σ_{-j}, we conclude that c_j is optimal for j, given what i thinks that j believes about his opponents' choices. Hence, if i assigns positive probability to j's choice c_j, then c_j is optimal for j, given what i thinks that j believes about his opponents' choices. In other words, i believes in j's rationality. As this applies to every opponent j, player i believes in the opponents' rationality.

Question 4.1.7 *Explain why in the simple belief hierarchy above, i also believes that every opponent j believes in his opponents' rationality.*

By continuing in this fashion, we eventually conclude that player i expresses common belief in rationality with this simple belief hierarchy. Hence, we see that every simple belief hierarchy that is generated by a Nash equilibrium will express common belief in rationality. Altogether, we arrive at the following conclusion.

Theorem 4.1.1 (Relation with Nash equilibrium) *Consider the simple belief hierarchy for player i generated by a belief combination $(\sigma_1, ..., \sigma_n)$. Then, this belief hierarchy expresses common belief in rationality, if and only if, the belief combination $(\sigma_1, ..., \sigma_n)$ is a Nash equilibrium.*

As an illustration, consider the beliefs diagram in Figure 4.1.1 for "Movie for two". Let us focus on the belief hierarchy for you that starts at your choice *home.* This is a simple belief hierarchy generated by the belief $\sigma_1 = home$ about your choice, assigning probability 1 to you staying at *home*, and the belief $\sigma_2 = home$ about Barbara's choice, assigning probability 1 to Barbara staying at *home* as well. As we have seen that this simple belief hierarchy expresses common belief in rationality, we conclude from the theorem above that (σ_1, σ_2) must be a Nash equilibrium. Indeed, σ_1 assigns probability 1 to you staying at *home*, which is optimal for you under the belief $\sigma_{-1} = \sigma_2$ that Barbara stays at *home* as well. Also, σ_2 assigns probability 1 to Barbara staying at *home*, which is optimal for Barbara under the belief $\sigma_{-2} = \sigma_1$ that you stay at *home* as well.

With Theorem 4.1.1 at hand, we can now answer the following question: What choices can you rationally make if you hold a *simple* belief hierarchy that expresses common belief in rationality? Suppose player i holds a simple belief hierarchy, generated by the belief combination $(\sigma_1, ..., \sigma_n)$, that expresses common belief in rationality. Then, we know from Theorem 4.1.1 that $(\sigma_1, ..., \sigma_n)$ must be a Nash equilibrium. Also, every choice c_i that is optimal for player i under this simple belief hierarchy must be optimal under the belief σ_{-i}, as this is the belief that i holds about his opponents' choices. In other words, c_i is optimal for player i in the Nash equilibrium $(\sigma_1, ..., \sigma_n)$. We thus conclude that if player i holds a simple belief hierarchy that expresses common belief in rationality, then every choice that is optimal for player i must be optimal in some Nash equilibrium.

Definition 4.1.3 (Choice optimal in a Nash equilibrium) *A choice c_i is **optimal in a Nash equilibrium** if there is some Nash equilibrium $(\sigma_1, ..., \sigma_n)$ such that c_i is optimal for the belief σ_{-i}.*

The other direction is also true: If a choice is optimal in a Nash equilibrium, then it is optimal for a simple belief hierarchy that expresses common belief in rationality. To see this, suppose that choice c_i is optimal for player i in a Nash equilibrium $(\sigma_1, ..., \sigma_n)$, which means that c_i is optimal under the belief σ_{-i}. By Theorem 4.1.1 we know that the simple belief hierarchy for player i generated by $(\sigma_1, ..., \sigma_n)$ expresses common belief in rationality. In this belief hierarchy, player i holds the belief σ_{-i} about the opponents' choices. Since c_i is optimal under the belief σ_{-i}, we conclude that c_i is optimal for this simple belief hierarchy that expresses common belief in rationality. The choices that a player can rationally make with a simple belief hierarchy that expresses common belief in rationality can thus be characterized as follows.

Theorem 4.1.2 (Relation with Nash equilibrium choices) *A choice is optimal for a simple belief hierarchy that expresses common belief in rationality, if and only if, that choice is optimal in a Nash equilibrium.*

In other words, if we want to find all choices in a game that are optimal for a simple belief hierarchy that expresses common belief in rationality, then we must first find all Nash equilibria for this game, and subsequently all choices that are optimal in these Nash equilibria. Unfortunately, however, finding all Nash equilibia in a game can be rather difficult, as some of the examples in the next subsection will show.

4.1.3 Examples

We will now apply Theorem 4.1.2 to three examples, to find the choices you can rationally make under common belief in rationality with a simple belief hierarchy.

Example 4.2: Movie for two.

Consider the example "Movie for two" which has been introduced before, with the decision problems as depicted in Table 4.1.1. The question we wish to answer is: What choice(s) can you rationally make with a simple belief hierarchy that expresses common belief in rationality?

From the beliefs diagram of Figure 4.1.1, we can conclude that under common belief in rationality – without insisting on a simple belief hierarchy – you can rationally go to *Cinema Palace, The Movie Corner,* or stay at *home*. Indeed, in that beliefs diagram your choices *Palace* and *Corner* are supported by non-simple belief hierarchies that express common belief in rationality, while your choice to stay at *home* is even supported by a *simple* belief hierarchy that expresses common belief in rationality. The question thus remains: Can you also rationally go to *Cinema Palace* or *The Movie Corner* with a *simple* belief hierarchy that expresses common belief in rationality? If such simple belief hierarchies exist, they must be part of some different beliefs diagram. Instead of checking for all alternative beliefs diagrams – which is anyhow an impossible task since there are infinitely many – we rely on Theorem 4.1.2 to answer this question. That is, we will first try to find all Nash equilibria in this game, and then check whether your choice *Palace* or *Corner* is optimal in some of these Nash equilibria.

We will show that the only Nash equilibrium in this game is $(\sigma_1, \sigma_2) = (home, home)$, in which it is believed with probability 1 that you and Barbara stay at *home.* To prove this, consider a Nash equilibrium (σ_1, σ_2) for this game and assume, contrary to what we want to show, that $\sigma_1(Palace) > 0$. Then, since (σ_1, σ_2) is Nash equilibrium, *Palace* must be optimal for you under the belief σ_2, which is only possible if $\sigma_2(Palace) > 0$. Indeed, if $\sigma_2(Palace) = 0$, then choosing *Palace* would yield a utility of 0, which would be less than what you get by choosing *home.* Since $\sigma_2(Palace) > 0$ and (σ_1, σ_2) is a Nash equilibrium, the choice *Palace* must be optimal for Barbara under the belief σ_1. This is only possible if $\sigma_1(Corner) > 0$, since otherwise choosing *Palace* would yield Barbara a utility of 0, which would be less than what she gets by choosing *home.* Hence, we conclude that $\sigma_1(Corner) > 0$ and $\sigma_1(Palace) > 0$. However, this means that *both Palace* and *Corner* must be optimal for you under the belief σ_2. In particular, *Palace* and *Corner* must give you the same expected utility, which is only possible if $\sigma_2(Palace) = \sigma_2(Corner)$. Please check this. As a consequence, both $\sigma_2(Palace)$ and $\sigma_2(Corner)$ cannot be larger than 0.5. But then, the expected utility you get from choosing *Palace* can be at most $(0.5) \cdot 4 = 2$, which is less than what you would get by choosing *home.* We thus conclude that in a Nash equilibrium, $\sigma_1(Palace) > 0$ is impossible, and hence $\sigma_1(Palace) = 0$.

This implies, however, that choosing *Corner* can no longer be optimal for Barbara, as it would yield her a utility of 0. Since (σ_1, σ_2) is a Nash equilibrium, we thus conclude that $\sigma_2(Corner) = 0$. But then, choosing *Corner* can no longer be optimal for you, as it would yield you a utility of 0. As such, $\sigma_1(Corner) = 0$. Since $\sigma_1(Palace) = 0$ and $\sigma_1(Corner) = 0$, it follows that $\sigma_1(home) = 1$. But then, only *home* can be optimal for Barbara, which means that $\sigma_2(home) = 1$ as well. We thus conclude that $(\sigma_1, \sigma_2) = (home, home)$, and hence this is the only Nash equilibrium in this game. Since the only optimal choice for you in this Nash equilibrium is to stay at *home,* we see that under common belief in rationality with a *simple* belief hierarchy, you can only rationally choose to stay at *home*, and expect Barbara to do the same.

The intuitive reason why under common belief in rationality with a *simple* belief hierarchy you cannot rationally go to *Cinema Palace* is the following. *Cinema Palace* is only optimal for you if you believe, with high probability, that Barbara chooses *Cinema Palace* as well. For Barbara, in turn, it can only be optimal to choose *Cinema Palace* if she assigns a high probability to you choosing *The Movie Corner.* Finally, it can only be optimal for you to choose *The Movie Corner* if you believe, with high probability, that Barbara chooses *The Movie Corner* as well. Hence, you can only rationally go to *Cinema Palace* under common belief in rationality if you assign a high probability to (a) Barbara choosing *Cinema Palace*, a high probability to the event that (b) Barbara assigns a high probability to you choosing *The Movie Corner*, and a high probability to the event that (c) Barbara assigns a high probability to the event that you assign a high probability to Barbara choosing *The Movie Corner.* Because of (a) and (c), you must then necessarily believe, with high probability, that Barbara is incorrect about your first-order belief, and hence you cannot hold a simple belief hierarchy. In the same fashion, one can also intuitively explain why under common belief in rationality with a simple belief hierarchy, you cannot rationally go to *The Movie Corner.* That is, under common belief in rationality with a simple belief hierarchy, you can only rationally choose to stay at *home*.

Example 4.3: When Chris joins the party.

Reconsider the example "When Chris joins the party" from Section 3.3.2, with the decision problems as depicted in Table 3.2.1. Which choices can you rationally make here under common belief in rationality with a simple belief hierarchy?

In the beliefs diagram from Figure 3.2.2, consider the belief hierarchy that starts at your choice *green.* In Question 4.1.6 we have already seen that this belief hierarchy is *simple,* because it is generated by the belief combination ($\sigma_1 = green$, $\sigma_2 = blue$, $\sigma_3 = yellow$). Since your choice *green* is optimal under this belief hierarchy, we can conclude that you can rationally wear *green* under common belief in rationality with a simple belief hierarchy. Another way to see this is to observe that ($\sigma_1 = green$, $\sigma_2 = blue$, $\sigma_3 = yellow$) is a Nash equilibrium, and that your choice *green* is optimal in this Nash equilibrium. Please check this. In light of Theorem 4.1.2 we then know that you can rationally choose *green* under common belief in rationality with a simple belief hierarchy.

What about your choices *red* and *yellow*? Can you rationally make these choices as well under common belief in rationality with a simple belief hierarchy? Since we have seen in Section 3.4.3 that *yellow* can never be optimal for you for *any* belief, we know that you cannot rationally choose *yellow* under common belief in rationality with a simple belief hierarchy. It remains to explore your choice *red.* In the beliefs diagram from Figure 3.2.2, the belief hierarchy that starts at your choice *red* expresses common belief in rationality, since the diagram only involves solid arrows. Since your choice *red* is rational for that belief hierarchy, we see that you can rationally choose *red* under common belief in rationality. However, the belief hierarchy starting at your choice *red* is not simple. To see this, note that you believe that Barbara chooses *green,* yet at the same time you believe that Chris believes that Barbara chooses *yellow,* and not *green.* Hence, this belief hierarchy cannot be simple.

Therefore, the beliefs diagram from Figure 3.2.2 does not yet tell us whether or not you can rationally make your choice *red* under common belief in rationality with a simple belief hierarchy. To answer this question we rely on Theorem 4.1.2, and ask whether there are Nash equilibria in which your choice *red* is optimal. In fact, there is a Nash equilibrium in this game in which both *green* and *red* are optimal for you. Consider the belief combination

$$(\sigma_1 = \tfrac{1}{2} \cdot green + \tfrac{1}{2} \cdot red,\ \sigma_2 = \tfrac{1}{3} \cdot green + \tfrac{2}{3} \cdot yellow,\ \sigma_3 = blue).$$

Question 4.1.8 *Show that this belief combination is a Nash equilibrium, and that your choices green and red are both optimal in this Nash equilibrium.*

Hence, we conclude by Theorem 4.1.2 that under common belief in rationality with a simple belief hierarchy, you can also rationally choose *red.* Summarizing, we see that in this game you can rationally choose *green* and *red* under common belief in rationality with a simple belief hierarchy, but not *yellow.*

What choices can Barbara and Chris rationally make under common belief in rationality with a simple belief hierarchy? To answer this question, it is sufficient to look at the two Nash equilibria above. Let us start with Barbara's choices. We know from Section 3.4.3 that under common belief in rationality, Barbara can only rationally make the choices *blue, green* and *yellow,* but not *red.* Hence, Barbara can definitely not rationally choose *red* under common belief in rationality with a simple belief hierarchy. Note that her choice *blue* is optimal in the Nash equilibrium ($\sigma_1 = green$, $\sigma_2 = blue$, $\sigma_3 = yellow$), and that her choices *green* and *yellow* are optimal in the Nash equilibrium

$$(\sigma_1 = \tfrac{1}{2} \cdot green + \tfrac{1}{2} \cdot red,\ \sigma_2 = \tfrac{1}{3} \cdot green + \tfrac{2}{3} \cdot yellow,\ \sigma_3 = blue).$$

Please check this. In light of Theorem 4.1.2, we thus conclude that Barbara can rationally choose *blue, green* and *yellow* under common belief in rationality with a simple belief hierarchy.

For Chris' choices, observe that his choice *yellow* is optimal in the Nash equilibrium ($\sigma_1 = green$, $\sigma_2 = blue$, $\sigma_3 = yellow$), and that his choice *blue* is optimal in the Nash equilibrium

$$(\sigma_1 = \tfrac{1}{2} \cdot green + \tfrac{1}{2} \cdot red,\ \sigma_2 = \tfrac{1}{3} \cdot green + \tfrac{2}{3} \cdot yellow,\ \sigma_3 = blue).$$

By Theorem 4.1.2 we therefore know that Chris can rationally choose *yellow* and *blue* under common belief in rationality with a simple belief hierarchy.

Summarizing, we see that under common belief in rationality with a simple belief hierarchy, you can rationally choose *green* and *red* but not *yellow,* that Barbara can rationally choose *blue, green* and *yellow* but not *red,* and that Chris can rationally choose *blue* and *yellow.* Since we know from Section 3.4.3 that these are exactly the choices that you, Barbara and Chris can rationally make under common belief in rationality, the focus on simple belief hierarchies does not further restrict the choices that the players can rationally make under common belief in rationality here.

We thus see that in the example "When Chris joins the party", it does not matter for your choices whether, in addition to common belief in rationality, we also insist on simple belief hierarchies or not. We will now consider a new example with three players where simple belief hierarchies do make a difference for the choices you can rationally make.

Example 4.4: Opera for three.

This evening there will be the première of a new opera in the local theater. Although you are not too fond of opera, you are willing to go if both Barbara and Chris would join. However, last night Barbara and Chris had a fierce fight and would therefore rather avoid each other. In fact, Barbara would only consider going to the opera if you would join but not Chris. Similarly, Chris would only go to the opera if you would join but not Barbara. Every person thus has two choices: to go to the *opera*, or to stay at *home*. The utilities for you are as follows: If you stay at home you would watch your favorite movie and have a utility of 3. If you go to the *opera* and both Barbara and Chris are there, you would have a good time and enjoy a utility of 4. However, if you go the *opera* but either Barbara or Chris does not show up, you will not enjoy the evening and have a utility of 0. The utilities for Barbara and Chris are similar. The only difference is that Barbara will only have a good time at the opera if you join but not Chris, whereas Chris will only have a good time if you join but not Barbara. This story gives rise to the decision problems in Table 4.1.3. Here, the state (*home, opera*) in your decision problem means that Barbara stays at home and Chris goes to the opera. The state (*home, opera*) in Barbara's decision problem means that you stay at home and Chris goes to the opera. The state (*home, opera*) in Chris' decision problem means that you stay at home and Barbara goes to the opera.

The question is: What choice(s) can you rationally make under common belief in rationality with a simple belief hierarchy? A partial answer is given by the beliefs diagram in Figure 4.1.2.

Here, the choices *home* and *opera* are abbreviated by *h* and *o*, respectively. Note that all arrows in this beliefs diagram are solid, which means that every choice in this diagram is optimal for the belief hierarchy that starts at that choice. Please check this. Hence, by Theorem 3.3.2, all belief hierarchies in this diagram express common belief in rationality. Since it is optimal for you to go to the *opera* under the belief hierarchy that starts at your choice *opera,* and it is optimal for you to stay at *home* for the belief hierarchy that starts at your choice *home,* we conclude that under common belief in rationality you can rationally go to the *opera* or stay at *home.*

You	(*opera*, *opera*)	(*home*, *opera*)	(*opera*, *home*)	(*home*, *home*)
opera	4	0	0	0
home	3	3	3	3

Barbara	(*opera*, *opera*)	(*home*, *opera*)	(*opera*, *home*)	(*home*, *home*)
opera	0	0	4	0
home	3	3	3	3

Chris	(*opera*, *opera*)	(*home*, *opera*)	(*opera*, *home*)	(*home*, *home*)
opera	0	0	4	0
home	3	3	3	3

Table 4.1.3 Decision problems in "Opera for three"

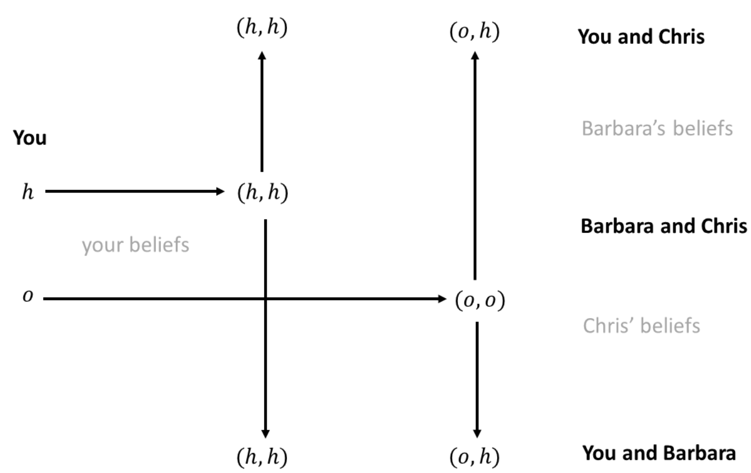

Figure 4.1.2 Beliefs diagram for "Opera for three"

Moreover, the belief hierarchy that starts at your choice *home* is generated by the beliefs combination ($\sigma_1 = home$, $\sigma_2 = home$, $\sigma_3 = home$), and is therefore simple. Hence, you can rationally choose to stay at *home* under common belief in rationality with a simple belief hierarchy.

What about going to the *opera*? Can you rationally make this choice under common belief in rationality with a simple belief hierarchy? The beliefs diagram does not give an answer to that question. The belief hierarchy that starts at your choice *opera* is not simple, since you believe that Chris goes to the *opera*, but at the same time you believe that Barbara believes that Chris stays at *home*. Hence, you believe that Barbara does not share your belief about Chris. But maybe there is some other beliefs diagram in which going to the opera *is* supported by a simple belief hierarchy that expresses common belief in rationality.

To answer this question we rely on Theorem 4.1.2 and try to find all Nash equilibria in this game. In fact, we will show that ($\sigma_1 = home$, $\sigma_2 = home$, $\sigma_3 = home$) is the only Nash equilibrium here. Suppose that $(\sigma_1, \sigma_2, \sigma_3)$ is a Nash equilibrium and assume, contrary to what we want to show, that $\sigma_1(opera) > 0$. Then, *opera* must be optimal for you under the belief σ_{-1} about the opponents' choices. Let us denote by $u_1(opera,\ \sigma_{-1})$ the expected utility for you of going to the opera under the belief σ_{-1}, and by $u_1(home,\ \sigma_{-1})$ the expected utility of staying at home. From your decision problem in Table 4.1.3 we see that

$$u_1(opera,\ \sigma_{-1}) = 4 \cdot \sigma_2(opera) \cdot \sigma_3(opera)$$

and

$$u_1(home,\ \sigma_{-1}) = 3.$$

Since *opera* must be optimal for you under the belief σ_{-1}, we must have that $u_1(opera,\ \sigma_{-1}) \geq u_1(home,\ \sigma_{-1})$, which is only possible if $\sigma_2(opera) \geq 0.75$ and $\sigma_3(opera) \geq 0.75$. In particular, $\sigma_2(opera) > 0$. Since $(\sigma_1, \sigma_2, \sigma_3)$ is a Nash equilibrium, the choice *opera* must be optimal for Barbara under the belief σ_{-2} about her opponents' choices. From Barbara's decision problem in Table 4.1.3 we see that

$$u_2(opera,\ \sigma_{-2}) = 4 \cdot \sigma_1(opera) \cdot \sigma_3(home)$$

and

$$u_2(home,\ \sigma_{-2}) = 3.$$

However, we have just seen that $\sigma_3(opera) \geq 0.75$, which means that $\sigma_3(home) \leq 0.25$. But then, $u_2(opera,\ \sigma_{-2}) \leq 4 \cdot (0.25) < 3$. Since Barbara can guarantee a utility of 3 by staying at *home*, we conclude that going to the *opera* cannot be optimal for Barbara under the belief σ_{-2}, which is a contradiction. We thus conclude that our initial assumption, that $\sigma_1(opera) > 0$, must be wrong. Hence, in every Nash equilibrium we must have that $\sigma_1(opera) = 0$, and hence $\sigma_1 = home$. But if Barbara and Chris expect you to stay at *home* with probability 1, it can never be optimal for Barbara and Chris to go to the *opera,* and hence we must also have that $\sigma_2 = home$ and $\sigma_3 = home$. Summarizing, we see that the only Nash equilibrium in this game is ($\sigma_1 = home$, $\sigma_2 = home$, $\sigma_3 = home$). In this Nash equilibrium, your only rational choice is to stay at *home,* and hence we conclude that you can only rationally stay at *home* under common belief in rationality with a simple belief hierarchy.

In particular, there is no *simple* belief hierarchy expressing common belief in rationality that would make you go to the *opera.* The intuitive reason is as follows. Consider a belief hierarchy that expresses common belief in rationality and that makes you go to the *opera.* Then, you must assign a high probability to Barbara going to the *opera*, and a high probability to Chris going to the *opera.* However, for Barbara it is only optimal to go to the *opera* if she assigns a low probability to Chris going to the *opera.* Hence, you must (a) assign a high probability to Chris going to the *opera*, and

(b) assign a high probability to the event that Barbara assigns a low probability to Chris going to the *opera*. But then, you must believe, with high probability, that Barbara does not share your belief about Chris, and therefore this belief hierarchy cannot be simple.

4.1.4 Nash Equilibria Always Exist

In each of the examples above, we were always able to find *simple* belief hierarchies that express common belief in rationality. But is this always the case? That is, can we always find, for every game and every player, at least one *simple* belief hierarchy that expresses common belief in rationality? The answer is not that obvious. We have seen, in Theorem 3.4.2, that for every game and for every player we can always find at least one belief hierarchy that expresses common belief in rationality. But such a belief hierarchy need not necessarily be simple.

Nevertheless, we can show that simple belief hierarchies that express common belief in rationality always exist for every player. To show this, we prove that for every game there is always at least one Nash equilibrium. Since, by Theorem 4.1.1, every simple belief hierarchy that is generated by a Nash equilibrium will expresses common belief in rationality, this will guarantee the existence of simple belief hierarchies expressing common belief in rationality.

Theorem 4.1.3 (Nash equilibria always exist) *For every game there is at least one Nash equilibrium.*

Unfortunately, the proof for this existence theorem is not as intuitive as the proof for Theorem 3.4.2, where we showed that for every game there is at least one belief hierarchy that expresses common belief in rationality. The details can be found in the proofs section at the end of this chapter.

Now, consider an arbitrary game and a player i within that game. Theorem 4.1.3 guarantees that there is a Nash equilibrium $(\sigma_1, ..., \sigma_n)$ in this game. Consider the simple belief hierarchy for player i generated by the Nash equilibrium $(\sigma_1, ..., \sigma_n)$. Then, we know by Theorem 4.1.1 that this belief hierarchy will express common belief in rationality. Therefore, we can construct for every player i at least one *simple* belief hierarchy that expresses common belief in rationality.

Theorem 4.1.4 (Existence) *For every game, and every player i, we can always find a simple belief hierarchy for player i that expresses common belief in rationality.*

In particular, for every game it is always possible to simultaneously reason in accordance with common belief in rationality, and to believe that your opponents are correct about your beliefs.

4.1.5 How Reasonable is Nash Equilibrium?

We know from Theorem 4.1.1 that Nash equilibrium is obtained if we combine the reasoning of common belief in rationality with the logic of a simple belief hierarchy. Suppose we accept common belief in rationality as a meaningful way of reasoning. Then, the question whether Nash equilibrium is a meaningful concept or not reduces to the following: To what extent are the conditions imposed by a simple belief hierarchy natural?

Recall that a simple belief hierarchy imposes the following restrictions on your beliefs: (a) you believe that your opponents are correct about your beliefs, and if there are at least two other players j and k, (b) you believe that player j holds the same belief about player k as you do, and (c) your belief about j's choice is independent from your belief about k's choice. But how reasonable are these conditions (a), (b) and (c)? Of course this is highly subjective, but let me at least try to argue why I personally do not find these conditions very compelling.

First, it seems highly artificial to impose that you must believe that your opponents are *correct* about the beliefs you hold. After all, the opponents cannot read your mind, and hence there is no particular reason why we should expect others to be correct about our own beliefs. Consider, for instance, the beliefs diagram in Figure 4.1.1 for "Movie for two". In the belief hierarchy that justifies your choice to go to *Cinema Palace* you believe that Barbara goes to *Cinema Palace* while at the same time you believe that Barbara believes that you believe that Barbara goes to *The Movie Corner* (and not to *Cinema Palace*). Hence, you believe that Barbara is wrong about your first-order belief. In my view there is nothing wrong with this belief hierarchy, since it expresses common belief in rationality, but this belief hierarchy is excluded by the concept of Nash equilibrium. Even more, we have seen that Nash equilibrium completely excludes the choice to go to *Cinema Palace,* although it can be justified by a perfectly meaningful belief hierarchy that expresses common belief in rationality.

Also, I find the restriction (b) above, stating that you must believe that player j has the same belief about player k's choice as you do, hard to justify. In case there are at least two reasonable choices for player k, say a and b, it seems perfectly fine to believe that player k chooses a, while at the same time believing that player j believes that player k chooses b. Consider, for instance, the beliefs diagram in Figure 4.1.2 for "Opera for three". According to common belief in rationality, both staying at *home* and going to the *opera* seem reasonable choices for Chris. Hence, there does not seem to be a problem with your belief hierarchy that starts at your choice *opera,* in which you believe that Chris goes to the *opera*, while at the same time believing that Barbara believes that Chris will stay at *home*. Nevertheless, this belief hierarchy, and your choice *opera* it supports, are excluded by the concept of Nash equilibrium.

Finally, condition (c) above, which states that your belief about j's choice must be independent from your belief about k's choice, also seems problematic to me. Consider the example "Opera for three", and suppose the première of the opera will be tomorrow evening. Assume you know that there is a 50% chance of a thunderstorm tomorrow evening. With this particular background information, it seems reasonable to believe that, with probability 0.5, both Barbara and Chris will stay at *home* (because there will be a thunderstorm), and with probability 0.5, both Barbara and Chris will go to the *opera* (because the weather will be good). Such beliefs, however, are excluded by Nash equilibrium, because your belief about Barbara's choice is not independent from your belief about Chris' choice.

My personal conclusion is thus that Nash equilibrium is based on rather problematic epistemic assumptions. This despite the fact that Nash equilibrium has been the dominant concept in game theory for a very long time. At the same time, this also shows the power of an epistemic approach to game theory: It reveals the – often implicit – epistemic assumptions behind various concepts in game theory, such as Nash equilibrium. And by doing so we can discuss the appeal of these concepts by critically analyzing the epistemic assumptions on which they are based.

4.2 Symmetric Beliefs

In this section we focus on *symmetric* belief hierarchies. We first explain what we mean by a symmetric belief hierarchy, and show that symmetric belief hierarchies can be characterized by a *common prior on choice-type combinations.* We use this insight to demonstrate that common belief in rationality in combination with symmetric belief hierarchies can be characterized by the concept of *correlated* *equilibrium.* This result, in turn, is applied to various examples to identify those choices that can rationally be made under common belief in rationality with a symmetric belief hierarchy. We finally discuss how reasonable the notion of correlated equilibrium is from a conceptual point of view.

4.2.1 Symmetric Belief Hierarchies

Consider the beliefs diagram in Figure 3.2.1 for "Going to a party", and concentrate on your belief hierarchy that starts at your choice *blue.* One way to read this belief hierarchy is as follows: You choose *blue* because you believe that Barbara chooses *red,* and you believe that Barbara chooses *red* because you believe that Barbara believes that you choose *blue.* And so on. This belief hierarchy can be viewed as symmetric, because the second piece of this sentence can be interpreted as the symmetric counterpart to the first piece. This symmetry can also be detected visually, by noting that the second arrow, which goes from Barbara's choice *red* to your choice *blue,* is literally symmetric to the first arrow, which goes from your choice *blue* to Barbara's choice *red.*

The same cannot be said about your other belief hierarchies in this beliefs diagram. Take, for instance, the belief hierarchy that starts at your choice *green.* There, you choose *green* because you believe that Barbara chooses *blue,* but you do not believe that Barbara chooses *blue* because she believes that you choose *green.* Even stronger, you believe that Barbara chooses *blue* because you believe that she holds a belief in which she assigns probability zero to you choosing *green.* Hence, this belief hierarchy is certainly not symmetric.

The distinction between symmetric and non-symmetric belief hierarchies can also be made in games with more than two players. Consider, for instance, the example "Opera for three", with the beliefs diagram in Figure 4.1.2. Your belief hierarchy that starts at your choice *home* is symmetric. To see this, note that according to this belief hierarchy you choose *home* because you believe that both Barbara and Chris choose *home.* Moreover, you believe that Barbara chooses *home* because you believe that Barbara believes that both you and Chris choose *home.* Similarly, you believe that Chris chooses *home* because you believe that Chris believes that both you and Barbara choose *home.* Here, the last two sentences can be viewed as the symmetric counterparts – from Barbara's and Chris' view, respectively – of the first sentence. Or, in visual terms, the belief hierarchy is generated by the arrow from your choice h to Barbara's and Chris' choice pair (h, h), the arrow from Barbara's choice h to your and Chris' choice pair (h, h), and the arrow from Chris' choice h to your and Barbara's choice pair (h, h). Since these three arrows are symmetric, the belief hierarchy generated by it may be viewed as symmetric as well.

In contrast, your belief hierarchy that starts at your choice *opera* is definitely not symmetric. Indeed, you choose *opera* because you believe that Barbara chooses *opera* and Chris chooses *opera,* yet at the same time you believe that Barbara chooses *opera* because you believe that Barbara believes that Chris chooses *home* (and not *opera*). This belief hierarchy thus assumes a clear asymmetry between Barbara and you when it comes to the belief about Chris' choice.

The symmetric belief hierarchies above display a particularly easy case of symmetry, for two reasons. First, they both involve probability 1 beliefs only. Moreover, both belief hierarchies are *simple* which, as we will see, constitutes a very special and strong case of symmetry. In general, however, symmetric belief hierarchies need not be restricted to probability 1 beliefs, nor need they be simple. The following example will illustrate this.

Example 4.5: Rock, paper, scissors.

During a long and boring train ride, Barbara and you decide to play the famous game of *rock, paper, scissors.* After a while, the game gets equally boring, however. To make it more interesting, you and

You	*rock*	*paper*	*scissors*	*diamond*
rock	1	3	4	1
paper	4	1	3	4
scissors	3	4	1	3
bomb	4	0	1	1

Barbara	*rock*	*paper*	*scissors*	*bomb*
rock	1	3	4	0
paper	4	1	3	4
scissors	3	4	1	1
diamond	1	3	4	1

Table 4.2.1 Decision problems in "Rock, paper, scissors"

Barbara are allowed to add one more object next to the rock, the paper and the scissors. After a long thought you decide to add a *bomb,* which can be visualized by an exploding hand. You are the only one who can use this object. Since the bomb would destroy the rock, it will win against a rock. However, the paper could wrap the bomb, and hence the bomb would loose against paper. Finally, it would tie against scissors, as neither of the two could destroy the other.

In response to your surprising contribution, Barbara decides to add a *diamond* to the game since it would be strong enough to survive an explosion. More precisely, there would be a tie between a bomb and a diamond, as neither would be able to destroy the other. Apart from this, the diamond has the same properties as the rock: It would tie against a rock, loose against paper, since it would be wrapped by it, and win against scissors, as it is strong enough to crush the scissors. Barbara is the only one who can use the diamond. As usual, the rock beats the scissors, the paper beats the rock, and the scissors beat the paper. And, of course, equal objects tie against each other.

