8. Common Knowledge

8.1 Individual knowledge

In extensive-form games with imperfect information we use information sets to represent what the players know about past choices (when it is their turn to move). An information set of Player i is a collection of nodes in the tree where it is Player i's turn to move and the interpretation is that Player i knows that she is making her choice at one of those nodes, but she does not know which of these nodes has actually been reached. In this chapter we extend the notion of information set to more general settings.

We start with an example. After listening to her patient's symptoms, a doctor reaches the conclusion that there are only five possible causes: (1) a bacterial infection, (2) a viral infection, (3) an allergic reaction to a drug, (4) an allergic reaction to food and (5) environmental factors. The doctor decides to do a lab test. If the lab test turns out to be positive then the doctor will be able to rule out causes (3)-(5), while a negative lab test will be an indication that causes (1) and (2) can be ruled out. To represent the doctor's possible states of information and knowledge we can use five states: a, b, c, d, and e. Each state represents a possible cause, as shown in Figure 8.1.

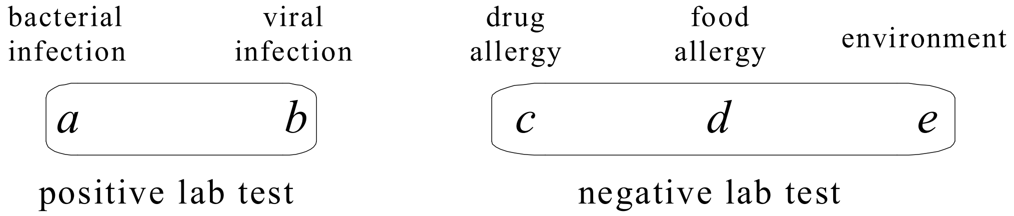


Figure 8.1: The information provided by the lab test.

We can partition the set $\{a,b,c,d,e\}$ into two sets: the set $\{a,b\}$, representing the state of knowledge of the doctor if she is informed that the lab test is positive, and the set $\{c,d,e\}$, representing the state of knowledge of the doctor if she is informed that the lab test is negative.

Consider the proposition "the cause of the patient's symptoms is either an infection or environmental factors". We can think of this proposition as the set of states $\{a,b,e\}$ where the proposition is in fact true; furthermore, we can ask the question "after receiving the result of the lab test, at which state would the doctor know the proposition represented by the set $\{a,b,e\}$?"

- If the true state is a, then – after viewing the result of the lab test – the doctor will think that it is possible that the state is either a or b and thus know that the cause of the patient's symptoms is an infection (hence she will also know the weaker proposition that the cause is either an infection or environmental factors); the same is true if the true state is b.
- On the other hand, if the true state is e then the doctor will consider c, d and e as possibilities and thus will not be able to claim that she knows that the cause of the patient's symptoms is either an infection or environmental factors.

Hence the answer to the question "after receiving the result of the lab test, at which state would the doctor know the proposition represented by the set $\{a,b,e\}$?" is "at states a and b only".
We can now turn to the general definitions.

Definition 8.1.1 Let W be a finite set of *states*, where each state is to be understood as a complete specification of the relevant *facts* about the world. An *information partition* is a partition $\mathscr{I}$ of W (that is, a collection of subsets of W that (1) are pairwise disjoint and (2) whose union covers the entire set W); the elements of the partition are called *information* sets. For every $w \in W$ we denote by $I(w)$ the information set that contains state w.

In the example of the doctor, $W = \{a,b,c,d,e\}$ and $\mathscr{I} = \{\{a,b\},\{c,d,e\}\}$; furthermore, $I(a) = I(b) = \{a,b\}$ and $I(c) = I(d) = I(e) = \{c,d,e\}$.

Definition 8.1.2 Let W be a set of states. We will call the subsets of W events. Let $\mathscr{I}$ be a partition of W, E an event (thus $E \subseteq W$) and $w \in W$ a state. We say that *at w the agent knows E* if and only if the information set to which w belongs is contained in E, that is, if and only if $I(w) \subseteq E$.

In the example of the doctor, where $W = \{a,b,c,d,e\}$ and $\mathscr{I} = \{\{a,b\},\{c,d,e\}\}$, let $E = \{a,b,d,e\}$; then at a and b the doctor knows E because $I(a) = I(b) = \{a,b\} \subseteq E$, but at d it is not true that the doctor knows E because $I(d) = \{c,d,e\} \nsubseteq E$ (since $c \in I(d)$ but $c \notin E$) and, for the same reason, also at c and e it is not true that the doctor knows E.

Note that it is possible that there is no state where the agent knows a given event. In the doctor example, if we consider event $F = \{a,c\}$ then there is no state where the doctor knows F.

Definition 8.1.3 Using Definition 8.1.2, we can define a *knowledge operator* K on events that, given as input any event E, produces as output the event KE defined as the set of states at which the agent knows E.

Let 2^W denote the set of events, that is the set of subsets of W.[a] Then the knowledge operator is the function $K : 2^W \to 2^W$ defined as follows: for every $E \subseteq W$, $KE = \{w \in W : I(w) \subseteq E\}$.

[a] If W contains n elements, then there are 2^n subsets of W, hence the notation 2^W. For example, if $W = \{a,b,c\}$ then $2^W = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a,b\}, \{a,c\}, \{b,c\}, \{a,b,c\}\}$. (Recall that $\emptyset$ denotes the empty set.)

In the example of the doctor, where $W = \{a,b,c,d,e\}$ and $\mathscr{I} = \{\{a,b\},\{c,d,e\}\}$, let $E = \{a,b,d,e\}$ and $F = \{a,c\}$; then $KE = \{a,b\}$ and $KF = \emptyset$.

Given an event $G \subseteq W$ we denote by $\neg G$ the complement of G, that is, the set of states that are not in G. For example, if $W = \{a,b,c,d,e\}$ and $G = \{a,b,d\}$ then $\neg G = \{c,e\}$. Thus while KG is the event that the agent knows G, $\neg KG$ is the event that the agent does not know G.

Note the important difference between event $\neg KG$ (the agent does not know G) and event $K\neg G$ (the agent knows that it is not the case that G):

- if $w \in \neg KG$ then at state w the agent does not know G but she might not know $\neg G$ either, that is, it may be that she considers G possible ($I(w) \cap G \neq \emptyset$) and she also considers $\neg G$ possible ($I(w) \cap \neg G \neq \emptyset$).[1]
- On the other hand, if $w \in K\neg G$ then at state w the agent knows that G is not true, because every state that she considers possible is in $\neg G$ ($I(w) \subseteq \neg G$).

Thus $K\neg G \subseteq \neg KG$ but the converse inclusion does not hold.

In the example of the doctor, where $W = \{a,b,c,d,e\}$ and $\mathscr{I} = \{\{a,b\},\{c,d,e\}\}$, again let $E = \{a,b,d,e\}$ (so that $\neg E = \{c\}$) and $F = \{a,c\}$ (so that $\neg F = \{b,d,e\}$); then $KE = \{a,b\}$, $\neg KE = \{c,d,e\}$, $K\neg E = \emptyset$, $KF = \emptyset$, $\neg KF = W$ and $K\neg F = \emptyset$. Note the interesting fact that, since we can apply the knowledge operator to any event, we can also compute the event that the agent knows that she knows E (that is, the event $K(KE)$, which we will denote more simply as KKE) and the event that the agent knows that she does not know E (that is, the event $K\neg KE$).

Continuing the example of the doctor, where $W = \{a,b,c,d,e\}$, $\mathscr{I} = \{\{a,b\},\{c,d,e\}\}$ and $E = \{a,b,d,e\}$, KKE is the set of states where the agent knows event $KE = \{a,b\}$; thus $KKE = \{a,b\}$. Furthermore, since $\neg KE = \{c,d,e\}$, $K\neg KE = \{c,d,e\}$. As noted in the following remark, this is not a coincidence.

R The knowledge operator $K : 2^W \to 2^W$ satisfies the following properties (which you are asked to prove in Exercise 8.5): for every event $E \subseteq W$,

- $KE \subseteq E$
- $KKE = KE$
- $K\neg KE = \neg KE$.

[1] In the example of the doctor, where $W = \{a,b,c,d,e\}$ and $\mathscr{I} = \{\{a,b\},\{c,d,e\}\}$, if $F = \{a,c\}$ (so that $\neg F = \{b,d,e\}$) then $KF = \emptyset$ and $K\neg F = \emptyset$; for instance, if the true state is a then the doctor considers F possible (because her information set is $I(a) = \{a,b\}$ and $I(a) \cap F = \{a\} \neq \emptyset$) but she also considers $\neg F$ possible (because $I(a) \cap \neg F = \{b\} \neq \emptyset$).

Test your understanding of the concepts introduced in this section, by going through the exercises in Section 8.4.1 at the end of this chapter.

8.2 Interactive knowledge

We can now extend our analysis to the case of several agents and talk about not only what an individual knows about relevant facts but also about what she knows about what other individuals know and what they know about what she knows, etc. There is an entertaining episode of the TV series *Friends* in which Phoebe and Rachel reason about whether Chandler and Monica know that they (= Phoebe and Rachel) know that Chandler and Monica are having an affair: see `https://www.youtube.com/watch?v=LUN2YN0bOi8` (or search for the string 'Friends-They Don't Know That We Know They Know We Know').

Again we start with a set of states W, where each state represents a complete description of the relevant *facts*. Let there be n individuals. To represent the possible states of mind of each individual we use an information partition: $\mathscr{I}_i$ denotes the partition of individual $i \in 1,\ldots,n$. As before, we call the subsets of W events. Using Definition 8.1.3 we can define a knowledge operator for every individual.

Let K_i be the knowledge operator of individual i; thus, for every event $E \subseteq W$, $K_iE = \{w \in W : I_i(w) \subseteq E\}$. Now consider an event E and an individual, say Individual 1; since K_1E is an event (it is the set of states where Individual 1 knows event E), we can compute the event K_2K_1E, which is the event that Individual 2 knows event K_1E, that is, the event that 2 knows that 1 knows E.

But there is no need to stop there: we can also compute the event $K_3K_2K_1E$ (the event that 3 knows that 2 knows that 1 knows E) and the event $K_1K_3K_2K_1E$ (the event that 1 knows that 3 knows that 2 knows that 1 knows E), etc. A few examples will be useful.

We begin with an abstract example, without interpreting the states in terms of specific facts. Let the set of states be $W = \{a,b,c,d,e,f,g,h\}$. There are three individuals, Ann, Bob and Carol, with the information partitions shown in Figure 8.2.

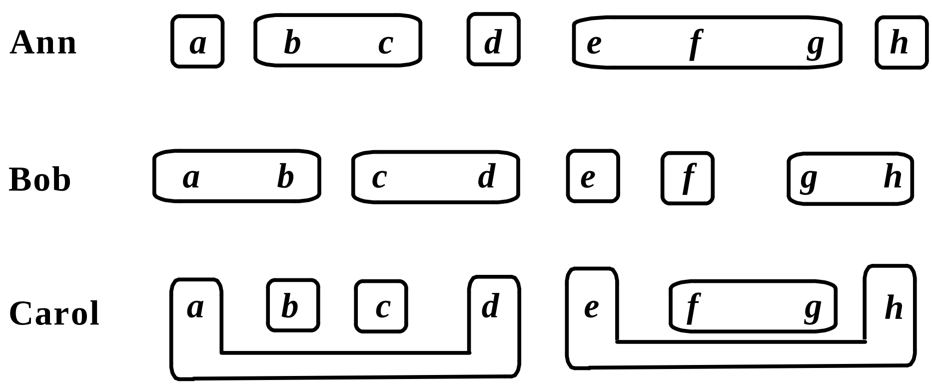


Figure 8.2: The information partitions of three individuals.

