9. Adding Beliefs to Knowledge

9.1 Sets and probability: brief review

We begin with a very brief review of definitions and concepts from set theory and probability theory.

9.1.1 Sets

We will focus on finite sets, that is, sets that have a finite number of elements. Let U be a finite set. The set of subsets of U is denoted by 2^U. The reason for this notation is that if U contains n elements then there are 2^n subsets of U. For example, if $U = \{a,b,c\}$ then the set of subsets of U is the following collection of $2^3 = 8$ sets:

$$2^U = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a,b\}, \{a,c\}, \{b,c\}, \{a,b,c\}\},$$

where $\emptyset$ denotes the empty set, that is, a set with no elements.

The following notation is used to denote membership in a set and to denote that one set is contained in another: $x \in A$ means that x is an element of the set A (capital letters are used to denote sets and lower-case letters to denote elements) and $A \subseteq B$ means that A is a subset of B, that is, every element of A is also an element of B. Note that $A \subseteq B$ allows for the possibility that $A = B$.

Next we review operations that can be performed on sets.

- Let $A \in 2^U$. The *complement of A in U*, denoted by $\neg A$, is the set of elements of U that are not in A. When the "universe of discourse" U is clear from the context, one simply refers to $\neg A$ as the *complement of A*.
  For example, if $U = \{a,b,c,d,e,f\}$ and $A = \{b,d,f\}$ then $\neg A = \{a,c,e\}$.
  Note that $\neg U = \emptyset$ and $\neg\emptyset = U$.

- Let $A, B \in 2^U$. The *intersection of A and B*, denoted by $A \cap B$, is the set of elements that belong to both A and B.
  For example, if $A = \{b, d, f\}$ and $B = \{a, b, d, e\}$ then $A \cap B = \{b, d\}$.
  If $A \cap B = \emptyset$ we say that A and B are *disjoint*.
- Let $A, B \in 2^U$. The *union of A and B*, denoted by $A \cup B$, is the set of elements that belong to either A or B (or both).
  For example, if $A = \{b, d, f\}$ and $B = \{a, b, d, e\}$ then $A \cup B = \{a, b, d, e, f\}$.

The above operations on sets are illustrated in Figure 9.1.

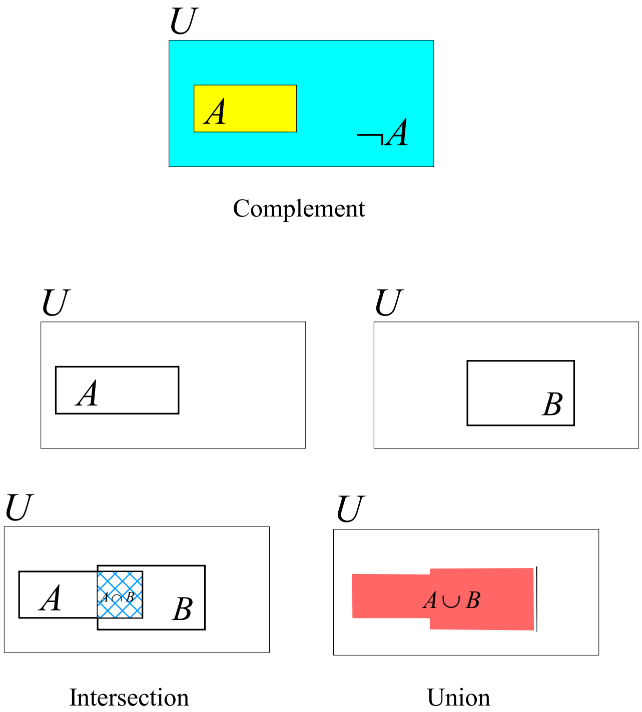


Figure 9.1: Operations on sets.

We denote by $A \setminus B$ the set of elements of A that are not in B. Thus, $A \setminus B = A \cap \neg B$. For example, if $A = \{b, d, f\}$ and $B = \{a, b, d, e\}$ then $A \setminus B = \{f\}$ and $B \setminus A = \{a, e\}$.

The following are known as *De Morgan's Laws*:

- $\neg(A \cup B) = \neg A \cap \neg B$
- $\neg(A \cap B) = \neg A \cup \neg B$

Let us verify De Morgan's Laws in the following example:
$U = \{a, b, c, d, e, f, g, h, i, j, k\}$, $A = \{b, d, f, g, h, i\}$ and $B = \{a, b, f, i, k\}$.
Then $\neg A = \{a, c, e, j, k\}$, $\neg B = \{c, d, e, g, h, j\}$, $A \cup B = \{a, b, d, f, g, h, i, k\}$
so that $\neg(A \cup B) = \{c, e, j\} = \neg A \cap \neg B$;
furthermore, $A \cap B = \{b, f, i\}$ so that $\neg(A \cap B) = \{a, c, d, e, g, h, j, k\} = \neg A \cup \neg B$.

9.1.2 Probability

In probability theory the "universal set" U is called the *sample space* and the subsets of U are called *events*. A *probability measure on* U is a function $P : 2^U \to [0,1]$ that assigns to every event $E \in 2^U$ a number greater than or equal to 0 and less than or equal to 1, as shown in Figure 9.2, with the following restrictions:

1. $P(U) = 1$.
2. For every two events $E, F \in 2^U$, if $E \cap F = \emptyset$ then $P(E \cup F) = P(E) + P(F)$.

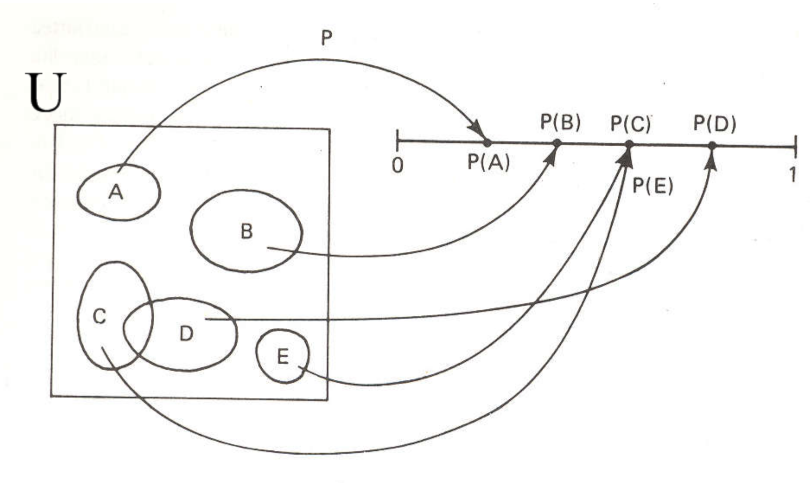


Figure 9.2: A probability measure.

From the above two properties one can obtain the following properties (the reader might want to try to prove them using Properties 1 and 2 above):

- $P(\neg E) = 1 - P(E)$, for every event E (this follows from the fact that E and $\neg E$ are disjoint and their union is equal to U).
- $P(\emptyset) = 0$ (this follows from the previous line and the fact that $\emptyset = \neg U$).
- For every two events $E, F \in 2^U$, $P(E \cup F) = P(E) + P(F) - P(E \cap F)$ (see Exercise 9.5).
- For every two events $E, F \in 2^U$, if $E \subseteq F$ then $P(E) \leq P(F)$.
- If $E_1, E_2, ..., E_m \in 2^U$ $(m \geq 2)$ is a collection of mutually disjoint sets (that is, for every $i, j = 1, \ldots, m$ with $i \neq j$, $E_i \cap E_j = \emptyset$) then $P(E_1 \cup E_2 \cup ... \cup E_m) = P(E_1) + P(E_2) + ... + P(E_m)$.

When the set U is finite, a *probability distribution* p on U is a function that assigns to each element $z \in U$ a number $p(z)$, with $0 \leq p(z) \leq 1$, and such that $\sum_{z \in U} p(z) = 1$.
Given a probability distribution $p : U \to [0,1]$ on U one can obtain a probability measure $P : 2^U \to [0,1]$ by defining, for every event A, $P(A) = \sum_{z \in A} p(z)$.
Conversely, given a probability measure $P : 2^U \to [0,1]$, one can obtain from it a probability distribution $p : U \to [0,1]$ by defining, for every $z \in U$, $p(z) = P(\{z\})$. Thus, the two notions are equivalent.

Test your understanding of the concepts introduced in this section, by going through the exercises in Section 9.8.1 at the end of this chapter.

9.2 Probabilistic beliefs

An information set contains all the states that an individual considers possible, that is, the states that the individual cannot rule out, given her information. However, of all the states that are possible, the individual might consider some to be more likely than others and might even dismiss some states as "extremely unlikely" or "implausible".

For example, suppose that there are only three students in a class: Ann, Bob and Carla. The professor tells them that in the last exam one of them got 95 points (out of 100), another 78 and the third 54. We can think of a state as a triple (a,b,c), where a is Ann's score, b is Bob's score and c is Carla's score. Then, based on the information given by the professor, Ann must consider all of the following states as possible: (95,78,54), (95,54,78), (78,95,54), (78,54,95), (54,95,78) and (54,78,95).
Suppose, however, that in all the previous exams Ann and Bob always obtained a higher score than Carla and often Ann outperformed Bob. Then Ann might consider states (95,78,54) and (78,95,54) much more likely than (78,54,95) and (54,78,95).
To represent such judgments of relative likelihood we add to an information set a probability distribution over the states in the information set. The probability distribution expresses the individual's *beliefs*, while the information set represents what the individual *knows*. In this example, Ann's beliefs could be represented by the following probability distribution:

$$\begin{pmatrix} (95,78,54) & (95,54,78) & (78,95,54) & (54,95,78) & (78,54,95) & (54,78,95) \\ \frac{9}{16} & \frac{4}{16} & \frac{2}{16} & \frac{1}{16} & 0 & 0 \end{pmatrix} \qquad (\blacklozenge)$$

According to these beliefs, Ann considers it very likely that she got the highest score, is willing to dismiss the possibility that Carla received the highest score as extremely unlikely (she assigns probability 0 to the two states where Carla's score is 95) and considers it much more likely that she, rather than Bob, received the highest score.

Recall that, given a probability distribution $p : U \to [0,1]$ on a set U (U can be thought of as an information set for the individual under consideration) and an event $E \subseteq U$, the probability of event E, denoted by $P(E)$, is defined as as the sum of the probabilities of the elements of E:

$$P(E) = \sum_{x \in E} p(x).$$

For instance, continuing the above example, the proposition "Ann received the highest score" corresponds to event

$$E = \{(95,78,54),(95,54,78)\}$$

and – according to Ann's beliefs – the probability of that event is

$$P(E) = p(95,78,54) + p(95,54,78) = \tfrac{9}{16} + \tfrac{4}{16} = \tfrac{13}{16} = 81.25\%.$$

On the other hand, the proposition "Bob's score is higher than Carla's score" corresponds to the event

$$F = \{(95,78,54),(78,95,54),(54,95,78)\}$$

and – according to Ann's beliefs – the probability of that event is

$$P(F) = p(95,78,54) + p(78,95,54) + p(54,95,78) = \tfrac{9}{16} + \tfrac{2}{16} + \tfrac{1}{16} = \tfrac{12}{16} = 75\%.$$

Definition 9.2.1 We say that an individual is *certain* of an event E if she attaches probability 1 to E (that is, if $P(E) = 1$).

In the above example, Ann is certain of event

$$G = \{(95,78,54), (95,54,78), (78,95,54), (54,95,78)\},$$

corresponding to the proposition "Carla did not get the highest score". She is also certain of every event H such that $G \subseteq H$.

R Note the important difference between knowledge and certainty: if at a state x the individual knows an event E (that is, $x \in KE$) then, at that state, E is indeed true ($x \in E$),[1] that is, it is never the case that an individual knows something which is false; on the other hand, an individual can be certain of something which is false, that is, it is possible that the individual assigns probability 1 to an event E even though the true state is not in E.[2]

Continuing the example of the exam scores, suppose that – before distributing the exams – the professor says "I was surprised to see that, this time, Ann did not get the highest score". This new announcement by the professor informs the students that the true state is neither (95,78,54) nor (95,54,78). Thus we can view the effect of the new piece of information as shrinking Ann's information set from

$$\{(95,78,54), (95,54,78), (78,95,54), (54,95,78), (78,54,95), (54,78,95)\}$$

to

$$\{(78,95,54), (54,95,78), (78,54,95), (54,78,95)\}.$$

How should Ann revise her beliefs in response to the new information?
The answer cannot be that we simply drop states (95,78,54) and (95,54,78) from the probability distribution (♦) because the result would be

state	$(78,95,54)$	$(54,95,78)$	$(78,54,95)$	$(54,78,95)$
probability	$\frac{2}{16}$	$\frac{1}{16}$	0	0

which is not a probability distribution, since the probabilities do not add up to 1 (they add up to $\frac{3}{16}$). The topic of belief revision is addressed in Section 9.4.

Test your understanding of the concepts introduced in this section, by going through the exercises in Section 9.8.2 at the end of this chapter.

[1] It was proved in Exercise 8.5 (Chapter 8) that, for every event E, $KE \subseteq E$.

[2] In the above example, if the true state is (78,54,95) then it does not belong to event G, representing the proposition "Carla did not get the highest score" and yet Ann assigns probability 1 to G, that is, she is certain of G.

9.3 Conditional probability and Bayes' rule

9.3.1 Conditional probability

Let $A, B \subseteq U$ be two events (where U is the universal set or sample space) and P a probability measure on U.
If $P(B) > 0$, the *conditional probability of A given B*, denoted by $P(A|B)$, is defined as follows:

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}} \tag{9.1}$$

For example, if $P(A \cap B) = 0.2$ and $P(B) = 0.6$ then $P(A|B) = \frac{0.2}{0.6} = \frac{1}{3}$.

One way to visualize conditional probability is to think of U as a geometric shape of area 1 (e.g. a square with each side equal to 1 unit of measurement).
For a subset A of the unit square, $P(A)$ is the area of A.
If B is a non-empty subset of the square then $A \cap B$ is that part of A that lies in B and $P(A|B)$ is the area of $A \cap B$ relative to the area of B, that is, as a fraction of the area of B. This is illustrated in Figure 9.3.

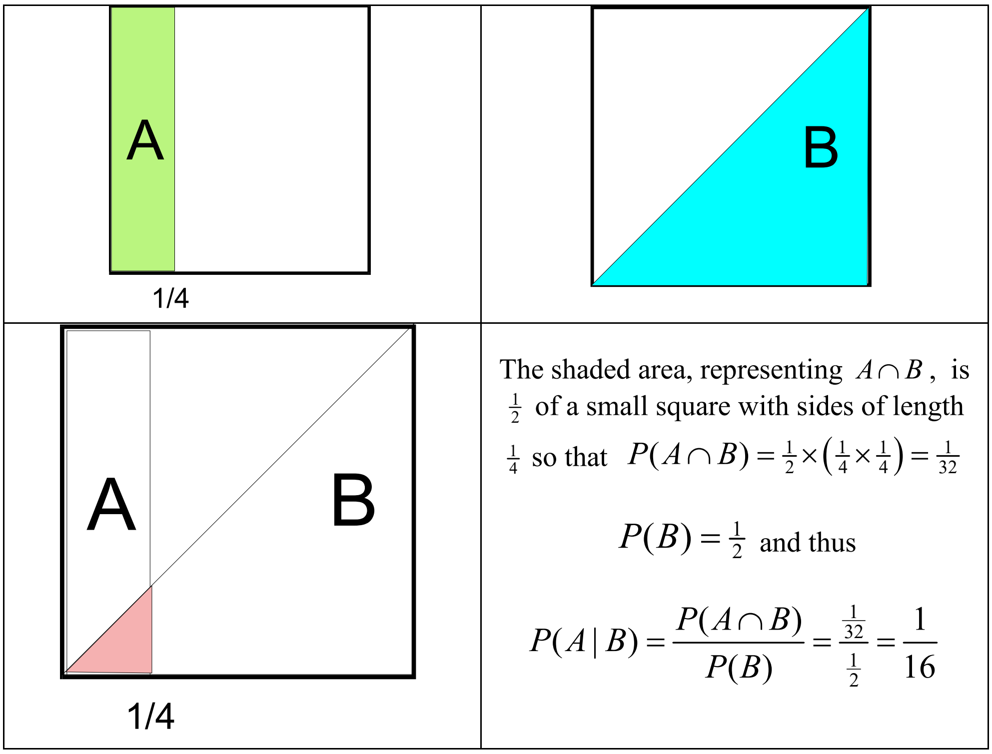


Figure 9.3: Geometric interpretation of the conditional probability $P(A|B)$.

Next we derive from the conditional probability formula (9.1) three versions of what is known as *Bayes' rule*.

