Developing inquiry skills

Which of the events in the first opening scenario are independent? Which are mutually exclusive?

7.4 Complete, concise and consistent representations

You can use diagrams as a rich source of information when solving problems. Choosing the correct way to represent a problem is a skill worth developing. For example, consider the following problem:

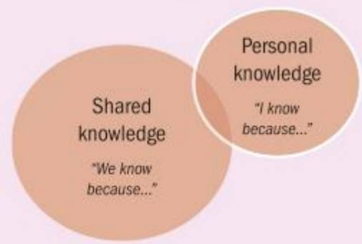
In a class of 15 students, 3 study art, 6 biology of whom 1 studies art. A student is chosen at random. How many simple probabilities can you find? How many combined probabilities can you find?

Let A represent the event "An art student is chosen at random from this group" and B "A biology student is chosen at random from this group".

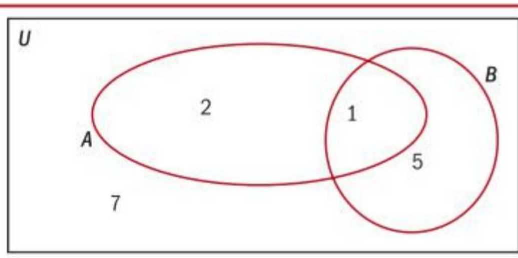
If you represent the problem only as **text**, the simple probabilities $P(A) = \frac{1}{5}$ and $P(B) = \frac{2}{5}$ can be found easily but calculating these do not show you the whole picture of how the sets relate to each other.

TOK

In TOK it can be useful to draw a distinction between shared knowledge and personal knowledge. The IB use a Venn diagram to represent these two types of knowledge. If you are to think about mathematics (or any subject, in fact) what could go in the three regions illustrated in the diagram?

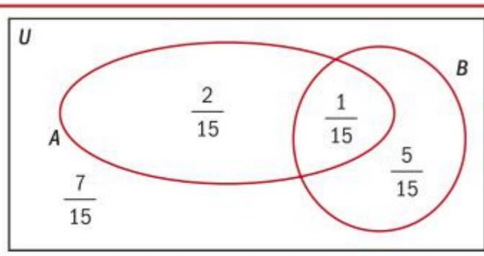


Represent this information as follows in a **Venn diagram** to see more detail:



The rectangle represents the sample space U for which $P(U) = 1$, the total probability.

The diagram allows us to find $P(B|A) = \frac{1}{3}$, $P(B'|A) = \frac{2}{3}$ etc easily.



The Venn diagram can be adapted to show the distribution of the total probability in **four** regions that represent mutually exclusive events: $P(A \cap B) = \frac{1}{15}$, $P(A' \cap B) = \frac{5}{15}$, $P(A \cap B') = \frac{2}{15}$ and $P(A' \cap B') = \frac{7}{15}$.

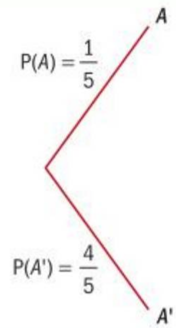
Hence the probability that a randomly chosen student studies neither biology nor art is $P(A' \cap B') = \frac{7}{15}$. The simple probability

$P(B) = P(A \cap B) + P(A' \cap B) = \frac{1}{15} + \frac{5}{15} = \frac{2}{5}$ is represented as a union of two mutually exclusive events in the Venn diagram.

The Venn diagram can be therefore be used to find all the simple, combined and conditional probabilities.

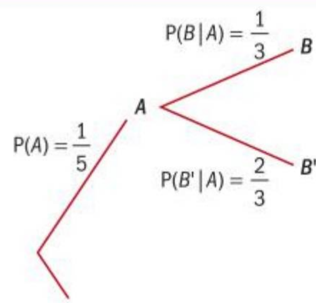
Represent the problem as a **tree diagram** by first choosing one student from the group and determining whether they study art or not.