Importantly, Barbara and you both find joy in mimicking the act of beating the other object, or even the act of being beaten by the other object. Indeed, the act of wrapping, crushing, cutting or exploding the other object creates a moment of intense laughter for the two of you during the boring train ride. Whenever an object beats the other object, then the person with the winning object enjoys a utility of 4, whereas the other person still obtains a utility of 3. If there is a tie between the two objects, then there is no act to enjoy, and both would get a utility of only 1.

There are two exceptions to this rule: If you choose the bomb and Barbara chooses the rock, then you tremendously enjoy the act of exploding the rock, and your utility will be 4, as expected. However, Barbara would feel cheated by this, and her utility would only be 0 in this case. If, on the other hand, you choose the bomb and Barbara chooses the paper, the bomb will be wrapped by the paper, and this would give Barbara a utility of 4, as expected. In contrast, you would feel terribly disappointed in this case, since your new invention would loose against something as simple as paper, and your utility would only be 0.

This story can be summarized by the decision problems in Table 4.2.1. Consider the beliefs diagram at the top of Figure 4.2.1, in which your choice *paper* appears twice. The second time it appears, we denote it by *paper'* as to distinguish it from the first time it appears. Concentrate on your belief hierarchy that starts at your choice *rock*.

At first sight this belief hierarchy does not appear symmetric, especially because of the different probabilities that are involved. However, the same belief hierarchy can also be represented by the *weighted beliefs diagram* at the bottom of Figure 4.2.1, which *is* symmetric.

This weighted beliefs diagram should be read as follows: The numbers $1, 2$ and 3 at the various arrows represent the *weights* that we assign to these arrows. We therefore speak of a *weighted* beliefs diagram. Note that the two outgoing arrows at your choice *rock* carry the weights 1 and 2. Therefore,

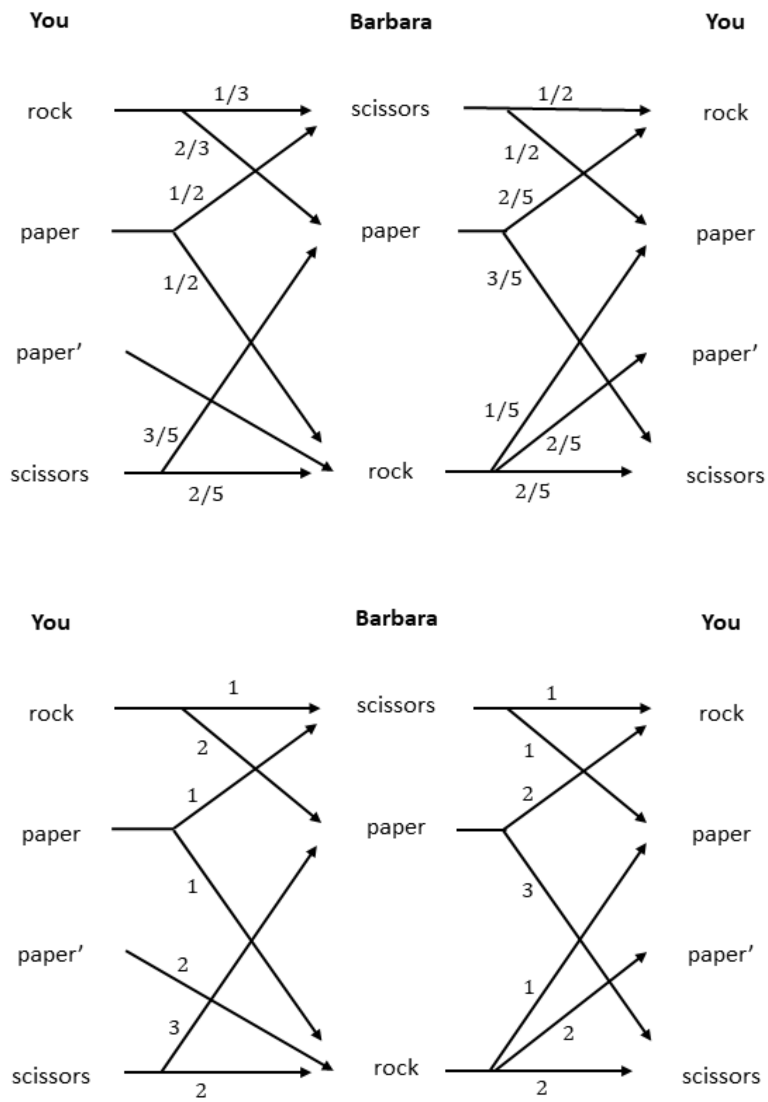

Figure 4.2.1 Beliefs diagram and an associated symmetric weighted beliefs diagram for "Rock, paper, scissors"

the relative weight of the first outgoing arrow is

$$\frac{1}{1+2} = 1/3,$$

whereas the relative weight of the second outgoing arrow is

$$\frac{2}{1+2} = 2/3.$$

These numbers 1/3 and 2/3 are exactly the probabilities of the same outgoing arrows in the non-weighted beliefs diagram on top. The weights 1 and 2 at the outgoing arrows leaving your choice *rock* therefore induce the probabilistic belief that assigns probability 1/3 to Barbara choosing *scissors* and probability 2/3 to Barbara choosing *paper.* These probabilities are obtained by taking the *relative weights.*

Similarly, the weights 2 and 3 at the outgoing arrows leaving Barbara's choice *paper* induce the probabilistic belief in which Barbara assigns probability

$$\frac{2}{2+3} = 2/5$$

to you choosing *rock, and* probability

$$\frac{3}{2+3} = 3/5$$

to you choosing *scissors*. Again, these probabilities are obtained by taking the relative weights at both outgoing arrows. Note that these probabilities coincide exactly with the probabilities in the non-weighted beliefs diagram on top.

Consider finally the unique outgoing arrow at your choice *paper'* in the weighted beliefs diagram, carrying a weight of 2. Of course, its relative weight is

$$\frac{2}{2} = 1$$

because there is no other outgoing arrow here. It thus induces the belief for you that assigns probability 1 to Barbara choosing *rock,* just like the corresponding arrow in the non-weighted beliefs diagram.

In the same fashion, it can be verified that the remaining relative weights in the weighted beliefs diagram at the bottom induce exactly the probabilities at the corresponding arrows in the beliefs diagram on top. We thus conclude that the weighted beliefs diagram at the bottom of Figure 4.2.1 induces exactly the same beliefs – and hence, the same belief hierarchies – as the non-weighted beliefs diagram on top of Figure 4.2.1. Therefore, the weighted beliefs diagram at the bottom *induces* the (non-weighted) beliefs diagram on top.

Note that the weighted beliefs diagram is *symmetric*, because every arrow carries the same weight as its symmetric counterpart. To see this, consider for instance the arrow from your choice *paper'* to Barbara's choice *rock,* which has a weight of 2. The symmetric counterpart would be the arrow from Barbara's choice *rock* to your choice *paper',* which also has a weight of 2. Similarly, the arrow from your choice *scissors* to Barbara's choice *paper* has a weight of 3. This is the same as the weight of its symmetric counterpart, which is the arrow from Barbara's choice *paper* to your choice *scissors.* The same form of symmetry can be verified for all other arrows in this weighted beliefs diagram as well.

Summarizing, we see that your belief hierarchy that starts at your choice *rock* is induced by a *symmetric weighted* beliefs diagram – the one at the bottom of Figure 4.2.1. For that reason, we call this belief hierarchy *symmetric.*

Note, however, that this belief hierarchy is not simple. Indeed, your first-order belief assigns probability 1/3 to Barbara choosing *scissors* and 2/3 to Barbara choosing *paper.* At the same time, you believe with probability 1/3 that Barbara assigns probability 1/2 to the event that you assign probability 1/2 to Barbara's choices *scissors* and *rock.* As such, you do not believe, with probability 1, that Barbara is correct about your first-order belief, and therefore this belief hierarchy cannot be simple. We thus see that a symmetric belief hierarchy need not be simple.

Question 4.2.1 *Consider the beliefs diagram for "Going to a party" in Figure 3.2.1, and concentrate on the belief hierarchy that starts at your choice red. Explain why this belief hierarchy cannot be generated by a symmetric weighted beliefs diagram. We therefore conclude that this belief hierarchy is not symmetric.*

The idea of symmetric belief hierarchies can also be extended to games with more than two players. As an illustration, consider the example "When Chris joins the party" from the previous chapter, and the beliefs diagram in the upper half of Figure 4.2.2.

We will argue that your belief hierarchy starting at your choice *green* is symmetric. To see this, observe first that the beliefs diagram in the upper half of Figure 4.2.2 is induced by the *weighted* beliefs diagram in the lower half of that same figure. For instance, the forked arrow from Chris' choice y to your and Barbara's choice combinations (g, b) and (r, g) carries the weights 4 and 1 in the weighted beliefs diagram. Therefore, the induced probability that Chris assigns to your and Barbara's choice combination (g, b) is equal to its relative weight, which is

$$\frac{4}{4+1} = 0.8.$$

This is exactly the probability assigned to the corresponding arrow in the beliefs diagram. In a similar way, it can be checked that all the relative weights in the weighted beliefs diagram correspond exactly to the probabilities in the beliefs diagram. We can therefore conclude that the weighted beliefs diagram in the lower half of Figure 4.2.2 induces the beliefs diagram in the upper half of that figure.

Moreover, the weighted beliefs diagram turns out to be symmetric. Take, for instance, the arrow from your choice g to Barbara and Chris' choice combination (b, y), which carries a weight of 4. The symmetric arrow from Barbara's perspective is the arrow from Barbara's choice b to your and Chris' choice combination (g, y), which carries the same weight of 4. The symmetric arrow from Chris' perspective is the arrow from his choice y to your and Barbara's choice combination (g, b), which also carries this weight of 4.

In a similar fashion it can be verified that for every arrow in the weighted beliefs diagram, the two symmetric arrows for the two opponents are also present in the weighted beliefs diagram, and these symmetric counterparts carry the same weight as the arrow we started from. As such, the weighted beliefs diagram can be called *symmetric.*

We thus conclude that the beliefs diagram from Figure 4.2.2 is induced by the symmetric weighted beliefs diagram from the same figure. In particular, your belief hierarchy starting at your choice *green* is induced by this particular symmetric weighted beliefs diagram, and therefore we call this belief hierarchy *symmetric.* In fact, *all* belief hierarchies from Figure 4.2.2 are symmetric, since they are all induced by the same symmetric weighted beliefs diagram in the lower half of that figure.

With these illustrations in mind, we can now give a general definition of symmetric belief hierarchies. We will do so in steps: First, we define a weighted beliefs diagram, we then explain when such a weighted beliefs diagram is symmetric, indicate how a weighted beliefs diagram induces a normal beliefs diagram, and finally call a belief hierarchy symmetric if it is part of a beliefs diagram that is induced by a symmetric weighted beliefs diagram.

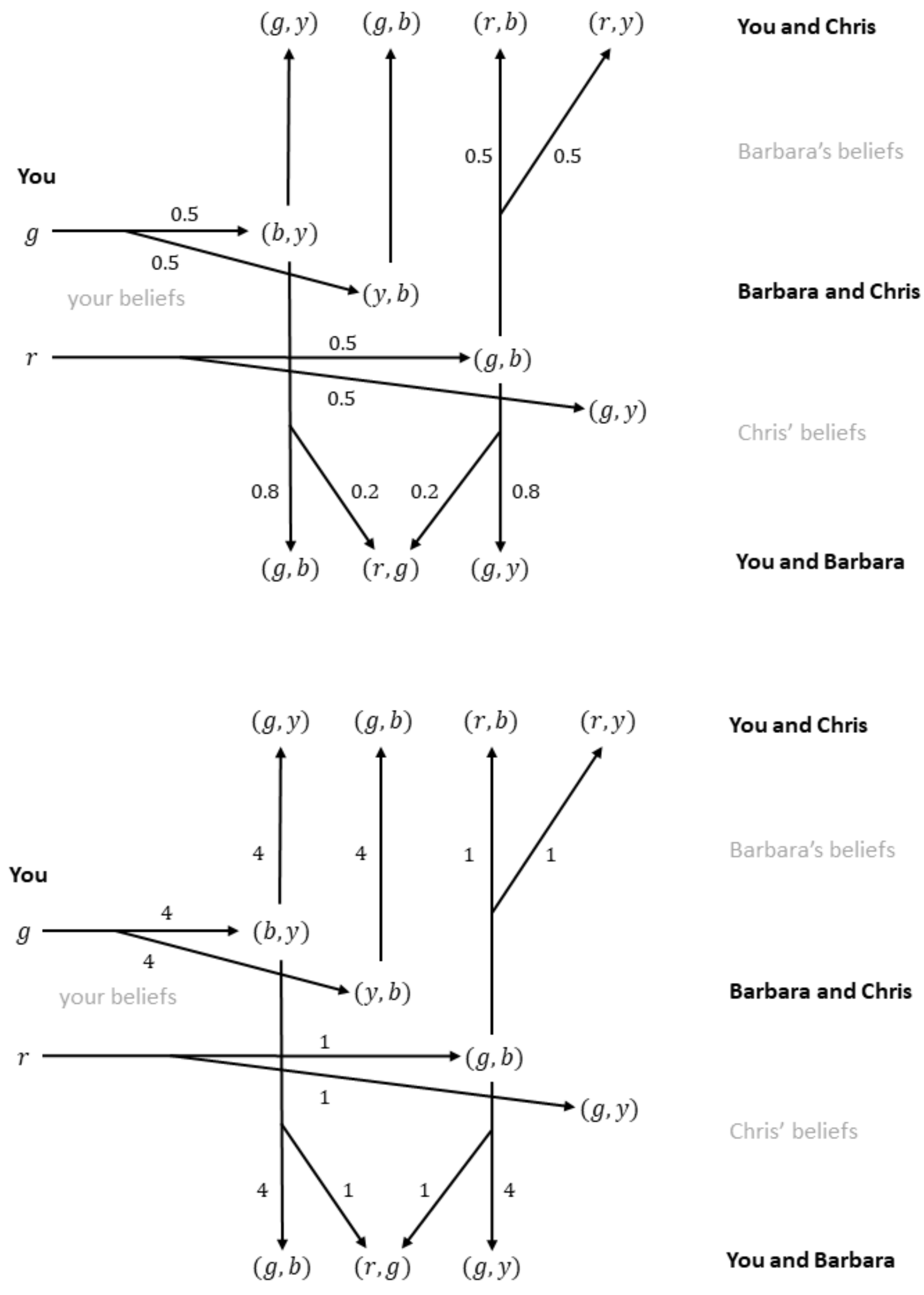

Figure 4.2.2 A beliefs diagram and an associated symmetric weighted beliefs diagram for "When Chris joins the party"

Definition 4.2.1 (Symmetric belief hierarchy) *(a) A* ***weighted beliefs diagram*** *starts from a beliefs diagram, removes the probabilities at the forked arrows (if there are any), and assigns to every arrow a from a choice c_i to an opponents' choice combination $(c_j)_{j \neq i}$ some positive weight, which we call $w(a)$.*

(b) Consider an arrow a from a choice c_i to an opponents' choice combination $(c_j)_{j \neq i}$. For every opponent j, the ***symmetric counterpart*** *to a is the arrow from the choice c_j to the opponents' choice combination $(c_k)_{k \neq j}$, using the same choices as a.*

(c) A weighted beliefs diagram is ***symmetric*** *if for every arrow a, each of its symmetric counterparts (one for every opponent) is also part of the diagram, and carries the same weight as a.*

(d) The weighted beliefs diagram induces a (normal) beliefs diagram in which the probability of an arrow a leaving a choice c_i is equal to

$$p(a) = \frac{w(a)}{\sum\limits_{\text{arrows } a' \text{ leaving } c_i} w(a')}.$$

(e) A belief hierarchy is ***symmetric*** *if it is part of a beliefs diagram that is induced by a symmetric weighted beliefs diagram.*

In part (a) we interpret a forked arrow, from a choice c_i to *several* opponents' choice combinations, as a *collection* of arrows. Consider, for instance, the beliefs diagram from Figure 4.2.2. The forked arrow from Chris' choice y to the choice combinations (g, b) and (r, g) by you and Barbara is interpreted as a pair of arrows: one arrow from Chris' choice y to the opponents' choice combination (g, b), which receives weight 4 in the weighted beliefs diagram, and one arrow from Chris' choice y to the opponents' choice combination (r, g), which receives weight 1 in the weighted beliefs diagram.

To illustrate parts (b) and (c), consider the weighted beliefs diagram from Figure 4.2.2, and the arrow a from your choice r to the opponents' choice combination (g, b). The symmetric counterpart for Barbara to a is the arrow from Barbara's choice g to the opponents' choice combination (r, b), whereas the symmetric counterpart for Chris is the arrow from Chris' choice b to the opponents' choice combination (r, g). Note that all these symmetric counterparts carry the same weight 1, because the weighted beliefs diagram is symmetric.

In case the same choice c_i of a player appears more than once in the (weighted) beliefs diagram, the different copies of c_i are formally viewed as different choices. Suppose, for instance, that the choice c_i appears twice, with the first copy denoted by c_i and the second copy denoted by c_i'. Then, c_i and c_i' are viewed as different. Fix an opponents' choice combination $(c_j)_{j \neq i}$. Then, also the arrow from c_i to $(c_j)_{j \neq i}$ and the arrow from c_i' to $(c_j)_{j \neq i}$ are viewed as different.

As an illustration, consider the beliefs diagram and the weighted beliefs diagram from Figure 4.2.1 for "Rock, paper, scissors", where your choice *paper* appears twice. The second time it appears, it is denoted by *paper'*. Formally, *paper* and *paper'* are then viewed as different choices within the beliefs diagram. Also, the arrow from your choice *paper* to Barbara's choice *rock* and the arrow from your choice *paper'* to Barbara's choice *rock* are viewed as different in the (weighted) beliefs diagram.

The reason your choices *paper* and *paper'* are viewed as different is that they induce different belief hierarchies. Indeed, if we start at your choice *paper* then, in the induced belief hierarchy, you assign probability 1/2 to Barbara's choices *scissors* and *rock.* If we start at your choice *paper'* instead then you assign probability 1 to Barbara's choice *rock.* This is true in general: Within a given beliefs diagram, two copies of the same choice will typically induce two different belief hierarchies. Therefore, we will treat these as two different choices.

4.2.2 Relation with Common Prior

We will now look for an easy way to characterize belief hierarchies that are symmeric. Remember that the same choice may appear more than once in a beliefs diagram. In the beliefs diagram from Figure 4.2.1, for instance, your choice *paper* appears twice. This enables us to model a belief hierarchy in which the same choice *paper* for you is justified by two different beliefs. Indeed, consider Barbara's belief hierarchy that starts at her choice *rock.* There, she assigns probability 1/5 to the event that you choose *paper* while assigning probability 1/2 to Barbara's choices *scissors* and *rock.* At the same time, she assigns probability 2/5 to the event that you choose *paper* (denoted by *paper'* in the beliefs diagram) while assigning probability 1 to Barbara's choice *rock.* Hence, Barbara justifies your choice *paper* by two different first-order beliefs that you can have about her. In general, it may be necessary to include two, or more, copies of the same choice c_i if we want to model a belief hierarchy in which the same choice c_i is justified by two, or more, different beliefs.

To distinguish between the various copies of c_i that may be present, we identify every copy of c_i with a choice-type pair (c_i, t_i), where t_i is some type. The intuition is that t_i represents the belief hierarchy that is obtained if we start at this particular copy of the choice c_i. As an illustration, consider again the beliefs diagram from Figure 4.2.1. In the beliefs diagram, the two copies of your choice *paper* are denoted by *paper* and *paper',* respectively. To distinguish between these two copies, we may identify the first copy with the choice-type pair (*paper*, t_1^p) and the second copy with (*paper*, $\hat{t}_1^p$). Here, t_1^p represents your belief hierarchy that starts at *paper* for you, and $\hat{t}_1^p$ represents the belief hierarchy that starts at *paper'* for you. In the same way, we can identify your choices *rock* and *scissors* with the choice-type pairs (*rock*, t_1^r) and (*scissors*, t_1^s), where t_1^r represents the unique belief hierarchy that starts at your choice *rock* and t_1^s represents the unique belief hierarchy that starts at your choice *scissors.* Similarly, Barbara's choices *scissors, paper* and *rock* can be identified with the choice-type pairs (*scissors*, t_2^s), (*paper*, t_2^p) and (*rock*, t_2^r). This results in the alternative representation of the beliefs diagram, and the associated symmetric weighted beliefs diagram, in Figure 4.2.3.

We call these two diagrams the *choice-type representation* of the beliefs diagram and the weighted beliefs diagram, respectively. The sets of types $T_1 = \{t_1^r, t_1^p, \hat{t}_1^p, t_1^s\}$ and $T_2 = \{t_2^s, t_2^p, t_2^r\}$ needed for the choice-type representation are called the *associated sets of types.*

On the basis of these choice-type representations, we can show that the entire beliefs diagram is induced by a *unique* probability distribution on choice-type combinations, which we call a *common prior.* To see how this works, consider the symmetric weighted beliefs diagram from Figure 4.2.3, and the choice-type combination $(c,t) = ((paper, \hat{t}_1^p), (rock, t_2^r))$ within that weighted beliefs diagram. Since the weighted beliefs diagram is symmetric, the arrow from $(paper, \hat{t}_1^p)$ to $(rock, t_2^r)$ and the arrow from $(rock, t_2^r)$ to $(paper, \hat{t}_1^p)$ both receive weight 2. As such, we can assign the unique weight $w(c,t) = 2$ to this particular choice-type combination.

If we assign such weights to all choice-type combinations, then we obtain the weights as displayed in Table 4.2.2. Here, the weights always correspond to the first number in the cell. All choice-type combinations that do not appear are assumed to receive weight 0.

If we divide every weight $w(c,t)$ by the sum of all weights, which is 12, we obtain the second number in each cell, which appears between brackets. As each of these second numbers is at least 0, and the sum of all second numbers is 1, these numbers correspond to a *probability distribution* π on choice-type combinations, which we call a *common prior.*

Moreover, this common prior π *induces* the beliefs diagram in the upper half of Figure 4.2.3, as follows. Consider, for instance, the arrow from your choice-type pair (*rock,* t_1^r) to Barbara's choice-type pair (*scissors,* t_2^s), which carries the probability 1/3. We are thus looking at the choice-type combination $(c,t) = ((rock,\ t_1^r), (scissors,\ t_2^s))$ with common prior probability $\pi(c,t) = 1/12$. The total

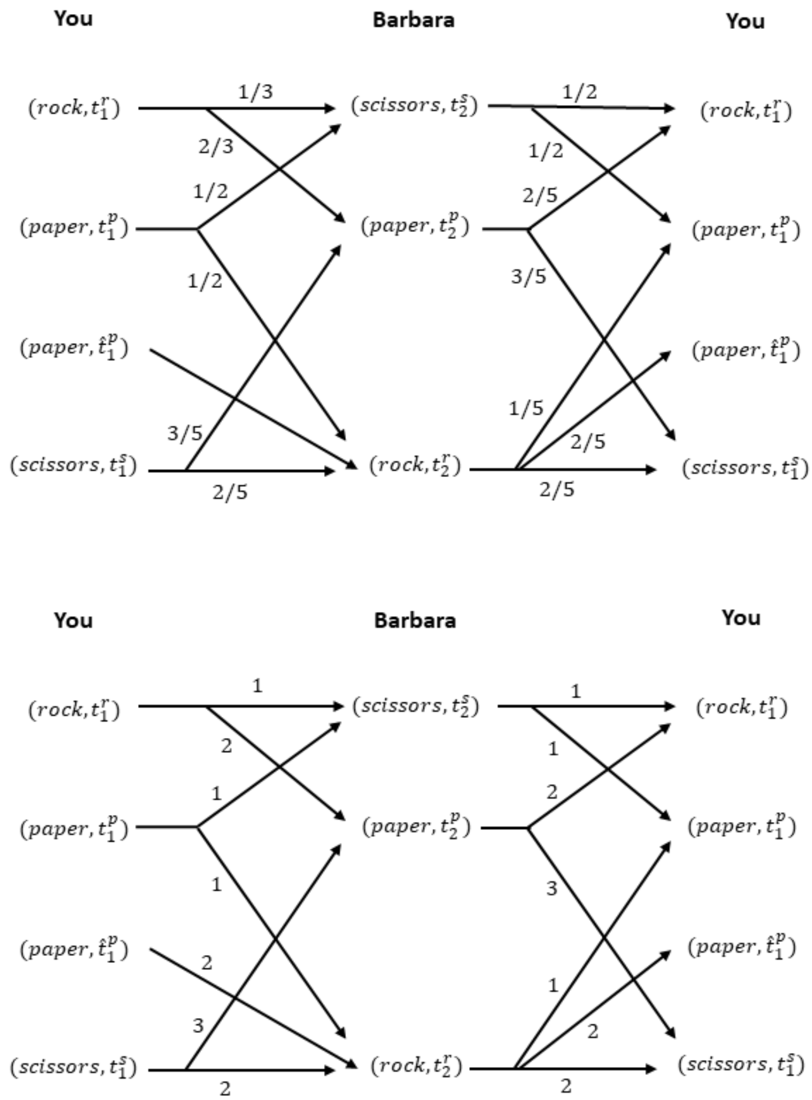

Figure 4.2.3 Choice-type representation of the beliefs diagram and weighted beliefs diagram for "Roch, paper, scissors"

(c,t)	$((rock,\ t_1^r),(scissors,\ t_2^s))$	$((rock,\ t_1^r),(paper,\ t_2^p))$	$((paper,\ t_1^p),(scissors,\ t_2^s))$
weight $w(c,t)$	1 (1/12)	2 (2/12)	1 (1/12)

(c,t)	$((paper,\ t_1^p),(rock,\ t_2^r))$	$((paper,\ \hat{t}_1^p),(rock,\ t_2^r))$
weight $w(c,t)$	1 (1/12)	2 (2/12)

(c,t)	$((scissors,\ t_1^s),(paper,\ t_2^p))$	$((scissors,\ t_1^s),(rock,\ t_2^r))$
weight $w(c,t)$	3 (3/12)	2 (2/12)

Table 4.2.2 Weights on choice-type combinations in "Rock, paper, scissors"

common prior probability assigned to the choice-type pair (*rock,* t_1^r) we started from is

$$\pi(\textit{rock, } t_1^r) = \pi((\textit{rock, } t_1^r), (\textit{scissors, } t_2^s)) + \pi((\textit{rock, } t_1^r), (\textit{paper, } t_2^p)) = 1/12 + 2/12 = 3/12.$$

As such,

$$\frac{\pi((\textit{rock, } t_1^r), (\textit{scissors, } t_2^s))}{\pi(\textit{rock, } t_1^r)} = \frac{1/12}{3/12} = 1/3,$$

which happens to be the probability attached to the arrow from (*rock,* t_1^r) to (*scissors,* t_2^s).

The fraction above has a natural interpretation: The enumerator may be viewed as the prior probability of the event that "you choose *rock,* your type is t_1^r, Barbara chooses *scissors,* and Barbara's type is t_2^s". In turn, the denominator represents the prior probability of the event that "you choose *rock* and your type is t_1^r". The fraction may thus be seen as the probability you assign to the event that "Barbara chooses *scissors* and Barbara's type is t_2^s", conditional on the event that "you choose *rock* and your type is t_1^r". In other words, the fraction is the probability that your type t_1^r assigns to Barbara's choice type pair (*scissors,* t_2^s), which is indeed 1/3.

More generally, it may be verified in Figure 4.2.3 that for every arrow a from a choice-type pair (c_1, t_1) of yours to a choice-type pair (c_2, t_2) of Barbara, we have that

$$p(a) = \frac{\pi((c_1, t_1), (c_2, t_2))}{\pi(c_1, t_1)}, \tag{4.2.1}$$

where $p(a)$ is the probability of arrow a in the beliefs diagram, and $\pi(c_1, t_1)$ is the total common prior probability assigned to (c_1, t_1). Please check this. Similarly as above, the fraction represents the prior probability you assign to the event that "Barbara chooses c_2 and Barbara's type is t_2", conditional on the event that "you choose c_1 and your type is t_1". This, in turn, should match the probability that your type t_1 assigns to (c_2, t_2), which is $p(a)$.

We can also view the situation from Barbara's perspective, and verify that for every arrow a from a choice-type pair (c_2, t_2) of Barbara to a choice-type pair (c_1, t_1) of yours, it holds that

$$p(a) = \frac{\pi((c_1, t_1), (c_2, t_2))}{\pi(c_2, t_2)}. \tag{4.2.2}$$

In this sense, the common prior π on choice-type combinations induces the beliefs diagram from Figure 4.2.3.

The reason why the beliefs diagram is induced by a common prior lies in its symmetry: Since there is an associated symmetric weighted beliefs diagram, we could build a common prior on the basis of the weights in that symmetric weighted beliefs diagram, and this common prior induces the beliefs diagram we started with.

This construction does not only work in this specific example, but it works in general: Whenever we have a symmetric beliefs diagram in choice-type representation (possibly with more than two players), we can look at the associated symmetric weighted beliefs diagram, and build a common prior on the basis of these weights. This common prior will then always induce the symmetric beliefs diagram we started with.

The other direction is also true: If the beliefs diagram is induced by a common prior, then the beliefs diagram will automatically be symmetric.

Hence, we reach the general conclusion that the symmetric beliefs diagrams are precisely those that are induced by a common prior on choice-type combinations. As a consequence, a belief hierarchy is symmetric (that is, part of a symmetric beliefs diagram) precisely when it is part of a beliefs diagram that is induced by a common prior on choice-type combinations.

To formally state this result, we first need to define what it means, in general, that a beliefs diagram is induced by a common prior. Consider a beliefs diagram in choice-type representation with possibly more than two players. Then, an arrow a will always go from a choice-type pair (c_i, t_i) for player i to an opponents' choice-type combination (c_{-i}, t_{-i}). Every such arrow a will have a probability $p(a)$ in the beliefs diagram.

Definition 4.2.2 (Common prior on choice-type combinations) *Consider a beliefs diagram in choice-type representation, with associated sets of types T_i for every player i. Let $C \times T$ be the corresponding set of all choice-type combinations.*

*(a) A **common prior on choice-type combinations** is a probability distribution π that assigns to every choice-type combination (c, t) in $C \times T$ a probability $\pi(c, t)$.*

*(b) The beliefs diagram is **induced by a common prior** π on choice-type combinations, if for every choice-type combination $((c_i, t_i), (c_{-i}, t_{-i}))$ and every player i, the corresponding arrow a from (c_i, t_i) to (c_{-i}, t_{-i}) is present exactly when $\pi((c_i, t_i), (c_{-i}, t_{-i})) > 0$, and the probability of this arrow a is equal to*

$$p(a) = \frac{\pi((c_i, t_i), (c_{-i}, t_{-i}))}{\pi(c_i, t_i)}.$$

*(c) A belief hierarchy is **induced by a common prior** π on choice-type combinations if it is part of a beliefs diagram that is induced by π.*

Above we have argued that a beliefs diagram is symmetric precisely when it is induced by a common prior. As a consequence, a belief hierarchy is symmetric precisely when it is induced by a common prior. This leads to the following result.

Theorem 4.2.1 (Relation with common prior) *A belief hierarchy is symmetric, if and only if, it is induced by a common prior on choice-type combinations.*

As an illustration, consider the beliefs diagram in choice-type representation on top of Figure 4.2.3, for "Rock, paper, scissors". Concentrate on your belief hierarchy that starts at (*rock*, t_1^r). Since we have shown that this belief hierarchy is symmetric, we know by Theorem 4.2.1 that this belief hierarchy can be derived from a common prior on choice-type combinations π. Moreover, we have seen that one such common prior on choice-type combinations is the probability distribution π from Table 4.2.2.

Question 4.2.2 *Consider the beliefs diagram in Figure 4.2.2 for "When Chris joins the party". Find a common prior on choice-type combinations that induces this beliefs diagram.*

4.2.3 Relation with Correlated Equilibrium

In the previous two subsections we have focused on symmetric belief hierarchies. First, we have defined what it means for a belief hierarchy to be symmetric, and subsequently we have characterized symmetric belief hierarchies as those that are induced by a common prior on choice-type combinations.

We now wish to combine *symmetric* belief hierarchies with the central notion of *common belief in rationality.* More precisely, we will zoom in on symmetric belief hierarchies that express common belief in rationality, and try to characterize those choices that can rationally be made while holding such belief hierarchies. As we shall see, this will lead us to the concept of *correlated equilibrium.*

Consider a belief hierarchy β_i that expresses common belief in rationality. Suppose it is part of some beliefs diagram in choice-type representation, where it starts at a choice-type pair (c_i^*, t_i^*) for player i. We can always design the beliefs diagram in such a way that the choice c_i^* is *optimal* for the first-order belief in the belief hierarchy. In this case, the outgoing arrows at (c_i^*, t_i^*) are all solid. As the belief hierarchy β_i believes in the opponents' rationality, all step 2 arrows following these outgoing arrows will be solid as well. Moreover, all step 3 arrows following the step 2 arrows will also be solid, because the belief hierarchy expresses 2-fold belief in rationality. And so on. By continuing in this fashion, we conclude that all arrows that are part of this belief hierarchy are solid. But then, we may restrict to the – possibly smaller – beliefs diagram that only contains the arrows that are present in the belief hierarchy β_i, and where all arrows are solid. We thus see that every belief hierarchy that expresses common belief in rationality is part of some beliefs diagram where all arrows are solid.