Let E and F be two events such that $P(E) > 0$ and $P(F) > 0$. Then, using the conditional probability formula (9.1) we get

$$P(E|F) = \frac{P(E \cap F)}{P(F)} \tag{9.2}$$

and

$$P(F|E) = \frac{P(E \cap F)}{P(E)}. \tag{9.3}$$

From (9.3) we get that

$$P(E \cap F) = P(F|E)\,P(E) \tag{9.4}$$

and replacing (9.4) in (9.2) we get

$$\textbf{Bayes' formula} \text{ version 1 : } \boxed{P(E|F) = \frac{P(F|E)\,P(E)}{P(F)}} \tag{9.5}$$

As an illustration of how one can use (9.5), consider the following example. You are a doctor examining a middle-aged man who complains of lower back pain. You know that 25% of men in the age group of your patient suffer from lower back pain.

There are various causes of lower back pain; one of them is chronic inflammation of the kidneys. This is not a very common disease: it affects only 4% of men in the age group that you are considering. Among those who suffer from chronic inflammation of the kidneys, 85% complain of lower back pain.

What is the probability that your patient has chronic inflammation of the kidneys? Let I denote inflammation of the kidneys and L denote lower back pain.

The information you have is that $P(I) = \frac{4}{100}$, $P(L) = \frac{25}{100}$ and $P(L|I) = \frac{85}{100}$.
Thus, using (9.5), we get that

$$P(I|L) = \frac{P(L|I)\,P(I)}{P(L)} = \frac{\frac{85}{100}\left(\frac{4}{100}\right)}{\frac{25}{100}} = 0.136 = 13.6\%.$$

We now derive a second version of Bayes' formula. According to Bayes' rule (9.5), $P(E|F) = \frac{P(F|E)\,P(E)}{P(F)}$.
From set theory we have that, given any two sets A and B, $A = (A \cap B) \cup (A \cap \neg B)$ and the two sets $A \cap B$ and $A \cap \neg B$ are disjoint. Thus, $P(A) = P(A \cap B) + P(A \cap \neg B)$.

Hence, in the denominator of Bayes' formula (9.5) we can replace $P(F)$ with $P(F \cap E) + P(F \cap \neg E)$.
Then, using the formula for conditional probability we get that $P(F \cap E) = P(F|E)P(E)$ and $P(F \cap \neg E) = P(F|\neg E)P(\neg E)$. Thus, $P(F) = P(F|E)P(E) + P(F|\neg E)P(\neg E)$.

Replacing this in Bayes' formula (9.5) we get

$$\textbf{Bayes' formula} \text{ version 2 :} \quad \boxed{P(E|F) = \frac{P(F|E)\,P(E)}{P(F|E)P(E) + P(F|\neg E)P(\neg E)}} \tag{9.6}$$

As an illustration of how one can use (9.6), consider the following example.

Enrollment in a Decision Making class is as follows: 60% economics majors (E), 40% other majors ($\neg E$). In the past, 80% of the economics majors passed and 65% of the other majors passed.
A student tells you that she passed the class. What is the probability that she is an economics major? Let A stand for "pass the class". Then, using (9.6),

$$P(E|A) = \frac{P(A|E)\,P(E)}{P(A|E)P(E) + P(A|\neg E)P(\neg E)} = \frac{\frac{80}{100}\left(\frac{60}{100}\right)}{\frac{80}{100}\left(\frac{60}{100}\right) + \frac{65}{100}\left(\frac{40}{100}\right)} = \frac{24}{37} = 64.86\%.$$

One more example: 0.3763% of the population (that is, approximately 4 in 100,000 individuals) is infected with the HIV virus.
Let H be the event "a randomly selected individual has the HIV virus".
Then $P(H) = 0.003763$ and $P(\neg H) = 0.996237$.

A blood test can be used to detect the virus. The blood test has a true positive rate (sensitivity) of 99.97% and a true negative rate (specificity) of 98.5%. Thus, (letting '+' denote a positive blood test and '$-$' a negative blood test) $P(+|H) = 0.9997$, $P(-|H) = 0.0003$, $P(+|\neg H) = 0.015$ and $P(-|\neg H) = 0.985$.

Now suppose that you pick an individual at random, administer the blood test and it turns out to be positive. What is the probability that the individual has the HIV virus? That is, what is $P(H|+)$? Using (9.6),

$$P(H|+) = \frac{P(+|H)\;P(H)}{P(+|H)\;P(H) + P(+|\neg H)\;P(\neg H)}$$

$$= \frac{0.9997\,(0.003763)}{0.9997\,(0.003763) + 0.015\,(0.996237)} = 0.201 = 20.1\%.$$

A generalization of (9.6) is as follows: If $\{E_1, \ldots, E_n\}$ is a partition of the sample space U,[3] then, for every event F, $P(F) = P(F|E_1)\;P(E_1) + \cdots + P(F|E_n)\;P(E_n)$.
Thus, using (9.5) we obtain that, for every $i = 1, \ldots, n$,

$$\textbf{Bayes' formula} \text{ version 3 :} \quad \boxed{P(E_i|F) = \frac{P(F|E_i)\,P(E_i)}{P(F|E_1)P(E_1) + \ldots + P(F|E_n)P(E_n)}} \tag{9.7}$$

[3] That is, the sets $E_1, \ldots, E_n$ (1) cover the set U (in the sense that $E_1 \cup \cdots \cup E_n = U$) and (2) are pairwise disjoint (in the sense that, for all $i, j = 1, \ldots, n$ with $i \neq j$, $E_i \cap E_j = \emptyset$).

Example: enrollment in a class is restricted to the following majors: economics (E), statistics (S) and math (M). Current enrollment is: 40% E, 35% S and 25% M.
Let A be the event "pass the class". According to past data, $P(A|E) = 60\%$, $P(A|S) = 50\%$ and $P(A|M) = 75\%$.
A student from this class tells you that she received a passing grade. What is the probability that she is an economics major? Using (9.7),

$$P(E|A) = \frac{P(A|E)\,P(E)}{P(A|E)\,P(E) + P(A|S)\,P(S) + P(A|M)\,P(M)}$$

$$= \frac{\frac{60}{100}\left(\frac{40}{100}\right)}{\frac{60}{100}\left(\frac{40}{100}\right) + \frac{50}{100}\left(\frac{35}{100}\right) + \frac{75}{100}\left(\frac{25}{100}\right)} = \frac{96}{241} = 39.83\%.$$

Test your understanding of the concepts introduced in this section, by going through the exercises in Section 9.8.3 at the end of this chapter.

9.4 Changing beliefs in response to new information

The issue of how to "rationally" modify one's initial beliefs – expressed as a probability measure P on a set U – after receiving an item of information (represented by a subset F of U) has been studied extensively by philosophers and logicians. Two different situations may arise:

- In one case, the item of information F was not ruled out by the initial beliefs, in the sense that event F was assigned positive probability ($P(F) > 0$). Information might still be somewhat surprising, in case $P(F)$ is small (close to zero), but it is not completely unexpected. We call this case *belief updating*.
- The other case is where the item of information was initially dismissed, in the sense that it was assigned zero probability ($P(F) = 0$). In this case the information received is completely surprising or completely unexpected. We call this case *belief revision*.

We shall first address the issue of belief updating.

9.4.1 Belief updating

It is generally agreed that the rational way to update one's beliefs is by conditioning the initial probability measure on the information received, that is, by using the conditional probability formula (9.1).

Definition 9.4.1 We use the expression *belief updating* or *Bayesian updating* to refer to the modification of initial beliefs (expressed by an initial probability distribution P) obtained by applying the conditional probability rule; this assumes that the belief change is prompted by the arrival of new information, represented by an event F such that $P(F) > 0$.

Thus, when receiving a piece of information $F \subseteq U$ such that $P(F) > 0$, one would change one's initial probability measure P into a new probability measure P_{new} by

- reducing the probability of every state in $\neg F$ (the complement of F) to zero (this captures the notion that the information represented by F is trusted to be correct), and
- setting $P_{new}(s) = P(s|F)$ for every state $s \in F$.

Thus, for every state $s \in U$,

$$P_{new}(s) = P(s|F) = \begin{cases} 0 & \text{if } s \notin F \\ \frac{P(s)}{P(F)} & \text{if } s \in F \end{cases} \tag{9.8}$$

(recall the assumption that $P(F) > 0$).
Thus, for every event $E \subseteq U$, $P_{new}(E) = \sum\limits_{s \in E} P_{new}(s) = \sum\limits_{s \in E} P(s|F) = P(E|F)$.

As an illustration, let us go back to the example of Section 9.2 concerning three students: Ann, Bob and Carla. The professor tells them that in the last exam one of them got 95 points (out of 100), another 78 and the third 54. We represented a state as a triple (a,b,c), where a is Ann's score, b is Bob's score and c is Carla's score. The initial information given by the professor is thus represented by the set

$$U = \{(95,78,54),(95,54,78),(78,95,54),(78,54,95),(54,95,78),(54,78,95)\}$$

Based on the results of previous exams, Ann forms the following probabilistic beliefs:

$$\begin{array}{cccccc} (95,78,54) & (95,54,78) & (78,95,54) & (54,95,78) & (78,54,95) & (54,78,95) \\ \frac{9}{16} & \frac{4}{16} & \frac{2}{16} & \frac{1}{16} & 0 & 0 \end{array}$$

We then supposed that – before distributing the exams – the professor made the additional remark "I was surprised to see that, this time, Ann did not get the highest score". This new announcement by the professor informs the students that the true state is neither (95,78,54) nor (95,54,78).
Thus the new piece of information is represented by the event

$$F = \{(78,95,54),(54,95,78),(78,54,95),(54,78,95)\}.$$

How should Ann revise her beliefs in response to this new piece of information?
Conditioning Ann's initial beliefs on the event F yields the following updated beliefs:

$$\begin{array}{cccccc} (95{,}78{,}54) & (95{,}54{,}78) & (78{,}95{,}54) & (54{,}95{,}78) & (78{,}54{,}95) & (54{,}78{,}95) \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} & 0 & 0 \end{array}$$

These updated beliefs can be represented more succinctly as a probability distribution on the set F (that is, by dropping the states that belong to the complement of F, which are zero-probability states in the updated beliefs):

$$\begin{array}{cccc} (78,95,54) & (54,95,78) & (78,54,95) & (54,78,95) \\ \frac{2}{3} & \frac{1}{3} & 0 & 0 \end{array}$$

9.4.2 Belief revision

[Note: the material of this section will not be needed until Chapter 13. It is presented here for completeness on the topic of beliefs.]

How should a rational individual revise her beliefs when receiving information that is completely surprising, that is, when informed of an event E to which her initial beliefs assigned zero probability ($P(E) = 0$)?

As we will see in Part IV, belief revision is very important in dynamic (or extensive-form) games. In such games a player may find herself at an information set that, according to her initial beliefs, had zero probability of being reached and thus will have to form new beliefs reflecting the unexpected information.

The best known theory of rational belief revision is the so-called *AGM theory*, which takes its name from its originators: Alchourrón (a legal scholar), Gärdenfors (a philosopher) and Makinson (a computer scientist); their pioneering contribution was published in 1985. Just like the theory of expected utility (Chapter 5), the AGM theory is an axiomatic theory: it provides a list of "rationality" axioms for belief revision and provides a representation theorem.[4]

Although the AGM theory was developed within the language of propositional logic, it can be restated in terms of a set of states and a collection of possible items of information represented as events. We first introduce the non-probabilistic version of the theory and then add graded beliefs, that is, probabilities.

Let U be a finite set of states and $\mathscr{E} \subseteq 2^U$ a collection of events (subsets of U) representing possible items of information; we assume that $U \in \mathscr{E}$ and $\emptyset \notin \mathscr{E}$. To represent initial beliefs and revised beliefs we introduce a function $f : \mathscr{E} \to 2^U$, which we call a belief revision function.

Definition 9.4.2 Let U be a finite set of states and $\mathscr{E}$ a collection of events such that $U \in \mathscr{E}$ and $\emptyset \notin \mathscr{E}$. A *belief revision function* is a function $f : \mathscr{E} \to 2^U$ that satisfies the following properties: for every $E \in \mathscr{E}$, (1) $f(E) \subseteq E$ and (2) $f(E) \neq \emptyset$.

The interpretation of a belief revision function is as follows. First of all, $f(U)$ represents the initial beliefs, namely the set of states that the individual initially considers possible.[5] Secondly, for every $E \in \mathscr{E}$, $f(E)$ is the set of states that the individual would consider possible if informed that the true state belongs to E; thus $f(E)$ represents the individual's *revised beliefs* after receiving information E.[6]

---

[4] We will not list and discuss the axioms here. The interested reader can consult `http://plato.stanford.edu/entries/formal-belief/` or, for a discussion which is closer to the approach followed in this section, Bonanno (2009).

[5] The universal set U can be thought of as representing minimum information: all states are possible. If the initial beliefs were to be expressed probabilistically, by means of a probability distribution P over U, then $f(U)$ would be the support of P, that is, the set of states to which P assigns positive probability. Thus, $f(U)$ would be the smallest event of which the individual would initially be certain (that is, to which she assigns probability 1): she would initially be certain of (assign probability 1 to) any event F such that $f(U) \subseteq F$.

[6] If the revised beliefs after receiving information E were to be expressed probabilistically, by means of a probability distribution P_E over U, then $f(E)$ would be the support of P_E, that is, the set of states to which P_E assigns positive probability. Thus, $f(E)$ would be the smallest event of which the individual would be certain *after having been informed that* E: according to her revised beliefs she would be certain of any event F such that $f(E) \subseteq F$. [Note that, since – by assumption – $f(E) \subseteq E$, the individual is assumed to be certain of the information received (e.g. because she trusts the source of the information).]

One of the implications of the AGM axioms for belief revision is the following condition, which is known as Arrow's Axiom (proposed by the Nobel laureate Ken Arrow in the context of rational choice, rather than rational belief revision):

$$if\ \ E,F \in \mathscr{E},\ \ E \subseteq F\ \ and\ \ E \cap f(F) \neq \emptyset\ \ then\ \ f(E) = E \cap f(F).$$

Arrow's Axiom says that if information E implies information F ($E \subseteq F$) and there are states in E that would be considered possible upon receiving information F ($E \cap f(F) \neq \emptyset$), then the states that the individual would consider possible if informed that E are precisely those that belong to both E and $f(F)$ ($f(E) = E \cap f(F)$).

Although necessary for a belief revision policy that satisfies the AGM axioms, Arrow's Axiom is not sufficient. Before stating the necessary and sufficient conditions for rational belief revision, we remind the reader of the notion of a complete and transitive relation on a set U (Chapter 2, Section 2.1).

- In Chapter 2 the relation was denoted by $\succsim$ and was interpreted in terms of *preference*: $o_1 \succsim o_2$ was interpreted as "the individual considers outcome o_1 to be at least as good as outcome o_2".
- In the present context the interpretation is in terms of "plausibility": $s \succsim s'$ means that the individual considers state s to be *at least as plausible* as state s'; $s \succ s'$ means that s is considered to be *more plausible* than s' and $s \sim s'$ means that s is considered to be *just as plausible as* s'.

Definition 9.4.3 A *plausibility order* on a set of states U is a binary relation $\succsim$ on U that is complete (for every two states s_1 and s_2, either $s_1 \succsim s_2$ or $s_2 \succsim s_1$, or both) and transitive
(if $s_1 \succsim s_2$ and $s_2 \succsim s_3$ then $s_1 \succsim s_3$). We define $s_1 \succ s_2$ as "$s_1 \succsim s_2$ and $s_2 \not\succsim s_1$" and we define $s_1 \sim s_2$ as "$s_1 \succsim s_2$ and $s_2 \succsim s_1$".

The following theorem is based on a result by Adam Grove.[7]

Theorem 9.4.1 Let U be a finite set of states, $\mathscr{E}$ a collection of events (representing possible items of information), with $U \in \mathscr{E}$ and $\emptyset \notin \mathscr{E}$, and $f : \mathscr{E} \to 2^U$ a belief revision function (Definition 9.4.2). Then the belief revision policy represented by the function f is compatible with the AGM axioms of belief revision if and only if there exists a plausibility order $\succsim$ on U that *rationalizes* f in the sense that, for every $E \in \mathscr{E}$, $f(E)$ is the set of most plausible states in E: $f(E) = \{s \in E : s \succsim s' \text{ for every } s' \in E\}$.

[7] Adam Grove, Two modelings for theory change, *Journal of Philosophical Logic*, 1988, Vol. 17, pages 157-170. That result was proved within the context of propositional logic. The version given here is proved in Giacomo Bonanno, Rational choice and AGM belief revision, *Artificial Intelligence*, 2009, Vol. 88, pages 221-241.

Definition 9.4.4 A belief revision function $f : \mathscr{E} \to 2^U$ which is rationalized by a plausibility order is called an *AGM belief revision function*.

R An AGM belief revision function satisfies Arrow's Axiom. The converse is not true: it is possible for a belief revision function $f : \mathscr{E} \to 2^U$ to satisfy Arrow's Axiom and yet fail to be rationalized by a plausibility order.