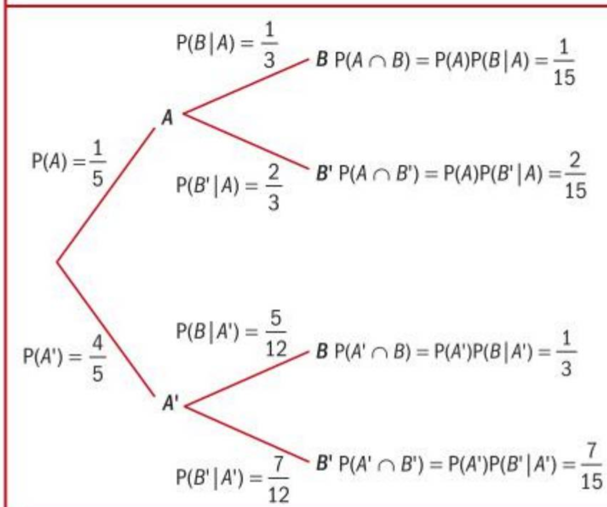
The probabilities in this process can be represented as a tree with the two events A and A'.



To construct the next part of the tree, imagine a student who **does** study art and consider whether this student studies biology or not.

This involves writing the same conditional probabilities as found in the Venn diagram.





Similarly, complete the rest of the tree as shown.

Then apply the multiplication law of probability to find the probability represented at the end of each "branch" of the tree. For example,

$$P(A \cap B) = P(A)P\left(B \middle| A\right) = \frac{1}{5} \times \frac{1}{3} = \frac{1}{15}.$$

The total probability of 1 is seen to be distributed along the branches of the tree by applying the multiplication law of probability for the other combined events.

Notice that the simple probability $P(B) = P(A \cap B) + P(A' \cap B) = \frac{1}{15} + \frac{1}{3} = \frac{2}{5}$ can be found from the probabilities at the end of two branches of the tree diagram.

A tree diagram is another way to represent all the possible outcomes of an event. The end of each branch represents a combined event.

International-mindedness

In 1933 Russian mathematician Andrey Kolmogorov built up probability theory from fundamental axioms in a way comparable with Euclid's treatment of geometry that forms the basis for the modern theory of probability. Kolmogorov's work is available in an English translation titled *The Foundations of Probability Theory*.

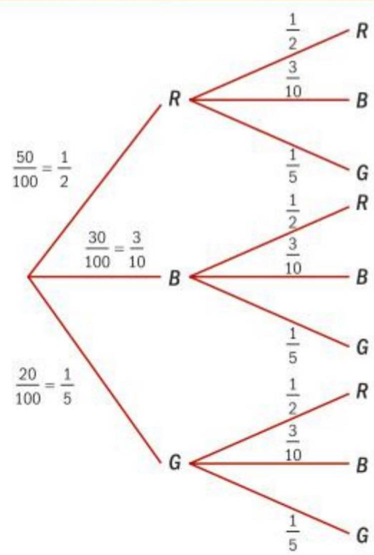
Investigation 12

Situation 1

A CAS project aims to raise funds by organizing a simple lottery. A number of coloured dice are placed in a box: 50 red dice, 30 blue and 20 green. To play the lottery, a die is chosen at random from the box and its colour noted. It is then replaced in the box and another die is chosen at random. If the two colours are the same, a small prize is awarded. The organizers of the project, Dani and Malena, want to know the probability of winning a prize in the most efficient way. They represent and interpret the information given by the lottery in two different ways.

Consider and reflect on the different approaches.

Dani's representation



The first set of branches of the tree represent the colour possibilities of the first die chosen. The second set of branches represent the second, so there is an implicit time axis from left to right.

Dani's application and interpretation

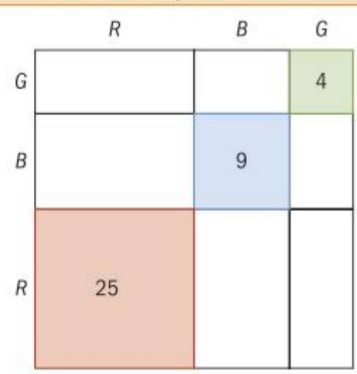
Since the dice are replaced in the box, the probability of choosing a red die on the second draw is the same as choosing a red on the first. The same argument applies to all the other pairs of combined events. Hence the probability of any particular choice on the second choice is **independent** of the first.