Hence, let us assume that the belief hierarchy β_i above is indeed part of a beliefs diagram in which all arrows are solid. Suppose now that the belief hierarchy β_i does not only express common belief in rationality, but is also *symmetric*. Then, by Theorem 4.2.1, the beliefs diagram is induced by a common prior π on choice-type combinations. That is, for every player j, every choice-type pair (c_j, t_j), and every opponents' choice-type combination (c_{-j}, t_{-j}), the arrow a from (c_j, t_j) to (c_{-j}, t_{-j}) is present in the beliefs diagram exactly when $\pi((c_j, t_j), (c_{-j}, t_{-j})) > 0$, and the arrow a has probability

$$p(a) = \frac{\pi((c_j, t_j), (c_{-j}, t_{-j}))}{\pi(c_j, t_j)}. \tag{4.2.3}$$

Let us denote by $\pi(\cdot \mid c_j, t_j)$ the belief of player j about the opponents' choice-type combinations given by

$$\pi(c_{-j}, t_{-j} \mid c_j, t_j) := \frac{\pi((c_j, t_j), (c_{-j}, t_{-j}))}{\pi(c_j, t_j)} \text{ for every } (c_{-j}, t_{-j}) \in C_{-j} \times T_{-j}. \tag{4.2.4}$$

We call $\pi(\cdot \mid c_j, t_j)$ the *belief of player j conditional on his choice-type pair* (c_j, t_j). Hence, by (4.2.3) and (4.2.4) we can write that

$$p(a) = \pi(c_{-j}, t_{-j} \mid c_j, t_j) \tag{4.2.5}$$

whenever a is an arrow from the choice-type pair (c_j, t_j) to the opponents' choice-type combination (c_{-j}, t_{-j}).

We say that a choice c_j^* is *optimal* for the belief $\pi(\cdot \mid c_j, t_j)$ conditional on (c_j, t_j) if

$$\sum_{(c_{-j}, t_{-j}) \in C_{-j} \times T_{-j}} \pi(c_{-j}, t_{-j} \mid c_j, t_j) \cdot u_j(c_j^*, c_{-j}) \geq \sum_{(c_{-j}, t_{-j}) \in C_{-j} \times T_{-j}} \pi(c_{-j}, t_{-j} \mid c_j, t_j) \cdot u_j(c_j', c_{-j}) \tag{4.2.6}$$

for all $c_j' \in C_j$. That is, c_j^* yields the highest possible expected utility under the belief $\pi(\cdot \mid c_j, t_j)$.

Fix a choice-type pair (c_j, t_j) for some player j with an outgoing arrow, which means that $\pi(c_j, t_j) > 0$. Remember from above that all arrows in the beliefs diagram are solid, and hence, in particular, every arrow leaving (c_j, t_j) is solid. This means that the choice c_j is optimal for the first-order belief represented by the arrow(s) leaving (c_j, t_j). Since, by (4.2.5), this first-order belief is induced by the belief $\pi(\cdot \mid c_j, t_j)$ for player j conditional on (c_j, t_j), we conclude that the choice c_j must be optimal for the belief $\pi(\cdot \mid c_j, t_j)$.

We thus see that every *symmetric* belief hierarchy β_i that expresses *common belief in rationality* is induced by a common prior π on choice-type combinations that satisfies the following optimality condition: for every player j, and every choice-type pair (c_j, t_j) that receives positive probability under π, the choice c_j must be optimal for the belief $\pi(\cdot \mid c_j, t_j)$ of player j conditional on (c_j, t_j). A common prior on choice-type combinations with this special property is called a *correlated equilibrium.*

Definition 4.2.3 (Correlated equilibrium) *A common prior π on choice-type combinations is a* ***correlated equilibrium*** *if, for every player i and every choice-type pair (c_i, t_i) with $\pi(c_i, t_i) > 0$, the choice c_i is optimal for the belief $\pi(\cdot \mid c_i, t_i)$ of player i conditional on his choice-type pair (c_i, t_i).*

By the arguments above we may thus conclude that every symmetric belief hierarchy that expresses common belief in rationality is induced by a correlated equilibrium. We can show, in fact, that the other direction is also true: Every belief hierarchy that is induced by a correlated equilibrium is symmetric and expresses common belief in rationality.

To see why, consider a belief hierarchy β_i that is induced by a correlated equilibrium π. Then, in particular, the belief hierarchy β_i is induced by a common prior on choice-type combinations, and hence it follows by Theorem 4.2.1 that β_i is symmetric.

To show that β_i expresses common belief in rationality, we look at the beliefs diagram in choice-type representation induced by the correlated equilibrium π, and argue that all arrows in this beliefs diagram must be solid. Consider an arbitrary choice-type pair (c_j, t_j) that has an outgoing arrow in this beliefs diagram. Then, we must have that $\pi(c_j, t_j) > 0$, and every outgoing arrow a from (c_j, t_j) to an opponents' choice-type combination (c_{-j}, t_{-j}) has the probability

$$p(a) = \frac{\pi((c_j, t_j), (c_{-j}, t_{-j}))}{\pi(c_j, t_j)} = \pi(c_{-j}, t_{-j} \mid c_j, t_j).$$

In other words, the first-order belief about the opponents' choices represented by the arrow(s) leaving (c_j, t_j) is induced by the belief $\pi(\cdot \mid c_j, t_j)$ on the opponents' choice-type combinations.

Since the common prior on choice-type combinations π is a correlated equilibrium, we know that c_j is optimal for the belief $\pi(\cdot \mid c_j, t_j)$ of player j conditional on (c_j, t_j). Hence, the choice c_j is optimal for the first-order belief that starts at (c_j, t_j) in the beliefs diagram. As such, all arrows that leave the choice-type pair (c_j, t_j) must be solid. As this applies to all choice-type pairs (c_j, t_j) with outgoing arrows, we conclude that all arrows in the induced beliefs diagram are solid.

We thus know, by Theorem 3.3.2, that every belief hierarchy in this beliefs diagram expresses common belief in rationality. In particular, the belief hierarchy β_i above, which is part of this beliefs diagram, expresses common belief in rationality. Altogether, we see that the belief hierarchy β_i is symmetric and expresses common belief in rationality. That is, every belief hierarchy β_i that is induced by a correlated equilibrium is symmetric and expresses common belief in rationality. Since we have already shown above that the other direction is also true, we obtain the following characterization of symmetric belief hierarchies that express common belief in rationality.

Theorem 4.2.2 (Relation with correlated equilibrium) *A belief hierarchy is symmetric and expresses common belief in rationality, if and only if, the belief hierarchy is induced by a correlated equilibrium.*

To illustrate this theorem, consider the beliefs diagram on top of Figure 4.2.3, for "Rock, paper, scissors". Focus on your belief hierarchy that starts at (*rock*, t_1^r). We have already seen that this belief hierarchy is symmetric, because it is induced by the symmetric weighted beliefs diagram at the bottom of Figure 4.2.3. Moreover, we have shown above that this symmetric belief hierarchy is induced by the common prior π on choice-type combinations given by Table 4.2.2.

Note that the belief hierarchy starting at (*rock*, t_1^r) expresses common belief in rationality, since all the arrows in the beliefs diagram are solid. We therefore know by Theorem 4.2.2 that this belief hierarchy must be induced by a correlated equilibrium. In fact, we can show that the common prior

π on choice-type combinations in Table 4.2.2 is a correlated equilibrium, by checking the optimality conditions in Definition 4.2.3.

Take, for instance, the choice-type pair $(paper,\ t_1^p)$ for you, with $\pi(paper,\ t_1^p) > 0$. Conditional on this choice-type pair $(paper,\ t_1^p)$, the induced conditional belief $\pi(\cdot \mid paper,\ t_1^p)$ about Barbara's choice-type combinations is given by

$$\pi(scissors,\ t_2^s \mid paper,\ t_1^p) = \frac{\pi((paper,\ t_1^p),\ (scissors,\ t_2^s))}{\pi(paper,\ t_1^p)} = \frac{1/12}{1/12 + 1/12} = 1/2,$$

and

$$\pi(rock,\ t_2^r \mid paper,\ t_1^p) = \frac{\pi((paper,\ t_1^p),\ (rock,\ t_2^r))}{\pi(paper,\ t_1^p)} = \frac{1/12}{1/12 + 1/12} = 1/2.$$

In particular, the belief $\pi(\cdot \mid paper,\ t_1^p)$ assigns probability 1/2 to Barbara's choices *scissors* and *rock.* Since your choice *paper* is optimal for this belief, we conclude that your choice *paper* is optimal for the belief $\pi(\cdot \mid paper,\ t_1^p)$ conditional on your choice-type pair $(paper,\ t_1^p)$.

Next, consider the choice-type pair $(rock,\ t_2^r)$ for Barbara, with $\pi(rock,\ t_2^r) > 0$. Conditional on this choice-type pair $(rock,\ t_2^r)$, the induced conditional belief $\pi(\cdot \mid rock,\ t_2^r)$ for Barbara about your choice-type combinations is given by

$$\begin{aligned}
\pi(paper,\ t_1^p \mid rock,\ t_2^r) &= \frac{\pi((paper,\ t_1^p),\ (rock,\ t_2^r))}{\pi(rock,\ t_2^r)} = \frac{1/12}{1/12 + 2/12 + 2/12} = 1/5, \\
\pi(paper,\ \hat{t}_1^p \mid rock,\ t_2^r) &= \frac{\pi((paper,\ \hat{t}_1^p),\ (rock,\ t_2^r))}{\pi(rock,\ t_2^r)} = \frac{2/12}{1/12 + 2/12 + 2/12} = 2/5
\end{aligned}$$

and

$$\pi(scissors,\ t_1^s \mid rock,\ t_2^r) = \frac{\pi((scissors,\ t_1^s),\ (rock,\ t_2^r))}{\pi(rock,\ t_2^r)} = \frac{2/12}{1/12 + 2/12 + 2/12} = 2/5.$$

In particular, Barbara's conditional belief $\pi(\cdot \mid rock,\ t_2^r)$ assigns probability $1/5 + 2/5 = 3/5$ to your choice *paper* and probability 2/5 to your choice *scissors.* Since Barbara's choice *rock* is optimal for this belief, we conclude that Barbara's choice *rock* is optimal for her conditional belief $\pi(\cdot \mid rock,\ t_2^r)$.

In a similar fashion it may be verified that for you, the choice *rock* is optimal for the belief $\pi(\cdot \mid rock,\ t_1^r)$, the choice *paper* is optimal for the belief $\pi(\cdot \mid paper,\ \hat{t}_1^p)$ and the choice *scissors* is optimal for the belief $\pi(\cdot \mid scissors,\ t_1^s)$. Also, it may be checked that for Barbara, the choice *scissors* is optimal for her belief $\pi(\cdot \mid scissors,\ t_2^s)$ and that the choice *paper* is optimal for her belief $\pi(\cdot \mid paper,\ t_2^p)$. Please verify this.

Summarizing, we conclude that the common prior π on choice-type combinations satisfies all optimality conditions, and is therefore a correlated equilibrium. Hence, your belief hierarchy that starts at your choice *rock* – and in fact *every* belief hierarchy from Figure 4.2.3 – is induced by the correlated equilibrium π above.

Question 4.2.3 *Consider the beliefs diagram in Figure 4.2.2 for "When Chris joins the party", and focus on your belief hierarchy that starts at your choice green. Is this belief hierarchy induced by a correlated equilibrium? If so, find a correlated equilibrium that induces it, and explain why it is a correlated equilibrium by checking the optimality conditions.*

On the basis of Theorem 4.2.2 we can now easily characterize those choices that you can rationally make under common belief in rationality with a symmetric belief hierarchy. Suppose that player i

holds a belief hierarchy β_i that is symmetric and expresses common belief in rationality, and assume that the choice c_i^* is optimal for the first-order belief b_i^1 in that belief hierarchy. By Theorem 4.2.2 we know that the belief hierarchy β_i is induced by a correlated equilibrium π. That is, β_i is part of a beliefs diagram in choice-type representation that is induced by the correlated equilibrium π on choice-type combinations. Let β_i be the belief hierarchy that starts at the choice-type pair (c_i, t_i). Then, $\pi(c_i, t_i) > 0$ and the first-order belief b_i^1 of β_i is induced by $\pi(\cdot \mid c_i, t_i)$. As the choice c_i^* is optimal for the first-order belief b_i^1, it immediately follows that c_i^* is optimal for the belief $\pi(\cdot \mid c_i, t_i)$ of player i conditional on (c_i, t_i). In this case, we say that c_i^* is *optimal in the correlated equilibrium* π.

Definition 4.2.4 (Choice optimal in a correlated equilibrium) *A choice c_i^* is **optimal in a correlated equilibrium** π if there is some choice-type pair (c_i, t_i) with $\pi(c_i, t_i) > 0$ such that c_i^* is optimal for the belief $\pi(\cdot \mid c_i, t_i)$ of player i conditional on (c_i, t_i).*

We have thus shown that every choice c_i^* that can rationally be made under common belief in rationality with a symmetric belief hierarchy must be optimal in a correlated equilibrium.

The other direction is also true: Every choice that is optimal in a correlated equilibrium can rationally be made under common belief in rationality with a symmetric belief hierarchy. To see this, consider a choice c_i^* that is optimal in a correlated equilibrium π. Then, there must be some choice-type pair (c_i, t_i) with $\pi(c_i, t_i) > 0$ such that c_i^* is optimal for the belief $\pi(\cdot \mid c_i, t_i)$. Consider the beliefs diagram induced by the correlated equilibrium π, and the belief hierarchy β_i in this beliefs diagram that starts at the choice-type pair (c_i, t_i). By Theorem 4.2.2 we know that the belief hierarchy β_i is symmetric and expresses common belief in rationality. Moreover, the first-order belief in the belief hierarchy β_i is induced by $\pi(\cdot \mid c_i, t_i)$ because β_i is the belief hierarchy that starts at (c_i, t_i). Since c_i^* is optimal for the belief $\pi(\cdot \mid c_i, t_i)$, it is also optimal for the first-order belief of the belief hierarchy β_i that starts at (c_i, t_i). Hence, c_i^* is optimal for the symmetric belief hierarchy β_i that expresses common belief in rationality.

Summarizing, we see that a choice c_i^* can rationally be made under common belief in rationality with a symmetric belief hierarchy, precisely when it is optimal in a correlated equilibrium. We thus obtain the following characterization result.

Theorem 4.2.3 (Relation with correlated equilibrium choices) *A choice is optimal for a symmetric belief hierarchy that expresses common belief in rationality, if and only if, the choice is optimal in a correlated equilibrium.*

In the next subsection we will use this result to find, in three of the examples we have seen, those choices you can rationally make under common belief in rationality with a symmetric belief hierarchy.

4.2.4 Examples

In Theorem 4.2.3, we have seen that the choices that can rationally be made under common belief in rationality with a symmetric belief hierarchy are precisely those that are optimal in some correlated equilibrium. This result, besides having an important conceptual value, can also be of great practical use to find, in a given game, those choices that are possible under common belief in rationality with a symmetric belief hierarchy. This will become clear below, where we apply Theorem 4.2.3 to the three examples we have seen so far in this chapter.

Example 4.6: Movie for two.

Recall the story from Section 4.1.1, and the decision problems from Table 4.1.1. Which choices can you rationally make here under common belief in rationality with a symmetric belief hierarchy? By looking at the beliefs diagram from Figure 4.1.1 we can easily see that under common belief in rationality with a symmetric belief hierarchy, you can rationally choose to stay at *home.* Indeed, consider your belief hierarchy that starts at your choice *home.* This belief hierarchy is clearly symmetric.

Question 4.2.4 *Find a common prior on choice-type combinations that induces your symmetric belief hierarchy that starts at your choice* home.

Moreover, this belief hierarchy expresses common belief in rationality because both arrows that constitute this belief hierarchy are solid. As your choice *home* is optimal under this belief hierarchy, you can rationally choose *home* under common belief in rationality with a symmetric belief hierarchy.

But what about your other two choices, to go to *Cinema Palace* or *The Movie Corner*? Can you rationally make these choices under common belief in rationality with a symmetric belief hierarchy? The beliefs diagram in Figure 4.1.1 does not answer this question, since the belief hierarchies that support your choices *Palace* and *Corner* are not symmetric.

Question 4.2.5 *Explain why these two belief hierarchies are not symmetric.*

But perhaps there is another beliefs diagram in which your choices *Palace* and *Corner* are supported by symmetric belief hierarchies that express common belief in rationality.

We will see, however, that this is not possible. To show this, we use Theorem 4.2.3 and prove that every correlated equilibrium must assign probability 1 to your choice *home* and Barbara's choice *home.* However, showing this is rather tedious and lengthy. In particular, at every correlated equilibrium you will always assign probability 1 to Barbara's choice *home,* and hence the only choice that is optimal for you in a correlated equilibrium is *home.* By Theorem 4.2.3 it would then follow that under common belief in rationality with a symmetric belief hierarchy, you can only rationally choose to stay at *home.*

Consider a correlated equilibrium π, which is a common prior on the set of choice-type combinations $C \times T$. We will show that π assigns probability 1 to your choice *home,* and probability 1 to Barbara's choice *home.* Along the way, we will use the following abbreviations:

$$\begin{aligned}
\pi(c_1,(c_2,t_2)) \;:&= \sum_{t_1\in T_1} \pi((c_1,t_1),(c_2,t_2)), \quad \pi((c_1,t_1),c_2) := \sum_{t_2 \in T_2} \pi((c_1,t_1),(c_2,t_2)), \\
\pi(c_1,c_2) \;:&= \sum_{t_1\in T_1}\sum_{t_2\in T_2} \pi((c_1,t_1),(c_2,t_2)), \\
\pi(c_1) \;:&= \sum_{c_2\in C_2} \pi(c_1,c_2) \text{ and } \pi(c_2) := \sum_{c_1\in C_1} \pi(c_1,c_2).
\end{aligned}$$

Assume, contrary to what we want to show, that $\pi(Pal_1, t_1) > 0$ for some type $t_1 \in T_1$. Here, Pal_1 stands for *Palace,* and we use the subscript 1 in Pal_1 to indicate that this choice belongs to player 1 (you). Then, by definition of a correlated equilibrium, *Palace* must be optimal for you for your belief $\pi(\cdot \mid Pal_1, t_1)$. In particular, your choice *Palace* must be at least as good as your choice *home* under that belief $\pi(\cdot \mid Pal_1, t_1)$, which means that

$$u_1(Pal_1,\ \pi(\cdot \mid Pal_1,t_1)) \geq u_1(home_1,\ \pi(\cdot \mid Pal_1,t_1)).$$

Note that

$$\begin{aligned} u_1(Pal_1,\ \pi(\cdot \mid Pal_1, t_1)) &= \sum_{(c_2,t_2)\in C_2\times T_2} \frac{\pi((Pal_1,t_1),(c_2,t_2))}{\pi(Pal_1,t_1)} \cdot u_1(Pal_1, c_2) \\ &= \frac{\pi((Pal_1,t_1),Pal_2)}{\pi(Pal_1,t_1)} \cdot 4. \end{aligned}$$

Moreover,

$$u_1(home_1,\ \pi(\cdot \mid Pal_1, t_1)) = 3.$$

As $u_1(Pal_1,\ \pi(\cdot \mid Pal_1, t_1)) \geq u_1(home,\ \pi(\cdot \mid Pal_1, t_1))$, we must have that

$$\frac{\pi((Pal_1,t_1),Pal_2)}{\pi(Pal_1,t_1)} \cdot 4 \geq 3$$

and hence

$$\pi((Pal_1,t_1),Pal_2) \geq 3/4 \cdot \pi(Pal_1,t_1).$$

Since this holds for every type $t_1 \in T_1$ with $\pi(Pal_1, t_1) > 0$, it follows that

$$\sum_{t_1\in T_1} \pi((Pal_1,t_1),Pal_2) \geq 3/4 \cdot \sum_{t_1\in T_1} \pi(Pal_1,t_1),$$

and therefore

$$\pi(Pal_1,\ Pal_2) \geq 3/4 \cdot \pi(Pal_1). \tag{4.2.7}$$

Since $\pi(Pal_1, t_1) > 0$ for some type t_1, this implies that $\pi(Pal_1, Pal_2) > 0$ and hence $\pi(Pal_2, t_2) > 0$ for some type $t_2 \in T_2$. Therefore, *Palace* must be optimal for Barbara for her belief $\pi(\cdot \mid Pal_2, t_2)$. In particular, *Palace* must be at least as good as *home* for Barbara under that belief. That is,

$$u_2(Pal_2,\ \pi(\cdot \mid Pal_2,t_2)) \geq u_2(home_2,\ \pi(\cdot \mid Pal_2,t_2)).$$

Since

$$u_2(Pal_2,\ \pi(\cdot \mid Pal_2,t_2)) = \frac{\pi(Cor_1,\ (Pal_2,t_2))}{\pi(Pal_2,t_2)} \cdot 4 \text{ and}$$

$$u_2(home_2,\ \pi(\cdot \mid Pal_2,t_2)) = 3$$

it follows that

$$\pi(Cor_1,\ (Pal_2,t_2)) \geq 3/4 \cdot \pi(Pal_2,t_2).$$

As this holds for every $t_2 \in T_2$ with $\pi(Pal_2, t_2) > 0$, we conclude that

$$\sum_{t_2\in T_2} \pi(Cor_1,\ (Pal_2,t_2)) \geq 3/4 \cdot \sum_{t_2\in T_2} \pi(Pal_2,t_2)$$

and hence

$$\pi(Cor_1,\ Pal_2) \geq 3/4 \cdot \pi(Pal_2). \tag{4.2.8}$$

Since $\pi(Pal_2, t_2) > 0$ for some type t_2 we have that $\pi(Cor_1, Pal_2) > 0$, and therefore $\pi(Cor_1, t_1) > 0$ for some type $t_1 \in T_1$. Hence, *Corner* must be optimal for you under your belief $\pi(\cdot \mid Cor_1, t_1)$, which implies that

$$u_1(Cor_1,\ \pi(\cdot \mid Cor_1,t_1)) \geq u_1(home_1,\ \pi(\cdot \mid Cor_1,t_1)).$$

As

$$u_1(Cor_1,\ \pi(\cdot \mid Cor_1, t_1)) = \frac{\pi((Cor_1, t_1), Cor_2)}{\pi(Cor_1, t_1)} \cdot 4 \text{ and}$$
$$u_1(home_1,\ \pi(\cdot \mid Cor_1, t_1)) = 3$$

we must have that

$$\pi((Cor_1, t_1), Cor_2) \geq 3/4 \cdot \pi(Cor_1, t_1).$$

As this holds for every $t_1 \in T_1$ with $\pi(Cor_1, t_1) > 0$, we have that

$$\sum_{t_1 \in T_1} \pi((Cor_1, t_1), Cor_2) \geq 3/4 \cdot \sum_{t_1 \in T_1} \pi(Cor_1, t_1)$$

and hence

$$\pi(Cor_1, Cor_2) \geq 3/4 \cdot \pi(Cor_1). \tag{4.2.9}$$

As $\pi(Cor_1, t_1) > 0$ for some type t_1, we conclude that $\pi(Cor_1, Cor_2) > 0$. Thus, $\pi(Cor_2, t_2) > 0$ for some $t_2 \in T_2$. Hence, *Corner* must be for Barbara under her belief $\pi(\cdot \mid Cor_2, t_2)$. In particular,

$$u_2(Cor_2,\ \pi(\cdot \mid Cor_2, t_2)) \geq u_2(home_2,\ \pi(\cdot \mid Cor_2, t_2)).$$

As

$$u_2(Cor_2,\ \pi(\cdot \mid Cor_2, t_2)) = \frac{\pi(Pal_1,\ (Cor_2, t_2))}{\pi(Cor_2, t_2)} \cdot 4 \text{ and}$$
$$u_2(home_2,\ \pi(\cdot \mid Cor_2, t_2)) = 3$$

it follows that

$$\pi(Pal_1,\ (Cor_2, t_2)) \geq 3/4 \cdot \pi(Cor_2, t_2).$$

Since this holds for every type t_2 with $\pi(Cor_2, t_2) > 0$, we conclude that

$$\sum_{t_2 \in T_2} \pi(Pal_1,\ (Cor_2, t_2)) \geq 3/4 \cdot \sum_{t_2 \in T_2} \pi(Cor_2, t_2)$$

and hence

$$\pi(Pal_1,\ Cor_2) \geq 3/4 \cdot \pi(Cor_2). \tag{4.2.10}$$

By combining (4.2.7), (4.2.8), (4.2.9) and (4.2.10) we get that

$$\begin{gathered}\pi(Pal_1, Pal_2) + \pi(Cor_1,\ Pal_2) + \pi(Cor_1, Cor_2) + \pi(Pal_1,\ Cor_2) \\ \geq 3/4 \cdot (\pi(Pal_1) + \pi(Pal_2) + \pi(Cor_1) + \pi(Cor_2)).\end{gathered}$$

As

$$\begin{aligned}\pi(Pal_1) &\geq \pi(Pal_1, Pal_2) + \pi(Pal_1,\ Cor_2), \\ \pi(Pal_2) &\geq \pi(Pal_1, Pal_2) + \pi(Cor_1,\ Pal_2), \\ \pi(Cor_1) &\geq \pi(Cor_1,\ Pal_2) + \pi(Cor_1, Cor_2) \text{ and} \\ \pi(Cor_2) &\geq \pi(Cor_1, Cor_2) + \pi(Pal_1,\ Cor_2)\end{aligned}$$

it follows that

$$\pi(Pal_1, Pal_2) + \pi(Cor_1,\ Pal_2) + \pi(Cor_1, Cor_2) + \pi(Pal_1,\ Cor_2)$$
$$\geq 3/4 \cdot 2 \cdot (\pi(Pal_1, Pal_2) + \pi(Cor_1,\ Pal_2) + \pi(Cor_1, Cor_2) + \pi(Pal_1,\ Cor_2)).$$

This, however, is only possible when

$$\pi(Pal_1, Pal_2) + \pi(Cor_1,\ Pal_2) + \pi(Cor_1, Cor_2) + \pi(Pal_1,\ Cor_2) = 0.$$

In particular, $\pi(Pal_1, Pal_2) = 0$, which contradicts our conclusion above that $\pi(Pal_1, Pal_2) > 0$.

Therefore, our assumption that $\pi(Pal_1, t_1) > 0$ for some type $t_1 \in T_1$ must be wrong. Hence, every correlated equilibrium π must have the property that $\pi(Pal_1, t_1) = 0$ for all types $t_1 \in T_1$, and hence $\pi(Pal_1) = 0$.

Now suppose that $\pi(Cor_2, t_2) > 0$ for some type t_2. Then, Barbara's belief $\pi(\cdot \mid Cor_2, t_2)$ must assign probability 0 to Pal_1 since $\pi(Pal_1) = 0$. This implies that Barbara's choice Cor_2 cannot be optimal for her belief $\pi(\cdot \mid Cor_2, t_2)$, which contradicts the assumption that π is a correlated equilibrium. Hence, the assumption $\pi(Cor_2, t_2) > 0$ cannot be true. Therefore, every correlated equilibrum π must satisfy $\pi(Cor_2) = 0$.

Suppose that $\pi(Cor_1, t_1) > 0$ for some type t_1. Since $\pi(Cor_2) = 0$, your belief $\pi(\cdot \mid Cor_1, t_1)$ must assign probability 0 to Barbara's choice Cor_2. But then, your choice Cor_1 cannot be optimal for your belief $\pi(\cdot \mid Cor_1, t_1)$. This contradicts the assumption that π is a correlated equilibrium. Hence, the assumption that $\pi(Cor_1, t_1) > 0$ cannot be true. Therefore, every correlated equilibrum π must satisfy $\pi(Cor_1) = 0$.

Suppose, finally, that $\pi(Pal_2, t_2) > 0$ for some t_2. Since $\pi(Pal_1) = 0$ and $\pi(Cor_1) = 0$, Barbara's belief $\pi(\cdot \mid Pal_2, t_2)$ must assign probability 1 to your choice $home_1$. But then, her choice Pal_2 cannot be optimal for her belief $\pi(\cdot \mid Pal_2, t_2)$. This contradicts the assumption that π is a correlated equilibrium. Hence, the assumption that $\pi(Pal_2, t_2) > 0$ cannot be true. Therefore, every correlated equilibrum π must satisfy $\pi(Pal_2) = 0$.

Summarizing, we see that every correlated equilibrium π must satisfy $\pi(Pal_1) = \pi(Cor_1) = 0$ and $\pi(Pal_2) = \pi(Cor_2) = 0$. Therefore, every correlated equilibrium has the property that $\pi(home_1, home_2) = 1$. In every such correlated equilibrium π, your conditional belief $\pi(\cdot \mid home,\ t_1)$ will always assign probability 1 to Barbara's choice *home,* and hence the only choice that is optimal for you in a correlated equilibrium is *home.*

Therefore, by Theorem 4.2.3, you can only rationally stay at *home* under common belief in rationality with a symmetric belief hierarchy. The intuitive reason, however, is not as easy as it was for common belief in rationality with a *simple* belief hierarchy, which also uniquely led to your choice to stay at *home.* But here is a possible intuition.

Suppose there would be a symmetric belief hierarchy for you that expresses common belief in rationality, and for which choosing *Palace* is optimal. Then, in your first-order belief you must assign a high probability to Barbara choosing *Palace,* and hence in the associated symmetric weighted beliefs diagram, the arrow from your choice *Palace* to Barbara's choice *Palace* must carry a weight a that is higher than the weights of the (possible) other arrows leaving your choice *Palace.* As the weighted beliefs diagram is symmetric, there must also be an arrow from Barbara's choice *Palace* to your choice *Palace,* with the same weight a. To make the choice *Palace* optimal for Barbara, we must assign an even higher weight b to the arrow from Barbara's choice *Palace* to your choice *Corner.* By symmetry, the arrow from your choice *Corner* to Barbara's choice *Palace* must carry the same weight b. To make your choice *Corner* optimal, we must assign an even higher weight c to the arrow from your choice

Corner to Barbara's choice *Corner.* Again, by symmetry, there must be an arrow from Barbara's choice *Corner* to your choice *Corner* with the same weight c. To make the choice *Corner* optimal for Barbara, we must assign an even higher weight d to the arrow from Barbara's choice *Corner* to your choice *Palace.* But then, by symmetry, we must assign the same, very high, weight d to the arrow from your choice *Palace* to Barbara's choice *Corner.* Consequently, this arrow from your choice *Palace* to Barbara's choice *Corner* would have a higher weight, d, than the arrow from your choice *Palace* to Barbara's choice *Palace,* which only has a weight of a. However, this means that in your first-order belief starting at your choice *Palace,* you assign a higher probability to Barbara choosing *Corner* than to Barbara choosing *Palace,* and thus your choice *Palace* cannot be optimal for the belief hierarchy that starts at your choice *Palace.* This cannot be. Therefore, there is no symmetric belief hierarchy expressing common belief in rationality for which it is optimal to choose *Palace.* A similar reasoning can be used to exclude your choice *Corner.*

Example 4.7: Opera for three.

Remember the story from Section 4.1.3, and the decision problems from Table 4.1.3. We again ask: Which choices can you rationally make under common belief in rationality with a symmetric belief hierarchy? From the beliefs diagram in Figure 4.1.2 we can immediately conclude that you can rationally stay at *home* under common belief in rationality with a symmetric belief hierarchy. The reason is that your belief hierarchy that starts at your choice *home* is symmetric.

Question 4.2.6 *Find a common prior on choice-type combinations that induces your symmetric belief hierarchy that starts at your choice home.*

This symmetric belief hierarchy also expresses common belief in rationality, because all arrows involved in this belief hierarchy are solid. As your choice *home* is optimal under this belief hierarchy, we conclude that you can rationally stay at *home* under common belief in rationality with a symmetric belief hierarchy.

What about your choice *opera*? Can you rationally go to the opera under common belief in rationality with a symmetric belief hierarchy? Again, the beliefs diagram does not answer this question, because the belief hierarchy in Figure 4.1.2 that supports your choice *opera* is not symmetric.

Question 4.2.7 *Explain why this belief hierarchy is not symmetric.*

But perhaps there is another beliefs diagram in which your choice *opera* is supported by a symmetric belief hierarchy that expresses common belief in rationality. We will show, however, that this is not possible. To prove this, we use Theorem 4.2.3 and show that every correlated equilibrium must assign probability 1 to the choice *home* for you, Barbara and Chris. Again, the proof is rather involved. Hence, in every correlated equilibrium you will always believe, with probability 1, that Barbara and Chris will stay at *home*, and hence the only optimal choice for you in a correlated equilibrium is to stay at *home* as well.