Within the context of probabilistic beliefs, let P be the probability distribution on a finite set of states U that represents the initial beliefs and P_E be the probability distribution representing the updated beliefs after receiving information E such that $P(E) > 0$ (thus the information is not surprising).

The *support* of a probability distribution P, denoted by $Supp(P)$, is the set of states to which P assigns positive probability: $Supp(P) = \{s \in U : P(s) > 0\}$.

The rule for *updating* beliefs upon receiving information E (Definition 9.4.1) implies the following:

$$if \quad E \cap Supp(P) \neq \emptyset \quad \text{(that is, } P(E) > 0\text{)} \quad then \quad Supp(P_E) = E \cap Supp(P). \qquad (9.9)$$

We call this the *qualitative belief updating rule* or qualitative Bayes' rule. It is easy to check that the qualitative belief updating rule is implied by Arrow's Axiom (see Exercise 9.22). Thus, by the above remark, an AGM belief revision function has incorporated in it the qualitative belief updating rule. In other words, *belief updating is included in the notion of AGM belief revision*.

A belief revision function, however, goes beyond belief updating because it also encodes new beliefs after receipt of surprising information (that is, after being informed of an event E such that $P(E) = 0$).

What is the probabilistic version of AGM belief revision? It turns out that in order to obtain probabilistic beliefs we only need to make a simple addition to an AGM belief revision function $f : \mathscr{E} \to 2^U$.

- Let P_0 be any full-support probability distribution on U (that is, P_0 is such that $P_0(s) > 0$, for every $s \in U$).
- Then, for every $E \in \mathscr{E}$, let P_E be the probability distribution obtained by conditioning P_0 on $f(E)$ (note: on $f(E)$, not on E):

$$P_E(s) = P_0(s|f(E)) = \begin{cases} \dfrac{P_0(s)}{\sum\limits_{s' \in f(E)} P_0(s')} & if\ s \in f(E) \\ \\ 0 & if\ s \notin f(E) \end{cases}$$

- Then P_U gives the initial probabilistic beliefs and, for every other $E \in \mathscr{E}$, P_E gives the revised probabilistic beliefs after receiving information E.
- The collection $\{P_E\}_{E \in \mathscr{E}}$ of probability distributions on U so obtained gives the individual's *probabilistic* belief revision policy (while the function $f : \mathscr{E} \to 2^U$ gives the individual's *qualitative* belief revision policy).

Definition 9.4.5 Let U be a finite set of states and $\mathscr{E}$ a collection of events such that $U \in \mathscr{E}$ and $\emptyset \notin \mathscr{E}$. A *probabilistic belief revision policy* is a collection $\{P_E\}_{E\in\mathscr{E}}$ of probability distributions on U such that, for every $E \in \mathscr{E}$, $Supp(P_E) \subseteq E$. P_U represents the initial beliefs and, for every other $E \in \mathscr{E}$, P_E represents the revised beliefs after receiving information E. The collection $\{P_E\}_{E\in\mathscr{E}}$ is called an *AGM probabilistic belief revision policy* if it satisfies the following properties:

1. there exists a plausibility order $\succsim$ on U such that, for every $E \in \mathscr{E}$, $Supp(P_E)$ is the set of most plausible states in E, that is,

$$Supp(P_E) = \left\{s \in E : s \succsim s' \text{ for every } s' \in E\right\},$$ [a]

2. there exists a full-support probability distribution P_0 on U such that, for every $E \in \mathscr{E}$, P_E is the probability distribution obtained by conditioning P_0 on $Supp(P_E)$.

[a] This condition says that if one defines the function $f : \mathscr{E} \to 2^U$ by $f(E) = Supp(P_E)$ then this function is an AGM belief revision function (see Definition 9.4.4).

Test your understanding of the concepts introduced in this section, by going through the exercises in Section 9.8.4 at the end of this chapter.

9.5 Harsanyi consistency of beliefs or like-mindedness

[Note: the material of this section will not be needed until Chapter 14. It is presented here for completeness on the topic of beliefs.]

We can easily extend the analysis to the case of two or more individuals. We already know how to model interactive knowledge by means of information partitions; the addition of beliefs is a simple step: we merely add, for every individual and for every information set, a probability distribution over the elements of that information set.[8] A two-person example is shown in Figure 9.4.

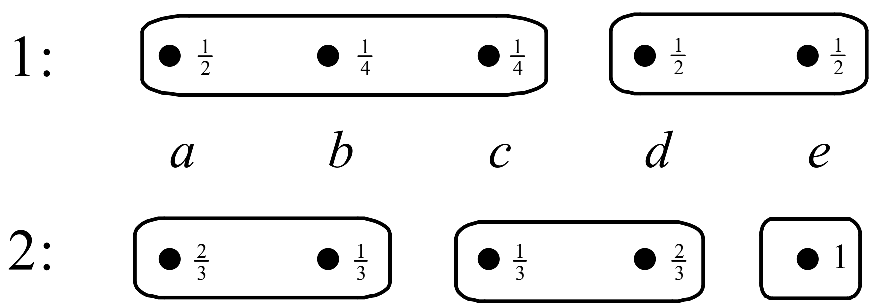


Figure 9.4: A two-person knowledge-belief structure.

[8] The probability distribution over an information set can be thought of as a probability distribution over the universal set U by assigning probability zero to every state which is not in the information set.

In this example, at every state, the two individuals hold different beliefs. For example, consider event $E = \{b, c\}$ and state a. Individual 1 attaches probability $\frac{1}{2}$ to E while Individual 2 attaches probability $\frac{1}{3}$ to E.

Can two "equally rational" individuals hold different beliefs? The answer is: of course! In the above example it is not surprising that the two individuals assign different probabilities to the same event E, because they have different information.

If the true state is a, then Individual 1's information is that the true state is either a or b or c, while Individual 2 considers only a and b possible (Individual 2 knows more than Individual 1).

Is there a precise way of expressing the fact that two individuals assign different probabilities to an event *exclusively because they have different information*?

In the above example we could ask the hypothetical question: if Individual 1 had the same information as Individual 2, would he agree with Individual 2's assessment that the probability of event $E = \{b, c\}$ is $\frac{1}{3}$?

This is, of course, a counterfactual question. The answer to this counterfactual question is affirmative: imagine giving Individual 1 the information that the true state is either a or b; then – according to Definition 9.4.1 – he would update his beliefs from $\begin{pmatrix} a & b & c \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$ to $\begin{pmatrix} a & b \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}$ and thus have the same beliefs as Individual 2.

We say that two individuals are *like-minded* if it is the case that they would have the same beliefs if they had the same information.

It is not straightforward how to turn this into a precise definition.

Consider, again, the example of Figure 9.4 and event $E = \{b, c\}$. Above we asked the question "what would Individual 1 believe if he knew as much as Individual 2?" This is a simple question because we can imagine giving more information to Individual 1 and have him update his beliefs based on that information.

However, we could also have asked the question "what would Individual 2 believe if he knew as little as Individual 1?" In this case we would have to imagine "taking away information" from Individual 2, by increasing his information set from $\{a,b\}$ to $\{a,b,c\}$. This is not the same as asking Individual 2 to update his beliefs based on information $\{a,b,c\}$, because updating on something you already know leaves your beliefs unchanged.

There is a sense in which the beliefs of the two individuals of Figure 9.4 are "in agreement": for example, they both consider state a twice as likely as state b. One could try to use this condition to define like-mindedness: for every two states x and y, whenever two individuals consider both x and y as possible (given their information) then they agree on the relative likelihood of x versus y.

Unfortunately, this condition is too weak. To see this, consider the three-individual example of Figure 9.5.

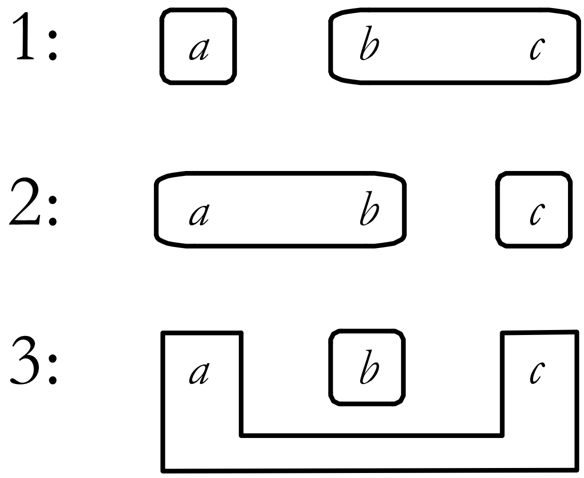


Figure 9.5: A three-person knowledge-belief structure.

No matter what the true state is, we cannot find two individuals and two states x and y that both individuals consider possible. Thus any beliefs would make the three individuals like-minded. For example, consider the following beliefs:

- Individual 1 at his information set $\{b,c\}$ assigns equal probability to b and c (thus considering b to be as likely as c),
- Individual 2 at her information set $\{a,b\}$ assigns probability $\frac{3}{4}$ to a and $\frac{1}{4}$ to b (thus considering a to be three times more likely than b), and
- Individual 3 at his information set $\{a,c\}$ assigns probability $\frac{1}{4}$ to a and $\frac{3}{4}$ to c (thus considering c to be three times more likely than a).

Then, putting together the beliefs of Individuals 1 and 2, we would have that a is judged to be three times more likely than c (according to Individual 2, a is three times more likely than b and, according to Individual 1, b is just as likely as c), while Individual 3 has the opposite judgment that c is three times more likely than a.

In order to give a precise definition of like-mindedness we need to introduce some notation.

- Let there be n individuals $(n \geq 2)$.
- Let U be a set of states and let $\mathscr{I}_i$ be the information partition of individual i $(i \in \{1, \ldots, n\})$.
- As usual, if s is a state, we denote by $I_i(s)$ the element (information set) of the partition $\mathscr{I}_i$ that contains s.
- Let $P_{i,s}$ denote the beliefs of individual i at state s, that is, $P_{i,s}$ is a probability distribution over $I_i(s)$.
- Clearly, we can think of $P_{i,s}$ as a probability distribution over the entire set of states U satisfying the property that if $s' \notin I_i(s)$ then $P_{i,s}(s') = 0$.
- Note that, for every individual i, there is just one probability distribution over $I_i(s)$ and thus if $s' \in I_i(s)$ then $P_{i,s} = P_{i,s'}$.[9]

Definition 9.5.1 A probability distribution P over U is called a *common prior* if, for every individual i and for every state s,

1. $P(I_i(s)) > 0$, and
2. updating P on $I_i(s)$ (see Definition 9.4.1) yields precisely $P_{i,s}$, that is, $P(s'|I_i(s)) = P_{i,s}(s')$, for every $s' \in I_i(s)$.

   When a common prior exists we say that the individuals are *like-minded* or that the individuals' beliefs are *Harsanyi consistent*.[a]

[a]John Harsanyi, who in 1994 won the Nobel Memorial prize in Economics (together with Reinhardt Selten and John Nash), introduced the *theory of games of incomplete information* which will be the object of Part V. In that theory the notion of Harsanyi consistency plays a crucial role.

[9]Thus, every individual knows his own probabilistic beliefs.

For instance, in the example of Figure 9.6, which reproduces Figure 9.4, a common prior exists and thus the two individuals are like-minded. Indeed, a common prior is

$$\begin{pmatrix} a & b & c & d & e \\ \frac{2}{8} & \frac{1}{8} & \frac{1}{8} & \frac{2}{8} & \frac{2}{8} \end{pmatrix}$$

(the reader should convince himself/herself that, indeed, updating this probability distribution on each information set yields the probability distribution written inside that information set). How can we determine if a common prior exists? The issue of existence

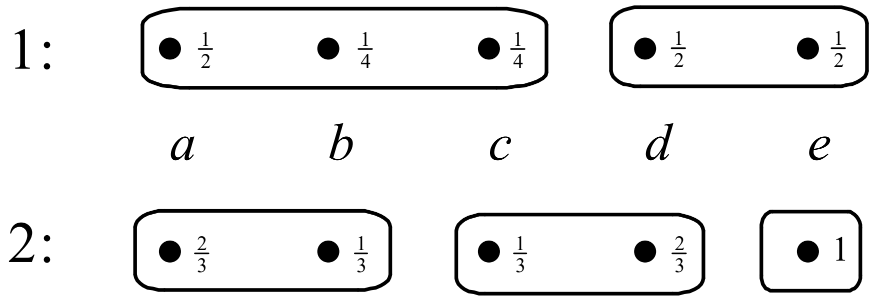


Figure 9.6: Copy of Figure 9.4.

of a common prior can be reduced to the issue of whether a system of equations has a solution. To see this, let us go back to the example of Figure 9.6. A common prior would be a probability distribution

$$\begin{pmatrix} a & b & c & d & e \\ p_a & p_b & p_c & p_d & p_e \end{pmatrix}$$

that satisfies the following conditions:

1. Updating on information set $\{a,b,c\}$ of Individual 1 we need

$$\frac{p_b}{p_a+p_b+p_c} = \frac{1}{4} \quad \text{and} \quad \frac{p_c}{p_a+p_b+p_c} = \frac{1}{4}.$$

   Note that from these two conditions the third condition follows, namely

$$\frac{p_a}{p_a+p_b+p_c} = \frac{1}{2}.$$

2. Updating on information set $\{d,e\}$ of Individual 1 we need $\frac{p_d}{p_d+p_e} = \frac{1}{2}$, from which it follows that $\frac{p_e}{p_d+p_e} = \frac{1}{2}$.

3. Updating on information set $\{a,b\}$ of Individual 2 we need $\frac{p_a}{p_a+p_b} = \frac{2}{3}$, from which it follows that $\frac{p_b}{p_a+p_b} = \frac{1}{3}$.
4. Updating on information set $\{c,d\}$ of Individual 2 we need $\frac{p_c}{p_c+p_d} = \frac{1}{3}$, from which it follows that $\frac{p_d}{p_c+p_d} = \frac{2}{3}$.

From the first condition we get $\boxed{p_b = p_c}$, from the second $\boxed{p_d = p_e}$, from the third $\boxed{p_a = 2p_b}$ and from the fourth $\boxed{p_d = 2p_c}$.
Adding to these three equalities the requirement that $p_a + p_b + p_c + p_d + p_e = 1$, we have a system of five equations in five unknowns, which admits a solution, namely $\left(\begin{array}{ccccc} a & b & c & d & e \\ \frac{2}{8} & \frac{1}{8} & \frac{1}{8} & \frac{2}{8} & \frac{2}{8} \end{array}\right)$.

It is not always the case that a common prior exists. For instance, if we add to the example of Figure 9.5 the beliefs shown in Figure 9.7, then we get a situation where the individuals' beliefs are not Harsanyi consistent. In this case, from the updating conditions of Individuals 1 and 2 one would get that $p_a = p_b$ and $p_b = p_c$, from which it follows that $p_a = p_c$; however, from the updating condition for Individual 3 we get that $p_a = 3p_c$, yielding a contradiction.

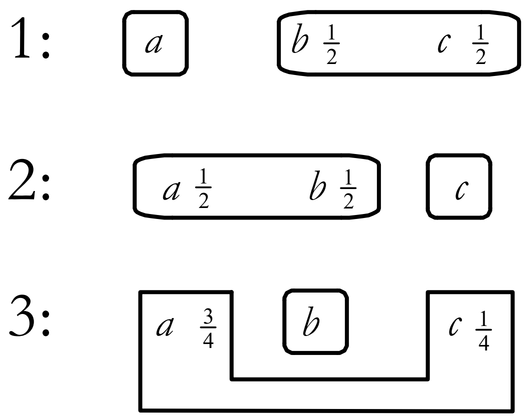


Figure 9.7: The structure of Figure 9.5 with the addition of beliefs.

Test your understanding of the concepts introduced in this section, by going through the exercises in Section 9.8.5 at the end of this chapter.

9.6 Agreeing to disagree

Can two rational and like-minded individuals agree to disagree? This question was raised in 1976 by Robert Aumann (who received the Nobel Memorial prize in Economics in 2005, together with Thomas Schelling). As remarked above, it is certainly quite possible for two rational individuals to have different beliefs about a particular event, because they might have different information. Let us go back to the example of Figure 9.4, reproduced in Figure 9.8 together with the common prior (showing that the individuals are like-minded).