Consider the event $E = \{a,b,c,f,g\}$. Let us compute the following events:

1. $K_{Ann}E$ (Ann knows E),
2. $K_{Bob}E$ (Bob knows E),
3. $K_{Carol}E$ (Carol knows E),
4. $K_{Carol}K_{Ann}E$ (Carol knows that Ann knows E),
5. $K_{Bob}K_{Carol}K_{Ann}E$ (Bob knows that Carol knows that Ann knows E),
6. $K_{Ann}\neg K_{Bob}K_{Carol}E$ (Ann knows that Bob does **not** know that Carol knows E).

All we need to do is apply Definition 8.1.3. First of all,

$$\textbf{1. } K_{Ann}E = \{a,b,c\}$$

(for example, $b \in K_{Ann}E$ because $I_{Ann}(b) = \{b,c\}$ and $\{b,c\} \subseteq E$,
while $f \notin K_{Ann}E$ because $I_{Ann}(f) = \{e,f,g\}$ and $\{e,f,g\}$ is not a subset of E).
Similarly,

$$\textbf{2}.\ K_{Bob}E = \{a,b,f\} \quad \text{and} \quad \textbf{3}.\ K_{Carol}E = \{b,c,f,g\}.$$

To compute $K_{Carol}K_{Ann}E$ we need to find the set of states where Carol knows event $\{a,b,c\}$, since $K_{Ann}E = \{a,b,c\}$. Thus,

$$\textbf{4. } K_{Carol}\underbrace{K_{Ann}E}_{=\{a,b,c\}} = \{b,c\}.$$

Hence

$$\textbf{5. } K_{Bob}\underbrace{K_{Carol}K_{Ann}E}_{=\{b,c\}} = \emptyset,$$

that is, there is no state where Bob knows that Carol knows that Ann knows E.

To compute $K_{Ann}\neg K_{Bob}K_{Carol}E$ first we start with $K_{Carol}E$, which we have already computed: $K_{Carol}E = \{b,c,f,g\}$; then we compute $K_{Bob}\underbrace{K_{Carol}E}_{=\{b,c,f,g\}}$, which is $\{f\}$: $K_{Bob}K_{Carol}E = \{f\}$; then we take the complement of this: $\neg K_{Bob}K_{Carol}E = \{a,b,c,d,e,g,h\}$ and finally we compute K_{Ann} of this event:

$$\textbf{6}.\ K_{Ann}\underbrace{\neg K_{Bob}K_{Carol}E}_{=\{a,b,c,d,e,g,h\}} = \{a,b,c,d,h\}.$$

Thus, for example, at state a it is true that Ann knows that Bob does not know that Carol knows E, while at state e this is not true.

Next we discuss a concrete example. The professor in a game theory class calls three students to the front of the room, shows them a large transparent box that contains many hats, some red and some white (there are no other colors) and tells them the following:

> "I will blindfold you and put a hat on each of you, then I will remove the box from the room and, after that, I will remove your blindfolds, so that each of you can see the color(s) of the hats worn by the other two students, but you will not be able to see the color of your own hat. Then I will ask you questions starting with Student 1 then Student 2 then Student 3 then back to Student 1 and so on."

After having placed the hats and removed the blindfolds, the professor asks Student 1 "Do you know the color of your own hat?" She replies "No." Then he asks Student 2 the same question: "Do you know the color of your own hat?" Student 2 says "No." Then he asks Student 3 and she says "No." Then he asks the same question again to Student 1 and again the answer is "No," and so on. After asking the same question over and over and always hearing the answer "No" he gets tired and tells them "I'll give you a piece of information: **I did not pick three white hats.**" He then resumes the questioning: first he asks Student 1 "Do you know the color of your own hat?" She replies "No." Then he asks Student 2 the same question: "Do you know the color of your own hat?" Student 2 says "No." Then he asks Student 3 and she says "Yes I do!" What color hat does she have? What color hats do Students 1 and 2 have?

To answer these questions, we begin by defining the set of possible states. We can think of a state as a triple (x_1, x_2, x_3), where $x_i \in \{R, W\}$ is the color of the hat of Student i (R means Red and W means White). Thus, for example, (R,W,R) is the state where Students 1 and 3 have a red hat, while Student 2 has a white hat. The possible states of information of the three students *before the professor announces that he did not pick three white hats* are represented by the information partitions shown in Figure 8.3, where we have connected with a line states that are in the same information set.

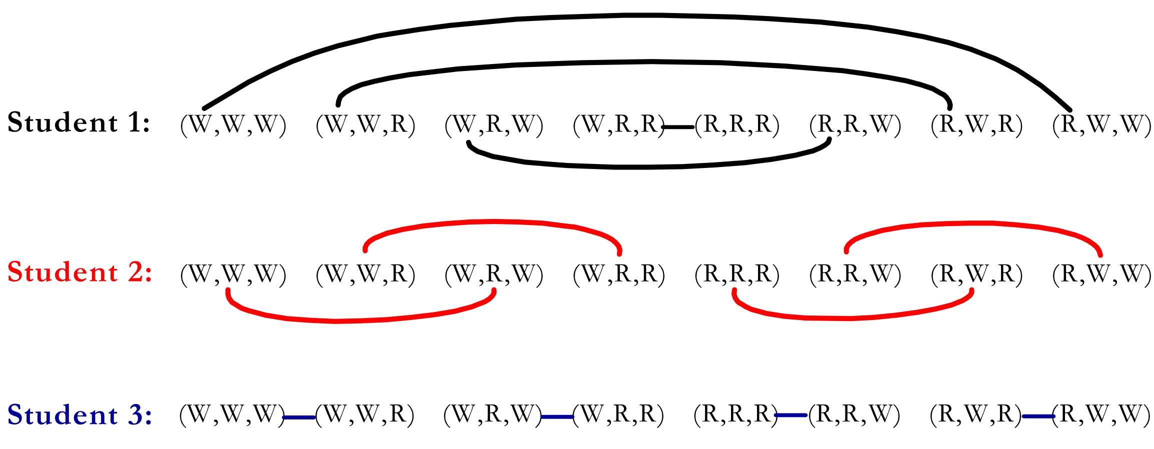


Figure 8.3: The initial information partitions for the Red/White hat example.

Whatever the state (that is, whatever hats the professor picks), each student is only uncertain about the color of her own hat: she can see, and thus knows, the colors of the hats of the other two students. Thus each information set contains only two elements.

Consider a particular state, say the state where all three hats are red: (R,R,R). At that state, obviously, each student knows that not all hats are white: he/she can actually see two red hats. Furthermore, each student knows that every other student knows that not all hats are white.[2]

Take, for example, Student 1. She sees that the hats of the other two students are red and thus she reasons that Student 2 also sees that the hat of Student 3 is red and hence Student 2 knows that not all hats are white (similarly, she reasons that Student 3 knows that not all hats are white). But does Student 1 know that Student 2 knows that Student 3 knows that not all hats are white? The answer is No. Seeing two red hats, Student 1 must consider it possible that her own hat is white, in which case Student 2 would, like Student 1, see a red hat on Student 3 but (unlike Student 1) would also see a white hat on Student 1; thus Student 2 would have to consider the possibility that her own hat is white in which case, putting herself in the shoes of Student 3, would reason that Student 3 would see two white hats and consider it possible that his own hat was also white, that is, consider it possible that all hats were white. We can see this more clearly by using the information partitions and the associated knowledge operators. To simplify matters, let us assign names to the states and rewrite the information partitions using these names, as shown in Figure 8.4.

(W,W,W)	(W,W,R)	(W,R,W)	(W,R,R)	(R,R,R)	(R,R,W)	(R,W,R)	(R,W,W)
a	b	c	d	e	f	g	h

Student 1's partition: $\{\{a,h\},\{b,g\},\{c,f\},\{d,e\}\}$

Student 2's partition: $\{\{a,c\},\{b,d\},\{e,g\},\{f,h\}\}$

Student 3's partition: $\{\{a,b\},\{c,d\},\{e,f\},\{g,h\}\}$

Figure 8.4: Assignment of names to the states in the Red/White hat example.

The proposition "not all hats are white" corresponds to event $E = \{b,c,d,e,f,g,h\}$ (the set of all states, excluding only state a).

[2] It is unfortunate that many people would use, incorrectly, the expression "all hats are not white" to mean that "it is not the case that all hats are white". The latter expression is equivalent to "at least one hat is red (possibly one, possibly two, possibly all three)", while the former is equivalent to "every hat is not white", that is, in this context, "every hat is red". [It is also unfortunate that when asked "how are you?" most people answer "I am good", instead of "I am well" or "I am fine"!]

Using Definition 8.1.3 we get the following events (recall that $E = \{b,c,d,e,f,g,h\}$ represents the proposition "not all hats are white"):

$$
\begin{aligned}
&E = \{b,c,d,e,f,g,h\},\\
&K_1E = \{b,c,d,e,f,g\},\\
&K_2E = \{b,d,e,f,g,h\},\\
&K_3E = \{c,d,e,f,g,h\}\\
&K_1K_2E = K_2K_1E = \{b,d,e,g\},\\
&K_1K_3E = K_3K_1E = \{c,d,e,f\},\\
&K_2K_3E = K_3K_2E = \{e,f,g,h\}
\end{aligned}
$$

Note that the intersection of all these events is the singleton set $\{e\}$. Thus at state e (where all the hats are red), and only at state e, everybody knows that not all hats are white and everybody knows that everybody knows that not all hats are white. Proceeding one step further, we have that

$$K_1K_2K_3E = K_1K_3K_2E = K_2K_1K_3E = K_2K_3K_1E = K_3K_1K_2E = K_3K_2K_1E = \emptyset.$$

Thus there is no state (not even e) where a student knows that another student knows that the third student knows that not all hats are white.
Let us now continue with the formal analysis of the story of the three hats. At some stage the professor makes the public announcement "I did not choose three white hats" (that is, he announces event $E = \{b,c,d,e,f,g,h\}$). This announcement makes it commonly known that state a is to be ruled out. Thus, after the announcement, the information partitions are reduced to the ones shown in Figure 8.5, obtained from Figure 8.3 by deleting the state (W,W,W).

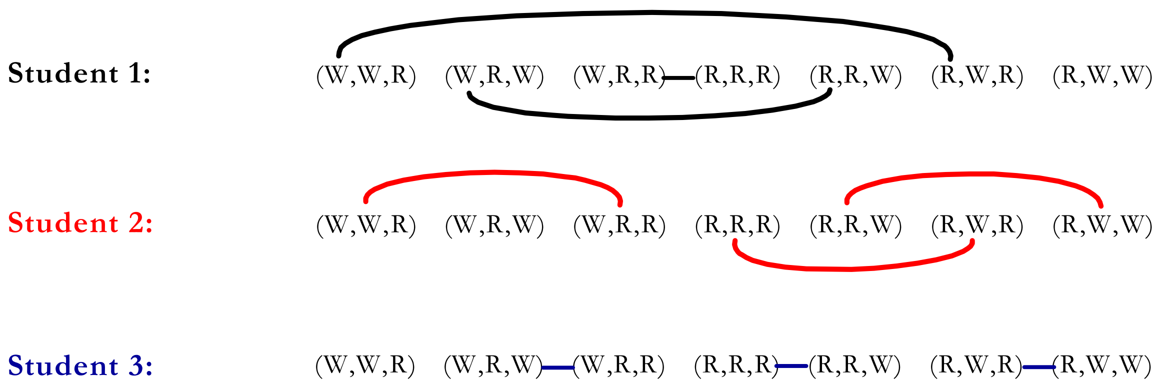


Figure 8.5: The reduced information partitions for the Red/White hat example after the announcement that not all hats are white: state *(W,W,W)* is removed from the partitions of Figure 8.3.