Consider an arbitrary correlated equilibrium π, which is a common prior on choice-type combinations. Assume, contrary to what we want to show, that $\pi(opera_1, t_1) > 0$ for some type $t_1 \in T_1$. Since π is a correlated equilibrium, your choice *opera* must be optimal for the conditional belief $\pi(\cdot \mid opera_1, t_1)$. That is, $u_1(opera_1,\ \pi(\cdot \mid opera_1, t_1)) \geq u_1(home_1,\ \pi(\cdot \mid opera_1, t_1))$. We have that

$$u_1(opera_1, \pi(\cdot \mid opera_1, t_1)) = \frac{\pi((opera_1, t_1),\ opera_2,\ opera_3)}{\pi(opera_1, t_1)} \cdot 4$$

where

$$\pi((opera_1, t_1),\ opera_2,\ opera_3) = \sum_{t_2 \in T_2} \sum_{t_3 \in T_3} \pi((opera_1, t_1),\ (opera_2, t_2),\ (opera_3, t_3)).$$

Moreover,

$$u_1(home_1,\ \pi(\cdot | opera_1, t_1)) = 3.$$

As $u_1(opera_1,\ \pi(\cdot \mid opera_1, t_1)) \geq u_1(home_1,\ \pi(\cdot \mid opera_1, t_1))$, it follows that

$$\frac{\pi((opera_1, t_1),\ opera_2,\ opera_3)}{\pi(opera_1, t_1)} \cdot 4 \geq 3,$$

and hence

$$\pi((opera_1, t_1),\ opera_2,\ opera_3) \geq 3/4 \cdot \pi(opera_1, t_1).$$

Since this holds for all types t_1 with $\pi(opera_1, t_1) > 0$, it follows that

$$\sum_{t_1 \in T_1} \pi((opera_1, t_1),\ opera_2,\ opera_3) \geq 3/4 \cdot \sum_{t_1 \in T_1} \pi(opera_1, t_1),$$

which means that

$$\pi(opera_1,\ opera_2,\ opera_3) \geq 3/4 \cdot \pi(opera_1). \tag{4.2.11}$$

As $\pi(opera_1) > 0$, it follows that $\pi(opera_1,\ opera_2,\ opera_3) > 0$, and hence $\pi(opera_2) > 0$ and $\pi(opera_3) > 0$.

Since $\pi(opera_2) > 0$, there is some type t_2 with $\pi(opera_2, t_2) > 0$. As π is a correlated equilibrium, the choice *opera* must be optimal for Barbara under the belief $\pi(\cdot \mid opera_2, t_2)$. That is, $u_2(opera_2,\ \pi(\cdot \mid opera_2, t_2)) \geq u_2(home_2,\ \pi(\cdot \mid opera_2, t_2)) = 3$. Since

$$u_2(opera_2, \pi(\cdot \mid opera_2, t_2)) = \frac{\pi(opera_1,\ (opera_2, t_2),\ home_3)}{\pi(opera_2, t_2)} \cdot 4$$

it follows that

$$\frac{\pi(opera_1,\ (opera_2, t_2),\ home_3)}{\pi(opera_2, t_2)} \cdot 4 \geq 3,$$

and hence

$$\pi(opera_1,\ (opera_2, t_2),\ home_3) \geq 3/4 \cdot \pi(opera_2, t_2).$$

Since this holds for all types t_2 with $\pi(opera_2, t_2) > 0$, it follows that

$$\sum_{t_2 \in T_2} \pi(opera_1,\ (opera_2, t_2),\ home_3) \geq 3/4 \cdot \sum_{t_2 \in T_2} \pi(opera_2, t_2),$$

and therefore

$$\pi(opera_1,\ opera_2,\ home_3) \geq 3/4 \cdot \pi(opera_2). \tag{4.2.12}$$

Also, since $\pi(opera_3) > 0$, there is some type t_3 with $\pi(opera_3, t_3) > 0$. As π is a correlated equilibrium, the choice *opera* must be optimal for Chris under the belief $\pi(\cdot \mid opera_3, t_3)$. That is, $u_3(opera_3,\ \pi(\cdot \mid opera_3, t_3)) \geq u_3(home_3,\ \pi(\cdot \mid opera_3, t_3)) = 3$. Since

$$u_3(opera_3, \pi(\cdot \mid opera_3, t_3)) = \frac{\pi(opera_1,\ home_2,\ (opera_3, t_3))}{\pi(opera_3, t_3)} \cdot 4$$

it follows that
$$\frac{\pi(opera_1,\ home_2,\ (opera_3, t_3))}{\pi(opera_3, t_3)} \cdot 4 \geq 3,$$
and hence
$$\pi(opera_1,\ home_2,\ (opera_3, t_3)) \geq 3/4 \cdot \pi(opera_3, t_3).$$
Since this holds for all types t_3 with $\pi(opera_3, t_3) > 0$, it follows that
$$\sum_{t_3 \in T_3} \pi(opera_1,\ home_2,\ (opera_3, t_3)) \geq 3/4 \cdot \sum_{t_3 \in T_3} \pi(opera_3, t_3),$$
and therefore
$$\pi(opera_1,\ home_2,\ opera_3) \geq 3/4 \cdot \pi(opera_3). \tag{4.2.13}$$

Note that
$$\begin{aligned}
\pi(opera_1) &\geq \pi(opera_1,\ opera_2,\ home_3) + \pi(opera_1,\ home_2,\ opera_3) \\
&\geq 3/4 \cdot \pi(opera_2) + 3/4 \cdot \pi(opera_3) \\
&\geq 3/4 \cdot \pi(opera_1,\ opera_2,\ opera_3) + 3/4 \cdot \pi(opera_1,\ opera_2,\ opera_3) \\
&= 3/2 \cdot \pi(opera_1,\ opera_2,\ opera_3) \\
&\geq 3/2 \cdot 3/4 \cdot \pi(opera_1) \\
&= 9/8 \cdot \pi(opera_1).
\end{aligned}$$

Here, the first inequality is true by definition, as
$$\pi(opera_1) = \sum_{c_2 \in C_2} \sum_{c_3 \in C_3} \pi(opera_1,\ c_2,\ c_3).$$

Moreover, the second inequality follows from (4.2.12) and (4.2.13), the third inequality is true by definition, whereas the fourth inequality follows from (4.2.11). Since $\pi(opera_1) \geq 9/8 \cdot \pi(opera_1)$, we conclude that $\pi(opera_1) = 0$, which contradicts our assumption that $\pi(opera_1, t_1) > 0$ for some type $t_1 \in T_1$.

Hence, we conclude that in every correlated equilibrium we must have that $\pi(opera_1) = 0$, and hence $\pi(home_1) = 1$. Therefore, in every correlated equilibrium, the conditional belief $\pi(\cdot \mid c_2, t_2)$ for Barbara must always assign probability 1 to you staying at *home*. Hence, only *home* can be optimal for Barbara for every conditional belief $\pi(\cdot \mid c_2, t_2)$. Since π is a correlated equilibrium, π can only assign positive probability to choice-type pairs (c_2, t_2) for Barbara where c_2 is optimal for $\pi(\cdot \mid c_2, t_2)$. But then, π can only assign positive probability to choice-type pairs $(home_2, t_2)$ for Barbara, and hence $\pi(home_2) = 1$.

Question 4.2.8 *Show, in a similar fashion, that every correlated equilibrium π must assign probability 1 to Chris' choice home.*

Since every correlated equilibrium π must satisfy $\pi(home_2) = 1$ and $\pi(home_3) = 1$, your conditional belief $\pi(\cdot \mid c_1, t_1)$ in a correlated equilibrium must always assign probability 1 to Barbara and Chris staying at *home*. Hence, the only choice that is optimal for you in a correlated equilibrium is to stay at *home* as well. By Theorem 4.2.3 we thus conclude that under common belief in rationality with a symmetric belief hierarchy, you can only rationally choose to stay at *home*. Recall that under common belief in rationality with a *simple* belief hierarchy, you can also only rationally stay at *home*.

The intuitive reason is as follows. Suppose there is a symmetric belief hierarchy for you that expresses common belief in rationality, and for which it is optimal to go to the *opera.* Then, in your first-order belief, you must assign a high probability to Barbara and Chris going to the *opera* as well. Therefore, in the associated symmetric weighted beliefs diagram, the weight a of the arrow from your choice *opera* to Barbara's and Chris' choice pair (*opera*, *opera*) must be higher than the weights of the (possible) other outgoing arrows at your choice *opera.* By symmetry, the arrow from Barbara's choice *opera* to your and Chris' choice pair (*opera*, *opera*) must carry the same weight a. To make the choice *opera* optimal for Barbara, the arrow from Barbara's choice *opera* to your and Chris' choice pair (*opera, home*) must carry an even higher weight b. But then, by symmetry, the arrow from your choice *opera* to Barbara's and Chris' choice pair (*opera, home*) must also be present, and must carry the same weight b, higher than a. However, this means that the arrow from your choice *opera* to Barbara's and Chris' choice pair (*opera, home*) has a higher weight, b, than the arrow from your choice *opera* to Barbara's and Chris' choice pair (*opera*, *opera*), which only carries a weight of a. Hence, in your first-order belief that starts at your choice *opera,* you deem it most likely that at least one of your friends stays at *home*. But then, it cannot be optimal to go to the *opera* under that first-order belief. This cannot be. Hence, we conclude that there is no symmetric belief hierarchy expressing common belief in rationality for which it is optimal to go to the *opera.*

Example 4.8: Rock, paper scissors.

Recall the story from Section 4.2.1 and the decision problems from Table 4.2.1. Which choices can you rationally make under common belief in rationality with a symmetric belief hierarchy? One way of answering this question is to focus on the common prior π on choice-combinations given by Table 4.2.2. We have seen in Section 4.2.3 that π is a correlated equilibrium.

Moreover, every choice for you is optimal in this correlated equilibrium. Indeed, your choice *rock* is optimal for the conditional belief $\pi(\cdot \mid \textit{rock}, t_1^r)$ which assigns probability 1/3 to Barbara's choice *scissors* and probability 2/3 to Barbara's choice *paper,* your choice *paper* is optimal for the conditional belief $\pi(\cdot \mid \textit{paper}, t_1^p)$ which assigns probability 1/2 to Barbara choosing *scissors* and probability 1/2 to Barbara choosing *rock,* your choice *scissors* is optimal for the conditional belief $\pi(\cdot \mid \textit{scissors}, t_1^s)$ which assigns probability 2/5 to Barbara's choice *rock* and probability 3/5 to Barbara's choice *paper,* whereas your choice *bomb* is optimal (together with your choice *paper*) for the conditional belief $\pi(\cdot \mid \textit{paper}, \hat{t}_1^p)$ that assigns probability 1 to Barbara choosing *rock.* Hence, we conclude by Theorem 4.2.3 that under common belief in rationality with a symmetric belief hierarchy, you can rationally choose *rock, paper, scissors* or *bomb.*

Another way to see this is by looking at the beliefs diagram at the top of Figure 4.2.1. We have seen that this beliefs diagram is induced by the symmetric weighted beliefs diagram at the bottom of Figure 4.2.1. Hence, every belief hierarchy present in this beliefs diagram is symmetric. Moreover, every belief hierarchy in this beliefs diagram also expresses common belief in rationality, because all arrows in the beliefs diagram are solid. As such, all belief hierarchies in this beliefs diagram express common belief in rationality and are symmetric. Note that your choice *rock* is optimal for the belief hierarchy that starts at your choice *rock,* your choice *paper* is optimal for the belief hierarchy that starts at your choice *paper,* your choice *bomb* is optimal for the belief hierarchy that starts at your choice *paper',* and your choice *scissors* is optimal for the belief hierarchy that starts at your choice *scissors.* Hence, you can rationally make each of your choices under common belief in rationality with a symmetric belief hierarchy.

4.2.5 How Reasonable is Correlated Equilibrium?

In Section 4.1.5 we have seen that the concept of Nash equilibrium imposes the following three problematic conditions: (a) player i believes that every opponent is correct about the beliefs player i has, (b) player i believes that opponent j has the same belief about a third player k's choice as player i has himself, and (c) player i's belief about opponent j's choice is independent from player i's belief about a third player k's choice. Here, the conditions (b) and (c) only apply to games with more than two players.

It turns out the concept of a symmetric belief hierarchy – and hence also the notion of correlated equilibrium – displays a weak version of the problematic properties (a) and (b). To see this, consider a symmetric belief hierarchy β_i for player i, induced by a symmetric weighted beliefs diagram. Suppose that β_i is obtained if we start at the choice-type combination (c_i, t_i). Now assume that the belief hierarchy β_i assigns positive probability to j's belief hierarchy β_j which starts at (c_j, t_j). Then, in the symmetric weighted beliefs diagram, there must be an arrow from (c_i, t_i) to an opponents' choice-type combination including (c_j, t_j). Since the weighted beliefs diagram is symmetric, there must also be an arrow from (c_j, t_j) to an opponents' choice-type combination including (c_i, t_i). In other words, player j's belief hierarchy β_j must assign positive probability to player i's actual belief hierarchy β_i.

We thus see that, whenever a symmetric belief hierarchy β_i assigns positive probability to an opponent's belief hierarchy β_j, then β_j must assign positive probability to i's actual belief hierarchy β_i. That is, with a symmetric belief hierarchy a player must believe that all opponents assign a *positive probability* to his actual belief hierarchy. This may be seen as a weak version of the correct beliefs condition (a) above. Indeed, condition (a) above states that a player with a *simple* belief hierarchy must believe that all opponents assign *full probability* to his actual belief hierarchy. But still, this weaker correct beliefs condition implied by a symmetric belief hierarchy seems problematic. Indeed, why must a player believe that his opponents assign a positive probability to his actual belief hierarchy? Why can a player not believe that his opponents are entirely wrong about his belief hierarchy? If there are several plausible belief hierarchies for a player, then in my opinion this player should be free to believe that his opponents assign probability 0 to his actual belief hierarchy.

As to property (b), suppose there are at least three players. Assume that the symmetric belief hierarchy β_i above assigns positive probability to k's choice c_k. Since β_i starts at the choice-type pair (c_i, t_i), there must be an arrow from (c_i, t_i) to an opponents' choice-type combination including c_k. Suppose that this opponents' choice-type combination includes the choice-type pair (c_j, t_j) for a third player j. Let β_j be the belief hierarchy for player j that starts at (c_j, t_j). Since the weighted beliefs diagram is symmetric, there must be an arrow from (c_j, t_j) to player k's choice c_k. That is, player j's belief hierarchy β_j must assign a positive probability to k's choice c_k.

Altogether, we see that with a symmetric belief hierarchy the following holds: If player i assigns positive probability to an opponent's choice c_k, then player i must assign positive probability to the event that every other player also assigns positive probability to c_k. This may be seen as a weak version of property (b) above. Still, I find this condition problematic as it excludes belief hierarchies where player i believes that player k chooses a, while believing that another player j believes that player k chooses b. Here, by "believe" we mean "assign probability 1 to". If both a and b are reasonable choices for player k, then such a belief hierarchy may be perfectly plausible, yet it is excluded by a symmetric belief hierarchy, and hence by correlated equilibrium as well.

Note that the concept of correlated equilibrium does not display a property similar to condition (c) above, since the concept of correlated equilibrium allows a player to have *correlated* beliefs about the choices of two different opponents. In that sense, correlated equilibrium is crucially different from Nash equilibrium.

4.3 One Theory per Choice

In the previous section we allowed for symmetric belief hierarchies in which the same choice is being explaining by two different beliefs. We will now restrict to belief hierarchies where every choice is explained by a *single* belief only. This condition is called *one theory per choice.* We show that symmetric belief hierarchies using one theory per choice can be characterized by a *common prior on choice combinations.* This is a simpler object than the common prior on *choice-type* combinations that was used to characterize symmetric belief hierarchies that possibly violate the one theory per choice condition. Building on this insight, we demonstrate how symmetric belief hierarchies that use one theory per choice and express common belief in rationality can be characterized by the concept of *canonical correlated equilibrium.* The latter is a common prior on choice combinations that satisfies certain optimality conditions, similar to those in *correlated equilibrium.* We apply this characterization to the example "Rock, paper, scissors", to show that your choice *bomb* cannot rationally be made with a symmetric belief hierarchy that expresses common belief in rationality *and uses one theory per choice.* Remember that you could rationally choose *bomb* with a symmetric belief hierarchy, violating the one theory per choice condition, that expresses common belief in rationality. Hence, the one theory per choice condition crucially matters if we focus on choices that can rationally be made with a symmetric belief hierarchy that expresses common belief in rationality. We finally investigate the relation with simple belief hierarchies from Section 4.1. We show that every simple belief hierarchy is symmetric and uses one theory per choice. Since we know that simple belief hierarchies expressing common belief in rationality always exist, it follows that we can always find symmetric belief hierarchies that use one theory per choice and express common belief in rationality.

4.3.1 One Theory per Choice Condition

Consider the beliefs diagram for "Rock, paper, scissors" at the top of Figure 4.2.1. Note that your choice *paper* appears twice in this beliefs diagram. The second time it appears it is denoted by *paper'*. What results is a belief hierarchy where the same choice *paper* for you is being explained by two different beliefs. Indeed, consider the belief hierarchy for you that starts at *paper'.* In that belief hierarchy, you believe that Barbara assigns probability 1/5 to the event that you choose *paper* while assigning probability 1/2 to Barbara's choices *scissors* and *rock.* But you also believe that Barbara assigns probability 2/5 to the event that you chose *paper* (denoted by *paper'*) while assigning probability 1 to Barbara choosing *rock.* That is, the same choice *paper* for you is being explained by two different theories (beliefs): one theory in which you assign probability 1/2 to Barbara's choices *scissors* and *rock,* and one theory in which you assign probability 1 to Barbara's choice *rock.* We say that this belief hierarchy of yours violates the *one theory per choice condition.*

On the other hand, *one theory per choice* is guaranteed if we consider beliefs diagrams in which every choice of a player only appears once. To see this, consider such a beliefs diagram. Since every choice c_i only appears once, there is a unique first-order belief, say $b_i^1[c_i]$, associated to every choice c_i in the beliefs diagram. But then, within a given belief hierarchy, every choice c_i can only be explained by a single first-order belief – the unique belief $b_i^1[c_i]$ attached to the choice c_i in the beliefs diagram. This naturally leads to the following definition.

Definition 4.3.1 (One theory per choice) *A belief hierarchy uses* ***one theory per choice*** *if it is generated by a beliefs diagram in which every choice of a player only appears once.*

It may be verified that all belief hierarchies considered in this book so far – except those that correspond to the beliefs diagram in Figure 4.2.1 – use one theory per choice. The reader may wonder why we did not explore the *one theory per choice condition* in the previous chapter, while investigating common belief in rationality. The reason is that for the choices you can rationally make under common belief in rationality, it does not matter whether we impose the one theory per choice condition or not.

To see this, remember from Theorem 3.4.1 that the choices which can rationally be made under common belief in rationality are exactly the choices that survive the *iterated elimination of strictly dominated choices*. In Section 3.4.5 we have shown how to construct an epistemic model, with a unique type $t_i^{c_i}$ associated to every choice c_i that survives the procedure, such that type $t_i^{c_i}$ expresses common belief in rationality and c_i is optimal for $t_i^{c_i}$. By construction, all the types $t_i^{c_i}$ within this epistemic model have the following property: The belief $b_i(t_i^{c_i})$ only assigns positive probability to choice-type pairs $(c_j, t_j^{c_j})$ for opponent j, where $t_j^{c_j}$ is the unique type associated to the choice c_j. This means that within the belief hierarchy of $t_i^{c_i}$, every choice c_j is being explained by a single belief hierarchy, which is the belief hierarchy induced by the unique type $t_j^{c_j}$ associated to the choice c_j. In other words, all belief hierarchies generated within this epistemic model use one theory per choice.

Question 4.3.1 *Explain how this epistemic model can be represented by a beliefs diagram in choice-type representation, where every choice only appears once.*

As a consequence, every choice that survives the *iterated elimination of strictly dominated choices* is optimal for a belief hierarchy that uses one theory per choice and expresses common belief in rationality. This, combined with Theorem 3.4.1, yields the following insight: Every choice that is optimal for a belief hierarchy that expresses common belief in rationality, is also optimal for a belief hierarchy that not only expresses common belief in rationality, but also uses one theory per choice. We thus obtain the following result.

Theorem 4.3.1 (One theory per choice does not matter for common belief in rationality) *A choice is optimal for a belief hierarchy that expresses common belief in rationality, if and only if, it is optimal for a belief hierarchy that expresses common belief in rationality and uses one theory per choice.*

Hence, for the choices that can rationally be made under common belief in rationality, it does not matter whether we additionally impose the one theory per choice condition or not. This is not true for symmetric belief hierarchies, however. We will see that for the choices that can rationally be made under common belief in rationality with a *symmetric* belief hierarchy, it is of crucial importance whether we additionally impose one theory per choice or not.

4.3.2 Relation with Common Prior

In Theorem 4.2.1 we have seen that every symmetric belief hierarchy – also those that violate the one theory per choice condition – can be characterized by a common prior on choice-type combinations. We will see that, if we additionally impose the one theory per choice condition, then a symmetric belief hierarchy can be characterized by an even simpler object: a common prior on *choice combinations*, without any mentioning of types.

Consider a symmetric belief hierarchy β that uses one theory per choice. Then, by Theorem 4.2.1, the belief hierarchy β is generated by a common prior π on choice-type combinations, assigning to every choice-type combination $(c_i, t_i)_{i \in I}$ some probability $\pi((c_i, t_i)_{i \in I})$. In other words, the belief hierarchy β is generated by a symmetric weighted beliefs diagram in choice-type representation, where the choice-type pairs (c_i, t_i) that appear are exactly the choice-type pairs that receive positive probability by

π, and the weight of the arrow from (c_i, t_i) to (c_{-i}, t_{-i}) is given by $\pi((c_i, t_i), (c_{-i}, t_{-i}))$. Moreover, since the belief hierarchy β uses one theory per choice, the symmetric weighted beliefs diagram can be chosen such that every choice only appears once. This means that for every choice c_i there is a unique type, say $t_i^{c_i}$, that figures together with c_i as a choice-type pair in the weighted beliefs diagram.

Now, consider the beliefs diagram that is induced by this symmetric weighted beliefs diagram. Since in the weighted beliefs diagram every choice only appears once, the same is true for the induced beliefs diagram. Moreover, in the beliefs diagram every arrow a from a choice c_i to an opponents' choice combination c_{-i} has the probability

$$p(a) = \frac{\pi((c_i, t_i^{c_i}), (c_{-i}, t_{-i}^{c_{-i}}))}{\pi(c_i, t_i^{c_i})}, \quad (4.3.1)$$

where $c_{-i}^{t_{-i}}$ is an abbreviation for $(c_j^{t_j})_{j \neq i}$.

We can now define, on the basis of π, a common prior $\hat{\pi}$ on *choice-combinations,* which assigns to every choice-combination $(c_i)_{i \in I}$ the probability

$$\hat{\pi}((c_i)_{i \in I}) := \pi((c_i, t_i^{c_i})_{i \in I}).$$

Together with (4.3.1) we conclude that within the induced beliefs diagram, every arrow a from c_i to c_{-i} has the probability

$$p(a) = \frac{\hat{\pi}(c_i, c_{-i})}{\hat{\pi}(c_i)}, \quad (4.3.2)$$

where

$$\hat{\pi}(c_i) = \sum_{c_{-i} \in C_{-i}} \hat{\pi}(c_i, c_{-i}).$$

When (4.3.2) is satisfied, we say that the beliefs diagram is induced by the common prior on choice combinations $\hat{\pi}$.

Definition 4.3.2 (Common prior on choice combinations) *(a) A* ***common prior on choice combinations*** *is a probability distribution $\hat{\pi}$ that assigns to every choice combination $(c_i)_{i \in I}$ a probability $\hat{\pi}((c_i)_{i \in I})$.*

(b) A beliefs diagram is ***induced by the common prior*** *on choice combinations $\hat{\pi}$ if every choice only appears once, if for every choice c_i and choice combination c_{-i}, the arrow a from c_i to c_{-i} is present exactly when $\hat{\pi}(c_i, c_{-i}) > 0$, and this arrow a has probability*

$$p(a) = \frac{\hat{\pi}(c_i, c_{-i})}{\hat{\pi}(c_i)}.$$

(c) A belief hierarchy is ***induced by a common prior*** *on choice combinations $\hat{\pi}$ if it is part of a beliefs diagram induced by $\hat{\pi}$.*

So far we have seen that every symmetric belief hierarchy that uses one theory per choice is induced by a common prior on choice combinations. It is easily seen that the other direction is also true.

Suppose that the belief hierarchy β is induced by a common prior $\hat{\pi}$ on choice combinations. Then, it is part of a beliefs diagram induced by $\hat{\pi}$. By construction, every choice in this beliefs diagram only appears once, and hence the belief hierarchy β uses one theory per choice. To show that the belief hierarchy β is symmetric, we will rely on Theorem 4.2.1 and construct a common prior on choice-type combinations that induces the beliefs diagram. For every choice c_i that receives positive probability under $\hat{\pi}$, define a single type $t_i^{c_i}$. Let the common prior π on choice-type combinations be given by

$$\pi((c_i, t_i^{c_i})_{i\in I}) := \hat{\pi}((c_i)_{i\in I})$$

for every choice combination $(c_i)_{i\in I}$. Then, the common prior π induces the same beliefs diagram as $\hat{\pi}$, since every choice is attached to a single type, and the probabilities of the choice combinations are the same under π and $\hat{\pi}$. Hence, the beliefs diagram is induced by the common prior on choice-type combinations π. But then, it follows from Theorem 4.2.1 that the belief hierarchy β is symmetric. We thus conclude that β uses one theory per choice and is symmetric. As such, every belief hierarchy that is induced by a common prior on choice combinations uses one theory per choice and is symmetric. Since the other direction is also true, as we have seen, we obtain the following characterization of symmetric belief hierarchies with one theory per choice.

Theorem 4.3.2 (Relation with common prior) *A belief hierarchy is symmetric and uses one theory per choice, if and only if, it is induced by a common prior on choice combinations.*

As an illustration, consider the beliefs diagram in Figure 4.2.2 for "When Chris joins the party". As every choice only appears once, every belief hierarchy in this beliefs diagram uses one theory per choice. Moreover, we have seen that the beliefs diagram is induced by the symmetric weighted beliefs diagram in the same figure, and hence every belief hierarchy is symmetric as well. By Theorem 4.3.2, each of these belief hierarchies is therefore induced by a common prior on choice combinations. From Question 4.2.2 we know that the beliefs diagram is induced by the common prior on choice-type combinations π given by

$$\begin{array}{rcl} \pi((g, t_1^g), (b, t_2^b), (y, t_3^y)) &=& 4/10,\ \pi((g, t_1^g), (y, t_2^y), (b, t_3^b)) = 4/10, \\ \pi((r, t_1^r), (g, t_2^g), (b, t_3^b)) &=& 1/10,\ \pi((r, t_1^r), (g, t_2^g), (y, t_3^y)) = 1/10. \end{array}$$

This, in turn, induces the common prior $\hat{\pi}$ on choice combinations given by

$$\hat{\pi}(g, b, y) = 4/10,\ \hat{\pi}(g, y, b) = 4/10,\ \hat{\pi}(r, g, b) = 1/10,\ \hat{\pi}(r, g, y) = 1/10.$$

It may be verified that the beliefs diagram, and hence each of its symmetric belief hierarchies using one theory per choice, are induced by this common prior $\hat{\pi}$ on choice combinations.

4.3.3 Relation with Canonical Correlated Equilibrium

As a next step, we wish to combine the idea of common belief in rationality with the notions of symmetry and one theory per choice. More precisely, we will characterize those symmetric belief hierarchies that use one theory per choice and express common belief in rationality. This will eventually lead us to the concept of *canonical correlated equilibrium.*

Consider a symmetric belief hierarchy β that uses one theory per choice and expresses common belief in rationality. Then we know from Theorem 4.2.2 that the belief hierarchy β is induced by a correlated equilibrium π, which is a common prior on choice-type combinations. That is, the belief hierarchy β is part of some beliefs diagram which is induced by the correlated equilibrium π. As β uses one theory per choice, every choice only appears once in this beliefs diagram, and hence for every choice c_i there is a unique type $t_i^{c_i}$ such that the choice-type pair $(c_i, t_i^{c_i})$ receives positive probability under π. But then, the beliefs diagram, and hence the belief hierarchy β, is induced by the common prior $\hat{\pi}$ on choice combinations which assigns to every choice-combination $(c_i)_{i\in I}$ the probability

$$\hat{\pi}((c_i)_{i\in I}) := \pi((c_i, t_i^{c_i})_{i\in I}). \tag{4.3.3}$$

As π is a correlated equilibrium, we know that for every player i, and every choice c_i with $\pi(c_i, t_i^{c_i}) > 0$, that the choice c_i is optimal for the conditional belief $\pi(\cdot \mid c_i, t_i^{c_i})$. By (4.3.3) we conclude that for every player i, and every choice c_i with $\hat{\pi}(c_i) > 0$, the choice c_i is optimal for the conditional belief $\hat{\pi}(\cdot \mid c_i)$. Here, the conditional belief $\hat{\pi}(\cdot \mid c_i)$ is the belief about the opponents' choice combinations given by

$$\hat{\pi}(c_{-i} \mid c_i) := \frac{\hat{\pi}(c_i, c_{-i})}{\hat{\pi}(c_i)} \text{ for all } c_{-i} \in C_{-i}.$$

A common prior $\hat{\pi}$ on choice combinations satisfying these optimality conditions is called a *canonical correlated equilibrium.*

Definition 4.3.3 (Canonical correlated equilibrium) *A common prior $\hat{\pi}$ on choice combinations is a* ***canonical correlated equilibrium*** *if for every player i, and every choice c_i with $\hat{\pi}(c_i) > 0$, the choice c_i is optimal for the conditional belief $\hat{\pi}(\cdot \mid c_i)$ of player i conditional on his choice c_i.*

By the arguments above, we have thus seen that every symmetric belief hierarchy β which uses one theory per choice and expresses common belief in rationality is induced by a canonical correlated equilibrium $\hat{\pi}$.

The other direction is also true, as we will show now. Indeed, consider a belief hierarchy β that is induced by a canonical correlated equilibrium $\hat{\pi}$, which is a common prior on choice combinations. Then, we immediately conclude from Theorem 4.3.2 that the belief hierarchy β is symmetric and uses one theory per choice. It remains to show that β also expresses common belief in rationality.

For every choice c_i that receives positive probability by $\hat{\pi}$, we can define a single type $t_i^{c_i}$. Then, $\hat{\pi}$ induces a common prior π on choice-type combinations given by

$$\pi((c_i, t_i^{c_i})_{i\in I}) := \hat{\pi}((c_i)_{i\in I}) \tag{4.3.4}$$

for every choice-type combination $(c_i, t_i^{c_i})_{i\in I}$. As $\hat{\pi}$ is a canonical correlated equilibrium, we know that for every player i, and every choice c_i with $\hat{\pi}(c_i) > 0$, the choice c_i is optimal for the conditional belief $\hat{\pi}(\cdot \mid c_i)$. But then, by (4.3.4) we conclude that for every player i, and every choice c_i with $\pi(c_i, t_i^{c_i}) > 0$, the choice c_i is optimal for the conditional belief $\pi(\cdot \mid c_i, t_i^{c_i})$. Therefore, the induced common prior π on choice-type combinations is a correlated equilibrium.

Since the belief hierarchy β is induced by the correlated equilibrium π, it follows from Theorem 4.2.2 that β expresses common belief in rationality. We thus have shown that every belief hierarchy that is induced by a canonical correlated equilibrium is symmetric, uses one theory per choice, and expresses common belief in rationality. We therefore arrive at the following characterization.

Theorem 4.3.3 (Relation with canonical correlated equilibrium) *A belief hierarchy is symmetric, uses one theory per choice, and expresses common belief in rationality, if and only if, it is induced by a canonical correlated equilibrium.*

With this result at hand, we can now also characterize the *choices* that you can rationally make while holding a symmetric belief hierarchy that uses one theory per choice and expresses common belief in rationality. Suppose that the choice c_i^* is optimal for a belief hierarchy β_i that is symmetric, uses one theory per choice, and expresses common belief in rationality. Then we know, by Theorem 4.3.3, that β_i is part of a beliefs diagram that is induced by a canonical correlated equilibrium $\hat{\pi}$. Remember that $\hat{\pi}$ is a common prior on choice combinations.