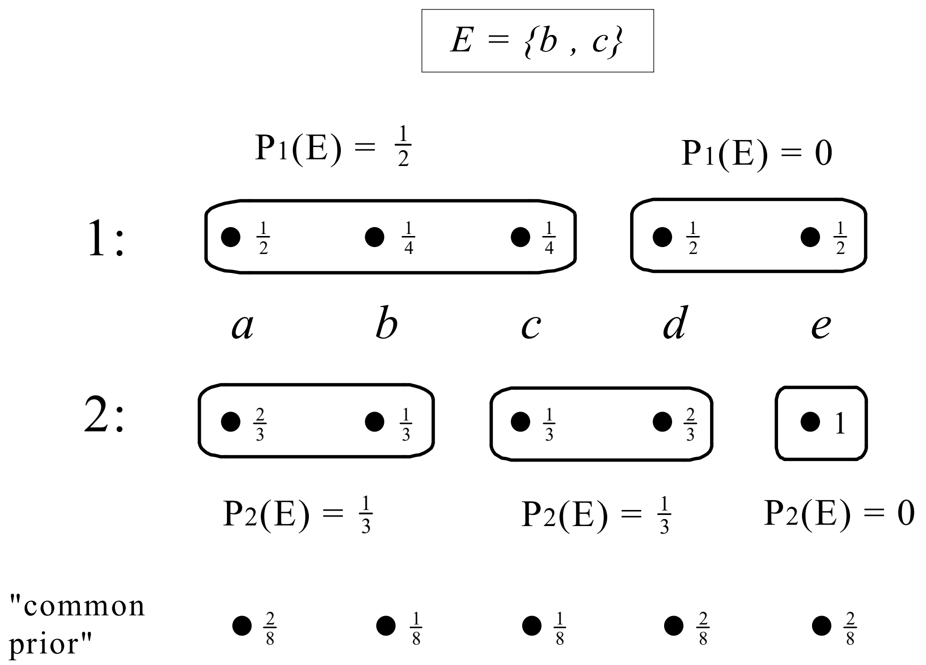


Figure 9.8: The structure of Figure 9.4 with a common prior.

Suppose that the true state is a and consider what the two individuals believe about event $E = \{b,c\}$.

- ◇ Individual 1's information set is $\{a,b,c\}$ and, given his beliefs at that information set, he attaches probability $\frac{1}{4}+\frac{1}{4}=\frac{1}{2}$ to E: let us denote this by "at state a, $P_1(E)=\frac{1}{2}$".
- ◇ Individual 2's information set is $\{a,b\}$ and, given her beliefs at that information set, she attaches probability $\frac{1}{3}$ to E: let us denote this by "at state a, $P_2(E)=\frac{1}{3}$".
- ◇ Thus the two individuals disagree about the probability of event E. Furthermore, they know that they disagree.

To see this, let $\left\|P_1(E)=\frac{1}{2}\right\|$ be the event that Individual 1 attaches probability $\frac{1}{2}$ to E; then

$$\left\|P_1(E)=\tfrac{1}{2}\right\| = \{a,b,c\}.$$

Similarly, let $\left\|P_2(E) = \frac{1}{3}\right\|$ be the event that Individual 2 attaches probability $\frac{1}{3}$ to E; then

$$\left\|P_2(E) = \tfrac{1}{3}\right\| = \{a,b,c,d\}.$$

These are events and thus we can check at what states the two individuals know them. Using Definition 8.1.3 (Chapter 8), we have that

$$K_1 \underbrace{\left\|P_2(E) = \tfrac{1}{3}\right\|}_{=\{a,b,c,d\}} = \{a,b,c\} \qquad \text{and} \qquad K_2 \underbrace{\left\|P_1(E) = \tfrac{1}{2}\right\|}_{=\{a,b,c\}} = \{a,b\}.$$

Hence

$$a \in \left\|P_1(E) = \tfrac{1}{2}\right\| \,\cap\, \left\|P_2(E) = \tfrac{1}{3}\right\| \,\cap\, K_1\left\|P_2(E) = \tfrac{1}{3}\right\| \,\cap\, K_2\left\|P_1(E) = \tfrac{1}{2}\right\|.$$

Thus at state a not only do the individuals disagree, but they know that they disagree. However, their disagreement is not common knowledge. Indeed, $K_1K_2\left\|P_1(E) = \frac{1}{2}\right\| = \emptyset$ and thus $a \notin K_1K_2\left\|P_1(E) = \frac{1}{2}\right\| = \emptyset$, that is, at state a it is not the case that Individual 1 knows that Individual 2 knows that Individual 1 assigns probability $\frac{1}{2}$ to event E. As the following theorem states, the opinions of two like-minded individuals about an event E cannot be in disagreement and, at the same time, commonly known.

The following theorem is proved in Section 9.7.

Theorem 9.6.1 — Agreement Theorem; Aumann, 1976. Let U be a set of states and consider a knowledge-belief structure with two individuals, 1 and 2. Let E be an event and let $p,q \in [0,1]$. Suppose that at some state s it is common knowledge that Individual 1 assigns probability p to E and Individual 2 assigns probability q to E. Then, if the individuals' beliefs are Harsanyi consistent (Definition 9.5.1), $p = q$. In other words, two like-minded individuals *cannot agree to disagree* about the probability of an event. Formally, if there exists a common prior and $s \in CK\left(\|P_1(E) = p\| \cap \|P_2(E) = q\|\right)$ then $p = q$.

Another way to think about this result is to imagine that the two individuals communicate their opinions to each other. Hearing that the other person has a different opinion is in itself a valuable piece of information, which ought to be incorporated (by updating) into one's beliefs. Thus sequential communication leads to changing beliefs. If, at the end of this process, the beliefs become common knowledge, then they must be identical.[10] We shall illustrate this with an example.

Imagine two scientists who agree that the laws of Nature are such that the true state of the world must be one of seven, call them *a, b, c, d, e, f, g*. They also agree on the relative likelihood of these possibilities, which they take to be as follows:

$$\begin{array}{ccccccc} a & b & c & d & e & f & g \\ \frac{4}{32} & \frac{2}{32} & \frac{8}{32} & \frac{5}{32} & \frac{7}{32} & \frac{2}{32} & \frac{4}{32} \end{array}$$

Experiments can be conducted to learn more. An experiment leads to a partition of the set of states. For example, if the true state is a and you perform an experiment then you might learn that the true state cannot be d or e or f or g but you still would not know which is the true state among the remaining ones. Suppose that the scientists agree that Scientist 1, from now on denoted by S_1, will perform experiment 1 and Scientist 2, denoted by S_2, will perform experiment 2. They also agree that each experiment would lead to a partition of the set of states as shown in Figure 9.9.

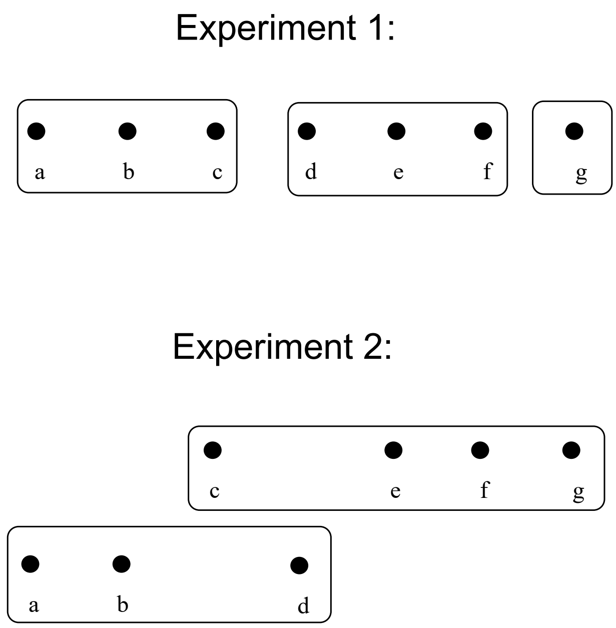


Figure 9.9: The partitions representing the two experiments.

[10] This line of reasoning was investigated by Geanakoplos and Polemarchakis (1982).

Suppose that the scientists are interested in establishing the truth of a proposition that is represented by the event $E = \{a, c, d, e\}$. Initially (given their shared probabilistic beliefs) they agree that the probability that E is true is 75%:

$$P(E) = P(a) + P(c) + P(d) + P(e) = \tfrac{4}{32} + \tfrac{8}{32} + \tfrac{5}{32} + \tfrac{7}{32} = \tfrac{24}{32} = \tfrac{3}{4} = 75\%.$$

Before they perform the experiments they also realize that, depending on what the true state is, after the experiment they will have an updated probability of event E conditional on what the experiment has revealed.

For example, they agree that if one performs Experiment 1 and the true state is b then the experiment will yield the information $F = \{a, b, c\}$ and $P(E|F)$ is given by

$$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{P(a) + P(c)}{P(a) + P(b) + P(c)} = \frac{\frac{4}{32} + \frac{8}{32}}{\frac{4}{32} + \frac{2}{32} + \frac{8}{32}} = \tfrac{12}{14} = \tfrac{6}{7} = 86\%.$$

[Note the interesting fact that sometimes experiments, although they are informative – that is, they reduce one's state of uncertainty – might actually induce one to become more confident of the truth of something that is false: in this case one would increase one's subjective probability that E is true from 75% to 86%, although E is actually false if the true state is b, as we hypothesized.]

We can associate with every cell of each experiment (that is, with every possible state of information yielded by the experiment) a new updated probability of event E, as shown in Figure 9.10.

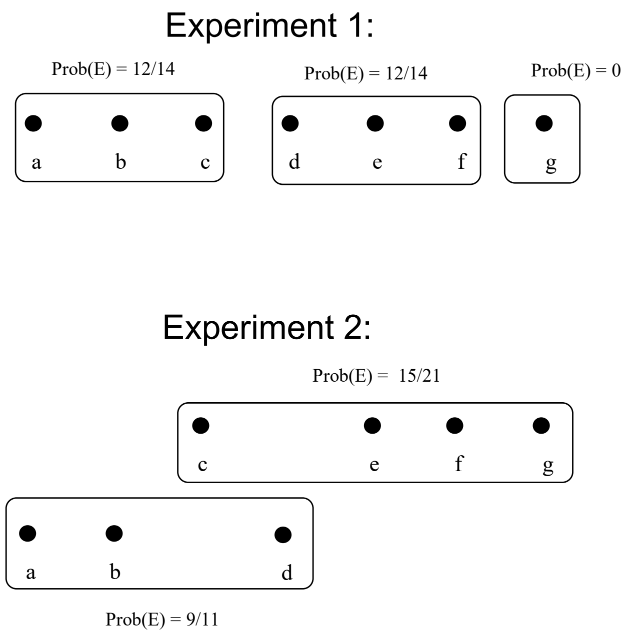
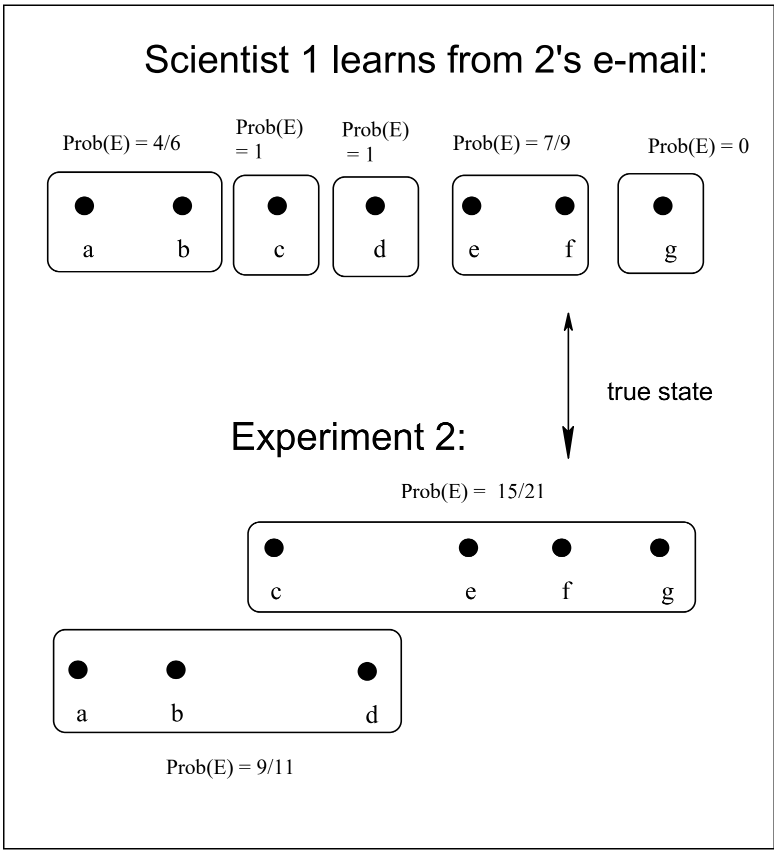


Figure 9.10: The probability of event E conditional on the result of each experiment.

Suppose now that each scientist goes to her laboratory and performs the respective experiment (Scientist 1 performs Experiment 1 and Scientist 2 performs Experiment 2). Assume also that

the true state is *f*

Suppose that the two scientists send each other an e-mail communicating their new subjective estimates of the truth of E. Scientist 1 writes that he now attaches probability $\frac{12}{14}$ to E, while Scientist 2 says that she attaches probability $\frac{15}{21}$ to E. So their estimates disagree (not surprisingly, since they have performed different experiments and have collected different information). Should they be happy with these estimates? The answer is negative.
Consider first S_1 (Scientist 1). He hears that S_2 has a revised probability of $\frac{15}{21}$ (recall our assumption that the true state is f). What does he learn from this? He learns that the true state cannot be d (if it had been, then he would have received an e-mail from S_2 saying "my new probability is $\frac{9}{11}$"). S_1's new state of knowledge and corresponding probabilities after receiving S_2's e-mail are then as shown in Figure 9.11.



If *g* is the true state, then experiment 1 reveals it.
If either *d*, *e* or *f* is the true state, then experiment 1 reveals $\{d,e,f\}$.
With *e* or *f* S_2's e-mail is "15/21", with *d* it is "9/11".
Thus with *e* or *f* one can narrow down to $\{e,f\}$ while with *d* one learns that the true state is *d*.
If the true state is *a*, *b* or *c*, then experiment 1 reveals $\{a,b,c\}$.
With *a* or *b* S_2 sends "9/11" while with *c* she sends "15/21".
Thus the e-mail splits $\{a,b,c\}$ into $\{a,b\}$ and $\{c\}$.

Figure 9.11: S_1's assessment of the probability of event E after receiving S_2's first e-mail.

Consider now Scientist 2. From S_1's e-mail she learns that S_1 has a new updated probability of $\frac{12}{14}$. What can she deduce from this? That the true state is *not* g (if it had been g then she would have received an e-mail from S_1 saying that her revised probability of E was zero). Thus she can revise her knowledge partition by eliminating g from her information set. A similar reasoning applies to the other states. S_2's new state of knowledge and corresponding probabilities after receiving S_1's e-mail are shown in Figure 9.12.

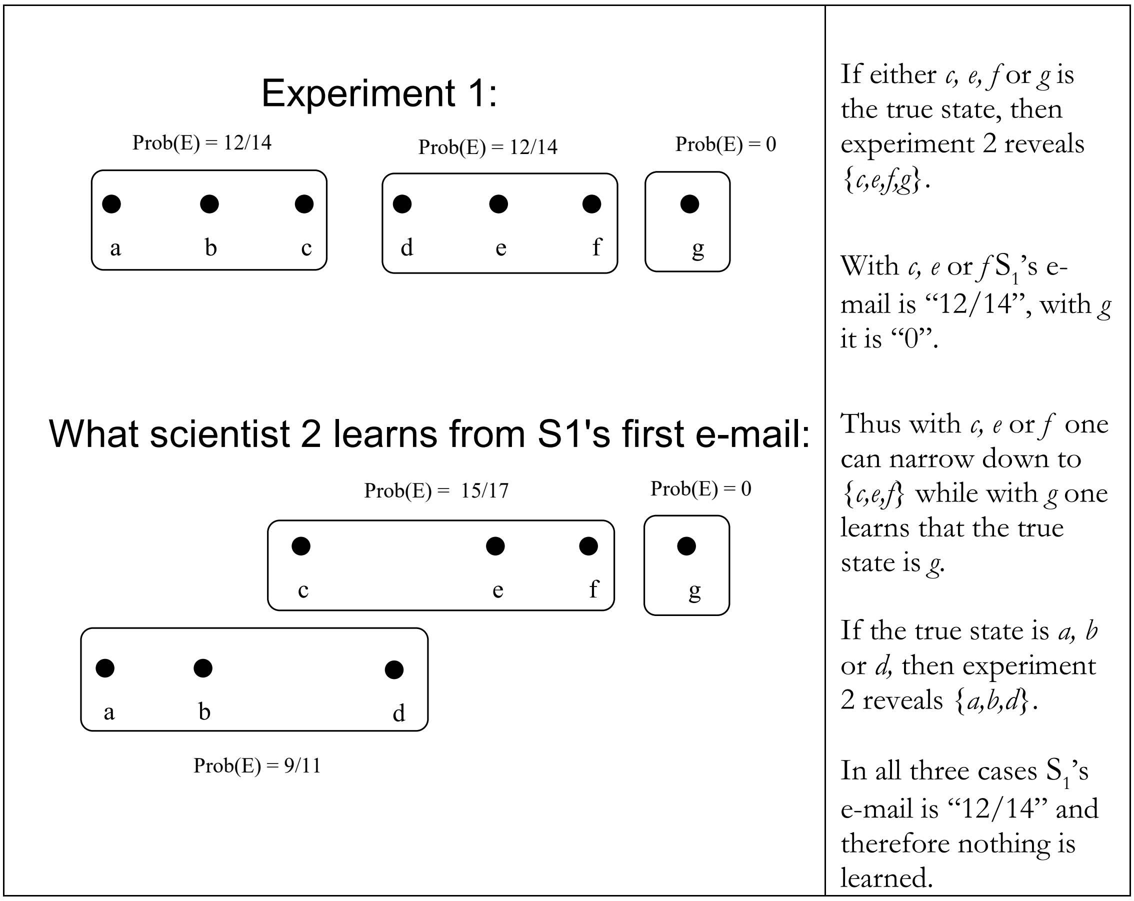


Figure 9.12: S_2's assessment of the probability of event E after receiving S_1's first e-mail.