Note that, at this stage, if the true state is (R,W,W) Student 1 knows that her hat is red (she sees two white hats and, having been informed that the professor did not choose three white hats, she can deduce that hers must be red). Similarly, if the true state is (W,R,W), Student 2 knows that her own hat is red and, if the true state is (W,W,R), Student 3 knows that her own hat is red. According to the story, after announcing that not all hats are white, the professor first asks Student 1 if she knows the color of her hat and she answers "No." From this answer everybody can deduce that the state is not (R,W,W) and thus this state can be deleted and the information partitions reduce to the ones shown in Figure 8.6.

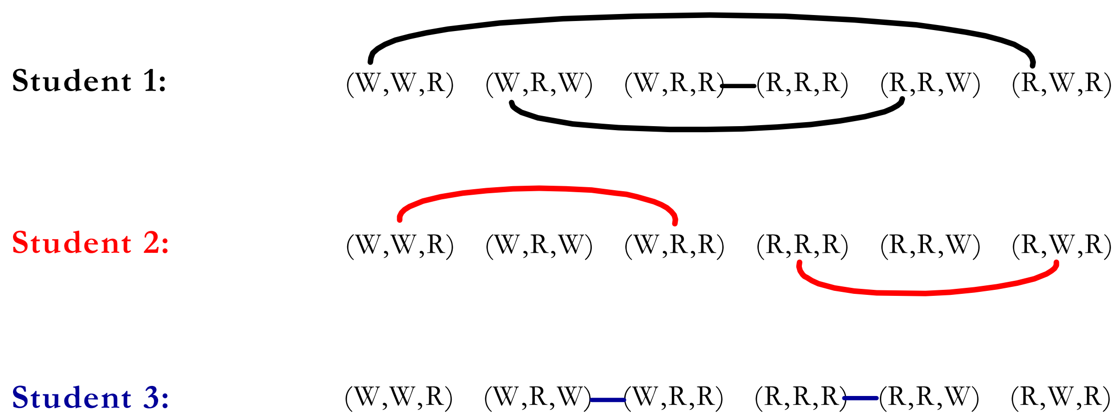


Figure 8.6: The reduced information partitions for the Red/White hat example after Student 1 says that she does not know the color of her own hat: state *(R,W,W)* is removed from the partitions of Figure 8.5.

Now, according to the story, Student 2 is asked whether she knows the color of her hat. Let us see what the possibilities are.

1. Student 2 will answer **Yes** if state is either (W,R,W) (in fact, in this case, she knew even before hearing Student 1's answer) or (R,R,W) (in this case, before hearing Student 1's answer, she thought the state might be either (R,R,W) or (R,W,W) but then, after hearing Student 1's answer, she was able to eliminate (R,W,W) as a possibility). In either of these two states Student 2's hat is red and thus she knows that her hat is red. Furthermore, upon hearing Student 2 say Yes, Student 3 learns that the state is either (W,R,W) or (R,R,W) and in both of these states his hat is white, thus he acquires the knowledge that his own hat is white.

2. In each of the remaining states, namely (W,W,R), (W,R,R), (R,R,R) and (R,W,R), Student 2 will answer **No**. Then everybody learns that the state is neither (W,R,W) nor (R,R,W) and thus the information partitions reduce to the ones shown in Figure 8.7 (obtained from Figure 8.6 by removing states (W,R,W) and (R,R,W)). Note that each of the remaining states is now a singleton information set for Student 3 and thus, upon hearing Student 2 say No, he learns what the state is: in particular he learns that his own hat is red (at each of these states Student 3's hat is red). In the original story, Students 1 and 2 answered No and Student 3 answered Yes and thus we fall in this second case.

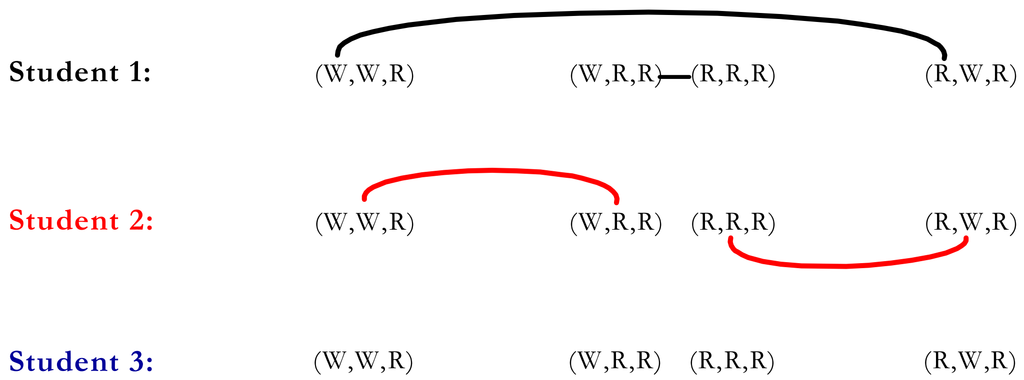


Figure 8.7: The reduced information partitions for the Red/White hat example after Student 2 also says that she does not know the color of her own hat: states *(W,R,W)* and *(R,R,W)* are removed from the partitions of Figure 8.6.

Now consider a blindfolded witness, who cannot see the color of anybody's hat, but knows that they are either red or white and hears the professor's announcement and subsequent questions, as well as the answers by the three students. What would the blindfolded witness learn about the true state?

The initial information set of the witness would consist of the set of all states; then,
(1) after the announcement of the professor, he would be able to eliminate state (W,W,W),
(2) after the negative answer of Student 1 he would be able to eliminate state (R,W,W) and
(3) after the negative answer of Student 2 he would be able to eliminate states (W,R,W) and (R,R,W).

The affirmative answer of Student 3 is not informative, because in each of these states Student 3 knows that the color of her own hat is red. Thus at the end of the exchange the witness's information set is $\{(W,W,R), (W,R,R), (R,R,R), (R,W,R)\}$. Hence all the witness knows is that the hat of Student 3 is red (on the other hand, Student 3 knows the color of all the hats, because she has figured out the color of her own hat and can see the hats of the other two students).

Let us now focus on state(R,R,R) where all hats are red. Initially, before the professor makes the announcement, no student is ever able to figure out the color of her own hat, no matter how many times the students are asked. However, as we saw, once the professor announces that not all hats are white, then after Students 1 and 2 reply negatively to the question whether they know the color of their own hat, Student 3 is able to deduce that her hat is red. Thus the professor's announcement provides crucial information. This, however, seems puzzling, because the professor merely tells the students *what they already knew*: each student, seeing two red hats, knows that not all hats are white (furthermore, as we saw above, each student also knows that every other students knows this). So how can giving the students a piece of information that they already possess make any difference? The answer is that the professor's *public* announcement makes it a matter of *common knowledge* that not all hats are white. Indeed, we saw above that – at the beginning – if the true state is (R,R,R), although everybody knows that not all hats are white and also everybody knows that everybody knows that not all hats are white, it is not the case that Student 1 knows that Student 2 knows that Student 3 knows that not all hats are white. Thus it is not common knowledge that not all hats are white. The notion of common knowledge is discussed in the next section.

Test your understanding of the concepts introduced in this section, by going through the exercises in Section 8.4.2 at the end of this chapter.

8.3 Common knowledge

Common knowledge is the strongest form of interactive knowledge: an event E is common knowledge if everybody knows E and everybody knows that everybody knows E and everybody knows that everybody knows that everybody knows E, and so on. For example, in the case of two individuals, we say that at state w event E is common knowledge if

$$w \in K_1E \cap K_2E \cap K_1K_2E \cap K_2K_1E \cap K_1K_2K_1E \cap K_2K_1K_2E \cap \dots$$

We denote by CKE the event that (that is, the set of states where) event E is common knowledge. Thus, in the case of two individuals,

$$CKE = K_1E \cap K_2E \cap K_1K_2E \cap K_2K_1E \cap K_1K_2K_1E \cap K_2K_1K_2E \cap \dots$$

Given the definition of common knowledge, it may seem impossible to check if an event is common knowledge, because it requires checking an infinite number of conditions. We will see that, on the contrary, it is very easy to determine if an event is common knowledge at any state. We begin with an example.

■ **Example 8.1** Abby proposes the following to Bruno and Caroline.

> "Tomorrow I will put you in two separate rooms, so that there will be no possibility of communication between you. I will then pick randomly an even number from the set $\{2,4,6\}$. Let n be that number. Then I will write the number $n-1$ on a piece of paper and the number $n+1$ on another piece of paper, shuffle the two pieces of paper and hand one to Bruno and the other to Caroline. For example, if I happen to pick the number 6, then I will write 5 on a piece of paper and 7 on another piece of paper, shuffle and give one piece of paper to each of you. After seeing the number handed to you, each of you will then write a pair of numbers on your piece of paper and return it to me. If
>
> 1. you write the same pair of numbers and
> 2. at least one of the two numbers is equal to the number that was actually given to Bruno then I will give \$1,000 to each of you, otherwise each of you will give me \$1,000."

■

Should Bruno and Caroline accept to play this game? They can agree today on how they should act tomorrow under various contingencies, bearing in mind that they will be unable to communicate with each other tomorrow.

We will see that Bruno and Caroline should indeed accept, because they have a strategy that guarantees that they will each get \$1,000 from Abby. The first step is to represent the set of possible states and the information partitions. We will describe a state by a triple abc, where a is the number picked by Abby, b is the number given to Bruno and c is the number given to Caroline. Bruno only observes b and Caroline only observes c. Thus the information partitions are as shown in Figure 8.8.

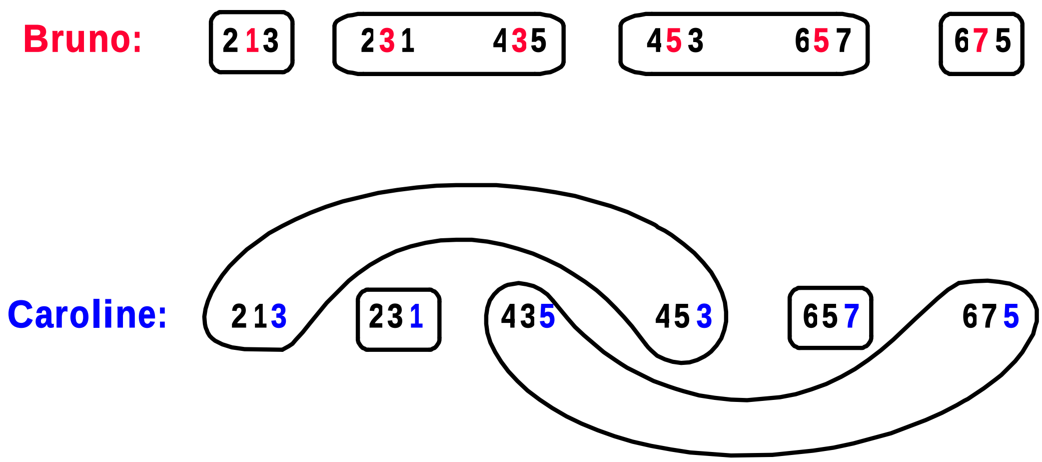


Figure 8.8: The information partitions for Example 8.1.