Let β_i be the belief hierarchy that starts at the choice c_i within the beliefs diagram. Since the belief hierarchy β_i expresses common belief in rationality, we can choose the beliefs diagram such that all arrows are solid. In particular, every arrow leaving c_i is solid, meaning that c_i is optimal for the first-order belief represented by the arrows leaving c_i. As the beliefs diagram is induced by $\hat{\pi}$, this first-order belief is the conditional belief $\hat{\pi}(\cdot \mid c_i)$. Hence, we conclude that the first-order belief of the belief hierarchy β_i is given by $\hat{\pi}(\cdot \mid c_i)$. Since the choice c_i^* is optimal for β_i, it must be the case that c_i^* is optimal for the conditional belief $\hat{\pi}(\cdot \mid c_i)$ in the canonical correlated equilibrium $\hat{\pi}$. In this case, we say that c_i^* is optimal in the canonical correlated equilibrium $\hat{\pi}$.

Definition 4.3.4 (Choice optimal in a canonical correlated equilibrium) *A choice c_i^* is* ***optimal in a canonical correlated equilibrium*** *$\hat{\pi}$ if there is some choice c_i with $\hat{\pi}(c_i) > 0$ such that c_i^* is optimal for the belief $\hat{\pi}(\cdot \mid c_i)$ of player i conditional on c_i.*

With our argument above, we concluded that every choice that is optimal for a symmetric belief hierarchy that uses one theory per choice and expresses common belief in rationality, must be optimal in a canonical correlated equilibrium. The other direction is also true.

To see this, consider a choice c_i^* that is optimal in a canonical correlated equilibrium $\hat{\pi}$. That is, there is some choice c_i with $\hat{\pi}(c_i) > 0$ such that c_i^* is optimal for the conditional belief $\hat{\pi}(\cdot \mid c_i)$. Now, consider the beliefs diagram induced by $\hat{\pi}$, and the belief hierarchy β_i that starts at c_i. By Theorem 4.3.3, this belief hierarchy β_i is symmetric, uses one theory per choice, and expresses common belief in rationality. By construction, the first-order belief of β_i is given by the conditional belief $\hat{\pi}(\cdot \mid c_i)$. Since c_i^* is optimal for $\hat{\pi}(\cdot \mid c_i)$, we conclude that c_i^* is optimal for the belief hierarchy β_i, which is symmetric, uses one theory per choice, and expresses common belief in rationality. Hence, every choice c_i^* that is optimal in a canonical correlated equilibrium will be optimal for a symmetric belief hierarchy that uses one theory per choice and expresses common belief in rationality. We thus obtain the following characterization.

Theorem 4.3.4 (Relation with canonical correlated equilibrium choices) *A choice is optimal for a symmetric belief hierarchy that uses one theory per choice and expresses common belief in rationality, if and only if, the choice is optimal in a canonical correlated equilibrium.*

We will now use this characterization to identify, in the example "Rock, paper, scissors", those choices you can rationally make with a symmetric belief hierarchy that uses one theory per choice and expresses common belief in rationality.

Example 4.9: Rock, paper, scissors.

We have seen in Example 4.8 in Section 4.2.4 that under common belief in rationality with a symmetric belief hierarchy, you can rationally make each of your choices *rock, paper, scissors* and *bomb.* Indeed, all belief hierarchies at the top of Figure 4.2.1 are symmetric, because they are generated by the symmetric weighted beliefs diagram at the bottom of that same figure. Moreover, all these belief hierarchies express common belief in rationality as all arrows are solid. Since your choice *rock* is optimal for the belief hierarchy that starts at *rock,* your choice *paper* is optimal for the belief hierarchies that start at *paper* and *paper',* your choice *scissors* is optimal for the belief hierarchy that starts at *scissors,* and your choice *bomb* is optimal for the belief hierarchy that starts at *paper',* we conclude that you can rationally make each of your choices under common belief in rationality with a symmetric belief hierarchy.

Note, however, that all the belief hierarchies in Figure 4.2.1 violate the one theory per choice condition, as your choice *paper* appears twice in the beliefs diagram. In fact, each of these belief hierarchies uses two different beliefs to justify the same choice *paper* for you. We will now show, on the basis of Theorem 4.3.4, that your choice *bomb* is not optimal for *any* belief hierarchy that is symmetric, expresses common belief in rationality, and uses *one theory per choice.* However, showing this will be far from easy, as will become clear below. Hence, in this example it is crucial for your optimal choices whether, in addition to symmetry and common belief in rationality, we add the one theory per choice condition or not.

Suppose, contrary to what we want to show, that your choice *bomb* is optimal for a symmetric belief hierarchy that expresses common belief in rationality and uses one theory per choice. Then, by Theorem 4.3.4, your choice *bomb* must be optimal in a canonical correlated equilibrium $\hat{\pi}$. That is, there must be a choice $c_1 \in C_1$ for you such that $\hat{\pi}(c_1) > 0$ and *bomb* is optimal under the conditional belief $\pi(\cdot \mid c_1)$. We show that this cannot be the case. We proceed by several steps.

(a) We first show that $\hat{\pi}(bomb_1) = 0$. Here, we use the subindex 1 to indicate that this choice belongs to you. Suppose, contrary to what we want to show, that $\hat{\pi}(bomb_1) > 0$. Since $\hat{\pi}$ is a canonical correlated equilibrium, it must then be the case that your choice $bomb_1$ is optimal for the conditional belief $\hat{\pi}(\cdot \mid bomb_1)$. Note that your choice $bomb_1$ is only optimal for the belief that assigns probability 1 to Barbara's choice $rock_2$. Indeed, if you assign a positive probability to any of the other choices for Barbara, your choice $paper_1$ would be strictly better than $bomb_1$. Hence, it must be that $\hat{\pi}(rock_2 \mid bomb_1) = 1$, where we use the subindex 2 to indicate that this choice belongs to Barbara. It then follows that $\hat{\pi}(rock_2) > 0$. Hence, $rock_2$ must be optimal for Barbara under the belief $\hat{\pi}(\cdot \mid rock_2)$, since $\hat{\pi}$ is a canonical correlated equilibrium. However, as $\hat{\pi}(rock_2 \mid bomb_1) = 1$ we have that $\hat{\pi}(bomb_1, rock_2) > 0$, and hence $\hat{\pi}(bomb_1 \mid rock_2) > 0$. But then, $diamond_2$ is strictly better for Barbara than $rock_2$, as $diamond_2$ yields Barbara a higher utility than her choice $rock_2$ if you choose $bomb_1$, whereas both choices yield the same utility for Barbara in all other cases. In particular, $rock_2$ cannot be optimal for Barbara under the belief $\hat{\pi}(\cdot \mid rock_2)$, which contradicts the fact that $\hat{\pi}$ is a canonical correlated equilibrium. We thus conclude that $\hat{\pi}(bomb_1) > 0$ cannot be the case, and hence $\hat{\pi}(bomb_1) = 0$.

(b) We next show that $\hat{\pi}(rock_2 \mid paper_1) = 1$. Recall from above that we assume there is a choice $c_1 \in C_1$ with $\hat{\pi}(c_1) > 0$ such that $bomb_1$ is optimal under the conditional belief $\hat{\pi}(\cdot \mid c_1)$. Since $bomb_1$ is only optimal for you under the belief that assigns probability 1 to Barbara's choice $rock_2$, we must have that $\hat{\pi}(rock_2 \mid c_1) = 1$. As $\hat{\pi}$ is a canonical correlated equilibrium, c_1 must be optimal for the conditional belief $\hat{\pi}(\cdot \mid c_1)$ which assigns probability 1 to Barbara's choice $rock_2$. Hence, c_1 can only be your choice $paper_1$ or your choice $bomb_1$. By (a) we know that $\hat{\pi}(bomb_1) = 0$, and hence it must be that c_1 is your choice $paper_1$. As $\hat{\pi}(rock_2 \mid c_1) = 1$, we conclude that $\hat{\pi}(rock_2 \mid paper_1) = 1$.

(c) Now we will show that $\hat{\pi}(scissors_2) = 0$. Suppose, contrary to what we want to show, that $\hat{\pi}(scissors_2) > 0$. Since $\hat{\pi}$ is a canonical correlated equilibrium, $scissors_2$ must be optimal for Barbara for the conditional belief $\hat{\pi}(\cdot \mid scissors_2)$. By (a) we know that $\hat{\pi}(bomb_1) = 0$, and hence $\hat{\pi}(bomb_1 \mid scissors_2) = 0$. Moreover, by (b) we know that $\hat{\pi}(rock_2 \mid paper_1) = 1$, which implies that $\hat{\pi}(paper_1, c_2) = 0$ for all of Barbara's choices c_2 other than $rock_2$. In particular, $\hat{\pi}(paper_1, scissors_2) = 0$, from which it follows that $\hat{\pi}(paper_1 \mid scissors_2) = 0$. But then, Barbara's choice $paper_2$ is better than her choice $scissors_2$ under the belief $\hat{\pi}(\cdot \mid scissors_2)$. This contradicts the assumption that $scissors_2$ is optimal for Barbara under the belief $\hat{\pi}(\cdot \mid scissors_2)$. Hence, we conclude that $\hat{\pi}(scissors_2) > 0$ cannot be the case, therefore $\hat{\pi}(scissors_2) = 0$.

(d) We continue by showing that $\hat{\pi}(rock_1) = 0$. Suppose, contrary to what we want to show, that

$\hat{\pi}(rock_1) > 0$. Since $\hat{\pi}$ is a canonical correlated equilibrium, the choice $rock_1$ must be optimal for you for the belief $\hat{\pi}(\cdot \mid rock_1)$. By (c) we know that $\hat{\pi}(scissors_2) = 0$, and hence $\hat{\pi}(scissors_2 \mid rock_1) = 0$. But then, your choice $scissors_1$ is better than your choice $rock_1$ under the belief $\hat{\pi}(\cdot \mid rock_1)$, which contradicts the assumption that $rock_1$ is optimal for the belief $\hat{\pi}(\cdot \mid rock_1)$. Hence, our assumption that $\hat{\pi}(rock_1) > 0$ cannot hold, which implies that $\hat{\pi}(rock_1) = 0$.

(e) Next, we prove that $\hat{\pi}(paper_2) = 0$. Suppose, contrary to what we want to show, that $\hat{\pi}(paper_2) > 0$. Since $\hat{\pi}$ is a canonical correlated equilibrium, Barbara's choice $paper_2$ must be optimal for the belief $\hat{\pi}(\cdot \mid paper_2)$. From (a) and (d) we know that $\hat{\pi}(bomb_1) = 0$ and $\hat{\pi}(rock_1) = 0$, which imply that $\hat{\pi}(rock_1, paper_2) = 0$ and $\hat{\pi}(bomb_1, paper_2) = 0$. Moreover, from (b) we know that $\hat{\pi}(rock_2 \mid paper_1) = 1$, which implies that $\hat{\pi}(paper_1, paper_2) = 0$. Hence, we conclude that $\hat{\pi}(rock_1, paper_2) = 0$, $\hat{\pi}(paper_1, paper_2) = 0$ and $\hat{\pi}(bomb_1, paper_2) = 0$, thus it must be that $\hat{\pi}(scissors_1 \mid paper_2) = 1$. But then, $rock_2$ and $diamond_2$ will be better for Barbara than $paper_2$ under the belief $\hat{\pi}(\cdot \mid paper_2)$, which contradicts the assumption that $paper_2$ is optimal for the belief $\hat{\pi}(\cdot \mid paper_2)$. Hence, the assumption that $\hat{\pi}(paper_2) > 0$ cannot be true, which means that $\hat{\pi}(paper_2) = 0$.

(f) We show that $\hat{\pi}(scissors_1) = 0$. Suppose, contrary to what we want to show, that $\hat{\pi}(scissors_1) > 0$. Since $\hat{\pi}$ is a canonical correlated equilibrium, your choice $scissors_1$ must be optimal for the belief $\hat{\pi}(\cdot \mid scissors_1)$. Since we know by (c) and (e) that $\hat{\pi}(scissors_2) = 0$ and $\hat{\pi}(paper_2) = 0$, it follows that $\hat{\pi}(scissors_2 \mid scissors_1) = 0$ and $\hat{\pi}(paper_2 \mid scissors_1) = 0$. But then, your choice $paper_1$ is better than $scissors_1$ under the belief $\hat{\pi}(\cdot \mid scissors_1)$, which contradicts the assumption that $scissors_1$ is optimal for $\hat{\pi}(\cdot \mid scissors_1)$. Therefore, the assumption that $\hat{\pi}(scissors_1) > 0$ cannot be true, which means that $\hat{\pi}(scissors_1) = 0$.

By (a), (b), (d) and (f) we thus see that $\hat{\pi}(bomb_1) = 0$, $\hat{\pi}(rock_2 \mid paper_1) = 1$, $\hat{\pi}(rock_1) = 0$ and $\hat{\pi}(scissors_1) = 0$. But then, it must be that $\hat{\pi}(paper_1, rock_2) = 1$ and $\hat{\pi}(c_1, c_2) = 0$ otherwise. As $\hat{\pi}$ is a canonical correlated equilibrium, Barbara's choice $rock_2$ must be optimal for the conditional belief $\hat{\pi}(\cdot \mid rock_2)$ which assigns probability 1 to your choice $paper_1$. This, however, is not true since Barbara's choice $scissors_2$ is better than $rock_2$ under that belief $\hat{\pi}(\cdot \mid rock_2)$. We thus obtain a contradiction.

Hence, there is a no canonical correlated equilibrium $\hat{\pi}$ for which there is a choice c_1 with $\hat{\pi}(c_1) > 0$ such that your choice *bomb* is optimal for the belief $\hat{\pi}(\cdot \mid c_1)$. In other words, your choice *bomb* is not optimal in a canonical correlated equilibrium. By Theorem 4.3.4 we therefore conclude that you cannot rationally choose *bomb* with a symmetric belief hierarchy that uses one theory per choice and expresses common belief in rationality.

What about your other three choices *rock, paper* and *scissors*? Can you rationally make each of those choices under common belief in rationality with a symmetric belief hierarchy that uses one theory per choice? The answer, as we will see, is "yes". The reason is that we can find a single canonical correlated equilibrium $\hat{\pi}$ for which your choices *rock, paper* and *scissors* are all optimal.

Consider the common prior on choice combinations $\hat{\pi}$ given by

$$\begin{aligned}
\hat{\pi}(rock_1,\ paper_2) &= 1/6,\ \hat{\pi}(rock_1,\ scissors_2) = 1/6, \\
\hat{\pi}(paper_1,\ rock_2) &= 1/6,\ \hat{\pi}(paper_1,\ scissors_2) = 1/6, \\
\hat{\pi}(scissors_1,\ rock_2) &= 1/6 \text{ and } \hat{\pi}(scissors_1,\ paper_2) = 1/6.
\end{aligned}$$

It can be shown that $\hat{\pi}$ is a canonical correlated equilibrium. Please verify this.

Question 4.3.2 *For this canonical correlated equilibrium $\hat{\pi}$, design the corresponding symmetric weighted beliefs diagram, and the beliefs diagram it induces. Explain why all belief hierarchies are symmetric, use one theory per choice, and express common belief in rationality.*

Since your choice *rock* is optimal for the belief $\hat{\pi}(\cdot \mid rock_1)$, your choice *paper* is optimal for the belief $\hat{\pi}(\cdot \mid paper_1)$ and your choice *scissors* is optimal for the belief $\hat{\pi}(\cdot \mid scissors_1)$, it follows that all of these three choices are optimal in a canonical correlated equilibrium. Hence, by Theorem 4.3.4, you can rationally choose *rock, paper* and *scissors* under common belief in rationality with a symmetric belief hierarchy that uses one theory per choice.

However, as we have seen, you can rationally choose *bomb* under common belief in rationality with a symmetric belief hierarchy, but *not* under common belief in rationality with a symmetric belief hierarchy that uses *one theory per choice.*

4.3.4 Relation with Simple Belief Hierarchies

An important question is whether we can always find, for every game and for every player, a belief hierarchy that meets all the conditions of this section – that is, a belief hierarchy that is symmetric, uses one theory per choice and expresses common belief in rationality. The answer is "yes". The reason is that every *simple* belief hierarchy, as defined in Section 4.1.1, is symmetric and uses one theory per choice. Since we have seen in Theorem 4.1.4 that we can always find, for every player, a *simple* belief hierarchy that expresses common belief in rationality, it follows that every player has at least one belief hierarchy that is *symmetric*, uses *one theory per choice*, and *expresses common belief in rationality.*

To see why a simple belief hierarchy is symmetric and uses one theory per choice, consider the example "When Chris joins the party", and the simple belief hierarchy for you generated by the belief combination

$$(\sigma_1 = \tfrac{1}{2} \cdot green + \tfrac{1}{2} \cdot red,\ \sigma_2 = \tfrac{1}{3} \cdot green + \tfrac{2}{3} \cdot yellow,\ \sigma_3 = blue).$$

We have seen in Question 4.1.8 that this belief combination is in fact a Nash equilibrium. How can we generate this simple belief hierarchy by a symmetric weighted beliefs diagram?

The answer is given by the weighted beliefs diagram in Figure 4.3.1. Consider, for instance, the arrow from your choice g to the choice combination (y, b) by Barbara and Chris. The weight assigned to this arrow a is

$$w(a) = \sigma_1(g) \cdot \sigma_2(y) \cdot \sigma_3(b) = 1/2 \cdot 2/3 \cdot 1 = 1/3.$$

That is, the weight is given by the product of the probabilities assigned to the choices involved in the arrow. The weights for the other arrows are determined in the same fashion. Please check this. It may be verified that this weighted beliefs diagram is indeed symmetric.

The induced beliefs diagram is the one depicted in Figure 4.3.2. The belief hierarchy for you that starts at your choice g, and the one that starts at your choice r, are both equal to the simple belief hierarchy generated by the belief combination $(\sigma_1, \sigma_2, \sigma_3)$ above. Please check this. Therefore, the symmetric weighted beliefs diagram from Figure 4.3.1 induces the simple belief hierarchy for you generated by $(\sigma_1, \sigma_2, \sigma_3)$.

Question 4.3.3 *Consider the example "Movie for two", and the simple belief hierarchy for you generated by the belief combination* $(\sigma_1 = (0.4)\cdot$ *Palace* $+(0.6)\cdot$ *home*, $\sigma_2 = (0.7)\cdot$ *Corner* $+(0.3)\cdot$ *home*$)$. *Create a symmetric weighted beliefs diagram that induces this simple belief hierarchy, using the construction above.*

This construction works in general. That is, by using the method above we can create, for every simple belief hierarchy, a symmetric weighted beliefs diagram that induces it. Moreover, since the symmetric weighted beliefs diagram so constructed only contains every choice once, it automatically uses one theory per choice. We thus obtain the following general relationship.

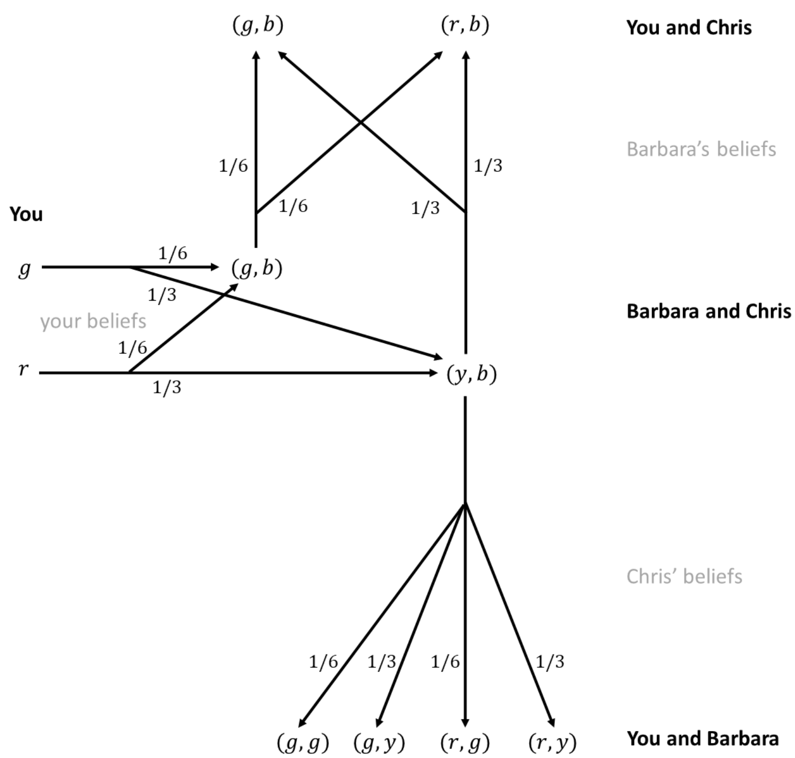

Figure 4.3.1 A symmetric weighted beliefs diagram inducing a simple belief hierarchy

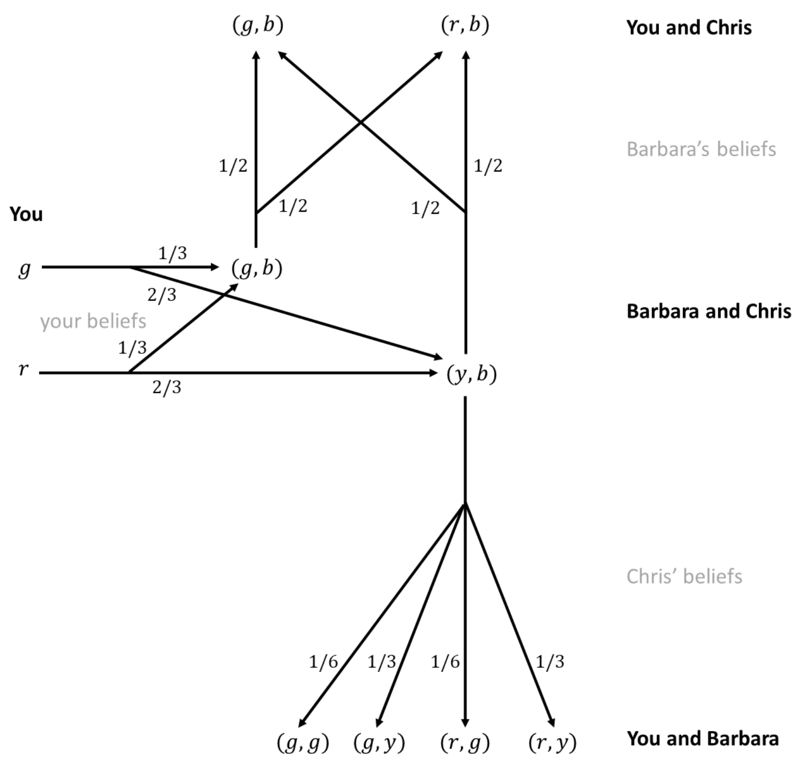

Figure 4.3.2 Beliefs diagram induced by the weighted beliefs diagram in Figure 4.3.1

Theorem 4.3.5 (Relation with simple belief hierarchies) *Every simple belief hierarchy is symmetric and uses one theory per choice.*

Recall, from Theorem 4.1.4, that we can always find a simple belief hierarchy that expresses common belief in rationality. Together with Theorem 4.3.5, we conclude that every game always contains belief hierarchies that are symmetric, use one theory per choice, and express common belief in rationality.

Corollary 4.3.1 (Existence) *For every game and every player, there is at least one belief hierarchy that expresses common belief in rationality, is symmetric, and uses one theory per choice.*

Theorem 4.3.5 reveals that simplicity of a belief hierarchy can be viewed as a particular form of symmetry. However, a simple belief hierarchy has additional properties that a symmetric belief hierarchy need not have. Remember that in a simple belief hierarchy β_i, player i believes that his opponents are correct about the actual belief hierarchy β_i he has. That is, if we follow the arrows in β_i, then all the choices c_i for player i that are reached will give rise to the same belief hierarchy β_i. As an illustration, consider the beliefs diagram of Figure 4.3.2 for "When Chris joins the party", and the simple belief hierarchy β_1 for you that starts at your choice *green.* It may be verified that the belief hierarchy for you that starts at your other choice *red* is exactly the same as β_1. Hence, there is only one belief hierarchy β_1 for you in this beliefs diagram, and the same holds for Barbara and Chris.

This need not be true in a symmetric belief hierarchy. Indeed, consider the beliefs diagram in Figure 4.2.2 for "When Chris joins the party", and the symmetric belief hierarchy β_1 for you that starts at your choice *green.* In that belief hierarchy, you believe with probability 0.5 that Chris assigns probability 0.2 to the event that you hold the belief hierarchy β_1' that starts at your choice *red,* which is different from your actual belief hierarchy β_1. Hence, with the symmetric belief hierarchy β_1 you do not believe, with certainty, that Chris is correct about your actual belief hierarchy.

Another property that a simple belief hierarchy has, but which is not necessarily shared by a symmetric belief hierarchy, is the following: In a game with more than two players, player i believes that opponent j has the same belief about the choice of a third player k as player i has himself. We have seen that simple belief hierarchies always display this feature. Symmetric belief hierarchies may violate this property, however. Consider again the beliefs diagram in Figure 4.2.2 for "When Chris joins the party", and the symmetric belief hierarchy for you that starts at your choice *red.* In that belief hierarchy, you assign probability 1 to the event that Barbara chooses *green,* yet at the same time you believe with probability 0.5 that Chris assigns probability 0.8 to Barbara choosing *yellow.* Hence, you do not believe with certainty that Chris shares your belief about Barbara's choice.

Remember also that in a game with more than two players, a player i with a simple belief hierarchy holds *independent* beliefs about the choices of his opponents. That is, player i's belief about opponent j's choice is independent from i's belief about a second opponent k's choice. Also this property need not hold within a symmetric belief hierarchy, as can be seen from the beliefs diagram in Figure 4.2.2. Consider again your symmetric belief hierarchy that starts at your choice *green,* which assigns probability 0.5 to the event that Barbara chooses *blue* and Chris chooses *yellow,* and assigns probability 0.5 to the event that Barbara chooses *yellow* and Chris chooses *blue.* This belief cannot be written as the product of a probabilistic belief σ_2 about Barbara's choice and a probabilistic belief σ_3 about Chris' choice, and is therefore not independent. We therefore conclude that simple belief hierarchies may be viewed as a particular instance of symmetry, but display additional properties that may not be shared by symmetric belief hierarchies.

A direct consequence of Theorem 4.3.5 is that, whenever a choice is optimal for a *simple* belief hierarchy that expresses common belief in rationality, then this choice will also be optimal for a

You	*gladiator*	*emperor*	*lion*
gladiator	5	4	2
emperor	4	5	0
lion	2	6	5

Barbara	*gladiator*	*emperor*	*lion*
gladiator	5	2	4
emperor	2	5	0
lion	4	5	5

Table 4.3.1 Choices and utilities for you and Barbara in "The masquerade ball"

symmetric hierarchy that uses *one theory per choice* and expresses common belief in rationality. The other direction, however, is not always true: A choice that is optimal for a symmetric belief hierarchy that uses one theory per choice and expresses common belief in rationality need not be optimal for a *simple* belief hierarchy that expresses common belief in rationality. This will become clear from the new example that follows.

Example 4.10: The masquerade ball.

This evening there will be a masquerade ball to which Barbara and you will be going. The theme of the ball is "The Roman empire", and both of you must choose which Roman outfit to wear. Because of past Carnival celebrations, you can both dress like a *gladiator*, an *emperor* or a *lion.* Since you will dance together at the ball, the combination of the outfits will be important. If you both choose the same oufit, the combination will be perfect and you both get a utility of 5. You like the combination of a gladiator and an emperor, which would give you a utility of 4 but Barbara a utility of only 2. Barbara, on the other hand, enjoys the combination of a gladiator and a lion, which would give her a utility of 4 but you a utility of 2 only. If you dress like a lion and Barbara like an emperor, that would be extremely funny for you, yielding you a utility of 6, but it would be embarassing for Barbara, who would get a utility of 0 in this case. Similarly, if Barbara dresses like a lion and you like an emperor, Barbara would find it very funny too (but not as funny as you deem it), yielding her a utility of 5, whereas this would be very embarrassing for you, giving you a utility of 0.

All the above can be summarized by the decision problems in Table 4.3.1. Consider the beliefs diagram at the top of Figure 4.3.3, and verify that it is induced by the *symmetric* weighted beliefs diagram at the bottom of that figure. Hence, all belief hierarchies in the beliefs diagram are symmetric. As every choice only appears once, and all arrows are solid, we may also conclude that all these belief hierarchies additionally use one theory per choice, and express common belief in rationality.

Question 4.3.4 *Find the canonical correlated equilibrium that induces this beliefs diagram.*

Note that your choice *gladiator* is optimal for the belief hierarchy that starts at your choice *gladiator,* your choice *emperor* is optimal for the belief hierarchy that starts at your choice *emperor,* and your choice *lion* is optimal for the belief hierarchy that starts at your choice *lion.* Hence, you can rationally make each of your three choices under common belief in rationality with a symmetric belief hierarchy that uses one theory per choice.

However, we will see that under common belief in rationality with a *simple* belief hierarchy, you cannot rationally choose to dress like an *emperor.* Suppose, contrary to what we want to show, that you could rationally choose *emperor* under common belief in rationality with a simple belief hierarchy. Then we know, by Theorem 4.1.2, that there must be a Nash equilibrium (σ_1, σ_2) in which your choice *emperor* is optimal.

This is only possible, however, if σ_2 assigns positive probability to Barbara's choices *gladiator* and *emperor.* Indeed, if σ_2 would assign probability 0 to Barbara's choice *gladiator,* then your choice

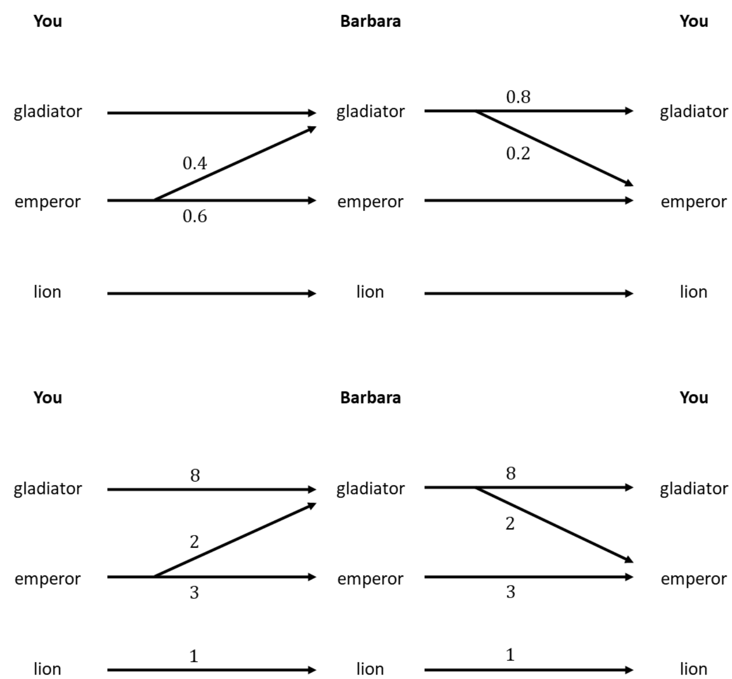

Figure 4.3.3 A beliefs diagram and an associated weighted beliefs diagram for "The masquerade ball"

lion would be better than your choice *emperor.* If, on the other hand, σ_2 would assign probability 0 to Barbara's choice *emperor,* then your choice *gladiator* would be better than your choice *emperor.* Therefore, *emperor* can only be optimal for you if $\sigma_2(gladiator) > 0$ and $\sigma_2(emperor) > 0$. Since (σ_1, σ_2) is a Nash equilibrium, both *gladiator* and *emperor* must be optimal for Barbara under the belief σ_1. This cannot be, however. To see this, note that *emperor* can only be optimal for Barbara if $\sigma_1(emperor) = 1$, since otherwise Barbara's choice *lion* would be better than her choice *emperor.* Please verify this. But if $\sigma_1(emperor) = 1$, Barbara's choice *gladiator* cannot be optimal for her under the belief σ_1. And hence we obtain a contradiction.

We must therefore conclude that your choice *emperor* cannot be optimal in a Nash equilibrium (σ_1, σ_2), and you can thus not rationally dress like an *emperor* under common belief in rationality with a simple belief hierarchy.

Question 4.3.5 *Find a Nash equilibrium in which your choice gladiator is optimal, and another Nash equilibrium in which your choice lion is optimal.*

In view of the question above, we know that under common belief in rationality with a simple belief hierarchy you can rationally dress like a *gladiator* or a *lion,* but not like an *emperor.*

The intuitive reason why a *simple* belief hierarchy, in combination with common belief in rationality, rules out your choice *emperor* is the following: In order for your choice *emperor* to be optimal, you must assign a positive probability to Barbara choosing *gladiator* and to Barbara choosing *emperor.* Since your belief hierarchy is simple, you believe that Barbara believes your are correct about her belief hierarchy, and hence you only deem possible one belief hierarchy β_2 for Barbara. Since you assign positive probability to Barbara choosing *gladiator* and to Barbara choosing *emperor,* these two choices must *both* be optimal for Barbara under the same belief hierarchy β_2. This, however, cannot be, since Barbara's choice *emperor* can only be optimal if she is certain you will choose *emperor* as well, in wich case it would be suboptimal for Barbara to choose *gladiator.*

Under common belief in rationality with a *symmetric* belief hierarchy, on the other hand, you may still assign positive probability to Barbara choosing *gladiator* and *emperor,* because you may support these two choices of Barbara by two *different* belief hierarchies. Consider, for instance, the beliefs diagram in Figure 4.3.3 and your symmetric belief hierarchy that starts at your choice *emperor.* This symmetric belief hierarchy supports your choice *emperor.* In that belief hierarchy, you assign positive probability to Barbara's choices *gladiator* and *emperor.* Importantly, you support Barbara's choice *gladiator* by the belief hierarchy that starts at her choice *gladiator,* whereas you support her choice *emperor* by the *different* belief hierarchy that starts at her choice *emperor.*

This is not possible in a *simple* belief hierarchy: There, you must use one and the same belief hierarchy β_2 for Barbara to support each of Barbara's choices that you assign positive probability to.