Thus the new situation after the first exchange of e-mails is as shown in Figure 9.13.

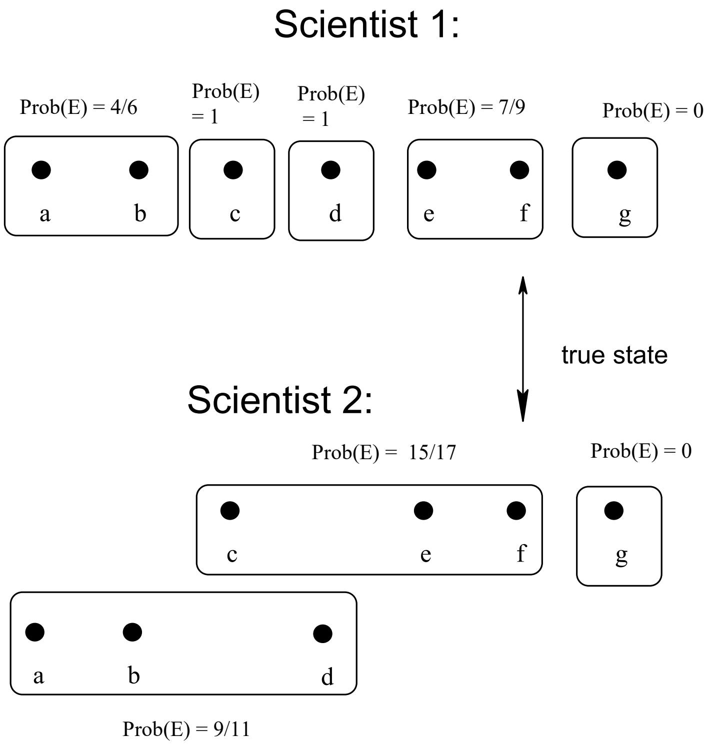


Figure 9.13: The new situation after the first e-mail exchange.

Now there is a second round of e-mails. S_1 communicates "in light of your e-mail, my new $P(E)$ is $\frac{7}{9}$", while S_2 writes "after your e-mail I changed my $P(E)$ to $\frac{15}{17}$". While S_1 learns nothing new from S_2's second e-mail, S_2 learns that the true state cannot be c $\big($the second e-mail would have been "$P(E) = 1$" if c had been the true state; in the hypothetical case where S_2's revised information was $\{a,b,d\}$ then after S_1's second e-mail it would split into $\{a,b\}$ and $\{d\}\big)$. Thus the new situation is a shown in Figure 9.14.

Now the two scientists have reached complete agreement: $P(E) = \frac{7}{9}$. Further exchanges do not convey any new information. Indeed it has become common knowledge (at state f) that both scientists estimate the probability of event E to be $\frac{7}{9} = 78\%$ (before the experiments the probability of E was judged to be $\frac{24}{32} = 75\%$; note, again, that with the experiments and the exchange of information they have gone farther from the truth than at the beginning!).

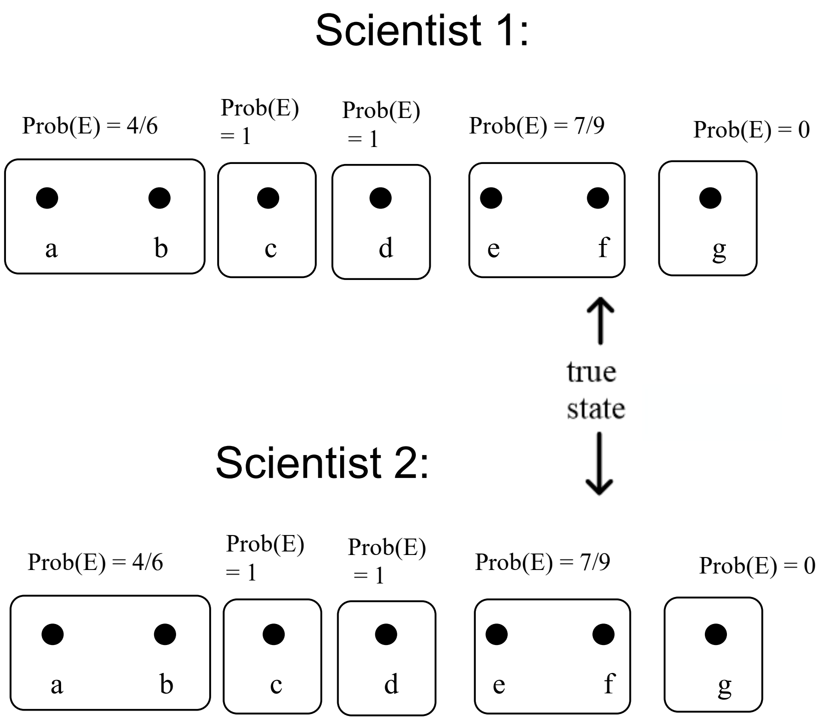


Figure 9.14: The new situation after the second e-mail exchange.

Notice that before the last step it was never common knowledge between the two what probability each scientist attached to E. When one scientist announced his subjective estimate, the other scientist found that announcement informative and revised her own estimate accordingly. After the exchange of two e-mails, the further announcement by one scientist of his/her estimate of the probability of E would not make the other scientist change his/her own estimate: the announcement would reveal nothing new.

Test your understanding of the concepts introduced in this section, by going through the exercises in Section 9.8.6 at the end of this chapter.

9.7 Proof of the Agreement Theorem

First we prove the following.

Lemma 9.1 Let U be a finite set of states, P a probability distribution on U and $E,F \subseteq U$ two events. Let $\{F_1,\ldots,F_m\}$ be a partition of F (thus $F = F_1 \cup \cdots \cup F_m$ and any two F_j and F_k with $j \neq k$ are non-empty and disjoint). Suppose that $P(E|F_j) = q$ for all $j = 1,\ldots,m$. Then $P(E|F) = q$.

Proof. By definition of conditional probability, $P(E|F_j) = \frac{P(E\cap F_j)}{P(F_j)}$. Hence, since $P(E|F_j) = q$, we have that $P(E \cap F_j) = q\,P(F_j)$. Adding over j, the left-hand side becomes $P(E \cap F)$ [because $E \cap F = (E \cap F_1) \cap \cdots \cap (E \cap F_m)$ and for any j and k with $j \neq k$, $(E \cap F_j) \cap (E \cap F_k) = \emptyset$, so that $P(E \cap F) = P(E \cap F_1) + \cdots + P(E \cap F_m)$] and the right-hand side becomes $qP(F)$. Hence $P(E|F) = \frac{P(E\cap F)}{P(F)} = \frac{qP(F)}{P(F)} = q$. ■

Proof of Theorem 9.6.1. Suppose that $CK\big(\|P_1(E) = p\| \cap \|P_2(E) = q\|\big) \neq \emptyset$. Let P be a common prior. Select an arbitrary $s \in CK\big(\|P_1(E) = p\| \cap \|P_2(E) = q\|\big)$ and let $I_{CK}(s)$ be the cell of the common knowledge partition containing s. Consider Individual 1. $I_{CK}(s)$ is equal to the union of a collection of cells (information sets) of 1's information partition. On each such cell 1's conditional probability of E, using the common prior P, is p. Hence, by Lemma 9.1, $P(E|I_{CK}(s)) = p$. A similar reasoning for Individual 2 leads to $P(E|I_{CK}(s)) = q$. Hence $p = q$. ■

9.8 Exercises

9.8.1 Exercises for Section 9.1: Sets and probability

The answers to the following exercises are in Section 9.9 at the end of this chapter.

Exercise 9.1

Let U be the universal set (or sample space) and E and F two events. Let the complement of E be denoted by $\neg E$ and the complement of F by $\neg F$.

Suppose that $P(E) = \frac{3}{10}$, $P(F) = \frac{3}{5}$ and $P(\neg E \cup \neg F) = \frac{4}{5}$. What is the probability of $E \cup F$?

Exercise 9.2

Consider the following probability distribution: $\begin{pmatrix} z_1 & z_2 & z_3 & z_4 & z_5 & z_6 & z_7 \\ \frac{3}{12} & \frac{1}{12} & 0 & \frac{3}{12} & \frac{2}{12} & 0 & \frac{3}{12} \end{pmatrix}$.

What is the probability of the event $\{z_2, z_3, z_5, z_6, z_7\}$?

Exercise 9.3

Let the universal set be $U = \{z_1, z_2, z_3, z_4, z_5, z_6, z_7, z_8\}$. Let $A = \{z_2, z_4, z_5, z_7\}$, $B = \{z_3, z_6, z_8\}$, $C = \{z_2, z_6\}$, $D = \{z_3, z_4\}$ and $E = \{z_7, z_8\}$.

You are given the following data: $P(A \cup B) = \frac{21}{24}$, $P(A \cap C) = \frac{5}{24}$, $P(B \cap C) = \frac{3}{24}$, $P(A \cap D) = \frac{2}{24}$, $P(B \cap D) = \frac{3}{24}$, $P(B) = \frac{7}{24}$ and $P(E) = \frac{2}{24}$.

(a) Find the probability $P(z_i)$ for each $i = 1, \ldots, 8$.

(b) Calculate $P\big((A \cup B) \cap (C \cup D)\big)$.

Exercise 9.4

Let $U = \{a, b, c, d, e, f, g, h, i\}$ and consider the following probability distribution:

$$\begin{pmatrix} a & b & c & d & e & f & g & h & i \\ \frac{11}{60} & 0 & \frac{7}{60} & \frac{9}{60} & \frac{16}{60} & \frac{5}{60} & \frac{4}{60} & \frac{8}{60} & 0 \end{pmatrix}$$

(a) Let $E = \{a, f, g, h, i\}$. What is the probability of E?

(b) List all the events that have probability 1.

Exercise 9.5

Let P be a probability measure on a finite set U and let A and B be two events (that is, subsets of U). Explain why $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. ■

Exercise 9.6

You plan to toss a fair coin three times and record the sequence of Heads/Tails.

(a) What is the set of possibilities (or universal set or sample space)?

(b) Let E be the event that you will get at least one Heads. What is E?

(c) What is the probability of event E?

(d) Let F be the event that you will get Tails either in the first toss or in the third toss? [Note: this is *not* an exclusive 'or'.] What is event F?

(e) What is the probability of event F?

■

9.8.2 Exercises for Section 9.2: Probabilistic beliefs

The answers to the following exercises are in Section 9.9 at the end of this chapter.

Exercise 9.7

Let the set of states be $W = \{a,b,c,d,e,f,g,h,i\}$.
Amy's initial beliefs are given by the following probability distribution:

state	a	b	c	d	e	f	g	h	i
probability	$\frac{11}{60}$	0	$\frac{7}{60}$	$\frac{9}{60}$	$\frac{16}{60}$	$\frac{5}{60}$	$\frac{4}{60}$	$\frac{8}{60}$	0

(a) Let $E = \{a, f, g, h, i\}$. What is the probability of E?

(b) Find all the events that Amy is certain of.

■

9.8.3 Exercises for Section 9.3: Conditional probability and Bayes' rule

The answers to the following exercises are in Section 9.9 at the end of this chapter.

Exercise 9.8

Let A and B be two events such that $P(A) > 0$ and $P(B) > 0$.
Prove that $P(A|B) = P(B|A)$ if and only if $P(A) = P(B)$. ■

Exercise 9.9

Two events A and B are *independent* if $P(A|B) = P(A)$.
Construct an example to show that $P(A|B) = P(B|A)$ and $P(A) = P(B)$ but A and B are **not** independent.

Exercise 9.10

There is an urn with 40 balls: 4 red, 16 white, 10 blue and 10 black. You close your eyes and pick a ball at random. Let E be the event "the selected ball is either red or white".

(a) What is the probability of E?

(b) Now somebody tells you: "the ball in your hand is not black". How likely is it now that you picked either a red or a white ball?

Exercise 9.11

Suppose there are 3 individuals. It is known that one of them has a virus. A blood test can be performed to test for the virus. If an individual does have the virus, then the result of the test will be positive.

However, the test will be positive also for an individual who does not have the virus but has a particular defective gene. It is known that exactly one of the three individuals has this defective gene: it could be the same person who has the virus or somebody who does not have the virus.

A test result will come up positive if and only if either the patient has the virus or the defective gene (or both).

Suppose that Individual 1 takes the blood test and the result is positive. Assuming that all the states are equally likely, what is the probability that he has the virus? [Hint: think of the universal set (or sample space) U as a list of states and each state tells you which individual has the virus and which individual has the defective gene.]

Exercise 9.12

Let A and B be two events such that $P(A) = 0.2, P(B) = 0.5$ and $P(B|A) = 0.1$.
Calculate $P(A|B)$.

Exercise 9.13

In a remote rural clinic with limited resources, a patient arrives complaining of low-abdomen pain. Based on all the information available, the doctor thinks that there are only four possible causes: a bacterial infection (b), a viral infection (v), cancer (c), internal bleeding (i).

Of the four, only the bacterial infection and internal bleeding are treatable at the clinic. In the past the doctor has seen 600 similar cases and they eventually turned out to be as follows:

b : bacterial infection	v : viral infection	c : cancer	i : internal bleeding
140	110	90	260

The doctor's probabilistic estimates are based on those past cases.

(a) What is the probability that the patient has a treatable disease?

There are two possible ways of gathering more information: a blood test and an ultrasound. A positive blood test will reveal that there is an infection, however it could be either bacterial or viral; a negative blood test rules out an infection and thus leaves cancer and internal bleeding as the only possibilities. The ultrasound, on the other hand, will reveal if there is internal bleeding.

(b) Suppose that the patient gets an ultrasound and it turns out that there is no internal bleeding. What is the probability that he does **not** have a treatable disease? What is the probability that he has cancer?

(c) If instead of getting the ultrasound he had taken the blood test and it had been positive, what would the probability that he had a treatable disease have been?

(d) Now let us go back to the hypothesis that the patient only gets the ultrasound and it turns out that there is no internal bleeding. He then asks the doctor: "if I were to take the blood test too (that is, in addition to the ultrasound), how likely is it that it would be positive?". What should the doctor's answer be?

(e) Finally, suppose that the patient gets both the ultrasound and the blood test and the ultrasound reveals that there is no internal bleeding, while the blood test is positive. How likely is it that he has a treatable disease?

■

Exercise 9.14

A lab technician was asked to mark some specimens with two letters, the first from the set $\{A,B,C\}$ and the second from the set $\{E,F,G\}$. For example, a specimen could be labeled as AE or BG, etc. He had a total of 220 specimens. He has to file a report to his boss by filling in the table shown in Figure 9.15. Unfortunately, he does not remember all the figures. He had written some notes to himself, which are reproduced below. Fill in the table with the help of his notes and conditional probabilities.
Here are the technician's notes:

LABEL	number
AE	
AF	
AG	
BE	
BF	
BG	
CE	
CF	
CG	

Figure 9.15: The specimen example of Exercise 9.14.

(a) Of all the ones that he marked with an E, $\frac{1}{5}$ were also marked with an A and $\frac{1}{5}$ were marked with a B.

(b) He marked 36 specimens with the label CE.

(c) Of all the specimens that he marked with a C, the fraction $\frac{12}{23}$ were marked with a G.

(d) Of all the specimens, the fraction $\frac{23}{55}$ were marked with a C.

(e) The number of specimens marked BG was twice the number of specimens marked BE.

(f) Of all the specimens marked with an A, the fraction $\frac{3}{20}$ were marked with an E.

(g) Of all the specimens marked with an A, $\frac{1}{10}$ were marked with a G.

■

9.8.4 Exercises for Section 9.4: Changing beliefs in response to new information

The answers to the following exercises are in Section 9.9 at the end of this chapter.

Exercises for Section 9.4.1: Belief updating

Exercise 9.15

Let the set of states be $U = \{a,b,c,d,e,f,g\}$. Fran's initial beliefs are as follows:

$$\begin{array}{cccccccc} \textit{state} & a & b & c & d & e & f & g \\ \textit{probability} & \frac{3}{20} & 0 & \frac{7}{20} & \frac{1}{20} & 0 & \frac{4}{20} & \frac{5}{20} \end{array}$$

Consider the event $E = \{a,d,e,g\}$.