Let us see if the following would be a successful strategy for Bruno and Caroline: "if Bruno gets a 1 or a 3 we will write the pair (1,3) and if Bruno gets a 5 or a 7 we will write the pair (5,7)."

Consider the event "Bruno gets a 1 or a 3": call this event E. Then $E = \{213,\ 231,\ 435\}$. If this event occurs (e.g. because the actual state is 213), will it be common knowledge between Bruno and Caroline that it occurred? It is straightforward to check that (B stands for Bruno and C for Caroline) $K_B E = E$, $K_C E = \{231\}$ and thus $K_B K_C E = \emptyset$. Hence, while Bruno, of course, will know if he gets a 1 or a 3, Caroline might know (if the state is 231) or she might not know (if the state is 213 or 435), but Bruno will never know that Caroline knows. Thus event E is far from being common knowledge, if it occurs. It is easy to check that the same is true of the event "Bruno gets a 5 or a 7." In order to be successful, a coordination strategy must be based on events that, when they occur, are commonly known.

Let us now consider an alternative strategy, namely

$$\begin{cases} \text{if Bruno gets a 1 or a 5 let us write the pair (1,5), and} \\ \\ \text{if Bruno gets a 3 or a 7 let us write the pair (3,7)} \end{cases} \qquad (\bigstar)$$

Let F be the event "Bruno gets a 1 or a 5", that is, $F = \{213,\ 453,\ 657\}$. Then $K_B F = F$ and $K_C F = F$, so that $K_B \underbrace{K_C F}_{=F} = F$, $K_C \underbrace{K_B F}_{=F} = F$, $K_B \underbrace{K_C K_B F}_{=F} = F$, $K_C \underbrace{K_B K_C F}_{=F} = F$ and so on. Hence $CKF = F$, that is, if event F occurs then it is common knowledge between Bruno and Caroline that it occurred.[3] Hence strategy ($\bigstar$) will be a successful coordination strategy, since the conditioning events, when they occur, are common knowledge between Bruno and Caroline.

In the above example we were able to show directly that an event was common knowledge at a given state (we showed that $CKF = F$, that is, that, for every $w \in F$, $w \in CKF$). We now show a faster method for computing, for every event E, the event CKE. The crucial step is to derive from the individuals' information partitions a new partition of the set of states which we call the *common knowledge partition*.

Definition 8.3.1 Consider a set of states W and n partitions $\mathscr{I}_1, \mathscr{I}_2, ..., \mathscr{I}_n$ of W. As usual, if w is a state, we denote by $I_i(w)$ the element (information set) of the partition $\mathscr{I}_i$ that contains w. Given two states $w, w' \in W$, we say that

- w' is *reachable from w in one step* if there exists an $i \in \{1, \ldots, n\}$ such that $w' \in I_i(w)$.
- w' is *reachable from w in two steps* if there exists a state $x \in W$ such that x is reachable from w in one step and w' is reachable from x in one step.[a]
- In general, w' is *reachable from w in m steps* ($m \geq 1$) if there is a sequence of states $\langle w_1, w_2, \ldots, w_m \rangle$ such that (1) $w_1 = w$, (2) $w_m = w'$ and (3) for every $k = 2, \ldots, m$, w_k is reachable from w_{k-1} in one step. Finally, we say that w' is *reachable from w* if, for some $m \geq 1$, w' is reachable from w in m steps.

[a] Thus w' is *reachable from w in two steps* if there exist $x \in W$ and $i, j \in \{1, \ldots, n\}$ such that $x \in I_i(w)$ and $w' \in I_j(x)$.

[3] Similarly, letting $G = \{231,\ 435,\ 675\}$ be the event "Bruno gets a 3 or a 7" we have that $CKG = G$.

Definition 8.3.2 Consider a set of states W and n partitions $\mathscr{I}_1, \mathscr{I}_2, \ldots, \mathscr{I}_n$ of W. Given a state w, *the common knowledge information set that contains* w, denoted by $I_{CK}(w)$, is the set of states reachable from w. The *common knowledge information partition* is the collection of common knowledge information sets.

Example 8.1 continued. Let us go back to Example 8.1, where there are two individuals, Bruno and Caroline, with the information partitions shown in Figure 8.9:

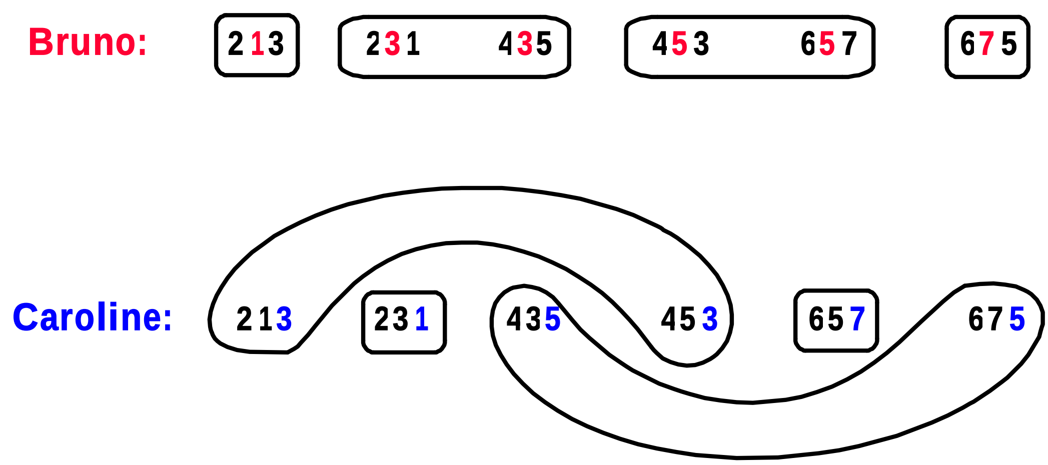


Figure 8.9: Copy of Figure 8.8.

Applying Definition 8.3.1 we have that $I_{CK}(213) = I_{CK}(453) = I_{CK}(657) = \{213, 453, 657\}$[4] and $I_{CK}(231) = I_{CK}(435) = I_{CK}(675) = \{231, 435, 675\}$. Thus the common knowledge partition is as shown in Figure 8.10.

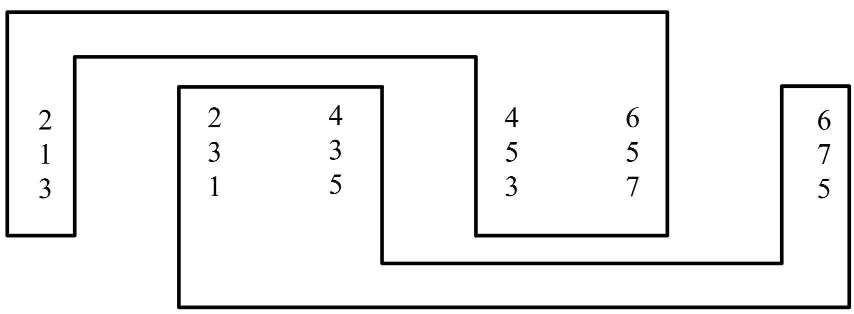


Figure 8.10: The common knowledge partition for the information partitions of Figure 8.9.

[4] In fact, 213 is reachable from itself in one step (through either Bruno or Caroline), 453 is reachable in one step from 213 (through Caroline) and 657 is reachable in two steps from 213 (the first step – to 453 – through Caroline and the second step – from 453 – through Bruno).

As a second example, consider the information partitions shown in Figure 8.11 (which reproduces Figure 8.2). The corresponding common knowledge partition is shown in Figure 8.12.

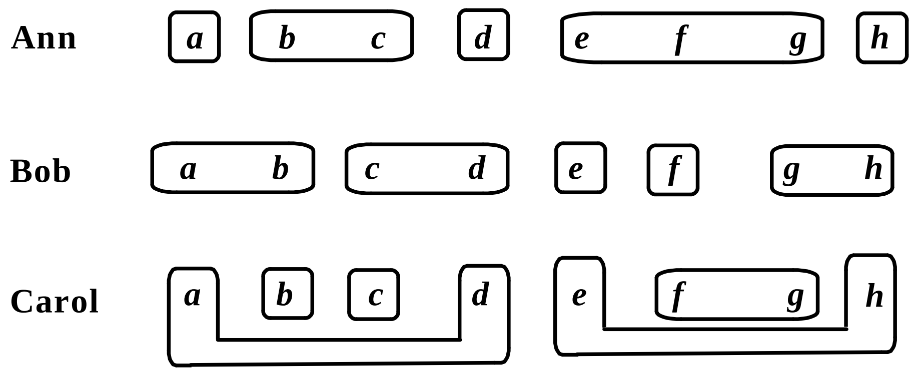


Figure 8.11: Copy of Figure 8.2.

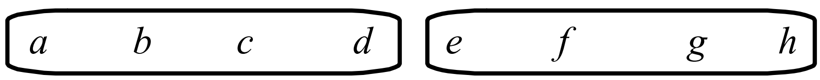


Figure 8.12: The common knowledge partition for the information partitions of Figure 8.11.

The following theorem states that, determining whether an event E is common knowledge at a state w, is equivalent to determining whether an individual whose information partition coincided with the common knowledge partition would know E at w.

Theorem 8.3.1 At state $w \in W$, event $E \subseteq W$ is common knowledge (that is, $w \in CKE$) if and only if $I_{CK}(w) \subseteq E$. Thus we can define the common knowledge operator $CK : 2^W \to 2^W$ as follows: $CKE = \{w \in W : I_{CK}(w) \subseteq E\}$.

Example 8.1 continued. Let us go back to Example 8.1 about Bruno and Caroline. Let F be the event that Bruno gets a 1 or a 5: $F = \{213,\ 453,\ 657\}$. Then, using Theorem 8.3.1, $CKF = F$ because $I_{CK}(213) = I_{CK}(453) = I_{CK}(657) = \{213, 453, 657\}$; thus – confirming what we found in Example 8.1 – at any state where Bruno gets a 1 or a 5 it is common knowledge between Bruno and Caroline that Bruno got a 1 or a 5.

Now let H be the event "Bruno did not get a 5", that is, $H = \{213, 231, 435, 675\}$. Then, using Theorem 8.3.1 we have that $CKH = \{231, 435, 675\}$. Thus while at state 231 Bruno does not get a 5 and this is common knowledge between Bruno and Caroline, at state 213 Bruno does not get a 5 but this is not common knowledge between Bruno and Caroline (in fact $213 \notin K_{Caroline}H = \{231, 435, 675\}$).

As a last example, consider again the information partitions of Figure 8.11, whose common knowledge partition was shown in Figure 8.12 and is reproduced in Figure 8.13.

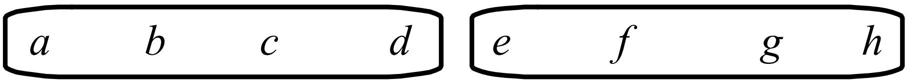


Figure 8.13: Copy of Figure 8.12.

Let $E = \{a,b,c,d,e,f\}$. Then $CKE = \{a,b,c,d\}$. Let $F = \{a,b,f,g,h\}$.. Then $CKF = \emptyset$.