The probability of two reds is $\frac{1}{2}\times\frac{1}{2}=\frac{1}{4}$, similarly the probability of two blues and two greens are $\frac{9}{100}$ and $\frac{1}{25}$ respectively.

These three events are **mutually exclusive**, so you can add them to find the combined probability required.

$$\text{P(the colours of the two dice are the same)}$$
$$=\frac{1}{4}+\frac{9}{100}+\frac{1}{25}=\frac{38}{100}=0.38$$

Malena's representation



The horizontal *R*, *B*, *G* represent the colour possibilities of the first die chosen, the vertical letters the second.

Malena's application and interpretation

Malena reasoned that she would expect 50 red dice on the first choice, and of these 50, 25 would be followed by a second red since the second choice is **independent** of the first. Similarly, she reasoned that out of an expected 30 blue die from the first throw, 9 would be followed by a second blue.

Malena reasoned that a total of $25+9+4$ trials in which two identical colours were chosen would lead to a theoretical probability of 38 out of 100.

Dani was puzzled by this at first, but he could see after a while that Malena had created a slightly unusual looking Venn diagram. It was clear to them both from Malena's diagram that the sets are all **mutually exclusive.**

Continued on next page

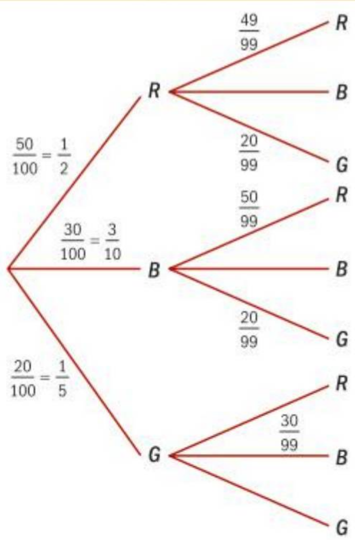
1 **Factual** Which representation involves less drawing and writing?

2 **Conceptual** Which diagram best represents how the stages of the game develop over time?

Situation 2

Then the CAS project organizers **changed the rules** as follows. A number of coloured dice are placed in a box: 50 red dice, 30 blue and 20 green. To play the lottery, a die is chosen at random from the box and its colour noted. It is **not** replaced in the box and another die is chosen at random. If the two colours are the same, a small prize is awarded.

Fill in the missing probabilities to find the probability of winning a prize.

Dani's representation	Dani's application and interpretation
 The probabilities on the second set of branches represent the fact that since the die chosen in the first choice has not been replaced, the sample space has changed in size. Hence the probability of a red on the second choice given that a red was chosen at first can be represented as $P(R_2 \mid R_1) = \frac{49}{99}$ and so on.	Since the dice are not replaced in the box, the probability of choosing any particular colour on the second draw is **dependent** on what the first choice was. For example, $P(R_2 \mid R_1) = \frac{49}{99}$ since one red die has been chosen from the 50 red dice and **not** replaced, hence there are 99 remaining to choose from. The probability of two reds is $\frac{1}{2} \times \frac{49}{99} = \frac{49}{198} \simeq 0.247$, similarly the probability of two blues and two greens are $\frac{??}{330}$ and $\frac{19}{???}$ respectively. These three events are **mutually exclusive**, so we can add them to find the combined probability required. $\text{P(the colours of the two dice are the same)}$ $= \frac{49}{198} + \frac{??}{330} + \frac{19}{???} = \frac{??}{99} \simeq 0.???$ The answer $\frac{??}{99}$ is exact. However, to make the comparison with 0.38 we must approximate to three significant figures. This example shows that changing from replacing the dice to not replacing the dice in fact reduces the probability of choosing two dice of the same colour in this problem.