4.4 Comparison of the Concepts

To conclude this chapter, we will compare the four concepts we have discussed in Chapters 3 and 4: common belief in rationality, common belief in rationality with a simple belief hierarchy, common belief in rationality with a symmetric belief hierarchy, and common belief in rationality with a symmetric belief hierarchy using one theory per choice. In Table 4.4.1, we compare the concepts in terms of the characterization of the choices that can rationally be made under the concept at hand. In this

Common belief in rationality with ...	**Optimal choices are those that ...**
...	survive iterated elimination of strictly dominated choices
symmetric belief hierarchy	are optimal in a correlated equilibrium
symmetric belief hierarchy using one theory per choice	are optimal in a canonical correlated equilibrium
simple belief hierarchy	are optimal in a Nash equilibrium

Table 4.4.1 Comparison of the concepts in Chapters 2 and 3

table, we have listed the four concepts from least restrictive (common belief in rationality) to most restrictive (common belief in rationality with a simple belief hierarchy). Recall from the previous section that every simple belief hierarchy is symmetric and uses one theory per choice, and hence the fourth concept is indeed more restrictive than the third.

We next compare the four concepts according to the choices they select for you in the various examples of Chapters 3 and 4. See Table 4.4.2. In the first column, we indicate for every example the section in which it was first introduced. In the other four columns, we indicate the choice(s) selected for you by the associated concept in that example. We also indicate the section in which it was shown that these were precisely the choices you can rationally make under the concept at hand. Consider, for instance, the example "Going to a party", and the second column that corresponds to common belief in rationality without further restrictions. In the corresponding cell, we state that under common belief in rationality you can only rationally choose the color *blue,* and this has been shown in Section 3.4.3. This implies that under the other three concepts, which are all more restrictive than common belief in rationality, the only choice you can rationally make must also be *blue.* Hence, no section needs to be specified at the other three columns for the example "Going to a party", because these results follow from the result of common belief in rationality.

Or consider the example "When Chris joins the party". In Section 3.4.3 it was shown that under common belief in rationality, you can rationally choose *green* and *red.* Later, in Section 4.1.3, we have shown that under common belief in rationality with a simple belief hierarchy, you can also rationally choose *green* and *red.* As a consequence, you can also rationally choose *green* and *red,* but no other colors, under common belief in rationality with a symmetric belief hierarchy, or under common belief in rationality with a symmetric belief hierarchy that uses one theory per choice. The reason is that the latter two concepts are both more restrictive than common belief in rationality, but less restrictive than common belief in rationality with a simple belief hierarchy. Therefore, no sections need to be specified for the third and fourth column at this particular example.

In the same fashion, it can be verified that all the results at cells without a section number follow from the results of cells with a section number. Take, for instance, the example "The masquerade ball", where we have shown in Section 4.3.4 that under common belief in rationality with a symmetric belief hierarchy that uses one theory per choice, you can rationally make each of your choices *gladiator, emperor* and *lion.* This implies that under common belief in rationality, or common belief in rationality with a symmetric belief hierarchy, you can also rationally make each of your choices, because these two concepts are less restrictive than common belief in rationality with a symmetric belief hierarchy that uses one theory per choice.

There is only one exception in the table: For the example "Rock, paper, scissors" we have not shown yet that under common belief in rationality with a simple belief hierarchy, you can rationally

Example	**Choices you can rationally make under common belief in rationality with ...** ...	a symmetric belief hierarchy	a symmetric belief hierarchy using one theory per choice	a simple belief hierarchy
Going to a party (Section 3.1)	blue (Section 3.4.3)	blue	blue	blue
When Chris joins party (Section 3.2.2)	green, red (Section 3.4.3)	green, red	green, red	green, red (Section 4.1.3)
Movie for two (Section 4.1.1)	Palace, Corner, home (Section 4.1.1)	home (Section 4.2.4)	home	home (Section 4.1.3)
Opera for three (Section 4.1.3)	home, opera (Section 4.1.3)	home (Section 4.2.4)	home	home (Section 4.1.3)
Rock, paper, scissors (Section 4.2.1)	rock, paper, scissors, bomb	rock, paper, scissors, bomb (Section 4.2.4)	rock, paper, scissors (Section 4.3.3)	rock, paper, scissors (Question 4.4.1)
Masquerade ball (Section 4.3.4)	gladiator, emperor, lion	gladiator, emperor, lion	gladiator, emperor, lion (Section 4.3.4)	gladiator, lion (Section 4.3.4)

Table 4.4.2 The four concepts in the various examples

choose *rock, paper* and *scissors.* This will be the task in the following question.

Question 4.4.1 *In the example "Rock, paper, scissors", find a single Nash equilibrium in which each of your choices rock, paper and scissors is optimal.*

Hence, in light of Theorem 4.1.2, you can rationally choose *rock, paper* and *scissors* under common belief in rationality with a simple belief hierarchy.

Table 4.4.2 shows, among other things, that the four concepts may yield different optimal choices in a game. In "Movie for two", for instance, you can rationally choose *Palace* under common belief in rationality, but not under common belief in rationality with a symmetric belief hierarchy. In "Rock, paper, scissors", you can rationally choose *bomb* under common belief in rationality with a symmetric belief hierarchy, but not under common belief in rationality with a symmetric belief hierarchy that uses one theory per choice. In "The masquerade ball ", finally, you can rationally choose *emperor* under common belief in rationality with a symmetric belief hierarchy that uses one theory per choice, but not under common belief in rationality with a simple belief hierarchy. That is, the four concepts are fundamentally different not only in terms of the restrictions they impose on a belief hierarchy, but also when it comes to the optimal choices they induce.

4.5 Proofs

4.5.1 Proofs of Section 4.1

Proof of Theorem 4.1.1. See the arguments in Section 4.1.2. ■

Proof of Theorem 4.1.2. See the arguments in Section 4.1.2. ■

To prove Theorem 4.1.3, we need some additional definitions and a mathematical result known as *Kakutani's fixed point theorem* (Kakutani (1941)). This result states that, under some conditions, a correspondence is guaranteed to have a *fixed point.* Before we can state Kakutani's fixed point theorem formally, we must first define what we mean by a correspondence, a fixed point, and we must formally introduce the conditions assumed in the theorem.

Let X be some finite set, and let $A \subseteq \mathbf{R}^X$ be some subset of the linear space $\mathbf{R}^X$. Recall from Section 2.8.1.1 that $\mathbf{R}^X$ contains all vectors v that assign a number $v(x) \in \mathbf{R}$ to every element $x \in X$. A *correspondence* from A to A is a mapping F that assigns to every vector $a \in A$ a nonempty *set* of vectors $F(A) \subseteq A$. Hence, a *function* is a special case of a correspondence where $F(a)$ consists of a single vector for every $a \in A$. A vector $a^* \in A$ is called a *fixed point* of the correspondence F if $a^* \in F(a^*)$. That is, the image of the vector a^* under F contains the vector a^* itself. Not every correspondence F from A to A has a fixed point. The question we wish to answer is: Can we find conditions on the correspondence F which guarantee that F has at least one fixed point? Kakutani's fixed point theorem presents one such set of conditions. We will now present these conditions.

Recall from Section 2.8.5 that a set $A \subseteq \mathbf{R}^X$ is *convex* if for every two points $a, b \in A$, and every number $\lambda \in [0, 1]$, the convex combination $(1 - \lambda) \cdot a + \lambda \cdot b$ is also in A. Geometrically, the convex combination $(1 - \lambda) \cdot a + \lambda \cdot b$ is a point on the line segment between a and b. Hence, in geometric terms, a set A is convex if the line segment between any two points $a, b \in A$ is completely contained in A. We call the correspondence F from A to A *convex-valued* if $F(a)$ is a non-empty convex set for every $a \in A$. Recall from Section 2.8.5 what it means for the set A to be *closed* and *bounded.* A set $A \subseteq \mathbf{R}^X$ is *compact* if it is both *closed* and *bounded.*

Finally, the correspondence F from A to A is *upper-semicontinuous* if for every two sequences $(a^k)_{k\in\mathbf{N}}$ and $(b^k)_{k\in\mathbf{N}}$ in A the following holds: if $b^k \in F(a^k)$ for all k, the sequence $(a^k)_{k\in\mathbf{N}}$ converges to $a \in A$, and the sequence $(b^k)_{k\in\mathbf{N}}$ converges to $b \in A$, then $b \in F(a)$. If the correspondence F is a function, then upper-semicontinuity is just the same as continuity. We are now ready to present Kakutani's fixed point theorem.

Theorem 4.5.1 (Kakutani's fixed point theorem) *Let X be a finite set, and A a nonempty, compact and convex subset of $\mathbf{R}^X$. Moreover, let F be a correspondence from A to A which is convex-valued and upper-semicontinuous. Then, the correspondence F has at least one fixed point.*

The proof can be found in the original paper by Kakutani (1941), but also in books like Border (1985). We are now ready to prove Theorem 4.1.3.

Proof of Theorem 4.1.3. For every player i, let $\Delta(C_i)$ denote the set of probability distributions on C_i. So, every combination of beliefs $(\sigma_1, ..., \sigma_n)$ belongs to the set $\Delta(C_1) \times ... \times \Delta(C_n)$. By

$$A := \Delta(C_1) \times ... \times \Delta(C_n)$$

we denote the set of all such belief combinations. Hence, A is a subset of some linear space $\mathbf{R}^X$. Moreover, it may easily be verified that the set A is nonempty, compact and convex.

For every $(\sigma_1, ..., \sigma_n) \in A$ and every player i, let $C_i^{opt}(\sigma_1, ..., \sigma_n)$ be the set of choices $c_i \in C_i$ that are optimal under the belief σ_{-i}. By $\Delta(C_i^{opt}(\sigma_1, ..., \sigma_n))$ we denote the set of probability distributions in $\Delta(C_i)$ that only assign positive probability to choices in $C_i^{opt}(\sigma_1, ..., \sigma_n)$. Define now the correspondence C^{opt} from A to A, which assigns to every belief combination $(\sigma_1, ..., \sigma_n) \in A$ the set of belief combinations

$$C^{opt}(\sigma_1, ..., \sigma_n) := \Delta(C_1^{opt}(\sigma_1, ..., \sigma_n)) \times ... \times \Delta(C_n^{opt}(\sigma_1, ..., \sigma_n)),$$

which is a subset of $\Delta(C_1) \times ... \times \Delta(C_n)$, and hence is a subset of A.

It may easily be verified that the set $C^{opt}(\sigma_1, ..., \sigma_n)$ is nonempty and convex for every $(\sigma_1, ..., \sigma_n)$. It thus follows that the correspondence C^{opt} is convex-valued. We now show that the correspondence C^{opt} is upper-semicontinuous. That is, we must show that for every sequence $(\sigma_1^k, ..., \sigma_n^k)_{k\in\mathbf{N}}$ converging to some $(\sigma_1, ..., \sigma_n)$, and every sequence $(\hat{\sigma}_1^k, ..., \hat{\sigma}_n^k)_{k\in\mathbf{N}}$ converging to some $(\hat{\sigma}_1, ..., \hat{\sigma}_n)$ with $(\hat{\sigma}_1^k, ..., \hat{\sigma}_n^k) \in C^{opt}(\sigma_1^k, ..., \sigma_n^k)$ for every k, it holds that $(\hat{\sigma}_1, ..., \hat{\sigma}_n) \in C^{opt}(\sigma_1, ..., \sigma_n)$.

Suppose, contrary to what we want to prove, that $(\hat{\sigma}_1, ..., \hat{\sigma}_n) \notin C^{opt}(\sigma_1, ..., \sigma_n)$. Then, there is some player i such that $\hat{\sigma}_i$ assigns positive probability to some c_i, whereas c_i is not optimal under σ_{-i}. But then, if k is large enough, $\hat{\sigma}_i^k$ assigns positive probability to c_i, and c_i is not optimal under σ_{-i}^k. However, this contradicts the assumption that $(\hat{\sigma}_1^k, ..., \hat{\sigma}_n^k) \in C^{opt}(\sigma_1^k, ..., \sigma_n^k)$. So, we conclude that $(\hat{\sigma}_1, ..., \hat{\sigma}_n) \in C^{opt}(\sigma_1, ..., \sigma_n)$, and hence the correspondence C^{opt} is upper-semicontinuous.

Summarizing, we see that the set $A = \Delta(C_1) \times ... \times \Delta(C_n)$ is nonempty, compact and convex, and that the correspondence C^{opt} from A to A is upper-semicontinuous and convex-valued. By Kakutani's fixed point theorem, it then follows that C^{opt} has at least one fixed point $(\sigma_1^*, ..., \sigma_n^*) \in A$. That is, there is some $(\sigma_1^*, ..., \sigma_n^*) \in A$ with

$$(\sigma_1^*, ..., \sigma_n^*) \in C^{opt}(\sigma_1^*, ..., \sigma_n^*).$$

By definition of C^{opt} this means that for every player i, we have that $\sigma_i^* \in \Delta(C_i^{opt}(\sigma_1^*, ..., \sigma_n^*))$. So, for every player i, the probability distribution σ_i^* only assigns positive probability to choices c_i that are optimal under σ_{-i}^*. This means, however, that $(\sigma_1^*, ..., \sigma_n^*)$ is a Nash equilibrium. So, a Nash equilibrium always exists. ■

4.5.2 Proofs of Section 4.2

Proof of Theorem 4.2.1. (a) Suppose first that a belief hierarchy β_i is symmetric. We will show that β_i is induced by a common prior π on choice-type combinations.

Since β_i is symmetric, it is part of a beliefs diagram which is induced by a *symmetric* weighted beliefs diagram. Consider some arrow a from (c_i, t_i) to $(c_j, t_j)_{j\neq i}$ in this symmetric weighted beliefs diagram. Since the weighted beliefs diagram is symmetric, every symmetric counterpart to a is present in the weighted beliefs diagram, and has the same weight as a. That is, for every choice-type combination $(c, t) = (c_i, t_i)_{i\in I}$ we can find a unique weight, call it $w(c, t)$, such that every arrow from a choice-type pair in (c, t) to the opponents' choice-type combination in (c, t) receives the same weight $w(c, t)$. Here, we assume that every arrow that is not present in the weighted beliefs diagram receives weight zero.

Now, consider an arbitrary arrow a in the symmetric weighted beliefs diagram, from the choice-type pair (c_i, t_i) to an opponents' choice-type combination (c_{-i}, t_{-i}). In the induced beliefs diagram, the probability assigned to arrow a is then equal to

$$p(a) = \frac{w(a)}{\sum_{\text{arrows } a' \text{ leaving } (c_i,t_i)} w(a')}. \tag{4.5.1}$$

Note that $w(a) = w((c_i, t_i), (c_{-i}, t_{-i}))$. Moreover, every arrow a' leaving (c_i, t_i) is an arrow from (c_i, t_i) to a (possibly different) opponents' choice-type combination (c'_{-i}, t'_{-i}), which has weight $w((c_i, t_i), (c'_{-i}, t'_{-i}))$. Together with (4.5.1) this leads to the insight that

$$p(a) = \frac{w((c_i, t_i), (c_{-i}, t_{-i}))}{\sum_{(c'_{-i}, t'_{-i}) \in C_{-i} \times T_{-i}} w((c_i, t_i), (c'_{-i}, t'_{-i}))},$$

where T_{-i} is the set of opponents' type combinations. As an abbreviation, define

$$w(c_i, t_i) := \sum_{(c'_{-i}, t'_{-i}) \in C_{-i} \times T_{-i}} w((c_i, t_i), (c'_{-i}, t'_{-i}))$$

as the total weight assigned to player i's choice-type pair (c_i, t_i). Then, we have that

$$p(a) = \frac{w((c_i, t_i), (c_{-i}, t_{-i}))}{w(c_i, t_i)} \tag{4.5.2}$$

for every arrow a from a choice-type pair (c_i, t_i) to an opponents' choice-type combination (c_{-i}, t_{-i}).

Let $C \times T$ be the set of all choice combinations $(c_i, t_i)_{i \in I}$, and let

$$W := \sum_{(c,t) \in C \times T} w(c, t)$$

be the sum of the weights assigned to all choice-type combinations. Define new weights $\pi(c, t)$ by

$$\pi(c, t) := \frac{1}{W} \cdot w(c, t) \tag{4.5.3}$$

for all choice-type combinations (c, t). That is, we take the original weight $w(c, t)$ and divide it by the sum of all weights. What is special about these new weights $\pi(c, t)$ is that the sum of all weights is equal to 1. This follows from the observation that

$$\sum_{(c,t) \in C \times T} \pi(c, t) = \sum_{(c,t) \in C \times T} \frac{1}{W} \cdot w(c, t) = \frac{1}{W} \cdot \sum_{(c,t) \in C \times T} w(c, t) = \frac{1}{W} \cdot W = 1.$$

But then, the new weights $\pi(c, t)$ constitute a probability distribution over the choice-type combinations in $C \times T$, as every number $\pi(c, t)$ is non-negative, and the sum of all numbers is equal to 1. Hence, π is a common prior on choice-type combinations.

In the induced beliefs diagram, consider an arrow a from a choice-type pair (c_i, t_i) to an opponents' choice-type combination (c_{-i}, t_{-i}). By (4.5.2) and (4.5.3) we see that the probability of the arrow a is equal to

$$\begin{aligned}
p(a) &= \frac{w((c_i, t_i), (c_{-i}, t_{-i}))}{w(c_i, t_i)} = \frac{w((c_i, t_i), (c_{-i}, t_{-i}))}{\sum_{(c'_{-i}, t'_{-i}) \in C_{-i} \times T_{-i}} w((c_i, t_i), (c'_{-i}, t'_{-i}))} \\
&= \frac{W \cdot \pi((c_i, t_i), (c_{-i}, t_{-i}))}{\sum_{(c'_{-i}, t'_{-i}) \in C_{-i} \times T_{-i}} \big(W \cdot \pi((c_i, t_i), (c'_{-i}, t'_{-i}))\big)} \\
&= \frac{W \cdot \pi((c_i, t_i), (c_{-i}, t_{-i}))}{W \cdot \sum_{(c'_{-i}, t'_{-i}) \in C_{-i} \times T_{-i}} \pi((c_i, t_i), (c'_{-i}, t'_{-i}))} = \frac{\pi((c_i, t_i), (c_{-i}, t_{-i}))}{\sum_{(c'_{-i}, t'_{-i}) \in C_{-i} \times T_{-i}} \pi((c_i, t_i), (c'_{-i}, t'_{-i}))}.
\end{aligned}$$

Similarly to $w(c_i, t_i)$, we define

$$\pi(c_i, t_i) := \sum_{(c'_{-i}, t'_{-i}) \in C_{-i} \times T_{-i}} \pi((c_i, t_i), (c'_{-i}, t'_{-i}))$$

to be the total probability assigned by π to player i's choice-type pair (c_i, t_i). By the above, we then conclude that

$$p(a) = \frac{\pi((c_i, t_i), (c_{-i}, t_{-i}))}{\pi(c_i, t_i)} \tag{4.5.4}$$

for every arrow a from a choice-type pair (c_i, t_i) to an opponents' choice-type combination (c_{-i}, t_{-i}).

This means, however, that the beliefs diagram so constructed is induced by the common prior π on choice-type combinations. In particular, the belief hierarchy β_i we started with is induced by this common prior π. This completes the proof for part (a).

(b) Suppose now that the belief hierarchy β_i is induced by a common prior π on choice-type combinations. We will show that β_i is symmetric.

Note that, by definition, β_i is part of a beliefs diagram that is induced by the common prior π. We will show that this beliefs diagram is induced by a symmetric weighted beliefs diagram. To see this, take an arrow a from a choice-type pair (c_i, t_i) to an opponents' choice-type combination (c_{-i}, t_{-i}), present in this beliefs diagram. By part (b) in Definition 4.2.2, it must then be the case that $\pi((c_i, t_i), (c_{-i}, t_{-i})) > 0$. Hence, for every player j, the symmetric counterpart of a will then be present also, since $((c_j, t_j), (c_{-j}, t_{-j})) = ((c_i, t_i), (c_{-i}, t_{-i}))$ and hence $\pi((c_j, t_j), (c_{-j}, t_{-j})) > 0$. That is, for every arrow in the beliefs diagram, every symmetric counterpart of this arrow will be present as well. Now, construct a weighted beliefs diagram in which every arrow a from a choice-type pair (c_i, t_i) to an opponents' choice-type combination (c_{-i}, t_{-i}) receives the weight

$$w(a) := \pi((c_i, t_i), (c_{-i}, t_{-i})). \tag{4.5.5}$$

We have already seen that for every opponent j, the symmetric counterpart to arrow a is also present. Moreover, by (4.5.5), the symmetric counterpart to a receives the same weight $\pi(c, t)$ as a, where $(c, t) = ((c_i, t_i), (c_{-i}, t_{-i}))$. We therefore conclude that the weighted beliefs diagram so constructed is symmetric.

From (4.5.5) it follows that

$$\sum_{\text{arrows } a' \text{ leaving } (c_i, t_i)} w(a') = \sum_{(c'_{-i}, t'_{-i}) \in C_{-i} \times T_{-i}} \pi((c_i, t_i), (c'_{-i}, t'_{-i})) = \pi(c_i, t_i). \tag{4.5.6}$$

Consider an arrow a from a choice-type pair (c_i, t_i) to an opponents' choice-type combination (c_{-i}, t_{-i}). By part (b) in Definition 4.2.2, and equations (4.5.5) and (4.5.6), we know that the probability of arrow a in the beliefs diagram is

$$p(a) = \frac{\pi((c_i, t_i), (c_{-i}, t_{-i}))}{\pi(c_i, t_i)} = \frac{w(a)}{\sum_{\text{arrows } a' \text{ leaving } (c_i, t_i)} w(a')}.$$

This means, however, that the beliefs diagram is induced by the symmetric weighted beliefs diagram constructed above. Since the belief hierarchy β_i is part of this beliefs diagram, we conclude that β_i is symmetric. This completes the proof of part (b). ■

Proof of Theorem 4.2.2. See the arguments in Section 4.2.3. ■

Proof of Theorem 4.2.3. See the arguments in Section 4.2.3. ■

4.5.3 Proofs of Section 4.3

Proof of Theorem 4.3.1. See the arguments in Section 4.3.1. ■

Proof of Theorem 4.3.2. See the arguments in Section 4.3.2. ■

Proof of Theorem 4.3.3. See the arguments in Section 4.3.3. ■

Proof of Theorem 4.3.4. See the arguments in Section 4.3.3. ■

Proof of Theorem 4.3.5. Consider a simple belief hierarchy β_i^* for player i generated by a combination of beliefs $(\sigma_1, ..., \sigma_n)$. Design a weighted beliefs diagram as follows: Every choice c_i of a player i appears at most once. Moreover, choice c_i is present in the weighted beliefs diagram precisely when $\sigma_i(c_i) > 0$. For every choice c_i and opponents' choice combination $(c_j)_{j \neq i}$, include the arrow a from c_i to $(c_j)_{j \neq i}$ in the diagram precisely when $\sigma_i(c_i) > 0$ and $\sigma_j(c_j) > 0$ for every opponent $j \neq i$. Moreover, set the weight of this arrow a equal to

$$w(a) = \sigma_i(c_i) \cdot \prod_{j \neq i} \sigma_j(c_j) = \sigma_1(c_1) \cdot \sigma_2(c_2) \cdot ... \cdot \sigma_n(c_n). \tag{4.5.7}$$

It can then be shown that this weighted beliefs diagram is symmetric, and that it induces a beliefs diagram that the belief hierarchy β_i^* is part of.

To see why it is symmetric, consider an arbitrary arrow a in the weighted beliefs diagram, from a choice c_i to an opponents' choice combination $(c_j)_{j \neq i}$. Then, it must be the case that $\sigma_i(c_i) > 0$ and $\sigma_j(c_j) > 0$ for every opponent $j \neq i$. That is, $\sigma_k(c_k) > 0$ for all players k. But then, for every opponent j, the symmetric arrow from c_j to $(c_k)_{k \neq j}$ must also be in the diagram. Moreover, both a and each of its symmetric counterparts carry the same weight, which by (4.5.7) is

$$\sigma_1(c_1) \cdot \sigma_2(c_2) \cdot ... \cdot \sigma_n(c_n).$$

We thus conclude that the weighted beliefs diagram is symmetric.

Now consider the beliefs diagram induced by this symmetric weighted beliefs diagram. Since every choice appears at most once, all belief hierarchies use one theory per choice. Look at the belief hierarchy β_i for player i that starts at some arbitrary choice c_i in the beliefs diagram. We will see that this belief hierarchy β_i is exactly the simple belief hierarchy β_i^* for player i generated by $(\sigma_1, ..., \sigma_n)$ – the belief hierarchy we started from.

Since the belief hierarchy β_i starts at choice c_i, it must be the case that $\sigma_i(c_i) > 0$. Now, consider an arrow a from c_i to an opponents' choice combination $(\sigma_j)_{j \neq i}$ in the induced beliefs diagram. Then, we must have that $\sigma_j(c_j) > 0$ for all opponents $j \neq i$, and by (4.5.7) this arrow a carries the weight

$$w(a) = \sigma_1(c_1) \cdot \sigma_2(c_2) \cdot ... \cdot \sigma_n(c_n) \tag{4.5.8}$$

in the weighted beliefs diagram. Therefore, by part (d) in Definition 4.2.1, the probability $p(a)$ of this arrow is equal to

$$p(a) = \frac{w(a)}{\sum_{\text{arrows } a' \text{ leaving } c_i} w(a')}. \tag{4.5.9}$$

By construction, the arrows leaving c_i are precisely the arrows a' from c_i to some opponents' choice combination $(c'_j)_{j \neq i}$ where $\sigma_j(c'_j) > 0$ for all $j \neq i$, and every such arrow a' carries the weight

$$w(a') = \sigma_i(c_i) \cdot \prod_{j \neq i} \sigma_j(c'_j).$$

Therefore,

$$\begin{aligned}\sum_{\text{arrows } a' \text{ leaving } c_i} w(a') &= \sum_{(c'_j)_{j\neq i}\in C_{-i}\,:\,\sigma_j(c'_j)>0 \text{ for all } j\neq i} \left[\sigma_i(c_i)\cdot\prod_{j\neq i}\sigma_j(c'_j)\right] \\ &= \sum_{(c'_j)_{j\neq i}\in C_{-i}} \left[\sigma_i(c_i)\cdot\prod_{j\neq i}\sigma_j(c'_j)\right] \\ &= \sigma_i(c_i)\cdot\sum_{(c'_j)_{j\neq i}\in C_{-i}}\left[\prod_{j\neq i}\sigma_j(c'_j)\right] \\ &= \sigma_i(c_i)\cdot\prod_{j\neq i}\left[\sum_{c'_j\in C_j}\sigma_j(c'_j)\right] \\ &= \sigma_i(c_i)\cdot 1 = \sigma_i(c_i).\end{aligned}$$

In the second equality we use the fact that summing over choice combinations $(c'_j)_{j\neq i}$ with $\sigma_j(c'_j)=0$ for some $j\neq i$ does not change the result, because the term $\sigma_i(c_i)\cdot\prod_{j\neq i}\sigma_j(c'_j)$ will be 0 in this case. In the third equality we take the constant term $\sigma_i(c_i)$ outside the sum. The fourth equality is based on elementary algebraic manipulations, whereas in the fifth equality we use the fact that σ_j is a probability distribution on C_j and therefore $\sum_{c'_j\in C_j}\ \sigma_j(c'_j)=1$. Hence, we conclude that

$$\sum_{\text{arrows } a' \text{ leaving } c_i} w(a') = \sigma_i(c_i). \tag{4.5.10}$$

By combining (4.5.9), (4.5.8) and (4.5.10) we see that in the induced beliefs diagram, the arrow a from c_i to the opponents' choice combination $(\sigma_j)_{j\neq i}$ has probability

$$p(a) = \frac{w(a)}{\sum_{\text{arrows } a' \text{ leaving } c_i} w(a')} = \frac{\sigma_1(c_1)\cdot\sigma_2(c_2)\cdot\ldots\cdot\sigma_n(c_n)}{\sigma_i(c_i)} = \prod_{j\neq i}\sigma_j(c_j). \tag{4.5.11}$$

Since this is independent from the choice c_i we started at, we conclude that in the induced beliefs diagram, *every* arrow from a choice of player i to the opponents' choice combination $(\sigma_j)_{j\neq i}$ has the same probability $\prod_{j\neq i}\sigma_j(c_j)$.

In the induced beliefs diagram, let us start at an arbitrary choice c_i for player i with $\sigma_i(c_i)>0$, and let us keep following the arrows to derive the induced belief hierarchy β_i for player i. By (4.5.11) above, every arrow from c_i to an opponents' choice combination $(\sigma_j)_{j\neq i}$ has probability $\prod_{j\neq i}\sigma_j(c_j)$. Hence, in the first-order belief, player i assigns probability $\prod_{j\neq i}\sigma_j(c_j)$ to every opponents' choice combination $(\sigma_j)_{j\neq i}$, just as in the simple belief hierarchy β_i^* generated by $(\sigma_1,...,\sigma_n)$.

Consider an arrow from c_i to an arbitrary opponent's choice c_j with $\sigma_j(c_j)>0$. By (4.5.11), every arrow from c_j to an opponents' choice combination $(c_k)_{k\neq j}$ has probability $\prod_{k\neq j}\sigma_k(c_k)$. That is, in the second-order belief, player i believes that player j assigns to every opponents' choice combination $(\sigma_k)_{k\neq j}$ probability $\prod_{k\neq j}\sigma_k(c_k)$, just as in the simple belief hierarchy β_i^* generated by $(\sigma_1,...,\sigma_n)$.

By continuing in this fashion, we can also show that in the third-order belief, player i believes that every opponent j believes that every other player k assigns probability $\prod_{l\neq k}\sigma_l(c_l)$ to every opponents' choice combination $(\sigma_l)_{l\neq k}$, just as in the simple belief hierarchy β_i^* generated by $(\sigma_1,...,\sigma_n)$, and similarly for all higher-order beliefs.

In other words, in the beliefs diagram induced by the weighted beliefs diagram above, every belief hierarchy for player i is the same as the simple belief hierarchy β_i^* generated by $(\sigma_1, ..., \sigma_n)$. As a consequence, there is only *one* belief hierarchy for every player i in the induced beliefs diagram, which is the simple belief hierarchy β_i^* generated by $(\sigma_1, ..., \sigma_n)$. Since the weighted beliefs diagram above is symmetric, the simple belief hierarchy β_i^* generated by $(\sigma_1, ..., \sigma_n)$ can be generated from a *symmetric* weighted beliefs diagram, and is therefore symmetric. As we have seen above that all belief hierarchies in the beliefs diagram use one theory per choice, we know that β_i^* has one theory per choice as well. This completes the proof. ■

Proof of Corollary 4.3.1. See the arguments in Section 4.3.4. ■

Solutions to In-Chapter Questions

Question 4.1.1. Your second-order belief is that you believe that Barbara believes that you stay at *home.* Since you believe that Barbara believes that you *indeed* believe that Barbara believes that you stay at *home*, you believe that Barbara is correct about your second-order belief.

Question 4.1.2. In the belief hierarchy that starts at your choice *blue,* you believe that Barbara chooses *red,* and believe that Barbara believes that you *indeed* believe that Barbara chooses *red.* Hence, you believe that Barbara is correct about your first-order belief. In the belief hierarchy that starts at your choice *green,* you believe that Barbara chooses *blue,* but at the same time you believe that Barbara assigns probability 0.4 to the event that you believe that Barbara chooses *yellow.* Hence, you do not believe that Barbara is correct about your first-order belief. In the belief hierarchy that starts at your choice *red,* your first-order belief is that you assign probability 0.6 to Barbara choosing *blue* and probability 0.4 to Barbara choosing *green.* At the same time, you assign probability 0.4 to the event that Barbara believes that you assign probability 1 to Barbara choosing *blue.* Hence, you do not believe that Barbara is correct about your first-order belief. Consider finally the belief hierarchy that starts at your choice *yellow.* There, you believe that Barbara chooses *yellow,* but you believe that Barbara believes that you assign probability 0.6 to Barbara choosing *blue* and probability 0.4 to Barbara choosing *green.* Hence, you do not believe that Barbara is correct about your first-order belief.