The smallest event that is common knowledge at state a is $I_{CK}(a) = \{a,b,c,d\}$ and the smallest event that is common knowledge at state g is $I_{CK}(g) = \{e,f,g,h\}$.

Test your understanding of the concepts introduced in this section, by going through the exercises in Section 8.4.3 at the end of this chapter.

8.4 Exercises

8.4.1 Exercises for Section 8.1: Individual knowledge

The answers to the following exercises are in Section 8.5 at the end of this chapter.

Exercise 8.1

You are in a completely dark room, where you cannot see anything. You open a drawer that you know to contain individual socks, all of the same size but of different colors. You know that 5 are blue and 7 are white.

(a) First use your intuition to answer the following question. What is the smallest number of socks you need to take out in order to know (that is, to be absolutely certain) that you have a matching pair (i.e. either a pair of blue socks or a pair of white socks)?

(b) Now represent this situation using states and information sets. Do this to

1. represent the situation after you have taken out one sock,
2. represent the situation after you have taken out two socks and
3. represent the situation after you have taken out three socks.

Remember that a state should encode a complete description of the relevant aspects of the situation; in particular, the state should tell us how many socks you have taken out and the color of each sock that you have taken out (thus the set of states changes over time as you take out more socks). ■

Exercise 8.2

Consider the situation described in Exercise 8.1: the room is dark and you have taken out three socks. Consider the following alternative scenarios.

(a) Somebody tells you the color of the third sock (but you still don't know the color of the other two socks). Represent your state of knowledge by means of an information set.

(b) Somebody tells you the color of the matching pair (but you don't know what socks you picked). Represent your state of knowledge by means of an information set.

■

Exercise 8.3

Let the set of states be $W = \{a,b,c,d,e,f,g,h,k,m,n\}$ and the information partition of an individual be $\{\{a,b,c\},\{d\},\{e,f,g,h\},\{k,m\},\{n\}\}$. Consider the following event: $E = \{a,b,d,k,n\}$.

(a) Does the individual know E at state a?

(b) Does the individual know E at state c?

(c) Does the individual know E at state d?

(d) Does the individual know E at state h?

(e) Does the individual know E at state k?

(f) Let KE denote the event that the individual knows E (that is, the set of states where the individual knows E). What is KE?

For the next question, recall that, given an event F, we denote by $\neg F$ the complement of F, that is, the set of states that are not in F.

(g) Once again, let $E = \{a,b,d,k,n\}$. What is the event $\neg KE$, that is the event that the individual does not know E? What is the event $K\neg KE$, that is, the event that the individual knows that she does not know E?

■

Exercise 8.4

The famous pirate Sandokan has captured you and put you in front of three numbered chests containing coins. Chest 1 is labeled "gold," Chest 2 is labeled "bronze," and Chest 3 is labeled "gold or silver." One chest contains gold coins only, another contains silver coins only, and the third bronze coins only.

(a) Represent the set of possible states in the case where the labels might or might not be correct (a state must describe the label and content of each box).

(b) Let E be the event "Chest 1 is mislabeled" (that is, what the label says is false). What states are in event E?

(c) Let F be the event "Chest 2 is mislabeled". What states are in event F?

(d) What is the event "both Chests 1 and 2 are mislabeled"?

(e) Suppose now that Sandokan tells you that **all** the chests are falsely labeled, that is, what the label says is false (for example, if it says "gold" then you can be sure that the chest does not contain gold coins). If you correctly announce the content of all the chests you will be given a total of \$1,000. If you make a mistake (e.g. state that a chest contains, say, gold coins while in fact it contains bronze coins) then you don't get any money at all. You can open any number of chests you like in order to inspect the content. However, the first time you open a chest, you have to pay \$500, the second time \$300, the third time \$100.

1. What is the set of possible states (assuming that Sandokan told you the truth)?
2. What is the maximum amount of money you can be absolutely certain to make?

■

Exercise 8.5

Prove each of the following properties of the knowledge operator. The proofs are straightforward applications of Definition 8.1.2.

- **Truth**: $KE \subseteq E$, that is, if at a state w one knows E, then E is indeed true at w.
- **Consistency**: $KE \cap K\neg E = \emptyset$, that is, one never simultaneously knows E and also $\neg E$ ($\neg E$ denotes the complement of E).
- **Positive introspection**: $KE \subset KKE$, that is, if one knows E then one knows that one knows E (one is fully aware of what one knows). [Note that it follows from this and Truth that $KE = KKE$, because from Truth we get that $KKE \subseteq KE$.]
- **Negative Introspection**: $\neg KE \subseteq K\neg KE$, that is, if one does not know E, then one knows that one does not know E (one is fully aware of what one does not know). [Note that it follows from this and Truth that $\neg KE = K\neg KE$, because from Truth we get that $K\neg KE \subseteq \neg KE$.]

- **Monotonicity**: If $E \subseteq F$, then $KE \subseteq KF$, that is, if E implies F then if one knows E then one knows F.
- **Conjunction**: $KE \cap KF = K(E \cap F)$, that is, if one knows E and one knows F, then one knows E and F, and *vice versa*.

8.4.2 Exercises for Section 8.2: Interactive knowledge

The answers to the following exercises are in Section 8.5 at the end of this chapter.

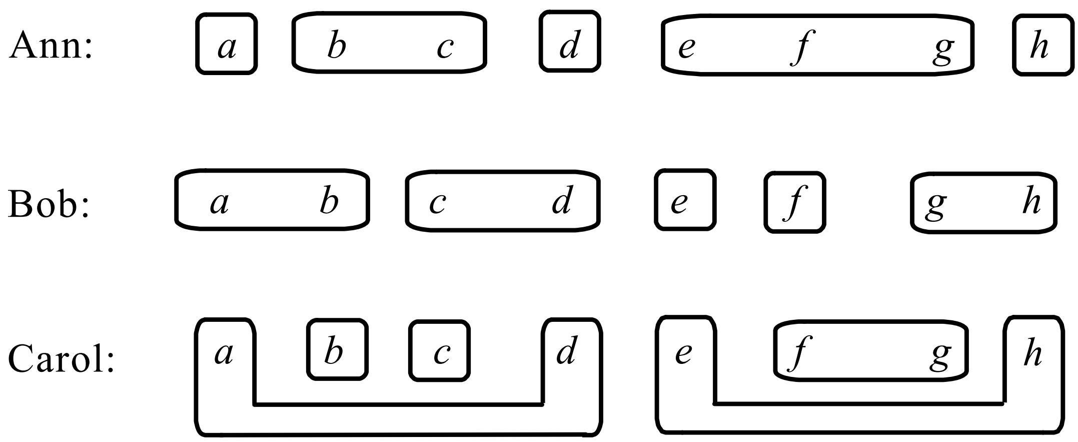


Figure 8.14: The information partitions for Exercise 8.6.

Exercise 8.6

Let the set of states be $W = \{a,b,c,d,e,f,g,h\}$. There are three individuals with the information partitions shown in Figure 8.14.

Consider the event $E = \{a,b,c,f,g\}$. Find the following events.

- $K_{Ann}E$ (the event that Ann knows E).
- $K_{Bob}E$,
- $K_{Carol}E$,
- $K_{Carol}K_{Ann}E$ (the event that Carol knows that Ann knows E),
- $K_{Bob}K_{Carol}K_{Ann}E$,
- $K_{Ann}\neg K_{Bob}K_{Carol}E$ (the event that Ann knows that Bob does not know that Carol knows E).

Exercise 8.7

Dan is at the airport. He calls his office and leaves a voice message that says: "My flight was cancelled and I am waiting to see if they can re-route me through Los Angeles or San Francisco. I will call one of you at home tonight at 8:00 sharp to let that person know whether I am in San Francisco or Los Angeles." Dan's office staff, consisting of Ann, Barb and Carol, were out for lunch. When they come back they listen to the message together. They leave the office at 5:00 pm and each goes to her home.

(a) Using information partitions, represent the possible states of knowledge of Ann, Barb and Carol concerning Dan's whereabouts at 8:15 pm, after Dan's call (there has been no communication among Ann, Barb and Carol after they left the office).

(b) Let E be the event that Dan calls either Barb or Carol. What states are in E?

(c) For the event E of Part **(b)**, find $K_A E$ (Ann knows E), $K_B K_A E$ (Barb knows that Ann knows E) and $\neg K_C E$ (Carol does not know E or it is not the case that Carol knows E).

(d) For the event E of Part **(b)**, find a state x where all of the following are true:
(1) at x Ann knows E,
(2) at x Barb knows that Ann knows E,
(3) at x it is not the case that Carol knows E.

■

Exercise 8.8

A set of lights is controlled by two switches, each of which can be in either the Up position or in the Down position. One switch is in room number 1, where Ann is; the other switch is in room number 2, where Bob is. The lights are in room number 3, where Carla is.
There are two lights: one red and one green. The red light is on if the two switches are in different positions (one up and the other down: it doesn't matter which is up and which is down), while the green light is on if the two switches are in the same position (both up or both down).
All this is common knowledge among Ann, Bob and Carla.

(a) Represent the possible states (you need to specify the position of each switch and which light is on).

(b) Represent the possible states of information of Ann, Bob and Carla by means of information partitions.

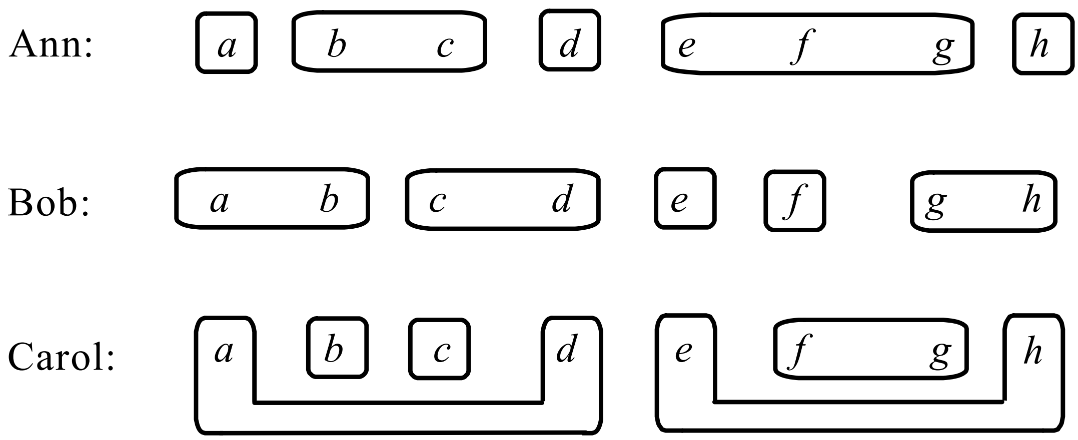
(c) Let G be the event "the green light is on". Find the events G, $K_A G$ (Ann knows G), $K_B G$ (Bob knows G), $K_C G$ (Carla knows G).

(d) Let L be the event "either the green light is on or the red light is on". Find the events L, $K_A L$ (Ann knows L), $K_B L$ (Bob knows L), $K_C L$ (Carla knows L).

■

8.4.3 Exercises for Section 8.3: Common knowledge

The answers to the following exercises are in Section 8.5 at the end of this chapter.



Exercise 8.9

Consider again the information partitions of Exercise 8.6, reproduced above.

(a) Find the common knowledge partition,

(b) Let $E = \{a, b, c, f, g\}$.
Find the event CKE, that is, the event that E is common knowledge.