3 **Conceptual** What feature of choosing without replacement makes a Venn diagram a poor choice as a representation?

4 **Conceptual** What feature of choosing without replacement makes a tree diagram a good choice as a representation?

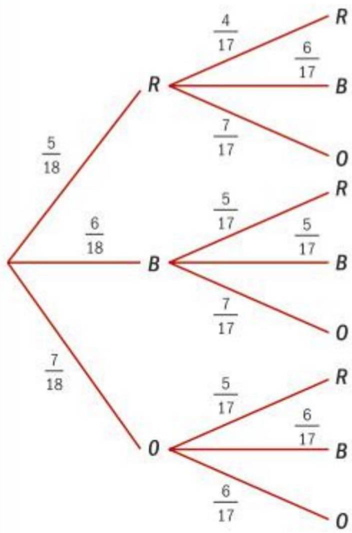
5 **Conceptual** How can we know when best to represent a problem with a tree diagram?

You can choose a diagram to make your calculation easier and more efficient. Also, sometimes you can choose how you calculate a probability to make it easier and more efficient, as shown in the following example:

Example 15

A box in an electronics store contains nine volt batteries of different colours. There are five red, six blue and seven orange. Two batteries are chosen to power up a radio that requires two batteries. Find the probability that both the batteries are different colours.

The batteries are selected **without** replacement since two batteries are needed to power up the radio.



Interpret the context to make sure you are applying the maths correctly.

Draw a tree diagram, labelling all the branches. Take care to fill in the probabilities correctly to show that the batteries are not replaced.

The answer can be found in two ways:

P(Batteries are different colours)
= P(RB or RO or BR or BO or OR or OB)

$$= \frac{5}{18}\times\frac{6}{17}+\frac{5}{18}\times\frac{7}{17}+\frac{6}{18}\times\frac{5}{17}+$$

$$\frac{6}{18}\times\frac{7}{17}+\frac{7}{18}\times\frac{5}{17}+\frac{7}{18}\times\frac{6}{17}=\frac{107}{153}$$

Or

P(Batteries are different colours)
= 1 − P(Batteries are identical colours)

$$= 1-\frac{5}{18}\times\frac{4}{17}-\frac{6}{18}\times\frac{5}{17}-\frac{7}{18}\times\frac{6}{17}=\frac{107}{153}$$

Notice that the second method involves finding only **three** combined probabilities whereas the first method involves finding **six**.

The complementary probability law $P(A) = 1 - P(A')$ can give you a short way to solve problems.

Example 16

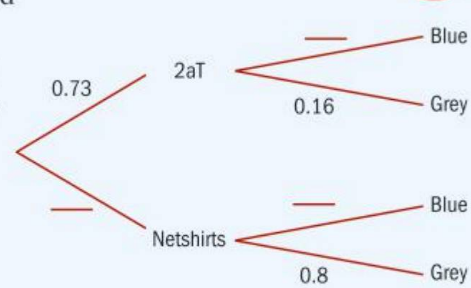
Mark buys his t-shirts from exactly two suppliers: 2aT and Netshirts, and he only buys two colours—grey and blue.

He buys 73% of his t-shirts from 2aT. Moreover, 16% of his 2aT t-shirts are grey, and 80% of his Netshirts t-shirts are grey.

a Copy and complete the tree diagram.

b Calculate the probability that a t-shirt randomly chosen by Mark one morning is grey.

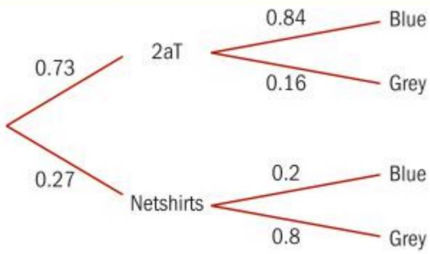
c Given that Mark chooses a grey t-shirt, determine the probability it was supplied by Netshirts.