(a) What probability does Fran attach to event E?

(b) Suppose that Fran is now informed that E is indeed true (that is, that the true state belongs to E). What are her updated beliefs?

(c) Consider the event $D = \{a,b,c,f,g\}$.
What probability does Fran assign to event D (1) initially and (2) after she is informed that E?

■

Exercise 9.16

Consider again the example where there are only three students in a class: Ann, Bob and Carla and the professor tells them that in the last exam one of them got 95 points (out of 100), another 78 and the third 54. Ann's initial beliefs are as follows (where the triple (a,b,c) is interpreted as follows: a is Ann's score, b is Bob's score and c is Carla's score):

$$\begin{array}{cccccc} (95,78,54) & (95,54,78) & (78,95,54) & (54,95,78) & (78,54,95) & (54,78,95) \\ \frac{16}{32} & \frac{8}{32} & \frac{4}{32} & \frac{2}{32} & \frac{1}{32} & \frac{1}{32} \end{array}$$

(a) Suppose that (before distributing the exams) the professor tells the students that Carla received a lower score than Bob. Let E be the event that represents this information. What is E?

(b) How should Ann update her beliefs in response to information E?

■

Exercise 9.17

Let the set of states be $U = \{a,b,c,d,e,f,g\}$.

Bill's initial beliefs are as follows: $\begin{array}{ccccccc} a & b & c & d & e & f & g \\ \frac{3}{20} & \frac{2}{20} & \frac{5}{20} & \frac{1}{20} & \frac{1}{20} & \frac{3}{20} & \frac{5}{20} \end{array}$

(a) Suppose that Bill receives information $E = \{a,c,e,f,g\}$. What are his updated beliefs?

(b) Suppose that, after receiving information E, he later learns a new piece of information, namely $F = \{b,d,e,f,g\}$.
What are his final beliefs (that is, after updating first on E and then on F)?

■

Exercise 9.18

Inspector Gethem has been put in charge of a museum robbery that took place yesterday. Two precious items were stolen: a statuette and a gold tiara, which were displayed in the same room. Surveillance cameras show that only three people visited the room at the time the items disappeared: call them suspect A, suspect B and suspect C.
Let a state be a complete specification of who stole what (including the possibility that the same person stole both items).

(a) List all the states.

(b) Inspector Gethem recognizes the suspects and, based on what he knows about them, initially believes that the probability that suspect A stole both items is $\frac{1}{20}$, the probability that suspect B stole both items is $\frac{3}{20}$ and the probability that suspect C stole both items is $\frac{4}{20}$. Furthermore, he assigns equal probability to every other state.
What are his initial beliefs?

(c) Suppose now that the inspector receives reliable information that suspect B did not steal the statuette and suspect C did not steal the tiara.
What are his beliefs after he updates on this information?

■

Exercise 9.19

Let the set of states be $U = \{a,b,c,d,e,f,g\}$ and let $E = \{a,d,e,g\}$.
The individual's initial beliefs are given by the following probability distribution, call it P:

$$\begin{array}{ccccccc} a & b & c & d & e & f & g \\ \frac{3}{20} & 0 & \frac{7}{20} & \frac{1}{20} & 0 & \frac{4}{20} & \frac{5}{20} \end{array}$$

(a) Calculate $P(E), P(b|E)$ and $P(d|E)$.

(b) Calculate the updated beliefs in response to information E.

■

Exercise 9.20

The instructor of a class has the following data on enrollment:

major	Economics	Mathematics	Philosophy	Psychology	Statistics
enrollment	35%	22%	18%	16%	9%

(a) A student in her class, Jim, tells her that his major is neither Math nor Statistics.
What are the instructor's beliefs about Jim's major upon learning this?

(b) After a while Jim further informs the instructor that he is not an Economics major.
What are the instructor's beliefs about Jim's major upon learning this second fact?

(c) Finally, Jim tells the instructor that he is not a Philosophy major.
What are the instructor's beliefs about Jim's major upon learning this third fact?

■

Exercises for Section 9.4.2: Belief revision

The answers to the following exercises are in Section 9.9 at the end of this chapter.

Exercise 9.21

Prove that an AGM belief revision function (Definition 9.4.4) satisfies Arrow's Axiom: if $E,F \in \mathcal{E}$, $E \subseteq F$ and $E \cap f(F) \neq \emptyset$ then $f(E) = E \cap f(F)$. ■

Exercise 9.22

Prove that the qualitative belief updating rule (9.9) (page 307) is implied by Arrow's Axiom. ■

Exercise 9.23

Let $U = \{a,b,c,d,e,g,h,k,m\}$ and let $\succsim$ be the following plausibility order on U (as usual, we use the convention that if the row to which state s belongs is above the row to which state s' belongs then $s \succ s'$, and if s and s' belong to the same row then $s \sim s'$).

most plausible	b,g
	c,k,m
	d,h
	e
least plausible	a

Let $\mathscr{E} = \{\{a,e\},\{d,e,k,m\},\{b,d,e,k\},U\}$.
Find the belief revision function $f : \mathscr{E} \to 2^U$ that is rationalized by $\succsim$. ■

Exercise 9.24

As in Exercise 9.23, let $U = \{a,b,c,d,e,g,h,k,m\}$ and

$$\mathscr{E} = \left\{ \underbrace{\{a,e\}}_{E}, \underbrace{\{d,e,k,m\}}_{F}, \underbrace{\{b,d,e,k\}}_{G}, U \right\}$$

Using the plausibility order of Exercise 9.23, namely

most plausible	b,g
	c,k,m
	d,h
	e
least plausible	a

find a collection of probability distributions $\{P_E, P_F, P_G, P_W\}$ that provides an AGM probabilistic belief revision policy (Definition 9.4.5). [There are many; find one.] ■

9.8.5 Exercises for Section 9.5: Harsanyi consistency of beliefs or like-mindedness

The answers to the following exercises are in Section 9.9 at the end of this chapter.

Exercise 9.25

Consider the knowledge-belief structure shown in Figure 9.16. Are the individuals' beliefs Harsanyi consistent? ■

Exercise 9.26

Consider the knowledge-belief structure shown in Figure 9.17.

(a) Show that if $p = \frac{1}{13}$ then the beliefs are Harsanyi consistent.

(b) Show that if $p \neq \frac{1}{13}$ then the beliefs are **not** Harsanyi consistent.

■

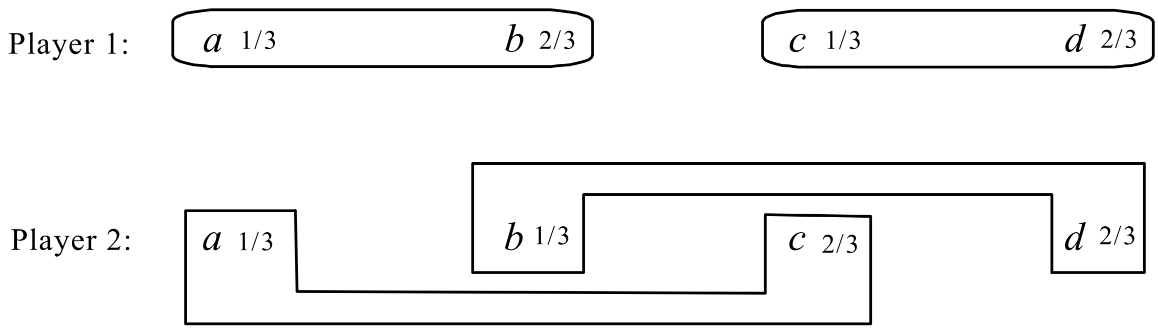


Figure 9.16: The knowledge-belief structure for Exercise 9.25.

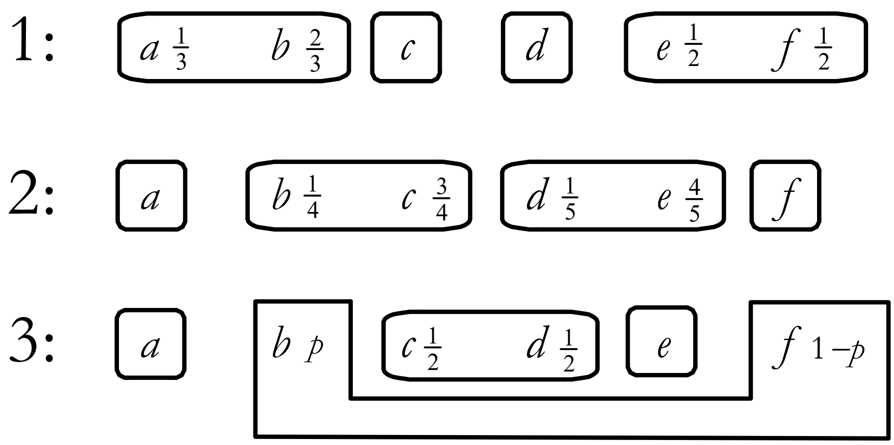


Figure 9.17: The knowledge-belief structure for Exercise 9.26.

9.8.6 Exercises for Section 9.6: Agreeing to disagree

The answers to the following exercises are in Section 9.9 at the end of this chapter.

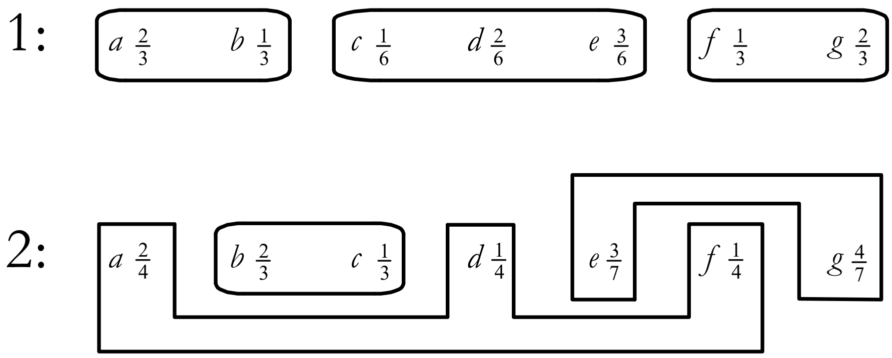


Figure 9.18: The knowledge-belief structure for Exercise 9.27.

Exercise 9.27 Consider the knowledge-belief structure shown in Figure 9.18.

(a) Find the common knowledge partition.

(b) Find a common prior.

(c) Let $E = \{b, d, f\}$.

(1) Is the probability that Individual 1 assigns to E common knowledge?

(2) Is the probability that Individual 2 assigns to E common knowledge?

(d) Let $E = \{b, d, f\}$.

(1) At state b does Individual 1 know what probability Individual 2 assigns to E?

(2) At state c does Individual 1 know what probability Individual 2 assigns to E?

(3) At state f does Individual 1 know what probability Individual 2 assigns to E?

■

Exercise 9.28 — ★★★ Challenging Question ★★★.

This is known as the *Monty Hall problem*.

You are a contestant in a show. You are shown three doors, numbered 1, 2 and 3. Behind one of them is a new car, which will be yours if you choose to open that door.

The door behind which the car was placed was chosen randomly with equal probability (a die was thrown, if it came up 1 or 2 then the car was placed behind Door 1, if it came up 3 or 4 then the car was placed behind Door 2 and if it came up 5 or 6 then the car was placed behind Door 3).

You have to choose one door. Suppose that you have chosen door number 1. Before the door is opened the host tells you that he knows where the car is and, to help you, he will open one of the other two doors, making sure that he opens a door behind which there is no car; if there are two such doors, then he will choose randomly with equal probability.

Afterwards he will give you a chance to change your mind and switch to the other closed door, but you will have to pay \$20 if you decide to switch.
Suppose that initially you chose Door 1 and the host opens Door 3 to show you that the car is not there. Assume that, if switching increases the probability of getting the car (relative to not switching), then you find it worthwhile to pay \$20 to switch.
Should you switch from Door 1 to Door 2?
Answer the question using two different approaches.

(a) **Method 1.** Let D_n denote the event that the car is behind door n and let O_n denote the event that the host opens door n ($n \in \{1,2,3\}$).
The prior probabilities are $P(D_1) = P(D_2) = P(D_3) = \frac{1}{3}$. Compute $P(D_1|O_3)$ using Bayes' rule (if $P(D_1|O_3) \geq \frac{1}{2}$ then you should not switch, since there is a cost in switching).

(b) **Method 2.** Draw (part of) an extensive form with imperfect information where Nature moves first and chooses where the car is, then you choose one door and then the host chooses which door to open (of course, the host's choice is made according to the rules specified above).
Reasoning about the information set you are in after you have pointed to door 1 and the host has opened door 3, determine if you should switch from door 1 to door 2.

■

9.9 Solutions to Exercises

Solution to Exercise 9.1. The general formula is $P(E \cup F) = P(E) + P(F) - P(E \cap F)$.

By The Morgan's Law, $\neg E \cup \neg F = \neg(E \cap F)$.

Thus, since $P(\neg(E \cap F)) = \frac{4}{5}$, we have that $P(E \cap F) = 1 - \frac{4}{5} = \frac{1}{5}$.

Hence, $P(E \cup F) = \frac{3}{10} + \frac{3}{5} - \frac{1}{5} = \frac{7}{10}$. □

Solution to Exercise 9.2.

$$P(\{z_2, z_3, z_5, z_6, z_7\}) = \sum_{i \in \{2,3,5,6,7\}} P(\{z_i\}) = \frac{1}{12} + 0 + \frac{2}{12} + 0 + \frac{3}{12} = \frac{1}{2}. \quad \square$$

Solution to Exercise 9.3.

(a) Since $\{z_1\}$ is the complement of $A \cup B$, $P(z_1) = 1 - \frac{21}{24} = \frac{3}{24}$.

Since $\{z_2\} = A \cap C$, $P(z_2) = \frac{5}{24}$.

Similarly, $P(z_6) = P(B \cap C) = \frac{3}{24}$, $P(z_3) = P(B \cap D) = \frac{3}{24}$ and

$P(z_4) = P(A \cap D) = \frac{2}{24}$.

Thus, $P(z_8) = P(B) - P(z_3) - P(z_6) = \frac{7}{24} - \frac{3}{24} - \frac{3}{24} = \frac{1}{24}$.

Hence, $P(z_7) = P(E) - P(z_8) = \frac{2}{24} - \frac{1}{24} = \frac{1}{24}$. Finally, $P(z_5) = 1 - \sum_{i \neq 5} P(z_i) = \frac{6}{24}$.

Thus, the probability distribution is:

$$\begin{pmatrix} z_1 & z_2 & z_3 & z_4 & z_5 & z_6 & z_7 & z_8 \\ \frac{3}{24} & \frac{5}{24} & \frac{3}{24} & \frac{2}{24} & \frac{6}{24} & \frac{3}{24} & \frac{1}{24} & \frac{1}{24} \end{pmatrix}$$

(b) $A \cup B = \{z_2, z_3, z_4, z_5, z_6, z_7, z_8\}$, $C \cup D = \{z_2, z_3, z_4, z_6\}$.

Hence, $(A \cup B) \cap (C \cup D) = C \cup D = \{z_2, z_3, z_4, z_6\}$ so that

$P\big((A \cup B) \cap (C \cup D)\big) = P(z_2) + P(z_3) + P(z_4) + P(z_6) = \frac{5}{24} + \frac{3}{24} + \frac{2}{24} + \frac{3}{24} = \frac{13}{24}$. □

Solution to Exercise 9.4. The given probability distribution is:

$$\begin{pmatrix} a & b & c & d & e & f & g & h & i \\ \frac{11}{60} & 0 & \frac{7}{60} & \frac{9}{60} & \frac{16}{60} & \frac{5}{60} & \frac{4}{60} & \frac{8}{60} & 0 \end{pmatrix}.$$

(a) Let $E = \{a, f, g, h, i\}$.

Then $P(E) = P(a) + P(f) + P(g) + P(h) + P(i) = \frac{11}{60} + \frac{5}{60} + \frac{4}{60} + \frac{8}{60} + 0 = \frac{28}{60} = \frac{7}{15}$.

(b) The events that have probability 1 are:

$\{a, c, d, e, f, g, h\} = U \setminus \{b, i\}$, $\{a, b, c, d, e, f, g, h\} = U \setminus \{i\}$,

$\{a, c, d, e, f, g, h, i\} = U \setminus \{b\}$ and $\{a, b, c, d, e, f, g, h, i\} = U$. □

Solution to Exercise 9.5. Since $P(A) = \sum_{w \in A} P(w)$ and $P(B) = \sum_{w \in B} P(w)$, when adding $P(A)$ to $P(B)$ the elements that belong to both A and B (that is, the elements of $A \cap B$) are added twice and thus we need to subtract $\sum_{w \in A \cap B} P(w)$ from $P(A) + P(B)$ in order to get $\sum_{w \in A \cup B} P(w) = P(A \cup B)$. □

Solution to Exercise 9.6.