(c) Let $F = \{a, b, c, d, e, g\}$.
Find CKF, that is, the event that F is common knowledge.

Exercise 8.10

In Exercise 8.7,

(a) Find the common knowledge partition,

(b) Find the event CKE (where E is the event that Dan calls either Barb or Carol).

Exercise 8.11

In Exercise 8.8,

(a) Find the common knowledge partition,

(b) Find the event CKG (where G is the event “the green light is on”),

(c) Find the event CKL (where L is the event “either the green light is on or the red light is on”).

Exercise 8.12

The set of states is $W = \{a,b,c,d,e,f,g,h\}$. There are four individuals with the following information partitions:

Individual 1: $\{\{a,\ b\},\ \{c\},\ \{d\},\ \{e,\ f\},\ \{g\},\ \{h\}\}$
Individual 2: $\{\{a\},\ \{b,\ c\},\ \{d\ ,\ e\},\ \{f\},\ \{g\},\ \{h\}\}$
Individual 3: $\{\{a,\ c\},\ \{b\},\ \{d\},\ \{e\},\ \{g\},\ \{f\ ,\ h\}\}$
Individual 4: $\{\{a\},\ \{b\ ,\ c\},\ \{d\ ,\ e\},\ \{f\ ,\ g\},\ \{h\}\}$

(a) Let $E = \{a,c,d,e\}$.
Find the following events: K_1E, K_2E, K_3E, K_4E and $K_1K_2\neg K_3E$
(Recall that $\neg$ denotes the complement of a set, that is, $\neg F$ is the set of all states that are not in F).

(b) Find the common knowledge partition.

(c) At what states is event $E = \{a,c,d,e\}$ common knowledge?

(d) Let $F = \{a,b,c,d,g,h\}$. Find the event that F is common knowledge.

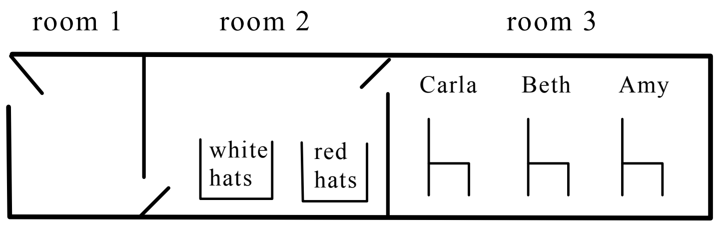


Figure 8.15: The situation described in Exercise 8.13.

Exercise 8.13

Amy, Beth and Carla are now in Room 1 as shown in Figure 8.15. They are asked to proceed, one at a time, to Room 3 through Room 2. In Room 2 there are two large containers, one with many red hats and the other with many white hats. They have to choose one hat, put it on their head and then go and sit in the chair in Room 3 that has their name on it.

Amy goes first, then Beth then Carla. The chairs are turned with the back to the door. Thus a person entering Room 3 can see whomever is already seated there but cannot be seen by them. Beth and Carla don't know how many hats there were in each box.

(a) Use information partitions to represent the possible states of knowledge of Amy, Beth and Carla after they are seated in Room 3.

(b) Suppose that Amy chose a white hat, Beth a red hat and Carla a white hat. Find the smallest event that is common knowledge among them. Give also a verbal description of this event.

(c) Repeat Parts **(a)** and **(b)** for the modified setting where there is a mirror in Room 3 that allows Amy to see the hat of Carla (but not that of Beth).

Exercise 8.14 — ⋆⋆⋆Challenging Question. ⋆⋆⋆.

Francis and his three daughters Elise, Sophia and Justine are in the same room. Francis gives a sealed envelope to Elise and tells her (in a loud voice, so that everybody can hear)

> "In this envelope I put a sum of money; I don't remember how much I put in it, but I know for sure that it was either \$4 or \$8 or \$12. Now, my dear Elise, go to your room by yourself and open it. Divide the money into two equal amounts, put the two sums in two different envelopes, seal them and give one to Sophia and one to Justine."

Unbeknownst to her sisters, Elise likes one of them more than the other and decides to disobey her father: after dividing the sum into two equal parts, she takes \$1 from one envelope and puts it in the other envelope. She then gives the envelope with more money to her favorite sister and the envelope with the smaller amount to the other sister. Sophia and Justine go to their respective rooms and privately open their envelopes, to discover, to their surprise, an odd number of dollars. So they realize that Elise did not follow their father's instructions. Neither Sophia nor Justine suspect that Elise kept some money for herself; in fact, it is common knowledge between them that Elise simply rearranged the money, without taking any for herself. Of course, neither Sophia nor Justine know in principle how much money Elise took from one envelope (although in some cases they might be able to figure it out). Thus it is **not** common knowledge between Sophia and Justine that Elise took only \$1 from one of the two envelopes. Your answers should reflect this.

(a) Use states and information partitions to represent the possible states of knowledge of Sophia and Justine.

(b) Let E be the event that Sophia is Elise's favorite sister. Find the events $K_S E$ (the event that Sophia knows it), $K_J E$ (the event that Justine knows it), $K_S K_J E$ and $K_J K_S E$.

(c) Find the common knowledge partition.

(d) Is there a state at which it is common knowledge between Sophia and Justine that Elise's favorite sister is Sophia?

(e) The night before, Francis was looking through his digital crystal ball and saw what Elise was planning to do. However the software was not working properly (the screen kept freezing) and he could not tell whether Elise was trying to favor Sophia or Justine. He knew he couldn't stop Elise and wondered how much money he should put in the envelope to make sure that the mistreated sister (whether it be Sophia or Justine) would not know that she had been mistreated. How much money should he put in the envelope? ■

8.5 Solutions to Exercises

Solution to Exercise 8.1

(a) The answer is that you need to take out three socks.

(b) Describe a state by the number of socks you have taken out and the color of each sock that you have taken out. After you have taken out only one sock:

1	1
B	W

(the information set captures the fact that you cannot see the color of the sock because it is dark). After you have taken out two socks:

2	2	2	2
B	B	W	W
B	W	B	W

After you have taken out three socks:

3	3	3	3	3	3	3	3
B	B	B	B	W	W	W	W
B	B	W	W	B	B	W	W
B	W	B	W	B	W	B	W

Now at every state there are (at least) two socks of the same color, thus you know that you have a matching pair, even though you don't know what state you are in and hence you don't know the color of the matching pair. An alternative (and equivalent) way to proceed would be to describe the state as a quadruple of numbers as follows:

number of blue socks in drawer	x_1
number of white socks in drawer	x_2
number of blue socks in your hand	x_3
number of white socks in your hand	x_4

Clearly, it must always be that $x_1 + x_3 = 5$ and $x_2 + x_4 = 7$.
The initial state is $x_1 = 5, x_2 = 7, x_3 = x_4 = 0$.
Taking one sock from the drawer will modify the state as shown in Figure 8.16.

Now taking a second sock will modify the state as shown in Figure 8.17.

Now taking a third sock will modify the state as shown in Figure 8.18.

Now you know that you have a matching pair because in every possible state you have at least 2 socks of the same color. Thus the answer is indeed that you need to take out three socks. □

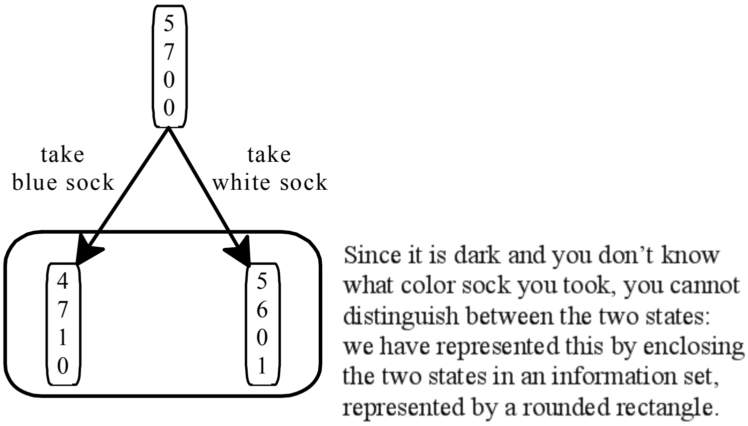


Figure 8.16: The possible states after taking one sock.

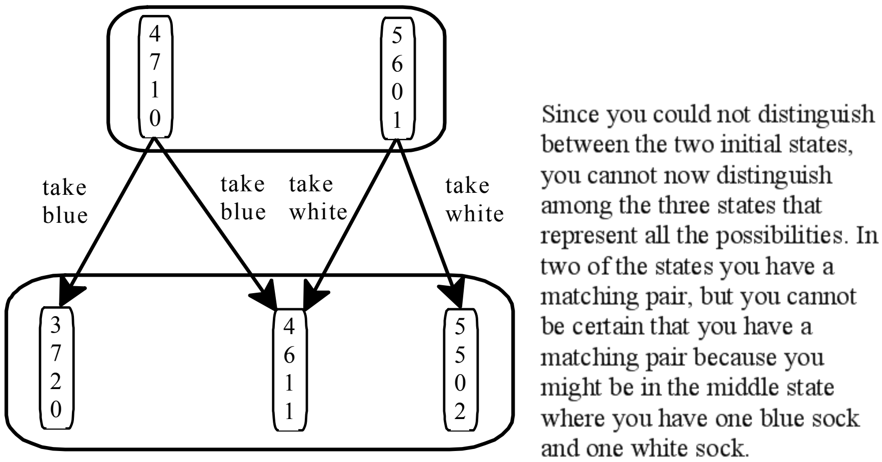


Figure 8.17: The possible states after taking two socks.

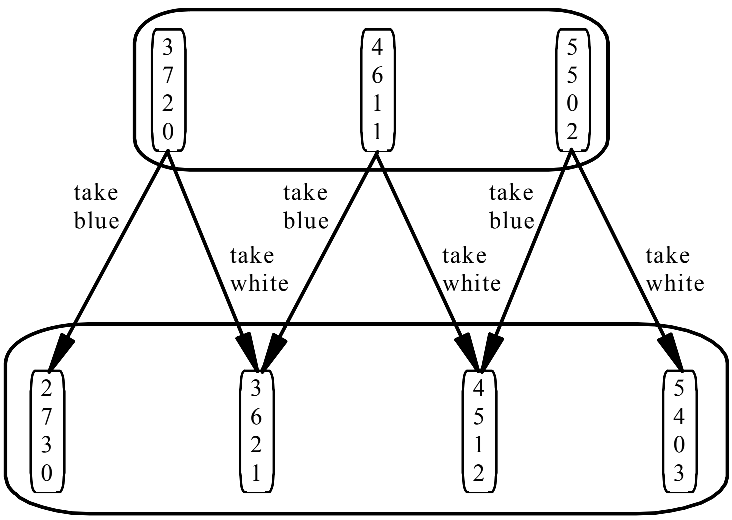


Figure 8.18: The possible states after taking three socks.

Solution to Exercise 8.2

(a)

3	3	3	3
B	B	W	W
B	W	B	W
B	B	B	B

3	3	3	3
B	B	W	W
B	W	B	W
W	W	W	W

(b)

3	3	3	3
B	B	W	B
B	W	B	B
B	B	B	W

You have a blue matching pair

3	3	3	3
W	B	W	W
W	W	B	W
B	W	W	W

You have a white matching pair

$\square$

Solution to Exercise 8.3

(a) No, because the information set containing a is $\{a,b,c\}$ which is not contained in $E = \{a,b,d,k,n\}$.