Question 4.1.3. The probability you believe that Chris assigns to you wearing *green* and Barbara wearing *blue* is

$$\sigma_1(\mathit{green}) \cdot \sigma_2(\mathit{blue}) = 1 \cdot (0.3) = 0.3.$$

The probability you believe that Chris assigns to you wearing *green* and Barbara wearing *red* is

$$\sigma_1(\mathit{green}) \cdot \sigma_2(\mathit{red}) = 1 \cdot (0.7) = 0.7.$$

You believe that Chris assigns probability zero to all other choice combinations by you and Barbara. You also believe that Barbara believes that Chris has exactly this first-order belief.

Question 4.1.4. Player i believes that j has belief σ_k about k's choice. At the same time, i believes that j believes that every other player l believes that j has belief σ_k about k's choice.

Question 4.1.5. Suppose that the probability distribution $p = (0.5) \cdot (\mathit{blue, blue}) + (0.5) \cdot (\mathit{red, yellow})$ could be written as the product of a probabilistic belief σ_2 about Barbara's choice and a probabilistic belief σ_3 about Chris' choice. Since p assigns positive probability to (*blue, blue*) and (*red, yellow*), the belief σ_2 must assign positive probability to Barbara's choices *blue* and *red,* and the belief σ_3 must assign positive probability to Chris' choices *blue* and *yellow.* But then, $p(\mathit{blue, yellow}) = \sigma_2(\mathit{blue}) \cdot \sigma_3(\mathit{yellow}) > 0$, which is a contradiction.

Question 4.1.6. Consider first your belief hierarchy that starts at your choice *green.* That belief hierarchy is simple because it is generated by the combination of beliefs ($\sigma_1 = \mathit{green}$, $\sigma_2 = \mathit{blue}$, $\sigma_3 = \mathit{yellow}$). Consider next Barbara's belief hierarchy that starts at her choice *yellow.* In that belief hierarchy, Barbara assigns probability 0.3 to you choosing *red* and probability 0.7 to you choosing *green.* At the same time, she believes that Chris assigns probability 1 to you choosing *red.* Hence, Barbara believes that Chris has a different belief about your choice as she has herself. Therefore, this belief hierarchy cannot be simple.

Question 4.1.7. Consider an opponent j of player i, and an opponent k of player j. Note that k may be i. Suppose that i believes that j assigns positive probability to k's choice c_k. Then, $\sigma_k(c_k) > 0$, as i believes that j has the belief σ_k about k's choice. Since $(\sigma_1, ..., \sigma_n)$ is a Nash equilibrium, the choice c_k must be optimal for player k under the belief σ_{-k}. Note that i believes that j believes that k has belief σ_{-k} about his opponents' choices. Hence, we conclude that if i believes that j assigns positive probability to k's choice c_k, then c_k is optimal for k given what i believes that j believes that k believes about his opponents' choices. In other words, i believes that j believes in k's rationality. As this holds for every opponent k of j, we conclude that i believes that j believes in his opponents' rationality.

Question 4.1.8. We must show that ($\sigma_1 = \frac{1}{2}\cdot$ *green* $+\frac{1}{2}\cdot$ *red,* $\sigma_2 = \frac{1}{3}\cdot$ *green* $+\frac{2}{3}\cdot$ *yellow*, $\sigma_3 =$ *blue*) is a Nash equilibrium. Note that σ_1 assigns positive probability to your choices *green* and *red.* Hence, we must verify that your choices *green* and *red* are both optimal under the belief σ_{-1}. The expected utilities of your three possible choices under the belief σ_{-1} are $u_1(\textit{green}) = \frac{2}{3} \cdot 3 = 2$, $u_1(\textit{red}) = 2$ and $u_1(\textit{yellow}) = \frac{1}{3} \cdot 1 = \frac{1}{3}$. Hence, indeed, your choices *green* and *red* are both optimal under the belief σ_{-1}. The belief σ_2 assigns probability $\frac{1}{3}$ to Barbara's choice *green* and probability $\frac{2}{3}$ to Barbara's choice *yellow.* We must verify that *green* and *yellow* are optimal for Barbara under the belief σ_{-2}. The expected utilities for Barbara's four choices under the belief σ_{-2} are $u_2(\textit{blue}) = 0$, $u_2(\textit{green}) = \frac{1}{2} \cdot 4 = 2$, $u_2(\textit{red}) = \frac{1}{2} \cdot 1 = \frac{1}{2}$ and $u_2(\textit{yellow}) = 2$. Hence, *green* and *yellow* are indeed optimal for Barbara under the belief σ_{-2}. Finally, the belief σ_3 assigns probability 1 to Chris' choice *blue.* Hence, we must verify that *blue* is optimal for Chris under the belief σ_{-3}. The expected utilities for both of Chris' choices under the belief σ_{-3} are $u_3(\textit{blue}) = 2$ and $u_2(\textit{yellow}) = \frac{1}{3} \cdot 1 = \frac{1}{3}$. Hence, *blue* is optimal for Chris under the belief σ_{-3}. We thus conclude that $(\sigma_1, \sigma_2, \sigma_3)$ is a Nash equilibrium. Above we have already shown that for you, both *green* and *red* are optimal under the belief σ_{-1}, and hence both *green* and *red* are optimal for you in this Nash equilibrium.

Question 4.2.1. Note that in the belief hierarchy that starts at your choice *red,* there is an arrow from your choice *red* to Barbara's choice *green,* but there is no "symmetric" arrow from Barbara's choice *green* to your choice *red.* Hence, there can be no symmetric weighted beliefs diagram that generates this belief hierarchy.

Question 4.2.2. Note first that every choice only appears once in the beliefs diagram. Hence, for the choice-type representation of the beliefs diagram we only need one type for every choice. The accociated sets of types thus become $T_1 = \{t_1^g, t_1^r\}$ for you, $T_2 = \{t_2^b, t_2^y, t_2^g\}$ for Barbara and $T_3 = \{t_3^y, t_3^b\}$ for Chris. We know that the beliefs diagram is induced by the symmetric weighted beliefs diagram in Figure 4.2.2. The weights in this symmetric weighted beliefs diagram induce the following weights on choice-type combinations:

$$\begin{aligned} w((g,t_1^g),(b,t_2^b),(y,t_3^y)) &= 4, \; w((g,t_1^g),(y,t_2^y),(b,t_3^b)) = 4, \\ w((r,t_1^r),(g,t_2^g),(b,t_3^b)) &= 1, \; w((r,t_1^r),(g,t_2^g),(y,t_3^y)) = 1. \end{aligned}$$

The sum of all weights is therefore 10. The induced common prior on choice-type combinations π is

$$\begin{aligned} \pi((g,t_1^g),(b,t_2^b),(y,t_3^y)) &= 4/10, \; \pi((g,t_1^g),(y,t_2^y),(b,t_3^b)) = 4/10, \\ \pi((r,t_1^r),(g,t_2^g),(b,t_3^b)) &= 1/10, \; \pi((r,t_1^r),(g,t_2^g),(y,t_3^y)) = 1/10. \end{aligned}$$

This common prior on choice-type combinations π induces the beliefs diagram from Figure 4.2.2.

Question 4.2.3. We know that this belief hierarchy is symmetric, because it is induced by the symmetric weighted beliefs diagram in Figure 4.2.2. Moreover, the belief hierarchy satisfies common belief in rationality because all arrows in the beliefs diagram are solid. Hence, we conclude by Theorem 4.2.2 that this belief hierarchy is induced by some correlated equilibrium π. It remains to find a correlated equilibrium that induces this belief hierarchy. We have seen in Question 4.2.2 that all belief hierarchies in Figure 4.2.2 are induced by the common prior on choice-type combinations π, where

$$\begin{aligned} \pi((g,t_1^g),(b,t_2^b),(y,t_3^y)) &= 4/10,\ \pi((g,t_1^g),(y,t_2^y),(b,t_3^b)) = 4/10, \\ \pi((r,t_1^r),(g,t_2^g),(b,t_3^b)) &= 1/10,\ \pi((r,t_1^r),(g,t_2^g),(y,t_3^y)) = 1/10. \end{aligned}$$

This common prior π on choice-type combinations is in fact a correlated equilibrium. To see why, we can verify all the optimality conditions. We check only one optimality condition here, as checking the other conditions can be done in a similar fashion.

Consider your choice-type pair (g,t_1^g) with $\pi(g,t_1^g)>0$. Conditional on (g,t_1^g), the induced belief $\pi(\cdot\mid g,t_1^g)$ on the opponents' choice-type combinations is given by

$$\begin{aligned} \pi((b,t_2^b),(y,t_3^y))\mid g,t_1^g) &= \frac{\pi((g,t_1^g),(b,t_2^b),(y,t_3^y))}{\pi(g,t_1^g)} = \frac{4/10}{4/10+4/10} = 0.5 \text{ and} \\ \pi((y,t_2^y),(b,t_3^b))\mid g,t_1^g) &= \frac{\pi((g,t_1^g),(y,t_2^y),(b,t_3^b))}{\pi(g,t_1^g)} = \frac{4/10}{4/10+4/10} = 0.5. \end{aligned}$$

Hence, the conditional belief $\pi(\cdot\mid g,t_1^g)$ assigns probability 0.5 to the opponents' choice combinations (b,y) and (y,b). Since the expected utilities of your three choices under this belief are given by

$$\begin{aligned} u_1(g,\ \pi(\cdot\mid g,t_1^g)) &= (0.5)\cdot 3+(0.5)\cdot 3 = 3, \\ u_1(r,\ \pi(\cdot\mid g,t_1^g)) &= (0.5)\cdot 2+(0.5)\cdot 2 = 2 \text{ and} \\ u_1(y,\ \pi(\cdot\mid g,t_1^g)) &= (0.5)\cdot 0+(0.5)\cdot 0 = 0, \end{aligned}$$

it follows that your choice *green* is optimal for your belief $\pi(\cdot\mid g,t_1^g)$.

In a similar way, the other optimality conditions for you, Barbara and Chris may be verified. We thus see that π satisfies all optimality conditions, and is therefore a correlated equilibrium. Moreover, this correlated equilibrium π induces all belief hierarchies from Figure 4.2.2, and hence in particular the belief hierarchy for you that starts at your choice *green.*

Question 4.2.4. Consider a single type t_1^{home} for you and a single type t_2^{home} for Barbara. Then, the common prior π that assigns probability 1 to the choice-type combination ((*home*, t_1^{home}), (*home*, t_2^{home})) induces your symmetric belief hierarchy that starts at your choice *home.*

Question 4.2.5. Consider your belief hierarchy that starts at your choice *Palace.* There is an arrow from your choice *Palace* to Barbara's choice *Palace.* However, the symmetric counterpart to it, which is the arrow from Barbara's choice *Palace* to your choice *Palace,* is not even present. Hence, this belief hierarchy is not symmetric. A similar argument holds for your belief hierarchy that starts at your choice *Corner.*

Question 4.2.6. Consider a single type t_1^{home} for you, a single type t_2^{home} for Barbara and a single type t_3^{home} for Chris. Then, the common prior π that assigns probability 1 to the choice-type combination ((*home*, t_1^{home}), (*home*, t_2^{home}), (*home*, t_3^{home})) induces your symmetric belief hierarchy that starts at your choice *home.*

Question 4.2.7. In the belief hierarchy, there is an arrow from your choice *opera* to Barbara's and Chris' choice combination (*opera, opera*). However, the symmetric counterpart for Barbara, which

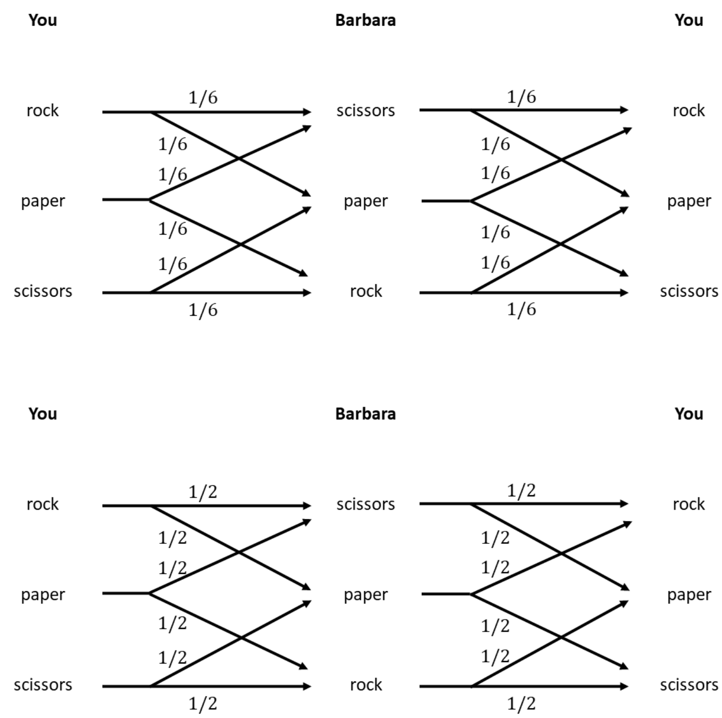

Figure 4.5.1 A symmetric weighted beliefs diagram, and induced beliefs diagram, for Question 4.3.2

would be the arrow from Barbara's choice *opera* to your and Chris' choice combination (*opera, opera*), is not even present. Hence, this belief hierarchy is not symmetric.

Question 4.2.8. We have seen that in every correlated equilibrium we must have that $\pi(home_1) = 1$. Therefore, in every correlated equilibrium, the conditional belief $\pi(\cdot \mid (c_3, t_3))$ for Chris must always assign probability 1 to you staying at *home.* Hence, only *home* can be optimal for Chris for every conditional belief $\pi(\cdot \mid (c_3, t_3))$. Since π is a correlated equilibrium, π can only assign positive probability to choice-type pairs (c_3, t_3) for Chris where c_3 is optimal for $\pi(\cdot \mid (c_3, t_3))$. But then, π can only assign positive probability to choice-type pairs $(home_3, t_3)$ for Chris, and hence $\pi(home_3) = 1$.

Question 4.3.1. Make a beliefs diagram in choice-type representation, where the choice-type pairs are exactly the pairs $(c_i, t_i^{c_i})$ such that c_i survives the iterated elimination of strictly dominated choices, and $t_i^{c_i}$ is the unique type that corresponds to c_i. Every arrow from $(c_i, t_i^{c_i})$ to an opponents' choice-type combination $(c_j, t_j^{c_j})_{j \neq i}$ has probability $b_i(t_i^{c_i})((c_j, t_j^{c_j})_{j \neq i})$. Then, this beliefs diagram induces, for every choice-type pair $(c_i, t_i^{c_i})$, exactly the same belief hierarchy as $t_i^{c_i}$ induces in the epistemic model. Moreover, every choice c_i only appears once in this beliefs diagram, as it only appears in the unique choice-type pair $(c_i, t_i^{c_i})$. Therefore, all belief hierarchies in this beliefs diagram use one theory per choice.

Question 4.3.2. Consider the symmetric weighted beliefs diagram, and the induced beliefs diagram, in Figure 4.5.1. Since the beliefs diagram is induced by a symmetric weighted beliefs diagram, all belief hierarchies are symmetric. As every choice only appears once, all belief hierarchies use one theory per choice. Finally, all belief hierarchies express common belief in rationality because all arrows are solid.

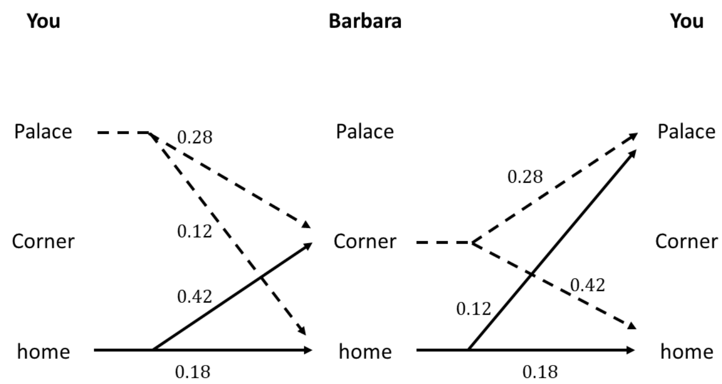

Figure 4.5.2 A symmetric weighted beliefs diagram for Question 4.3.3

Question 4.3.3. See Figure 4.5.2. Like in a usual beliefs diagram, we use solid and dashed arrows. If the arrow starts at a choice c_i, and c_i is not optimal for the belief represented by the arrow, then we use a dashed arrow instead of a solid arrow. It may be verified that this weighted beliefs diagram is symmetric. The two belief hierarchies for you, starting at your choices *Palace* and *home,* are identical, as they are both equal to the simple belief hierarchy generated by ($\sigma_1 = (0.4)\cdot$ *Palace* $+(0.6)\cdot$ *home*, $\sigma_2 = (0.7)\cdot$ *Corner* $+(0.3)\cdot$ *home*).

Question 4.3.4. This canonical correlated equilibrium $\hat{\pi}$ is given by

$$\begin{array}{rcl} \hat{\pi}(gladiator_1,\ gladiator_2) & = & 8/14,\ \hat{\pi}(emperor_1,\ gladiator_2) = 2/14, \\ \hat{\pi}(emperor_1,\ emperor_2) & = & 3/14 \text{ and } \hat{\pi}(lion_1,\ lion_2) = 1/14\ . \end{array}$$

Question 4.3.5. Your choice *gladiator* is optimal in the Nash equilibrium ($\sigma_1 =$ *gladiator*, $\sigma_2 =$ *gladiator*), and your choice *lion* is optimal in the Nash equilibrium ($\sigma_1 =$ *lion*, $\sigma_2 =$ *lion*).

Question 4.4.1. It may be verified that the combinations of beliefs (σ_1, σ_2), where

$$\sigma_1 = \tfrac{1}{3} \cdot rock + \tfrac{1}{3} \cdot paper + \tfrac{1}{3} \cdot scissors$$

and

$$\sigma_2 = \tfrac{1}{3} \cdot rock + \tfrac{1}{3} \cdot paper + \tfrac{1}{3} \cdot scissors$$

is a Nash equilibrium. Your choices *rock, paper* and *scissors* are all optimal under the belief $\sigma_{-1} = \sigma_2$.

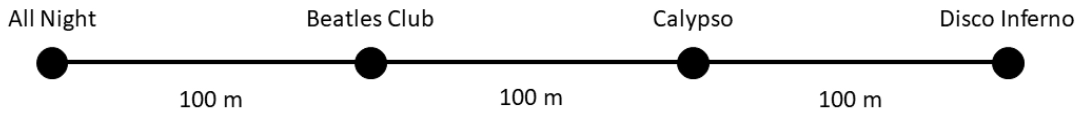

Figure 4.5.3 A map of the dancing places in "Dancing with Barbara"

Problems

Problem 4.1: Dancing with Barbara.

This evening, both Barbara and you feel like dancing. In town there are four places which are perfect for dancing: *All Night, Beatles Club, Calypso* and *Disco Inferno*. In Figure 4.5.3 you find a map of these four places. As you can see on the map, the distance between the two closest dancing places is always 100 meters.

Both you and Barbara can choose between staying at home and studying for the epistemic game theory exam, or going to one of these four places for an endless evening of dancing, not knowing where the other person will go. The problem, however, is that you are a terrible dancer whereas Barbara is a real dancing queen. For that reason, you would like to be as close as possible to Barbara when you go out dancing, whereas Barbara would like to be as far away from you as possible.

More precisely, if both you and Barbara go out dancing, then your utility would be 300 minus the distance (in meters) between the two places you and Barbara go to. Barbara's utility, on the other hand, would be exactly the distance (in meters) between these two places.

If you stay at home and study for the exam your utility will be 250. However, if you go out dancing and Barbara stays at home studying for the exam, then you feel guilty and your utility will be 0. The same holds for Barbara.

(a) Write down the decision problems for you and Barbara.

(b) Find all the choices that you and Barbara can rationally make under common belief in rationality. Which procedure do you use?

(c) Construct a beliefs diagram for this game with only solid arrows, using all of the choices you found in (b). For every belief hierarchy of you in this beliefs diagram, discuss whether it (i) expresses common belief in rationality, (ii) is symmetric, (iii) uses one theory per choice, and (iv) whether it is simple or not. Motivate your answers.

(d) Construct an epistemic model for this game such that for every choice c_i found in (b) there is a type t_i in the epistemic model that expresses common belief in rationality, and for which c_i is optimal.

(e) Explain why, in general, every correlated equilibrium can only assign positive probability to choices that can rationally be made under common belief in rationality.

(f) Show that there is only one choice you can rationally make under common belief in rationality with a symmetric belief hierarchy. (**Hint:** Use part (e)). For this choice, find a symmetric belief hierarchy that expresses common belief in rationality and for which that choice is optimal. To show that the belief hierarchy is symmetric, construct a symmetric weighted beliefs diagram that induces this belief hierarchy.

(g) Find all the choices you can rationally make under common belief in rationality with a simple belief hierarchy, and those you can rationally make under common belief in rationality with a symmetric belief hierarchy using one theory per choice.

Problem 4.2: The flat tyre.

You and Barbara are driving to a ballet performance, which starts at 8.00 pm. Today is not your lucky day, since at 7.30 pm you get a flat tyre, just in front of Deborah's house. The question is: What should you do? Essentially, there are three options for you and Barbara: You can either call Deborah, or call Chris, who lives a few kilometers away, on you can decide to change the wheel yourself. Since you must act quickly, there is no time for long discussions, and you both decide independently what to do.

If you both call Deborah, or both call Chris, then this person will come and change the wheel. If you decide to do it yourself and Barbara calls another person, then this person will come for nothing since you will have changed the wheel by the time the person has arrived. Similarly if Barbara decides to change the wheel herself and you call another person. If you both want to do it yourself, then you will insist so long until Barbara gives in and lets you change the wheel. In all of these occasions, the car will be fixed and you will both make it in time for the ballet performance. However, if you call Deborah and Barbara calls Chris, or the other way around, then Deborah and Chris will eventually see each other at your car and leave, because they have had a fierce argument last week. In that case, you will definitely be too late for the performance.

There are some situations here that make one of you angry. If you both call Deborah and she comes to fix the car, then you know by experience that she will only talk to Barbara, which makes you angry. Moreover, if you call Chris and Barbara decides to change the wheel herself, then you feel angry because Chris has come all the way for nothing. Similarly, if Barbara calls Chris and you fix the car yourself, then Barbara will be angry because Chris has come all the way for nothing. Remember that Deborah lives exactly where the accident happened, hence if she has to come for nothing it is not such a big deal, and nobody will be angry. However, if you both want to change the wheel yourself, then Barbara will be angry because you will not let her.

The utilities for you are as follows: If you make it on time for the ballet performance, then this will increase your utility by 3. Changing the wheel yourself is quite tiring and stressful, and will decrease your utility by 1, whereas getting angry will decrease your utility by 3. For Barbara the utilities are similar.

(a) Write down the decision problems for you and Barbara.

(b) Which choices can you and Barbara rationally make under common belief in rationality?

(c) Make a beliefs diagram in which only the choices found in (b) appear, and in which all arrows are solid. Which belief hierarchies for you express common belief in rationality? Which are symmetric? Which are simple? Which use one theory per choice?

(d) Show that each of your choices found in (b) can rationally be made under common belief in rationality with a symmetric belief hierarchy. To show this, construct for each of these choices a symmetric belief hierarchy that expresses common belief in rationality and for which that choice is optimal. Represent these belief hierarchies within the same beliefs diagram. Moreover, construct a symmetric weighted beliefs diagram that induces each of these symmetric belief hierarchies.

(e) Find a correlated equilibrium that induces each of the symmetric belief hierarchies constructed in (d).

	Athens	*Rome*	*Toledo*
You	3	4	8
Barbara	8	4	3
Chris	8	3	1

Table 4.5.1 Utilities in "The summer holiday"

(f) Which of the symmetric belief hierarchies constructed in (d) use one theory per choice? For each of these, construct a canonical correlated equilibrium that induces it.

(g) Show that one of your choices found in (b) cannot rationally be made under common belief in rationality with a simple belief hierarchy. Which choice is it?

(h) For each of your other choices found in (b), construct a simple belief hierarchy that expresses common belief in rationality and for which that choice is optimal. For each of these simple belief hierarchies, find the Nash equilibrium that generates it.

Problem 4.3: The summer holiday.

Next summer, Barbara, Chris and you plan to go on a holiday together, visiting a historical city in southern Europe. Now it is time to agree on the final destination. The latest shortlist contains three candidates: *Athens, Rome* and *Toledo*. The last city has been suggested by you, because you have read some wonderful stories about it. Indeed, Toledo is a delightful medieval city close to Madrid, but only few people know about it. The utilities that you, Barbara and Chris enjoy when visiting these three cities are given by Table 4.5.1. Moreover, staying at home would give you a utility of 2, and similarly for Barbara and Chris. As you can see, Barbara and Chris are both rather sceptical about Toledo, mainly because they are not familiar with this city. Chris even prefers to stay at home rather than going there. In contrast, you would love to see the impressive medieval city walls of Toledo with your own eyes.

To decide where to go, you have agreed on the following voting procedure: Each person secretly writes down one of the destinations on a piece of paper. If all three persons vote for the same city, you will all go there next summer and spend the holiday together. If two persons vote for the same city, and a third person votes for a different city, then these two persons will go on holiday in the city they voted for, whereas the other person will sadly stay at home. If all three persons vote for a different city, then the holiday destination will be decided by a dice roll: If the dice lands on 1 or 2, you will all go to Athens. If the result is 3 or 4, the destination will be Rome. If, finally, the dice lands on 5 or 6, you will all visit Toledo.

(a) Write down the decision problems for you, Barbara and Chris.

(b) Which cities can you, Barbara and Chris rationally vote for under common belief in rationality?

(c) Set up a beliefs diagram with solid arrows only, in which each of the choices found in (b) has an outgoing arrow. For each of your belief hierarchies in this diagram, discuss whether it (i) expresses common belief in rationality, (ii) is symmetric, (iii) satisfies the one theory per choice condition, and (iv) whether it is simple.

(d) Translate this beliefs diagram into an epistemic model.

(e) Show that each of your choices found in (b) can rationally be made under common belief in rationality with a symmetric belief hierarchy. That is, for each of these choices construct a symmetric belief hierarchy that expresses common belief in rationality and for which that choice is optimal. For every symmetric belief hierarchy so constructed, give an associated symmetric weighted beliefs diagram that generates it. Which of these symmetric belief hierarchies satisfy the one theory per choice condition?

(f) For every symmetric belief hierarchy constructed in (e), find a correlated equilibrium that induces it. In case the belief hierarchy satisfies the one theory per choice condition, find also a canonical correlated equilibrium that induces it.

(g) Show that there is exactly one choice for you found in (b) that *cannot* rationally be made under common belief in rationality with a simple belief hierarchy. For the other choices of you, construct a simple belief hierarchy expressing common belief in rationality for which that choice is optimal.
(**Hint:** For the first part of the question, proceed by showing the following steps. (1) Show that, under common belief in rationality, voting for *Rome* can only be optimal for Barbara if she assigns at least probability 0.6 to the event that Chris votes for *Rome* as well. (2) Show that, under common belief in rationality, voting for *Rome* can only be optimal for Chris if he assigns at least probability 0.8 to the event that Barbara votes for *Rome* as well. (3) Show that, under common belief in rationality with a simple hierarchy, you will either (i) assign probability 0 to the event that Barbara or Chris will vote for *Rome*, or (ii) assign at least probability 0.48 to the event that both Barbara and Chris will vote for *Rome*.)

(h) Explain intuitively why you cannot rationally make this choice under common belief in rationality with a simple belief hierarchy.

Literature

Simple belief hierarchies. The concept of simple belief hierarchies has been borrowed from Perea (2012), the predecessor to this book. It also appears in Geanakoplos, Pearce and Stacchetti (1989), although they do not employ the term simple belief hierarchies. They use it to define the concept of a *psychological Nash equilibrium* for psychological games. See Chapter 9 of this book for more details.

We use simple belief hierarchies to formalize the idea that a player believes that his opponents are correct about the belief hierarchy he holds. In fact, Perea (2007) and Perea (2012) show that this condition essentially *characterizes* simple belief hierarchies in two-player games. More precisely, Lemma 4.4 in Perea (2007) and Theorem 4.4.3 in Perea (2012) state that in a two-player game, player i holds a simple belief hierarchy precisely when he believes that his opponent is correct about his belief hierarchy, and believes that his opponent believes that he (player i) is correct about the opponent's belief hierarchy.

Lemma 4.4 in Perea (2007) and Theorem 4.4.5 in Perea (2012) also indicate which additional conditions are needed to epistemically characterize simple belief hierarchies in games with more than two players. For such games, simple belief hierarchies can be characterized by imposing, next to the two conditions above, the following requirements: (a) player i holds *conditionally independent beliefs* (Brandenburger and Friedenberg (2008)) about the choices of the various opponents, that is, holds independent beliefs about the choices of his opponents if we condition on a fixed belief hierarchy for each of the opponents, (b) player i believes, for every two different opponents j and k, that j holds the same belief about k as i does, and (c) player i believes that every opponent satisfies the conditions (a) and (b). In other words, not only does the concept of a simple belief hierarchy *imply* that player i believes that his opponents are correct about his beliefs, that player i holds conditionally independent beliefs, and that player i believes that j holds the same belief about k as i does, but these three conditions, and the belief that the opponents satisfy these conditions, also epistemically *characterize* the notion of a simple belief hierarchy.

Nash equilibrium. The notion of Nash equilibrium has been introduced by Nash (1950, 1951), and has had an enormous influence on the development of game theory during many decades. Indeed, for a very long time the concept of Nash equilibrium has been the central concept for studying the behavior of players in a static game. Moreover, it has been the basis for several *refinements* of Nash equilibrium in the literature, such as *perfect equilibrium* (Selten (1975)) and *proper equilibrium* (Myerson (1978)) for static games, and *subgame perfect equilibrium* (Selten (1965)) and *sequential equilibrium* (Kreps and Wilson (1982)) for dynamic games. At the same time, it remained rather unclear for several decades which assumptions Nash equilibrium, and each of its refinements, imposes on the reasoning of players. These assumptions were only revealed during the rise of *epistemic game theory* as a discipline that explicitly studies the beliefs and reasoning of players before they make a choice. As we have seen in this chapter, many of these assumptions seem highly problematic. For instance, how reasonable is it to assume that your opponents are correct about the beliefs that you have? Nevertheless, this is what is implicitly assumed if we use Nash equilibrium, or any of its refinements, to analyze a game.

Weakening common belief in rationality. In Theorem 4.1.1 we have shown that a simple belief hierarchy expresses common belief in rationality precisely when it is induced by a Nash equilibrium. In other words, Nash equilibrium can be epistemically characterized by a simple belief hierarchy in combination with common belief in rationality. However, for proving this result we only used the first two layers of common belief in rationality. Indeed, for showing that the simple belief hierarchy is induced by a Nash equilibrium, we only used that player i believes in his opponents' rationality, and that player i believes that his opponents believe in i's rationality. We can therefore alternatively characterize Nash equilibrium by a simple belief hierarchy in combination with the first two layers of common belief in rationality only.

This insight, together with Theorem 4.1.1, leads to the following observation: If a simple belief hierarchy satisfies the first two layers of common belief in rationality, then it also satisfies all the other layers of common belief in rationality. This, of course, is not true if we do not assume a simple belief hierarchy to begin with.

Epistemic foundations for Nash equilibrium in two-player games. We have seen above that Nash equilibrium can be characterized by the following three conditions: (a) player i's belief hierarchy is simple, (b) player i believes in the opponents' rationality, and (c) player i believes that all other players believe in their opponents' rationality (and hence, in particular, believe in i's rationality). We have also seen that in a two-player game, a simple belief hierarchy can be characterized by the following two conditions: (a1) player i believes that his opponent is correct about his belief hierarchy, and (a2) player i believes that his opponent believes that i is correct about the opponent's belief hierarchy.

Altogether, we conclude that in a two-player game the concept of Nash equilibrium can be epistemically characterized by the following conditions: (a1) player i believes that j is correct about his beliefs, and (a2) player i believes that j believes that i is correct about j's beliefs, (b) player i believes in j's rationality, and (c) player i believes that j believes in i' rationality. This is essentially the content of the "intrapersonal theorem" of Spohn (1982) on page 253 – the first epistemic characterization of Nash equilibrium I am aware of. In this intrapersonal theorem, Spohn adopts a *one-person perspective* like we do in this book, by imposing all epistemic conditions on the beliefs of a single player. The intrapersonal theorem of Spohn corresponds exactly to Perea's (2007) characterization of Nash equilibrium, when restricted to two-player games.

Spohn also offers an "interpersonal" variant of his theorem, in which the epistemic conditions are imposed on the beliefs of *both* players simultaneously. More precisely, he shows that in a two-player game, Nash equilibrium can be characterized by the following interpersonal conditions: (1) player 1 is rational, (2) player 2 is rational, (3) player 1 has belief σ_2 about 2's choice, (4) player 2 has belief σ_1 about 1's choice, (5) player 1 believes (2) and (4), and (6) player 2 believes (1) and (3). This result also appears as Theorem A in Aumann and Brandenburger (1995).