(a) There are 8 possibilities:

$$HHH \quad HHT \quad HTH \quad HTT \quad THH \quad THT \quad TTH \quad TTT$$

Since the coin is fair, each possibility has the same probability, namely $\frac{1}{8}$.

(b) $E = U \setminus \{TTT\}$, where U is the universal set (the set of 8 possibilities listed above).

(c) $P(E) = P(U) - P(TTT) = 1 - \frac{1}{8} = \frac{7}{8}$.

(d) $F = U \setminus \{HHH, HTH\}$

(e) $P(F) = P(U) - P(\{HHH, HTH\}) = 1 - \frac{1}{8} - \frac{1}{8} = \frac{6}{8} = \frac{3}{4}$. □

Solution to Exercise 9.7. The probability distribution is

state	a	b	c	d	e	f	g	h	i
probability	$\frac{11}{60}$	0	$\frac{7}{60}$	$\frac{9}{60}$	$\frac{16}{60}$	$\frac{5}{60}$	$\frac{4}{60}$	$\frac{8}{60}$	0

(a) Let $E = \{a, f, g, h, i\}$.
Then $P(E) = P(a) + P(f) + P(g) + P(h) + P(i) = \frac{11}{60} + \frac{5}{60} + \frac{4}{60} + \frac{8}{60} + 0 = \frac{28}{60} = \frac{7}{15}$.

(b) Amy is certain of all events that have probability 1, namely
$\{a,c,d,e,f,g,h\}$, $\{a,b,c,d,e,f,g,h\}$, $\{a,c,d,e,f,g,h,i\}$ and $\{a,b,c,d,e,f,g,h,i\}$.
□

Solution to Exercise 9.8. Suppose that $P(A|B) = P(B|A)$.
Since $P(A|B) = \frac{P(A \cap B)}{P(B)}$ and $P(B|A) = \frac{P(A \cap B)}{P(A)}$ it follows that $P(A) = P(B)$.
Conversely, if $P(A) = P(B)$ then $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A \cap B)}{P(A)} = P(B|A)$. □

Solution to Exercise 9.9. Example 1. Let $P(A) = \frac{1}{2}$ and let $B = \neg A$.
Then $P(B) = 1 - P(A) = \frac{1}{2}$ and $A \cap B = \emptyset$ so that $P(A \cap B) = 0$
and thus $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0}{\frac{1}{2}} = 0 = P(B|A) = \frac{P(A \cap B)}{P(A)}$.
Thus $P(A|B) = P(B|A) = 0$ and $P(A) = P(B) = \frac{1}{2}$ and A and B are not independent since $P(A|B) \neq P(A)$.

Example 2. $U = \{a,b,c\}, P(a) = P(c) = \frac{2}{5}$ and $P(b) = \frac{1}{5}$. Let $A = \{a,b\}$ and $B = \{b,c\}$.
Then $P(A) = P(B) = \frac{3}{5}$ and $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(b)}{P(b)+P(c)} = \frac{\frac{1}{5}}{\frac{3}{5}} = \frac{1}{3} = P(B|A)$.
Thus A and B are not independent since $P(A|B) \neq P(A)$. □

Solution to Exercise 9.10.

(a) $P(E) = \frac{4+16}{40} = \frac{1}{2}$.

(b) Let F be the event "the selected ball is not black". Then, initially, $P(F) = \frac{30}{40} = \frac{3}{4}$. Furthermore, $E \cap F = E$. Thus, $P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{P(E)}{P(F)} = \frac{\frac{1}{2}}{\frac{3}{4}} = \frac{2}{3}$. □

Solution to Exercise 9.11. First we list the possible states. A state is a complete description of the facts that are relevant: it tells you who has the virus and who has the gene.
Let us represent a state as a pair (x,y) interpreted as follows: individual x has the virus and individual y has the defective gene.
Then $U = \{a = (1,1),\ b = (1,2),\ c = (1,3),\ d = (2,1),\ e = (2,2),\ f = (2,3),\ g = (3,1),\ h = (3,2),\ i = (3,3)\}$.
Let V_1 be the event "Individual 1 has the virus". Then $V_1 = \{a,b,c\}$.
Let G_1 be the event "Individual 1 has the defective gene". Then $G_1 = \{a,d,g\}$.
Since every state is assumed to have probability $\frac{1}{9}$, $P(V_1) = P(G_1) = \frac{1}{9} + \frac{1}{9} + \frac{1}{9} = \frac{1}{3}$.
Let 1_+ be the event that a blood test administered to Individual 1 comes up positive.
Then $1_+ = \{a,b,c,d,g\}$ and $P(1_+) = \frac{5}{9}$.
Now we can compute the requested conditional probability as follows (note that $V_1 \cap 1_+ = V_1$):

$$P(V_1|1_+) = \frac{P(V_1 \cap 1_+)}{P(1_+)} = \frac{P(V_1)}{P(1_+)} = \frac{\frac{1}{3}}{\frac{5}{9}} = \frac{3}{5} = 60\%.$$

□

Solution to Exercise 9.12.
Using Bayes' rule, $P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{(0.1)(0.2)}{0.5} = 0.04 = 4\%$. □

Solution to Exercise 9.13. The probabilities are as follows:

b	v	c	i
$\frac{140}{600} = \frac{14}{60}$	$\frac{110}{600} = \frac{11}{60}$	$\frac{90}{600} = \frac{9}{60}$	$\frac{260}{600} = \frac{26}{60}$

(a) The event that the patient has a treatable disease is $\{b,i\}$.
$P(\{b,i\}) = P(b) + P(i) = \frac{14}{60} + \frac{26}{60} = \frac{2}{3}$.

(b) A negative result of the ultrasound is represented by the event $\{b,v,c\}$.
A non-treatable disease is the event $\{v,c\}$. Thus,

$$P(\{v,c\}|\{b,v,c\}) = \frac{P(\{v,c\} \cap \{b,v,c\})}{P(\{b,v,c\})} = \frac{P(\{v,c\})}{P(\{b,v,c\})} = \frac{\frac{11}{60} + \frac{9}{60}}{\frac{14}{60} + \frac{11}{60} + \frac{9}{60}} = \frac{10}{17} = 58.82\%.$$

$$P(c|\{b,v,c\}) = \frac{P(c)}{P(\{b,v,c\})} = \frac{\frac{9}{60}}{\frac{14}{60}+\frac{11}{60}+\frac{9}{60}} = \frac{9}{34} = 26.47\%.$$

(c) A positive blood test is represented by the event $\{b,v\}$.
A treatable disease is the event $\{b,i\}$. Thus,

$$P(\{b,i\}|\{b,v\}) = \frac{P(\{b,i\}\cap\{b,v\})}{P(\{b,v\})} = \frac{P(b)}{P(\{b,v\})} = \frac{\frac{14}{60}}{\frac{14}{60}+\frac{11}{60}} = \frac{14}{25} = 56\%.$$

(d) Here we want

$$P(\{b,v\}|\{b,v,c\}) = \frac{P(\{b,v\})}{P(\{b,v,c\})} = \frac{\frac{14}{60}+\frac{11}{60}}{\frac{14}{60}+\frac{11}{60}+\frac{9}{60}} = \frac{25}{34} = 73.53\%.$$

(e) We are conditioning on $\{b,v\}\cap\{b,v,c\} = \{b,v\}$;
thus, we want $P(\{b,i\}|\{b,v\})$ which was calculated in Part **(c)** as $\frac{14}{25} = 56\%$. □

Solution to Exercise 9.14.

Let $\#xy$ be the *number* of specimens that were marked xy (thus, $x \in \{A,B,C\}$ and $y \in \{D,E,F\}$) and $P(xy) = \frac{\#xy}{220}$ be the *fraction* of specimens that were marked xy; let $\#z$ be the number of specimens whose label contains a $z \in \{A,B,C,D,E,F\}$ and let $P(z) = \frac{\#z}{220}$; finally, let $P(xy|z) = \frac{\#xy}{\#z}$.[11]

With this notation we can re-write the information contained in the technician's notes as follows.

(a) $P(AE|E) = P(BE|E) = \frac{1}{5}$. It follows that the remaining three fifth were marked with a C, that is $P(CE|E) = \frac{3}{5}$

(b) $\#CE = 36$; thus, $P(CE) = \frac{36}{220}$. Since $P(CE|E) = \frac{P(CE)}{P(E)}$, using **(a)** we get $\frac{3}{5} = \frac{\frac{36}{220}}{P(E)}$, that is, $P(E) = \frac{36}{220}\left(\frac{5}{3}\right) = \frac{3}{11}$. Hence, the number of specimens marked with an E is $\frac{3}{11}\,220 = 60$. Furthermore, since $P(AE|E) = \frac{P(AE)}{P(E)}$, using **(a)** we get $\frac{1}{5} = \frac{P(AE)}{\frac{3}{11}}$, that is, $P(AE) = \frac{3}{55}$. Thus, the number of specimens marked AE is $\frac{3}{55}\,220 = 12$. The calculation for $P(BE|E)$ is identical; thus, the number of specimens marked BE is also 12. So far, we have:

[11] This is a conditional probability, since $P(xy|z) = \frac{\#xy}{\#z} = \frac{\frac{\#xy}{220}}{\frac{\#z}{220}} = \frac{P(xy)}{P(z)}$).

LABEL	number
AE	12
AF	
AG	
BE	12
BF	
BG	
CE	36
CF	
CG	

(c) $P(CG|C) = \frac{12}{23}$. Since $P(CG|C) = \frac{P(CG)}{P(C)}$, it follows that $\frac{12}{23} = \frac{P(CG)}{P(C)}$.

(d) $P(C) = \frac{23}{55}$. Thus, using **(c)** we get $\frac{12}{23} = \frac{P(CG)}{\frac{23}{55}}$, that is, $P(CG) = \frac{12}{55}$.

Hence, the number of specimens marked CG is $\frac{12}{55}\,220 = 48$.

Since $P(C) = \frac{23}{55}$, the total number of specimens marked with a C is $\frac{23}{55}\,220 = 92$.

Since 36 were marked CE (see the above table) and 48 were marked CG, it follows that the number of specimens marked CF is $92 - 48 - 36 = 8$. Up to this point we have:

LABEL	number
AE	12
AF	
AG	
BE	12
BF	
BG	
CE	36
CF	8
CG	48

(e) The number of BGs is twice the number of BEs.

Since the latter is 12 (see the above table), the number of BGs is 24.

(f) $P(AE|A) = \frac{3}{20}$. Since $P(AE|A) = \frac{P(AE)}{P(A)}$ and, from **(b)**, $P(AE) = \frac{3}{55}$, we have that $\frac{3}{20} = \frac{\frac{3}{55}}{P(A)}$. Hence, $P(A) = \frac{3}{55}\left(\frac{20}{3}\right) = \frac{4}{11}$.

Thus, the number of specimens marked with an A is $\frac{4}{11}220 = 80$.

Since $P(A) = \frac{4}{11}$ and, from **(d)**, $P(C) = \frac{23}{55}$, it follows that $P(B) = 1 - \frac{4}{11} - \frac{23}{55} = \frac{12}{55}$.

Thus, the number of specimens marked with a B is $\frac{12}{55}\,220 = 48$. Of these, 12 were marked BE and 24 were marked BG. Thus, the number of specimens marked BF is $48 - 12 - 24 = 12$. So far, we have:

LABEL	number
AE	12
AF	
AG	
BE	12
BF	12
BG	24
CE	36
CF	8
CG	48

(g) $P(AG|A) = \frac{1}{10}$. Since $P(AG|A) = \frac{P(AG)}{P(A)}$, and from **(f)** we have that $P(A) = \frac{4}{11}$ it follows that $\frac{1}{10} = \frac{P(AG)}{\frac{4}{11}}$, that is, $P(AG) = \frac{1}{10}\left(\frac{4}{11}\right) = \frac{4}{110}$.

Thus, the number of specimens marked AG is $\frac{4}{110}220 = 8$.

Since the number marked with an A is $\frac{4}{11}220 = 80$ and the number of those marked AE is 12 and the number ot those marked AG is 8, we get that the number of specimens marked AF is $80 - 12 - 8 = 60$.

Thus, we have completed the table:

LABEL	number
AE	12
AF	60
AG	8
BE	12
BF	12
BG	24
CE	36
CF	8
CG	48

$\square$

Solution to Exercise 9.15. Fran's initial beliefs are:

state	a	b	c	d	e	f	g
probability	$\frac{3}{20}$	0	$\frac{7}{20}$	$\frac{1}{20}$	0	$\frac{4}{20}$	$\frac{5}{20}$

The events under consideration are: $E = \{a,d,e,g\}$ and $D = \{a,b,c,f,g\}$.

(a) $P(E) = P(a)+P(d)+P(e)+P(g) = \frac{3}{20}+\frac{1}{20}+0+\frac{5}{20} = \frac{9}{20}$.

(b) $P(b|E) = P(c|E) = P(f|E) = 0$ (since each of these states does not belong to E), $P(a|E) = \frac{\frac{3}{20}}{\frac{9}{20}} = \frac{3}{9}$, $P(d|E) = \frac{\frac{1}{20}}{\frac{9}{20}} = \frac{1}{9}$, $P(e|E) = \frac{0}{\frac{9}{20}} = 0$ and $P(g|E) = \frac{\frac{5}{20}}{\frac{9}{20}} = \frac{5}{9}$.
Thus the operation of conditioning on event E yields the following updated probability distribution:

state	a	b	c	d	e	f	g
probability	$\frac{3}{9}$	0	0	$\frac{1}{9}$	0	0	$\frac{5}{9}$

(c) **(1)** Initially Fran attaches the following probability to event D:
$P(D) = P(a)+P(b)+P(c)+P(f)+P(g) = \frac{3}{20}+0+\frac{7}{20}+\frac{4}{20}+\frac{5}{20} = \frac{19}{20} = 95\%$.
(2) After updating on information E Fran attaches the following probability to event D:
$P(D|E) = \frac{P(D\cap E)}{P(E)} = \frac{P(\{a,g\})}{P(E)} = \frac{\frac{8}{20}}{\frac{9}{20}} = \frac{8}{9} = 88.89\%$. □

Solution to Exercise 9.16.

(a) $E = \{(95,78,54),(78,95,54),(54,95,78)\}$. Thus $P(E) = \frac{16}{32}+\frac{4}{32}+\frac{2}{32} = \frac{22}{32}$.

(b) Conditioning on E yields the following beliefs:

$(95,78,54)$	$(95,54,78)$	$(78,95,54)$	$(54,95,78)$	$(78,54,95)$	$(54,78,95)$
$\frac{\frac{16}{32}}{\frac{22}{32}} = \frac{8}{11}$	0	$\frac{\frac{4}{32}}{\frac{22}{32}} = \frac{2}{11}$	$\frac{\frac{2}{32}}{\frac{22}{32}} = \frac{1}{11}$	0	0

□

Solution to Exercise 9.17.

(a) Updating on information $E = \{a,c,e,f,g\}$ yields the following beliefs:

a	b	c	d	e	f	g
$\frac{3}{17}$	0	$\frac{5}{17}$	0	$\frac{1}{17}$	$\frac{3}{17}$	$\frac{5}{17}$

(b) Updating the beliefs of Part **(a)** on information $F = \{b,d,e,f,g\}$ yields the following beliefs:

a	b	c	d	e	f	g
0	0	0	0	$\frac{1}{9}$	$\frac{3}{9}$	$\frac{5}{9}$

□

Solution to Exercise 9.18. Represent a state as a pair (x,y) where x is the suspect who stole the statuette and y is the suspect who stole the tiara.

(a) The set of states is $U = \{(A,A),(A,B),(A,C),(B,A),(B,B),(B,C),(C,A),(C,B),(C,C)\}$.

(b) The inspector's initial beliefs are:

$$\begin{array}{ccccccccc} (A,A) & (A,B) & (A,C) & (B,A) & (B,B) & (B,C) & (C,A) & (C,B) & (C,C) \\ \frac{1}{20} & \frac{2}{20} & \frac{2}{20} & \frac{2}{20} & \frac{3}{20} & \frac{2}{20} & \frac{2}{20} & \frac{2}{20} & \frac{4}{20} \end{array}$$

[Explanation: $P(A,A)+P(B,B)+P(C,C) = \frac{1}{20}+\frac{3}{20}+\frac{4}{20} = \frac{8}{20}$; thus $\frac{12}{20}$ remains to be distributed equally among the remaining six states, so that each receives $\frac{2}{20}$.]

(c) The information is $F = \{(A,A),(A,B),(C,A),(C,B)\}$.
Updating on this information yields the following beliefs:

$$\begin{array}{ccccccccc} (A,A) & (A,B) & (A,C) & (B,A) & (B,B) & (B,C) & (C,A) & (C,B) & (C,C) \\ \frac{1}{7} & \frac{2}{7} & 0 & 0 & 0 & 0 & \frac{2}{7} & \frac{2}{7} & 0 \end{array}$$ □

Solution to Exercise 9.19.