(b) No, because the information set containing c is $\{a,b,c\}$ which is not contained in $E = \{a,b,d,k,n\}$.

(c) Yes, because the information set containing d is $\{d\}$ which is contained in $E = \{a,b,d,k,n\}$.

(d) No, because the information set containing h is $\{e,f,g,h\}$ which is not contained in $E = \{a,b,d,k,n\}$.

(e) No, because the information set containing k is $\{k,m\}$ which is not contained in $E = \{a,b,d,k,n\}$.

(f) $KE = \{d,n\}$.

(g) $\neg KE = \{a,b,c,e,f,g,h,k,m\}$. $K\neg KE = \{a,b,c,e,f,g,h,k,m\} = \neg KE$. $\square$

Solution to Exercise 8.4

(a) Describe a state by a triple $((x_1,y_1),(x_2,y_2),(x_3,y_3))$ where x_i is the label on box number i $(i = 1,2,3)$ and y_i is the content of box number i.

Then the set of possible states is

$$z_1 = ((G,B),(B,G),(G \text{ or } S,S)),\ z_2 = ((G,B),(B,S),(G \text{ or } S,G))$$

$$z_3 = ((G,G),(B,B),(G \text{ or } S,S)),\ z_4 = ((G,G),(B,S),(G \text{ or } S,B))$$

$$z_5 = ((G,S),(B,B),(G \text{ or } S,G)),\ z_6 = ((G,S),(B,G),(G \text{ or } S,B)).$$

Thus, for example, state z_4 is one where box number 1 (which is labeled 'gold') in fact contains gold coins, box number 2 (which is labeled 'bronze') as a matter of fact contains silver coins, and box number 3 (which is labeled 'gold or silver') as a matter of fact contains bronze coins. More simply, we could write a state as a triple (y_1, y_2, y_3) denoting the content of each box (since we are told what the labels are). In this simpler notation the states are:

$$z_1 = (B,G,S),\ z_2 = (B,S,G),\ z_3 = (G,B,S)$$

$$z_4 = (G,S,B),\ z_5 = (S,B,G),\ z_6 = (S,G,B))$$

(b) $E = \{(B,G,S),(B,S,G),(S,B,G),(S,G,B)\}$ (recall that the label says "gold").

(c) $F = \{(B,G,S),(B,S,G),(G,S,B),(S,G,B)\}$ (recall that the label says "bronze").

(d) $E \cap F = \{(B,G,S),(B,S,G),(S,G,B)\}$.

(e) Of all the states listed above, only state z_6 is such that all the labels are false. Hence Sandokan's statement reduces the set of states to only one: $z_6 = (S,G,B)$. This is because the label "gold or silver" must be on the chest containing the bronze coins. Hence we are only left with gold and silver. Then silver must be in the chest labeled "gold". Hence gold must be in the chest labeled "bronze". Thus, by looking at the labels you can correctly guess the content without opening any chests: you know that the true state is (S,G,B). So you will get \$1,000 (you are not guessing at random, you are deducing by reasoning and you don't need to open any chests). □

Solution to Exercise 8.5

- **Truth**: $KE \subseteq E$. Consider an arbitrary $w \in KE$. We have to show that $w \in E$.
  Since $w \in KE$, by Definition 8.1.2, $I(w) \subseteq E$.
  Since $w \in I(w)$, it follows that $w \in E$.

- **Consistency**: $KE \cap K\neg E = \emptyset$.
  Suppose that $w \in KE \cap K\neg E$ for some w and some E.
  Then, by Definition 8.1.2, $I(w) \subseteq E$ (because $w \in KE$) and $I(w) \subseteq \neg E$ (because $w \in \neg E$) and thus $I(w) \subseteq E \cap \neg E$.
  Since $E \cap \neg E = \emptyset$, this implies that $I(w) = \emptyset$, which is not true because $w \in I(w)$.

- **Positive introspection**: $KE \subseteq KKE$. Consider an arbitrary $w \in KE$.
  We need to show that $w \in KKE$, that is, that $I(w) \subseteq KE$ which means that $w' \in KE$ for every $w' \in I(w)$.
  Since $w \in KE$, by Definition 8.1.2, $I(w) \subseteq E$.
  Consider an arbitrary $w' \in I(w)$. By definition of partition, $I(w') = I(w)$.
  Thus $I(w') \subseteq E$, that is, by Definition 8.1.2, $w' \in KE$.

- **Negative introspection**: $\neg KE \subseteq K\neg KE$. Consider an arbitrary $w \in \neg KE$.
  We need to show that $w \in K\neg KE$, that is, that $I(w) \subseteq \neg KE$. By Definition 8.1.2, since $I(w) \subseteq \neg KE, I(w) \cap \neg E \neq \emptyset$.
  Consider an arbitrary $w' \in I(w)$; then, since (by definition of partition) $I(w') = I(w), I(w') \cap \neg E \neq \emptyset$ so that $w' \in \neg KE$.
  Thus we have shown that, for every $w' \in I(w), w' \in \neg KE$, that is, $I(w) \subseteq \neg KE$ which, by Definition 8.1.2, yields $w \in K\neg KE$.

- **Monotonicity**: if $E \subseteq F$, then $KE \subseteq KF$. Consider an arbitrary $w \in KE$. We need to show that $w \in KF$. Since $w \in KE$, by Definition 8.1.2, $I(w) \subseteq E$. Hence, since, by hypothesis, $E \subseteq F, I(w) \subseteq F$, that is, by Definition 8.1.2, $w \in KF$.
- **Conjunction**: $KE \cap KF = K(E \cap F)$. Let $w \in KE \cap KF$.
Then $w \in KE$ and $w \in KF$; by Definition 8.1.2, the former implies that $I(w) \subseteq E$ and the latter that $I(w) \subseteq F$,
so that $I(w) \subseteq E \cap F$ and hence, by Definition 8.1.2, $w \in K(E \cap F)$.
Conversely, suppose that $w \in K(E \cap F)$.
Then, by Definition 8.1.2, $I(w) \subseteq E \cap F$ and thus $I(w) \subseteq E$ and $I(w) \subseteq F$
so that, by Definition 8.1.2, $w \in KE$ and $w \in KF$; hence $w \in KE \cap KF$. □

Solution to Exercise 8.6

(a) $K_{Ann}E = \{a,b,c\}$,
(b) $K_{Bob}E = \{a,b,f\}$,
(c) $K_{Carol}E = \{b,c,f,g\}$,
(d) $K_{Carol}K_{Ann}E = K_{Carol}\{a,b,c\} = \{b,c\}$,
(e) $K_{Bob}K_{Carol}K_{Ann} = K_{Bob}\{b,c\} = \emptyset$,
(f) $K_{Ann}\neg K_{Bob}K_{Carol}E = K_{Ann}\neg K_{Bob}\{b,c,f,g\} = K_{Ann}\neg\{f\} = K_{Ann}\{a,b,c,d,e,g,h\} = \{a,b,c,d,h\}$. □

Solution to Exercise 8.7

(a) We can represent a state as a pair (x,y) where x is the city where Dan is (SF for San Francisco or LA for Los Angeles) and y is the person that Dan called (A for Ann, B for Barb and C for Carol). The information partitions are shown in Figure 8.19.

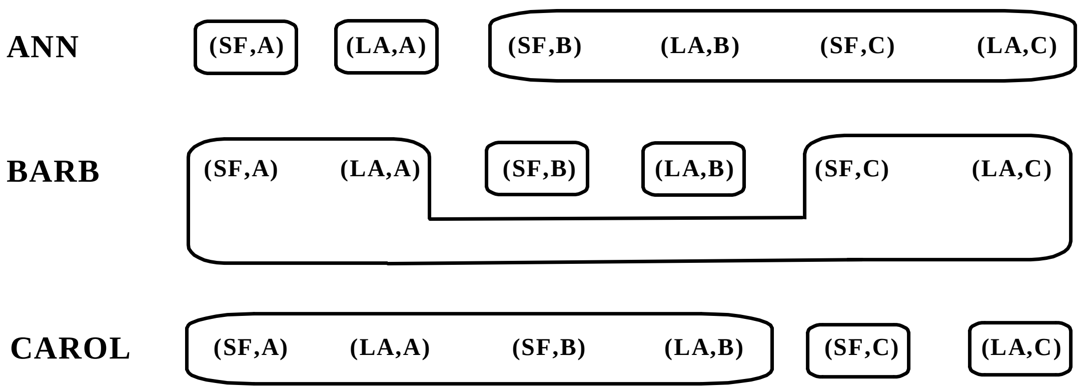


Figure 8.19: The information partitions for Exercise 8.7.

(b) $E = \{(SF,B),(LA,B),(SF,C),(LA,C)\}$
(c) $K_{Ann}E = E, K_{Barb}K_{Ann}E = \{(SF,B),(LA,B)\}$ and $K_{Carol}E = \{(SF,C),(LA,C)\}$ so that $\neg K_{Carol}E = \{(SF,A),(LA,A),(SF,B),(LA,B)\}$.
(d) We want a state x such that $x \in K_{Ann}E$, $x \in K_{Barb}K_{Ann}E$ and $x \in \neg K_{Carol}E$ (that is, $x \notin K_{Carol}E$). There are only two such states: *(SF,B)* and *(LA,B)*. Thus either x = *(SF,B)* or x = *(LA,B)*. □

Solution to Exercise 8.8

(a) We can represent a state as a triple $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ where x is the position of the switch in Room 1 (Up or Down), y is the position of the switch in Room 2 and z is the light which is on in Room 3 (Green or Red).

(b) The information partitions are shown in Figure 8.20

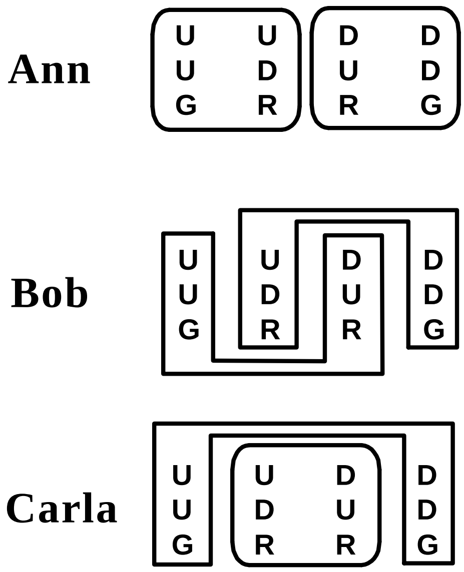


Figure 8.20: The information partitions for Exercise 8.8.

(c) $G = \left\{ \begin{array}{c|c} U & D \\ U & D \\ G & G \end{array} \right\}$, $K_A G = \emptyset$, $K_B G = \emptyset$, $K_C G = \left\{ \begin{array}{c|c} U & D \\ U & D \\ G & G \end{array} \right\}$.

(d) L is the set of all states. Hence $K_A L = K_B L = K_C L = L$. □

Solution to Exercise 8.9

(a) The common knowledge partition is shown in Figure 8.21.

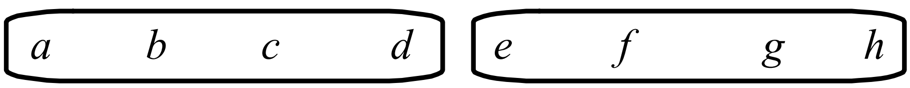


Figure 8.21: The common knowledge partition for Exercise 8.9.