These conditions need not imply common belief in rationality. Hence, this is different from our insight above, where we have seen that a *simple* belief hierarchy in combination with the first two layers of common belief in rationality implies all other layers of common belief in rationality. The reason is that the conditions (3), (4), (5) and (6) do not guarantee that the belief hierarchies for both players are simple.

Polak (1999) shows, however, that if in the theorem above the conditions are strengthened so that there is *common* belief in the first-order beliefs σ_2 and σ_1, then these conditions *would* imply common belief in rationality. The argument is that these new, stronger conditions *would* imply a simple belief hierarchy for both players, namely the belief hierarchy generated by (σ_1, σ_2). Brandenburger and Dekel (1989) provide an epistemic characterization of Nash equilibrium that goes in this direction: They show for two-player games that common belief in rationality, together with common belief in the first-order beliefs (σ_1, σ_2), guarantee that (σ_1, σ_2) is a Nash equilibrium, and *vice versa*. This result is thus very closely related to our epistemic characterization of Nash equilibrium in Theorem 4.1.1, where we have replaced "common belief in (σ_1, σ_2)" by "the simple belief hierarchy generated by (σ_1, σ_2)".

Other epistemic characterizations of Nash equilibrium for two-player games can be found in Asheim (2006, p.5) and Tan and Werlang (1988, Theorem 6.2.1).

Epistemic foundations for Nash equilibrium in games with more than two players. As we have seen above, a simple belief hierarchy in a game with more than two players can be characterized by the following conditions: (a) player i believes that the opponents are correct about his belief hierarchy, (b) player i holds conditionally independent beliefs about the choices of his various opponents, (c) player i believes that j holds the same belief about k's choice as i does, and (d) player i believes that every opponent satisfies (a), (b) and (c). Moreover, we have argued that Nash equilibrium can be characterized by a simple belief hierarchy and the first two layers of common belief in rationality. Put together, we conclude that Nash equilibrium in games with more than two players can be characterized by the conditions (a)–(d) above, together with the following two rationality restrictions: (e) player i believes in the opponents' rationality, and (f) player i believes that every opponent believes in the other players' rationality. This is essentially the message of Corollary 4.6 in Perea (2007).

Aumann and Brandenburger (1995) provide, in their Theorem B, a rather different foundation for Nash equilibrium in games with more than two players. They consider the following interpersonal conditions: (1) the belief hierarchies of all players are derived from a common prior on choice-type combinations, (2) there is common belief in the players actual belief hierarchies, and (3) every player believes in the opponents' rationality. Theorem B shows that under these conditions, each of i's opponents holds the same belief σ_i about i's choice, and these beliefs $(\sigma_1, ..., \sigma_n)$ constitute a Nash equilibrium.

Other epistemic foundations for Nash equilibrium in games with more than two players can be found in Brandenburger and Dekel (1987, Proposition 4.1), Barelli (2009, Proposition 6.1) and Bach and Tsakas (2014, Theorem 1).

Criticisms of Nash equilibrium. In Section 4.1.5 we have argued that Nash equilibrium imposes some potentially problematic conditions on the reasoning of players. Other, and relatively early, criticisms of the reasoning behind Nash equilibrium can be found, for instance, in Bernheim (1984) and Pearce (1984). Both Bernheim and Pearce argue that Nash equilibrium seems difficult to justify on the basis of rationality principles alone, and that some additional – hard to justify – assumptions must be made. They both claim that Nash equilibrium is certainly not a necessary consequence of plausible reasoning in games.

Common prior. The idea of a common prior goes back to Harsanyi (1967–1968). In this trilogy of papers, Harsanyi lays out a model for studying games with *incomplete information* (see Chapters 5 and 6 in this book) in which players may be uncertain about the opponents' utility functions. Harsanyi encodes the belief hierarchies about the players' utility functions and choices by means of *information vectors* (he also calls these *attribute vectors* or *types* at times) which specify, for a given player, a utility function, a choice and a probabilistic belief about the opponents' information vectors. Every such information vector induces an infinite belief hierarchy about the players' utility functions and choices, in the same way as a type in this and the previous chapter induces an infinite belief hierarchy about choices. From now on, we will refer to these information vectors as *Harsanyi types.*

Harsanyi then considers an important special case, which he calls the *consistent case,* in which the probabilistic beliefs of the various Harsanyi types about the opponents' Harsanyi types are all derived from a *single* probability distribution over the Harsanyi types. This is what we nowadays call a *common prior.* Since a Harsanyi type specifies a choice for the player, the common prior on Harsanyi types plays the same role as the common prior on choice-type combinations in this chapter.

There were several reasons why Harsanyi was especially interested in the consistent case, where all beliefs are derived from a common prior π on the set of Harsanyi-type combinations. First, it allowed him to mimick the game with incomplete information by a convenient, standard game with complete information, where the game starts with a chance move that randomly selects a Harsanyi-type combination according to the probability distribution π. Moreover, Harsanyi argued that the common prior still leaves sufficient room for modelling asymmetric beliefs between the players, by choosing a sufficiently asymmetric common prior. The assumption that all beliefs of all the players come from the same basic probability distribution is often called the *Harsanyi doctrine.*

Symmetric belief hierarchies. In Theorem 4.2.1 we use the idea of a common prior on choice-type combinations to characterize *symmetric belief hierarchies.* Hence, a common prior can be viewed as an expression of *symmetry* in the belief hierarchy. The concept of a symmetric belief hierarchy is, to the best of my knowledge, new.

However, Harsanyi (1967–1968) already hints at a form of symmetry that is implied by a common prior when looking at his *posterior-lottery model.* In this model it is assumed that first of all, every player specifies a choice for each of his possible Harsanyi-types, before his actual Harsanyi-type has been selected. Afterwards, the actual Harsanyi-type h_i for player i is selected, the associated choice is implemented, and his Harsanyi type h_i is matched with a combination of opponents' Harsanyi-types $(h_j)_{j \neq i}$ according to some subjective probability distribution. If there is a common prior π, then all these subjective probability distributions, for all the players, are equal to the probability distribution π. In particular, if player i with Harsanyi-type h_i believes to be matched with Harsanyi-type h_j with some positive probability, then h_j expects to be matched with h_i with some positive probability. Hence, under a common prior, partnerships between Harsanyi-types are *symmetric* in the posterior-lottery model.

Note that the idea of a common prior is much older than that of a symmetric belief hierarchy. One possible interpretation of our Theorem 4.2.1, from a historical perspective, is therefore that symmetry of a belief hierarchy can be viewed as an epistemic characterization of a common prior.

A crucial difference between our approach and that adopted in most papers and other books is that we interpret symmetry, and the common prior, from a one-person perspective, restricting the belief hierarchy of one player only. In most of the literature, it is assumed that the common prior determines the actual belief hierarchies of *all* players simultaneously. Interestingly, by carefully reading Harsanyi (1967–1968) one gets the impression that Harsanyi's approach was leaning more towards a one-person perspective.

Alternative interpretations of a common prior. Like Harsanyi (1967–1968), we interpret a common prior as a tool that restricts the belief hierarchies of the players in a game. More specifically, Harsanyi (1967–1968) views it as an expression of *mutually consistent* beliefs, whereas we use it as an expression of *symmetry* of a belief hierarchy.

There are other papers that adopt a different interpretation of a common prior. Aumann (1974), for instance, looks at a model with *states of the world,* where every state specifies a choice for each of the players, and where every player i, at each of the states ω, knows that the true state is in some set $I_i(\omega)$ containing the true state ω. In the objective version of *correlated equilibrium,* Aumann assumes that there is a common prior π on the set of all states, randomly selecting the true state ω, after which all players implement the choice associated to ω. As such, Aumann (1974) interprets the common prior as a *correlation device* that, as an output, may produce a *correlated* probability distribution over the players' choice combinations.

In Myerson (1986) and Forges (1986), the common prior may be interpreted as the outcome of a *communication game* in which the players send some input to a mediator, receive some probabilistic signal from the mediator and subsequently base their choice on the signal they receive. The correlation device in Aumann (1974) may be viewed as a special case of such a communication device in which players send no inputs.

Agreement theorems. Aumann (1976) has shown that, if the beliefs of two agents are derived from a common prior, then it cannot be common knowledge that agent 1 assigns probability p to an event A, and that agent 2 assigns a different probability $q \neq p$ to the same event A. In other words, the agents cannot agree to disagree about the probability of an event A. This result is known as Aumann's *agreement theorem,* and displays a remarkable property of beliefs that come from a common prior. Although Aumann (1976) proved the result for two agents only, the agreement theorem in fact applies to any number of agents.

In terms of belief hierarchies in games, Aumann's agreement theorem states the following: If i's belief hierarchy is induced by a common prior on choice-type combinations, then it cannot express common belief in the event that player j assigns probability p to an event A, and that another player k assigns a different probability $q \neq p$ to A. Here, player j or k may be equal to player i.

This property is not true if the belief hierarchy is not derived from a common prior, however. In that case, the belief hierarchy may express common belief in the event that player j assigns probability p to the event A and that player k assigns probability q to A, with $p \neq q$. As an illustration, consider the three-player game "When Chris joins the party" from this chapter. You may have a belief hierarchy in which you believe that Barbara chooses *blue,* believe that Chris believes that Barbara chooses *yellow,* believe that Barbara and Chris believe that you believe that Barbara chooses *blue,* believe that Barbara believes that Chris believes that Barbara chooses *yellow,* and so on. That is, you always believe that ... that you believe that Barbara chooses *blue,* whereas at the same time you always believe that ... that Chris believes that Barbara chooses *yellow.* Such a belief hierarchy would express common belief in the event that you assign probability 1 to Barbara choosing *blue* and that Chris assigns probability 0 to Barbara choosing *blue.* By Aumann's agreement theorem, this belief hierarchy cannot be derived from a common prior.

Subsequently, Aumann's agreement theorem has been extended and varied in different directions. Milgrom and Stokey (1982), for instance, consider a trading environment in which the agents have uncertainty about the allocations of others. Suppose we start from a Pareto optimal allocation. Milgrom and Stokey prove that if the beliefs of the agents come from a common prior, then there cannot be common knowledge of a trade that would make every agent weakly better off, and at least one agent strictly better off, than before. That is, the agents cannot agree on a trade that would be strictly beneficial for everyone. This result is known as Milgrom and Stokey's *no trade theorem.*

Bacharach (1985) proves a variant of Aumann's agreement theorem, which states that two "like-minded" agents cannot agree on two different rational decisions. Here, two agents are called like-minded if they would always reach the same rational decision if they had the same information. This can be seen as a generalization of the common prior assumption, since two agents whose beliefs come from the same common prior would always have the same belief – and hence reach the same decision – if they had the same information. Bacharach's result then states that between two like-minded agents, it cannot be common knowledge that agent 1 rationally reaches decision a and that agent 2 rationally reaches a different decision $b \neq a$.

Other papers that investigate Aumann's agreement theorem in variations of Aumann's original setting include Monderer and Samet (1989), Samet (1990), Hellman (2013), Bach and Perea (2013) and Bach and Cabessa (2017).

The theorems mentioned above thus display *necessary* conditions that follow from a common prior. In later years, people have found conditions that are not only necessary, but also *sufficient,* for a common prior. In other words, these people have specified conditions that *characterize* beliefs that are derived from a common prior. The first to do so was Morris (1994), who studied the trading environment by Milgrom and Stokey (1982). Like Milgrom and Stokey, he also started from a Pareto optimal allocation. Morris showed that the beliefs of the agents come from a common prior, if and only if, there is no trade that would make every agent weakly better off, and at least one agent strictly better off. In other words, a common prior can be *characterized* by the absence of a trade that is strictly beneficial for everyone. With this result, Morris (1994) thus essentially provides a converse to Milgrom and Stokey's no trade theorem.

Bonanno and Nehring (1999) prove a result for two agents that can be seen as a converse to Aumann's agreement theorem. They define a *proper belief index* as a function that assigns to every probabilistic belief some index, satisfying a certain regularity condition. It is then shown that the beliefs of the two agents come from a common prior, if and only if, for every proper belief index it cannot be common knowledge that agent 1's index is a and agent 2's index is $b \neq a$. That is, the two agents cannot agree to disagree on the value of any proper belief index. If the proper belief index is chosen to represent the probability assigned to a fixed event A, then we are back in Aumann's original setting of his agreement theorem.

Samet (1998a), Bonanno and Nehring (1999) and Feinberg (2000) prove that a common prior can also be characterized by the absence of a mutually acceptable bet. More precisely, it is shown that the beliefs of the agents are derived from a common prior, if and only if, there is no bet by which all agents would expect to make a strictly positive profit. This is often referred to as the *no bet theorem.*

Communicating beliefs. Geanakoplos and Polemarchakis (1982) investigate a dynamic communication process in which two agents alternatively communicate their subjective belief probability of a certain event to the other agent. The other agent then uses the communicated belief probability to update his own belief about the event, and communicates the updated belief probability to the other agent, and so on. Geanakoplos and Polemarchakis show that, if the agents' initial beliefs are derived from a common prior, then this communication process will terminate within finitely many steps, and the final belief probabilities of the two agents will coincide. This property will not generally hold if the beliefs are not derivable from a common prior.

Iterated expectations. Samet (1998b) studies the notion of *iterated expectations* induced by a belief hierarchy, and provides a characterization of the common prior based on this. To see what we mean by iterated expectations, consider the example "When Chris joins the party" and the beliefs diagram in Figure 4.2.2. Consider your belief hierarchy that starts at your choice *green.* In that belief hierarchy, you assign probability 0.5 to Barbara choosing *blue.* We call this your first-order expectation about Barbara's choice *blue.* At the same time, you assign probability 0.5 to the event that Chris assigns probability 0.8 to Barbara choosing *blue*, and you assign probability 0.5 to the event that Chris assigns probability 0 to Barbara choosing *blue.* Hence, the expected probability you believe that Chris assigns to Barbara choosing *blue* is $(0.5)\cdot(0.8)+(0.5)\cdot 0 = 0.4$. We call this a second-order expectation about Barbara's choice *blue,* summarizing what you believe that Chris believes about Barbara's choice *blue.* In a similar fashion we can derive a third-order expectation about Barbara's choice *blue,* which summarizes what you believe that Chris believes that you believe about Barbara's choice *blue.* It may be verified that this third-order expectation – and all higher-other expectations – all assign probability 0.4 to Barbara's choice *blue.* That is, the iterated expectations about Barbara's choice *blue* finally converge, and the limit expectation about her choice *blue* is 0.4. Moreover, this limit probability is precisely the probability that the associated common prior assigns to Barbara's choice *blue.* Indeed, we have seen in Question 4.2.3 that all belief hierarchies in Figure 4.2.2 are derived from the common prior

$$\begin{aligned}\pi((g,t_1^g),(b,t_2^b),(y,t_3^y)) &= 4/10,\ \pi((g,t_1^g),(y,t_2^y),(b,t_3^b)) = 4/10,\\ \pi((r,t_1^r),(g,t_2^g),(b,t_3^b)) &= 1/10,\ \pi((r,t_1^r),(g,t_2^g),(y,t_3^y)) = 1/10,\end{aligned}$$

which assigns probability 0.4 to Barbara choosing *blue.*

Samet (1998b) shows that this is not a coincidence – it always holds whenever the beliefs are derived from a common prior. More precisely, he considers a random variable f, and defines iterated expectations about the value of this random variable f, similarly to how we have defined iterated expectations above. If we order the agents along an infinite sequence $i_1, i_2, i_3, \ldots$ then the first-order expectation about f is i_1's expectation of the value of f, the second-order expectation is what i_1 expects of i_2's expectation of the value of f, and so on. Samet shows that, if the beliefs of the agents are derived from a common prior, then these iterated expectations about f converge to a limit expectation, and this limit expectation is exactly the expected value of f under the common prior.

In fact, Samet shows that the opposite direction is also true: If for every random variable f, the iterated expectations about f always converge to the same limit, no matter which order of the agents we choose, then the agents' beliefs are derivable from a common prior. Altogether, Samet thus shows that the beliefs of the agents come from a common prior, if and only if, the iterated expectations of any random variable always converge to the same limit, independent of the order of players we choose.

Criticisms of the common prior. In Section 4.2.5 we have argued that the notion of a symmetric belief hierarchy – and hence the concept of a common prior – display some properties that may be subject to criticism. Other criticisms of the common prior assumption can be found, for instance, in Morris (1995) and Gul (1998). Gul argues that a common prior, as a probability distribution on choice-type combinations or choice combinations, has no clear intrinsic meaning, and that it is the belief hierarchy that matters and not the common prior that induces it. Moreover, using his words, he finds that the common prior imposes a "complex and unintuitive restriction" on the belief hierarchy of a player. Both Morris and Gul explain, by means of a quote by Savage (1954), that Savage seemed to argue against the assumption of a common prior.

Correlated equilibrium. The concept of correlated equilibrium is equivalent to Harsanyi's (1967–1968) notion of *Bayesian equilibrium* if we restrict to games with *complete* information, in which no uncertainty about the opponents' utility functions is at play. Harsanyi defined a Bayesian equilibrium for a game with incomplete information as follows:

The players' belief hierarchies about choices and utility functions are encoded by means of Harsanyi-types h_i, which prescribe for player i a utility function $u_i(h_i)$ and a randomized choice $\sigma_i(h_i)$. Moreover, there is a common prior $\tilde{\pi}$ on the set of Harsanyi-type combinations which induces for every Harsanyi-type h_i the conditional belief $\tilde{\pi}(\cdot \mid h_i)$ about the opponents' Harsanyi-types. Since every opponent's Harsanyi-type prescribes a utility function and choice, the conditional belief $\tilde{\pi}(\cdot \mid h_i)$ induces a conditional belief for Harsanyi-type h_i about the opponents' utility functions and choices. In the same way as for epistemic models with types, every Harsanyi-type induces a full belief hierarchy on the players' utility functions and choices.

The common prior $\tilde{\pi}$ is called a *Bayesian equilibrium* if for every player i and every Harsanyi-type h_i selected with positive probability by $\tilde{\pi}$, the following optimality condition holds: Every choice c_i selected with positive probability by the prescribed randomized choice $\sigma_i(h_i)$ is optimal given the specified utility function $u_i(h_i)$ and given h_i's conditional belief $\tilde{\pi}(\cdot \mid h_i)$ about the opponents' Harsanyi-types.

Suppose now that there is no incomplete information, meaning that there is no uncertainty about the opponents' utility functions. Then, in Harsanyi's model we can forget about the prescribed utility functions $u_i(h_i)$, and hence every Harsanyi-type h_i only prescribes a randomized choice $\sigma_i(h_i)$, and induces a full belief hierarchy on choices. Moreover, we may assume without loss of generality that the randomized choice $\sigma_i(h_i)$ assigns probability one to one particular choice $c_i(h_i)$. Indeed, if $\sigma_i(h_i)$ assigns positive probability to two different choices c_i^1 and c_i^2, then we may "split" the Harsanyi-type h_i into two copies h_i^1 and h_i^2 that have the same belief hierarchy as h_i, but where h_i^1 prescribes the choice c_i^1 and h_i^2 prescribes the choice c_i^2. As such, a Harsanyi-type h_i can be identified with a choice-type pair (c_i, t_i) in our set-up, where c_i is the choice prescribed by h_i and t_i represents the belief hierarchy on choices induced by h_i.

But then, the common prior $\tilde{\pi}$ on Harsanyi-type combinations corresponds to a common prior π on choice-type combinations, and Harsanyi's optimality condition states that for every choice-type pair (c_i, t_i), the choice c_i is optimal given the conditional belief $\pi(\cdot \mid c_i, t_i)$ about the opponents' choice-type pairs. This, however, is exactly our definition of a correlated equilibrium. We therefore conclude that our definition of correlated equilibrium is indeed equivalent to Harsanyi's concept of Bayesian equilibrium if we restrict to complete information games.

This equivalence is formally shown in Bach and Perea (2017), who prove that a choice c_i is optimal for a utility function u_i in a Bayesian equilibrium precisely when c_i is optimal, within an epistemic model with types, for the utility function u_i under common belief in rationality with a common prior. Since we have seen that a common prior characterizes symmetric belief hierarchies, it follows that the choices that are optimal in a Bayesian equilibrium are precisely the choices that are optimal under common belief in rationality with a symmetric belief hierarchy. However, by Theorem 4.2.3, these are exactly the choices that are optimal in a correlated equilibrium. Hence, Bayesian equilibrium and our definition of correlated equilibrium deliver exactly the same optimal choices for every player, and may therefore be seen as behaviorally equivalent if we restrict to complete information games.

Aumann (1974) subsequently introduced a concept for games with complete information, which he called *correlated equilibrium,* that is equivalent to Bayesian equilibrium in such games, and hence is equivalent to our definition of correlated equilibrium as well. The difference lies in the languange he uses. Aumann defines a correlated equilibrium as follows:

He starts with a set Ω of states of the world, and assigns to every state ω a choice $c_i(\omega)$ for every player i. If the state ω is realized, then player i knows that the true state must be in the set $I_i(\omega)$ containing the true state ω. These sets $I_i(\omega)$ constitute a partition of the set Ω for every player i. The sets $I_i(\omega)$ are called *information cells.* A correlated equilibrium is then defined as a common prior $\tilde{\pi}$ on Ω satisfying the following optimality condition: For every state ω selected with positive probability by $\tilde{\pi}$ and for every player i, the prescribed choice $c_i(\omega)$ is optimal for player i, given the belief $\tilde{\pi}(\cdot \mid I_i(\omega))$ about the state, conditional on the information cell $I_i(\omega)$.

Similarly as for epistemic models with types, every state ω induces, for every player i, a full belief hierarchy on choices. To see this, first note that player i, at ω, holds the conditional belief $\tilde{\pi}(\cdot \mid I_i(\omega))$ about the state. Since every state prescribes a choice for every opponent, $\tilde{\pi}(\cdot \mid I_i(\omega))$ induces a first-order probabilistic belief about the opponents' choices. Moreover, as every state induces, for every opponent, a first-order belief about the other players' choices, the conditional belief $\tilde{\pi}(\cdot \mid I_i(\omega))$ about the state also induces a second-order belief about the opponents' first-order beliefs, and so on. In this way, every state ω induces for every player a full belief hierarchy on choices.

As such, every state ω can be indentified with a choice-type combination $(c_i, t_i)_{i \in I}$, where c_i is the choice for player i prescribed at ω and t_i represents i's belief hierarchy induced by ω. The common prior $\tilde{\pi}$ on states can thus be identified with a common prior π on choice-type combinations, and Aumann's optimality condition states that for every player i and every choice-type pair (c_i, t_i) that receives positive probability by π, the choice c_i must be optimal for the conditional belief $\pi(\cdot \mid c_i, t_i)$ about the opponents' choice-type pairs. This, however, is exactly our definition of a correlated equilibrium. We therefore conclude that Aumann's definition of correlated equilibrium is equivalent to ours, and hence is equivalent to Harsanyi's notion of Bayesian equilibrium when applied to games with complete information.

This equivalence has been proven formally in Bach and Perea (2020), who show that a choice c_i is optimal in Aumann's definition of a correlated equilibrium precisely when c_i is optimal, within an epistemic model with types, under common belief in rationality with a common prior. Above we have seen that the latter choices are exactly those that are optimal in a correlated equilibrium as defined in this chapter. Altogether, we see that a choice is optimal in Aumann's definition of a correlated equilibrium exactly when it is optimal in our definition of a correlated equilibrium. Hence, our definition of a correlated equilibrium is indeed equivalent to Aumann's.

The definition of correlated equilibrium we use in this chapter, employing epistemic models with types, is essentially the same as the definition of objective correlated equilibrium used in Dekel and Siniscalchi (2015).

Canonical correlated equilibrium. In Theorems 4.3.3 and 4.3.4 we have seen that adding the new condition of *one theory per choice* to common belief in rationality with a symmetric belief hierarchy, leads to a variant of correlated equilibium which we have called *canonical correlated equilibrium.* This concept is also sometimes called *correlated equilibrium distribution.* The condition of *one theory per choice* is taken from Bach and Perea (2020). Moreover, the two theorems mentioned above, which show that canonical correlated equilibrium is characterized by common belief in rationality with a symmetric belief hierarchy that uses one theory per choice, are also based on that paper.

In Section 4.3.3 we have shown, by means of the example "Rock, paper, scissors", that correlated equilibrium and canonical correlated equilibrium may differ in terms of the optimal choices they induce. Indeed, we have identified a choice in that example which can rationally be made under common belief in rationality with a symmetric belief hierarchy, but not under common belief in rationality with a symmetric belief hierarchy that uses *one theory per choice.* Hence, in view of Theorems 4.2.3 and 4.3.4, that choice is optimal is a correlated equilibrium, but not in a canonical correlated equilibrium. Our example "Rock, paper, scissors" is based on an example in Bach and Perea (2020). Some years earlier, Aumann and Drèze (2008) already showed by means of a different example that correlated equilibrium and canonical correlated equilibrium may differ in terms of the first-order beliefs they induce. However, in that example correlated equilibrium and canonical correlated equilibrium induce the same optimal choices for the players.

Despite the difference in terms of induced optimal choices and first-order beliefs, the literature often does not distinguish between correlated equilibrium and canonical correlated equilibrium, and employs both concepts under the same name *correlated equilibrium.* Many textbooks in game theory, for instance, only give the definition of canonical correlated equilibrium and call it correlated equilibrium. This may lead to some unfortunate confusion.

A possible reason for not distinguishing between correlated and canonical correlated equilibrium may be the well-known fact that both concepts are equivalent from an *ex-ante* perspective. That is, if we take a correlated equilibrium π, being a common prior on choice-type combinations, then the induced probability distribution $\hat{\pi}$ on the choice combinations will be a canonical correlated equilibrium. And conversely, every canonical correlated equilibrium $\hat{\pi}$ can be extended to a correlated equilibrium π that induces the same probability distribution on choice combinations as $\hat{\pi}$. Since most textbooks only concentrate on the *ex-ante* perspective, by considering only the induced probability distributions on choice combinations, it does not matter for their purpose whether one looks at correlated equilibrium or canonical correlated equilibrium.

However, from a decision-theoretic perspective it is not the *ex-ante* perspective that matters, but rather the *ad-interim* perspective that looks at the induced first-order beliefs and optimal choices. Indeed, ultimately we are interested in the choices a player can rationally make under a specific reasoning concept, and for this we must explore the possible first-order beliefs and optimal choices for that player. As an illustration, consider the example "Rock, paper, scissors" where we have seen that your choice *bomb* is optimal in a correlated equilibrium, but not in a canonical correlated equilibrium. From an *ex-ante* perspective, both concepts are equivalent because they induce the same set of probability distributions on the players' choice combinations. However, this *ex-ante* perspective is not sufficient to explain why you cannot optimally choose *bomb* in a canonical correlated equilibrium. For this, we must take the *ad-interim* perspective, and look at the possible first-order beliefs you can hold in a correlated equilibrium and a canonical correlated equilibrium. It turns out that in a canonical correlated equilibrium, you cannot hold the belief that assigns probability 1 to Barbara's choice *rock,* whereas in a correlated equilibrium you can. This is the reason why you cannot rationally choose *bomb* in a canonical correlated equilibrium, whereas you can in a correlated equilibrium. This, however, requires taking an *ad-interim* perspective.

Somewhat remarkably, also Aumann (1987) and Aumann and Drèze (2008) use the concept of canonical correlated equilibrium but call it correlated equilibrium. This despite the fact that Aumann and Drèze (2008) explicitly take an *ad-interim* perspective. By means of an example, they show that the first-order beliefs induced by canonical correlated equilibria may change if we duplicate the choices of all players, thereby transforming the game into the *doubled game.* However, duplicating a choice c_i is equivalent to considering two different choice-type pairs (c_i, t_i) and (c_i, t_i') involving the same choice c_i. A canonical correlated equilibrium in the doubled game would then correspond to a common prior over such choice-type combinations, and would thus qualify as a correlated equilibrium of the original (non-doubled) game. Hence, Aumann and Drèze (2008) effectively compare the first-order beliefs that are possible, for a given game, in a canonical correlated equilibrium and a correlated equilibrium, respectively.

Epistemic foundations for (canonical) correlated equilibrium. In Theorems 4.2.2 and 4.2.3 we have shown that correlated equilibrium can be epistemically characterized by common belief in rationality in combination with a symmetric belief hierarchy. As, in turn, a symmetric belief hierarchy is characterized by a common prior, correlated equilibrium can be characterized by common belief in rationality together with a common prior.

This is essentially the message in Aumann (1987), Dekel and Siniscalchi (2015, Theorem 4) and Bach and Perea (2020). There are a few differences between the first two papers and our characterization, though. First, Aumann (1987) and Dekel and Siniscalchi (2015) take an *ex-ante* perspective, and epistemically characterize the probability distributions on choice combinations, rather than the beliefs and the optimal choices as we do. Moreover, Aumann (1987) considers *canonical* correlated equilibrium rather than correlated equilibrium. However, since he takes an *ex-ante* perspective, it does not matter for his characterization. Finally, Aumann (1987) replaces common belief in rationality by a stronger condition, which requires that the players' choices are optimal at *all* states of the world. This change is inessential for the epistemic characterization, however.

In Theorems 4.3.3 and 4.3.4, we show that canonical correlated equilibrium can be characterized by common belief in rationality in combination with a symmetric belief hierarchy that uses one theory per choice. These results are based on Bach and Perea (2020).

Nash equilibria of games with communication. Forges (1986) and Myerson (1986), amongst

You	c	d
a	2	0
b	1	1

Barbara	a	b
c	0	0
d	1	0

Table 4.5.2 Choices versus beliefs in Nash equilibrium

others, study games with communication in which the players can send some input to a mediator, who then sends signals to the various players in a probabilistic fashion. Players can then base their choice upon the signal they receive. Forges (1986) characterizes the Nash equilibria of such games with communication. For the special case where the players do not send any input, she finds that the Nash equilibria of the game with communication correspond exactly to the canonical correlated equilibria of the original game, in terms of the induced probability distributions on choice combinations. In this sense, (canonical) correlated equilibria may be viewed as the result of adding a communication device to the game, provided one uses Nash equilibrium to explore the resulting game with communication.

Choices versus beliefs. In game theory, it is absolutely crucial to clearly distinguish between choices and beliefs about choices. This is especially important when investigating Nash equilibrium and (canonical) correlated equilibrium. Consider, for instance, the example "Rock, paper, scissors" with the utilities as depicted in Table 4.2.1. We have seen that your choice *bomb* is optimal in a correlated equilibrium. That is, you can rationally choose *bomb* with a symmetric belief hierarchy that expresses common belief in rationality.

At the same time, it may be verified that there is no correlated equilibrium that assigns positive probability to your choice *bomb*. Indeed, assume there were a correlated equilibrium π that assigns positive probability to your choice *bomb*. Then, as we have seen above, the induced probability distribution $\hat{\pi}$ on choice combinations would be a canonical correlated equilibrium. In particular, $\hat{\pi}$ would assign positive probability to your choice *bomb*. In Example 4.9 we have seen, however, that no canonical correlated equilibrium assigns positive probability to your choice *bomb*. We thus conclude that there is indeed no correlated equilibrium that assigns positive probability to your choice *bomb*. Or, equivalently put, there is no symmetric belief hierarchy that expresses common belief in rationality and that assigns positive probability to your choice *bomb* at any of its levels.

Hence, the fact that your choice *bomb* does not figure in any correlated equilibrium (or symmetric belief hierarchy that expresses common belief in rationality) does not mean that your choice *bomb* is not optimal in a correlated equilibrium (or optimal in a symmetric belief hierarchy that expresses common belief in rationality). It is thus of great importance to clearly distinguish between choices and beliefs about choices here.

A similar warning applies to the concept of Nash equilibrium. There are games in which a choice is optimal in a Nash equilibrium, but never receives positive probability by any Nash equilibrium. Or, equivalently, a choice may be optimal for a simple belief hierarchy that expresses common belief in rationality, while never appearing in any simple belief hierarchy that expresses common belief in rationality. Consider, for instance, the game in Table 4.5.2 between you and Barbara. It may be verified that the belief combination $(\sigma_1 = b,\, \sigma_2 = \frac{1}{2}c + \frac{1}{2}d)$ is a Nash equilibrium. Since your choice a is optimal (together with b) under the belief σ_2 about Barbara's choice, we conclude that a is optimal in a Nash equilibrium.

However, there is no Nash equilibrium that assigns a positive probability to your choice a. Suppose, on the contrary, that (σ_1, σ_2) is a Nash equilibrium with $\sigma_1(a) > 0$. Then, only d is optimal for Barbara under the belief σ_1, and hence σ_2 must assign probability 1 to d. But then, only b is optimal for you under the belief σ_2, which implies that σ_1 must assign probability 1 to your choice b. This is a contradiction, as we assumed that $\sigma_1(a) > 0$. Hence, we conclude that there is no Nash equilibrium that assigns positive probability to a. This despite the fact that a is optimal in a Nash equilibrium.