(a) $P(E) = P(a)+P(d)+P(e)+P(g) = \frac{3}{20}+\frac{1}{20}+0+\frac{5}{20} = \frac{9}{20}$, $P(b|E) = 0$
and $P(d|E) = \dfrac{\frac{1}{20}}{\frac{9}{20}} = \frac{1}{9}$.

(b) The updated beliefs are as follows: $\begin{array}{ccccccc} a & b & c & d & e & f & g \\ \frac{3}{9} & 0 & 0 & \frac{1}{9} & 0 & 0 & \frac{5}{9} \end{array}$ □

Solution to Exercise 9.20. The initial beliefs are:

Economics	Mathematics	Philosophy	Psychology	Statistics
$\frac{35}{100}$	$\frac{22}{100}$	$\frac{18}{100}$	$\frac{16}{100}$	$\frac{9}{100}$

(a) Updating on {Economics, Philosophy, Psychology} yields the following beliefs:

Economics	Mathematics	Philosophy	Psychology	Statistics
$\frac{35}{69}$	0	$\frac{18}{69}$	$\frac{16}{69}$	0

(b) Updating the beliefs of Part **(a)** on {Philosophy, Psychology}[12] yields the following beliefs:

Economics	Mathematics	Philosophy	Psychology	Statistics
0	0	$\frac{18}{34}$	$\frac{16}{34}$	0

(c) Updating the beliefs of Part **(b)** on {Psychology} yields the following beliefs:

Economics	Mathematics	Philosophy	Psychology	Statistics
0	0	0	1	0

that is, the instructor now knows that the student is a Psychology major. □

[12] This is the intersection of the initial piece of information, namely {Economics, Philosophy, Psychology}, and the new piece of information, namely {Mathematics, Philosophy, Psychology, Statistics}. Updating the updated beliefs on {Mathematics, Philosophy, Psychology, Statistics} yields the same result as updating on {Philosophy, Psychology}. Indeed, one would obtain the same result by updating the *initial* beliefs on {Philosophy, Psychology}.

Solution to Exercise 9.21.
Let $f : \mathscr{E} \to 2^U$ be an AGM belief revision function.
Let $E, F \in \mathscr{E}$ be such that $E \subseteq F$ and $E \cap f(F) \neq \emptyset$. We need to show that $f(E) = E \cap f(F)$.
By definition of AGM belief revision function (Definition 9.4.2), there is a plausibility order $\succsim$ on U such that

$$f(F) = \left\{s \in F : s \succsim s' \text{ for every } s' \in F\right\} \tag{9.10}$$

and

$$f(E) = \left\{s \in E : s \succsim s' \text{ for every } s' \in E\right\}. \tag{9.11}$$

Choose an arbitrary $s \in E \cap f(F)$.
Then, by (9.10) and the fact that $E \subseteq F$, $s \succsim s'$ for every $s' \in E$ and thus, by (9.11), $s \in f(E)$.
Hence, $E \cap f(F) \subseteq f(E)$.
Conversely, choose an arbitrary $s_1 \in f(E)$.
Then, since (by definition of belief revision function: Definition 9.4.2) $f(E) \subseteq E$, $s_1 \in E$.
We want to show that $s_1 \in f(F)$ [so that $s_1 \in E \cap f(F)$ and, therefore, $f(E) \subseteq E \cap f(F)$].
Suppose it is not true. Then, by (9.10), there exists an $s_2 \in F$ such that $s_2 \succ s_1$.
Select an $s_3 \in E \cap f(F)$ (recall that, by hypothesis, $E \cap f(F) \neq \emptyset$).
Then, by (9.10) (since $s_2, s_3 \in f(F)$), $s_3 \succsim s_2$, from which it follows (by transitivity of $\succsim$ and the fact that $s_2 \succ s_1$) that $s_3 \succ s_1$.
But then, since $s_3 \in E$, it is not true that $s_1 \succsim s'$ for every $s' \in E$, contradicting – by (9.11) – the hypothesis that $s_1 \in f(E)$. □

Solution to Exercise 9.22. For every event E (representing a possible item of information), let P_E be the probability distribution on E that represents the *revised* beliefs of the individual after receiving information E.
Let P be the probability distribution on U representing the individual's *initial* beliefs.
Define the following belief revision function f : $f(U) = Supp(P)$ and $f(E) = Supp(P_E)$.
Suppose that f satisfies Arrow's Axiom.
Then, for every event E, if $E \cap f(U) \neq \emptyset$ [that is, if $E \cap Supp(P) \neq \emptyset$ or $P(E) > 0$]
then $f(E) = E \cap f(U)$ [that is, $Supp(P_E) = E \cap Supp(P)$]. □

Solution to Exercise 9.23. We have that $\mathcal{E} = \{\{a,e\},\{d,e,k,m\},\{b,d,e,k\},U\}$ and $\precsim$ is given by

most plausible	b,g
	c,k,m
	d,h
	e
least plausible	a

Then the belief revision function rationalized by this plausibility order is given by: $f(\{a,e\}) = \{e\}$, $f(\{d,e,k,m\}) = \{k,m\}$, $f(\{b,d,e,k\}) = \{b\}$ and $f(U) = \{b,g\}$. $\square$

Solution to Exercise 9.24.
From Exercise 9.23 we get that $\{P_E,P_F,P_G,P_U\}$ must be such that
$Supp(P_E) = \{e\}$, $Supp(P_F) = \{k,m\}$, $Supp(P_G) = \{b\}$ and $Supp(P_U) = \{b,g\}$.
For every full-support probability distribution P_0 on U, there is a corresponding collection $\{P_E,P_F,P_G,P_U\}$. For example, if P_0 is the uniform distribution on U (that assigns probability $\frac{1}{9}$ to every state)
then the corresponding $\{P_E,P_F,P_G,P_U\}$ is given by:

state	a	b	c	d	e	g	h	k	m
P_E	0	0	0	0	1	0	0	0	0
P_F	0	0	0	0	0	0	0	$\frac{1}{2}$	$\frac{1}{2}$
P_G	0	1	0	0	0	0	0	0	0
P_U	0	$\frac{1}{2}$	0	0	0	$\frac{1}{2}$	0	0	0

As another example, if P_0 is the following probability distribution

state	a	b	c	d	e	g	h	k	m
P_0	$\frac{1}{50}$	$\frac{3}{50}$	$\frac{11}{50}$	$\frac{4}{50}$	$\frac{8}{50}$	$\frac{9}{50}$	$\frac{5}{50}$	$\frac{2}{50}$	$\frac{7}{50}$

then the corresponding $\{P_E,P_F,P_G,P_U\}$ is given by: P_E and P_G the same as above, and P_F and P_U as follows:

state	a	b	c	d	e	g	h	k	m
P_F	0	0	0	0	0	0	0	$\frac{2}{9}$	$\frac{7}{9}$ $\square$
P_U	0	$\frac{1}{4}$	0	0	0	$\frac{3}{4}$	0	0	0

Solution to Exercise 9.25. Yes. The following is a common prior: $\begin{pmatrix} a & b & c & d \\ \frac{1}{9} & \frac{2}{9} & \frac{2}{9} & \frac{4}{9} \end{pmatrix}$. $\square$

Solution to Exercise 9.26.

(a) Assume that $p = \frac{1}{13}$ (so that $1 - p = \frac{12}{13}$).

From the updating conditions for Individual 1 we get $p_b = 2p_a$ and $p_e = p_f$;

from the updating conditions for Individual 2 we get $p_c = 3p_b$ and $p_e = 4p_d$;

from the updating conditions for Individual 3 we get $p_c = p_d$ and $p_f = 12p_b$.

All these equations have a solution, which constitutes a common prior, namely

$$\begin{pmatrix} a & b & c & d & e & f \\ \frac{1}{63} & \frac{2}{63} & \frac{6}{63} & \frac{6}{63} & \frac{24}{63} & \frac{24}{63} \end{pmatrix}.$$

(b) From the equations of part **(a)** we have that $p_f = p_e, p_e = 4p_d, p_d = p_c$ and $p_c = 3p_b$, from which it follows that $p_f = 12p_b$.

Thus, updating on the information set $\{b, f\}$ of Individual 3 we get that we must have $1 - p = \frac{p_f}{p_b + p_f} = \frac{12p_b}{p_b + 12p_b} = \frac{12}{13}$ and thus $p = \frac{1}{13}$. Hence, if $p \neq \frac{1}{13}$ there is no common prior. □

Solution to Exercise 9.27.

(a) The common knowledge partition is the trivial partition shown in Figure 9.19.

(b) The common prior is also shown shown in Figure 9.19.

(c) **(1)** Individual 1 assigns probability $\frac{1}{3}$ to $E = \{b, d, f\}$ at every state. Thus his assessment of the probability of E is common knowledge at every state.

(2) At state a Individual 2 assigns probability $\frac{1}{2}$ to E, while at state b she assigns probability $\frac{2}{3}$ to E. Thus there is no state where her assessment of the probability of E is common knowledge.

(d) **(1)** At state b Individual 1's information set is $\{a, b\}$; since at a Individual 2 assigns probability $\frac{1}{2}$ to E, while at b she assigns probability $\frac{2}{3}$ to E, it follows that at b it is not the case that Individual 1 knows Individual 2's assessment of the probability of E.

(2) At state c Individual 1's information set is $\{c, d, e\}$; since at c Individual 2 assigns probability $\frac{2}{3}$ to E, while at d she assigns probability $\frac{1}{2}$ to E, it follows that at c it is not the case that Individual 1 knows Individual 2's assessment of the probability of E.

(3) At state f Individual 1's information set is $\{f, g\}$; since at f Individual 2 assigns probability $\frac{1}{2}$ to E, while at g she assigns probability 0 to E, it follows that at f it is not the case that Individual 1 knows Individual 2's assessment of the probability of E. □

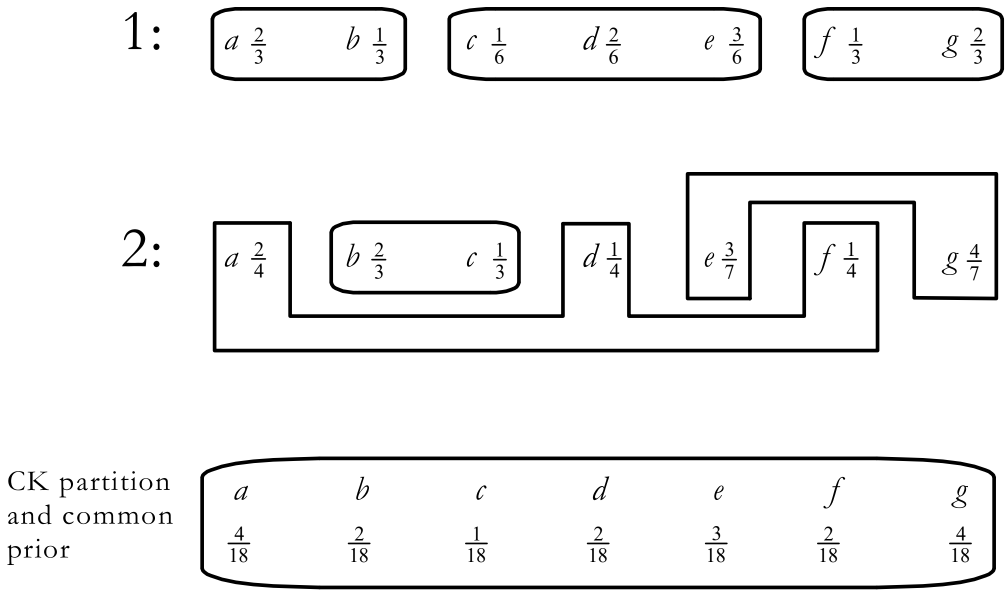


Figure 9.19: The information structure for Exercise 9.27.

Solution to Exercise 9.28.

(a) **Method 1**. We solve this problem using Bayes' formula.

For every $n \in \{1,2,3\}$, let D_n denote the event that the car is behind door n and let O_n denote the event that the host opens Door n.

The initial probabilities are $P(D_1) = P(D_2) = P(D_3) = \frac{1}{3}$.

We want to compute $P(D_1|O_3)$; if $P(D_1|O_3) \geq \frac{1}{2}$ then you should **not** switch, since there is a cost in switching (recall that Door 1 is your initial choice).

By Bayes' rule, $P(D_1|O_3) = \frac{P(O_3|D_1)\,P(D_1)}{P(O_3)}$.

We know that $P(D_1) = \frac{1}{3}$ and $P(O_3|D_1) = \frac{1}{2}$ (when the car is behind Door 1 then the host has a choice between opening Door 2 and opening Door 3 and he chooses with equal probability).

Thus,

$$P(D_1|O_3) = \frac{\frac{1}{2} \times \frac{1}{3}}{P(O_3)} = \frac{\frac{1}{6}}{P(O_3)}. \tag{9.12}$$

We need to compute $P(O_3)$:

$$\begin{aligned}
P(O_3) &= P(O_3|D_1)\,P(D_1) + P(O_3|D_2)\,P(D_2) + P(O_3|D_3)\,P(D_3) \\
&= P(O_3|D_1)\,\tfrac{1}{3} + P(O_3|D_2)\,\tfrac{1}{3} + P(O_3|D_3)\,\tfrac{1}{3} \\
&= \tfrac{1}{2}\left(\tfrac{1}{3}\right) + 1\left(\tfrac{1}{3}\right) + 0\left(\tfrac{1}{3}\right) = \tfrac{1}{6} + \tfrac{1}{3} = \tfrac{1}{2}
\end{aligned}$$

because $P(O_3|D_1) = \frac{1}{2}$, $P(O_3|D_2) = 1$ (if the car is behind Door 2 then the host has to open Door 3, since he cannot open the door that you chose, namely Door 1) and $P(O_3|D_3) = 0$ (if the car is behind Door 3 then the host cannot open that door).
Substituting $\frac{1}{2}$ for $P(O_3)$ in (9.12) we get that $P(D_1|O_3) = \frac{1}{3}$.
Hence, the updated probability that the car is behind the other door (Door 2) is $\frac{2}{3}$ and therefore you should switch.

(b) **Method 2.** The extensive form is shown in Figure 9.20
('cbn' means 'the car is behind door n', 'chn' means 'you choose door n').
The hypothesized sequence of events leads to either node x or node y (you first choose Door 1 and then the host opens Door 3).
The prior probability of getting to x is $\frac{1}{3} \times \frac{1}{2} = \frac{1}{6}$
while the prior probability of getting to y is $\frac{1}{3} \times 1 = \frac{1}{3}$.
The information you have is $\{x,y\}$ and, using the conditional probability rule,
$P(\{x\}|\{x,y\}) = \frac{P(\{x\}\cap\{x,y\})}{P(\{x,y\})} = \frac{P(\{x\})}{P(\{x,y\})} = \frac{\frac{1}{6}}{\frac{1}{6}+\frac{1}{3}} = \frac{1}{3}$.
Thus you should switch. □

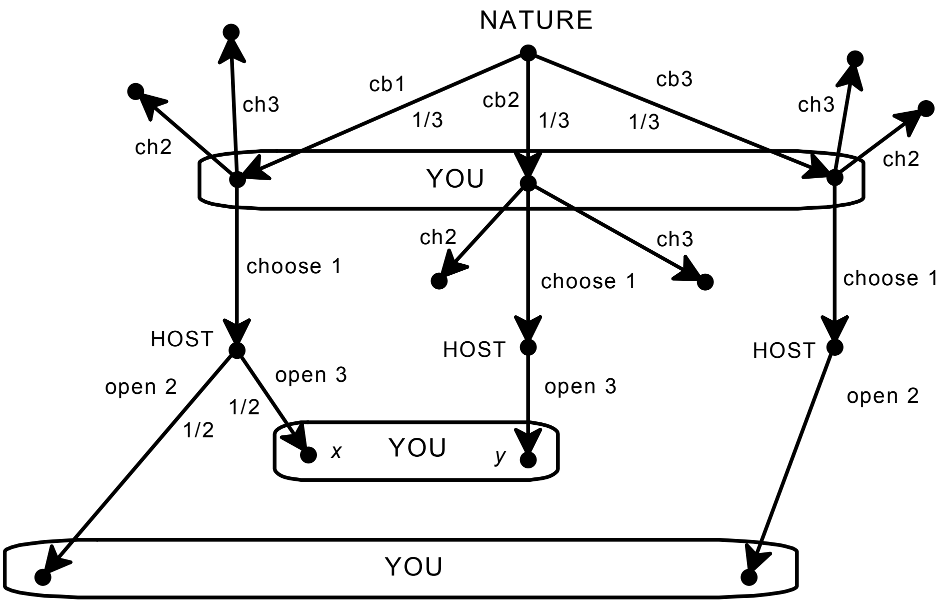


Figure 9.20: The Monty Hall problem represented using an imperfect-information frame.