(b) $CKE = \emptyset$ (where $E = \{a, b, c, f, g\}$).

(c) $CKF = \{a, b, c, d\}$ (where $F = \{a, b, c, d, e, g\}$). □

Solution to Exercise 8.10

(a) The common knowledge partition is the trivial partition shown in Figure 8.22.

(SF,A)	(LA,A)	(SF,B)	(LA,B)	(SF,C)	(LA,C)

Figure 8.22: The common knowledge partition for Exercise 8.10.

(b) $CKE = \emptyset$ □

Solution to Exercise 8.11

(a) The common knowledge partition is the trivial partition shown below:

$$\boxed{\begin{array}{cccc} U & U & U & U \\ U & D & U & D \\ G & R & R & G \end{array}}$$

(b) $CKG = \emptyset$.

(c) $CKL = L$. □

Solution to Exercise 8.12

(a) Let $E = \{a,c,d,e\}$. Then: $K_1E = \{c,d\}$, $K_2E = \{a,d,e\}$, $K_3E = \{a,c,d,e\}$, $K_4E = \{a,d,e\}$ and $K_1K_2\neg K_3E = \{g,h\}$ [in fact, $\neg K_3E = \{b,f,g,h\}$, so that $K_2\neg K_3E = \{f,g,h\}$ and $K_1K_2\neg K_3E = \{g,h\}$].

(b) The common knowledge partition is $\{\{a,b,c\},\{d,e,f,g,h\}\}$.

(c) At no state is event $E = \{a,c,d,e\}$ common knowledge: $CKE = \emptyset$.

(d) $CKF = \{a,b,c\}$ (where F=$\{a,b,c,d,g,h\}$). □

Solution to Exercise 8.13

(a) Represent a state as a triple of letters, where the top letter denotes the color of Amy's hat, the second letter the color of Beth's hat and the bottom letter the color of Carla's hat. Each of them knows which hat she chose; furthermore, Beth can see Amy's hat and Carla can see everything. Thus the partitions are as shown in Figure 8.23.

(b) When the true state is (W,R,W), the smallest event that is common knowledge among them is the first information set of the common knowledge partition, that is, the fact that Amy has a white hat.

(c) In the modified setting the information partitions are as shown in Figure 8.24. The common knowledge partition is the same as before, hence the smallest event that is common knowledge among all three of them is that Amy has a white hat. □

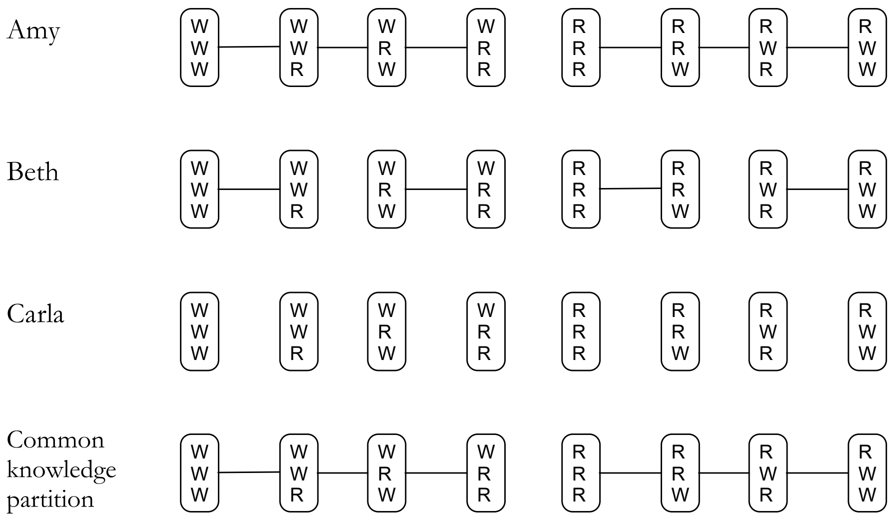


Figure 8.23: The information partitions for Part **(a)** of Exercise 8.13.

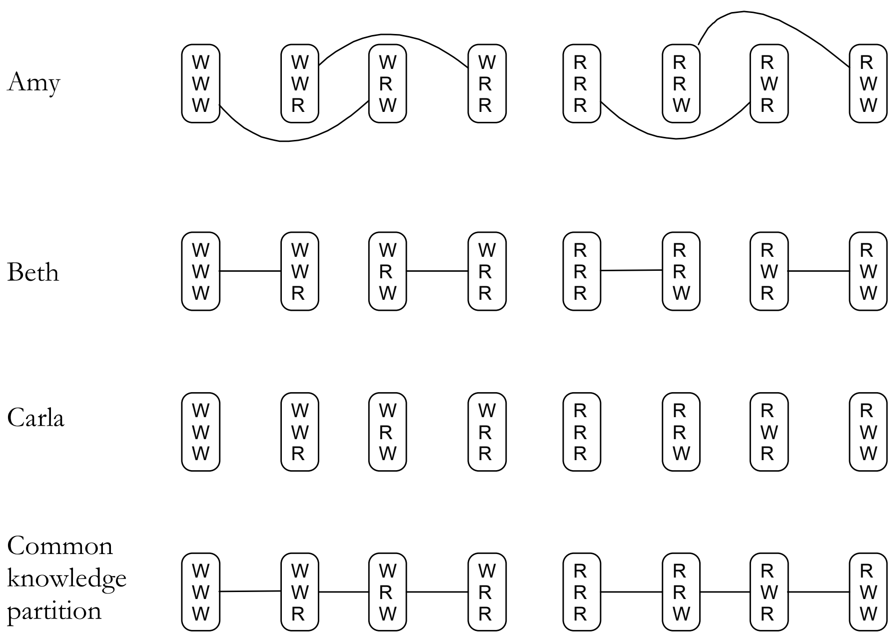


Figure 8.24: The information partitions for Part **(c)** of Exercise 8.13.

Solution to Exercise 8.14

(a) Describe a state by three numbers where the top number is the amount of money that Francis put in the envelope, the middle number is the amount given to Sophia and the bottom number is the amount given to Justine. As a matter of fact, each sister can only find either \$1 or \$3 or \$5 or \$7 in her own envelope. Thus the objectively possible states are:

$$\begin{pmatrix}4\\1\\3\end{pmatrix},\ \begin{pmatrix}4\\3\\1\end{pmatrix},\ \begin{pmatrix}8\\3\\5\end{pmatrix},\ \begin{pmatrix}8\\5\\3\end{pmatrix},\ \begin{pmatrix}12\\5\\7\end{pmatrix}\ \text{ and }\ \begin{pmatrix}12\\7\\5\end{pmatrix}.$$

However, besides these, there are also subjectively possible states, namely

$$\begin{pmatrix}8\\1\\7\end{pmatrix},\ \begin{pmatrix}8\\7\\1\end{pmatrix},\ \begin{pmatrix}12\\1\\11\end{pmatrix},\ \begin{pmatrix}12\\11\\1\end{pmatrix},\ \begin{pmatrix}12\\3\\9\end{pmatrix}\ \text{ and }\ \begin{pmatrix}12\\9\\3\end{pmatrix}.$$

(because Sophia and Justine don't know that Elise only transferred \$1 from one envelope to the other).

The information partitions are as follows:

SOPHIA:

$$\boxed{\begin{pmatrix}4\\1\\3\end{pmatrix}\begin{pmatrix}8\\1\\7\end{pmatrix}\begin{pmatrix}12\\1\\11\end{pmatrix}}\quad \boxed{\begin{pmatrix}4\\3\\1\end{pmatrix}\begin{pmatrix}8\\3\\5\end{pmatrix}\begin{pmatrix}12\\3\\9\end{pmatrix}}\quad \boxed{\begin{pmatrix}8\\5\\3\end{pmatrix}\begin{pmatrix}12\\5\\7\end{pmatrix}}$$

$$\boxed{\begin{pmatrix}8\\7\\1\end{pmatrix}\begin{pmatrix}12\\7\\5\end{pmatrix}}\quad \boxed{\begin{pmatrix}12\\9\\3\end{pmatrix}}\quad \boxed{\begin{pmatrix}12\\11\\1\end{pmatrix}}$$

JUSTINE:

$$\boxed{\begin{pmatrix}4\\1\\3\end{pmatrix}\begin{pmatrix}8\\5\\3\end{pmatrix}\begin{pmatrix}12\\9\\3\end{pmatrix}}\quad \boxed{\begin{pmatrix}4\\3\\1\end{pmatrix}\begin{pmatrix}8\\7\\1\end{pmatrix}\begin{pmatrix}12\\11\\1\end{pmatrix}}\quad \boxed{\begin{pmatrix}8\\3\\5\end{pmatrix}\begin{pmatrix}12\\7\\5\end{pmatrix}}$$

$$\boxed{\begin{pmatrix}8\\1\\7\end{pmatrix}\begin{pmatrix}12\\5\\7\end{pmatrix}}\quad \boxed{\begin{pmatrix}12\\3\\9\end{pmatrix}}\quad \boxed{\begin{pmatrix}12\\1\\11\end{pmatrix}}$$

(b) The event that Sophia is Elise's favorite sister is

$$E = \left\{ \begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 8 \\ 5 \\ 3 \end{pmatrix}, \begin{pmatrix} 8 \\ 7 \\ 1 \end{pmatrix}, \begin{pmatrix} 12 \\ 7 \\ 5 \end{pmatrix}, \begin{pmatrix} 12 \\ 9 \\ 3 \end{pmatrix}, \begin{pmatrix} 12 \\ 11 \\ 1 \end{pmatrix} \right\}$$

$$K_S E = \left\{ \begin{pmatrix} 8 \\ 7 \\ 1 \end{pmatrix}, \begin{pmatrix} 12 \\ 7 \\ 5 \end{pmatrix}, \begin{pmatrix} 12 \\ 9 \\ 3 \end{pmatrix}, \begin{pmatrix} 12 \\ 11 \\ 1 \end{pmatrix} \right\}$$

$$K_J E = \left\{ \begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 8 \\ 7 \\ 1 \end{pmatrix}, \begin{pmatrix} 12 \\ 11 \\ 1 \end{pmatrix} \right\}$$

$$K_S K_J E = \left\{ \begin{pmatrix} 12 \\ 11 \\ 1 \end{pmatrix} \right\} \quad \text{and} \quad K_J K_S E = \emptyset.$$

(c) The common knowledge partition is:

$$\boxed{\begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 8 \\ 1 \\ 7 \end{pmatrix}, \begin{pmatrix} 12 \\ 1 \\ 11 \end{pmatrix}, \begin{pmatrix} 8 \\ 5 \\ 3 \end{pmatrix}, \begin{pmatrix} 12 \\ 9 \\ 3 \end{pmatrix}, \begin{pmatrix} 12 \\ 5 \\ 7 \end{pmatrix}}$$

$$\boxed{\begin{pmatrix} 4 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 8 \\ 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 12 \\ 3 \\ 9 \end{pmatrix}, \begin{pmatrix} 8 \\ 7 \\ 1 \end{pmatrix}, \begin{pmatrix} 12 \\ 11 \\ 1 \end{pmatrix}, \begin{pmatrix} 12 \\ 7 \\ 5 \end{pmatrix}}$$

(d) At no state. In fact, $CKE = \emptyset$.

(e) Francis should put either \$8 or \$12 in the envelope. If he were to put \$4, then one sister would end up with \$1 and know that she was mistreated. In no other case does a mistreated sister know that she got less money than the other. $\square$