Continued on next page

Statistics and probability

a



The question states that the events are complementary so the diagram can be filled in.

b $P(\text{grey}) = P(\text{2aT and grey}) + P(\text{Netshirts and grey})$
$= 0.73 \times 0.16 + 0.27 \times 0.8 = 0.3328$

Two mutually exclusive events are added to find the probability that Mark chooses a grey shirt.

c $P(\text{Netshirts}|\text{grey}) = \dfrac{P(\text{Netshirts and grey})}{P(\text{grey})}$

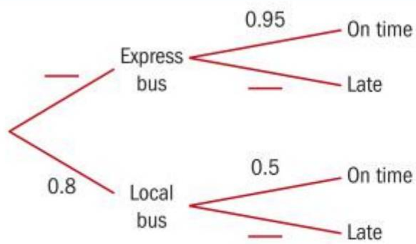
$= \dfrac{0.27 \times 0.8}{0.3328} \approx 0.649$

Apply the formula. Notice that you have already calculated the probabilities needed.

Exercise 7H

1 Denise can catch a local bus or an express bus to take her to work each day. The probability she catches the local bus is 0.8. If she catches the local bus, the probability that she is on time for work is 0.5. If Denise catches the express bus, the probability that she is on time for work is 0.95.

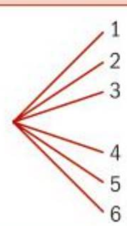
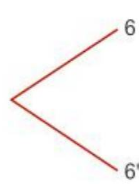
a Copy and complete the tree diagram below:



b Hence calculate the probability that Denise will be late for work.

2 A jewellery box contains 13 gold earrings, 10 silver earrings and 12 titanium earrings. Two earrings are drawn at random with replacement. Find the probability that they are made of different metals.

3 Chevy plays a game with four fair cubical dice numbered $\{1, 2, 3, 4, 5, 6\}$. She throws the four dice 160 times and finds that the event "throw at least one 6" occurs 77 times. She wonders what the theoretical probability is. She finds the answer after critically considering these four representations of the problem. Identify the best representation and justify your answer. Identify the worst representation and justify your answer.

Tree diagram with each of the four throws represented as below, giving a total of 1296 branches	A quick and easy calculation.	Tree diagram with each of the four throws represented as below, giving a total of 16 branches	A quick and easy application of probability laws.
	$4 \times \dfrac{1}{6} = \dfrac{2}{3}$		$1 - \left(\dfrac{5}{6}\right)^4$

4 A supermarket uses two suppliers, C and D, of strawberries. Supplier C supplies 70% of the supermarket's strawberries. Strawberries are examined in a quality control inspection (QCI); 90% of the strawberries supplied by C pass QCI and 95% of the strawberries from D pass QCI.

A strawberry is selected at random.

a Find the probability that the strawberry passes QCI.

b Given that a strawberry passes QCI, find the probability that it came from supplier D.

c In a sample of 2000 strawberries, find the expected number of strawberries that would fail QCI.

d The supermarket wants the probability that a strawberry passes QCI to be 0.93. Find the percentage of strawberries that should be supplied by D in order to achieve this.

5 Chevy plays a game in which she throws a pair of fair cubical dice numbered {1, 2, 3, 4, 5, 6} 24 times. Find the probability that she throws at least one double six.

6 A factory produces a large number of electric cars. A car is chosen at random from the production line as a prize in a competition. The probability that the car is blue is 0.5. The probability that the car has five doors is 0.3. The probability that the car is blue or has five doors is 0.6. Find the probability that the car chosen is not a blue car with five doors.

Pietro solves this problem with a Venn diagram but Maria solves it with a tree diagram. They both get the correct answer. Solve the problem both ways. Discuss and then state which is the most efficient method.

7 A choir contains 15 girls and 9 boys. The choirmaster randomly selects four singers from the choir to be interviewed by a newspaper. Find the probability that at least one boy is in the four singers selected.

Developing inquiry skills

Apply what you have learned in this section to represent the first opening problem with a tree diagram.

Hence find the probability that a cab is **identified** as yellow.

Apply the formula for conditional probability to find the probability that the cab was yellow **given that** it was identified as yellow.

How does your answer compare to your original subjective judgment?