Summary of key points

1 A **Venn diagram** can be used to represent events graphically. Frequencies or probabilities can be placed in the regions of the Venn diagram.

2 For **mutually exclusive** events, P(A or B) = P(A) + P(B).

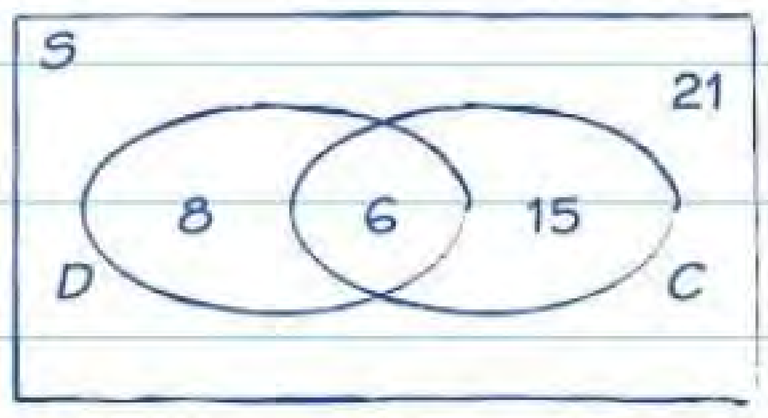
3 For **independent** events, P(A and B) = P(A) × P(B).

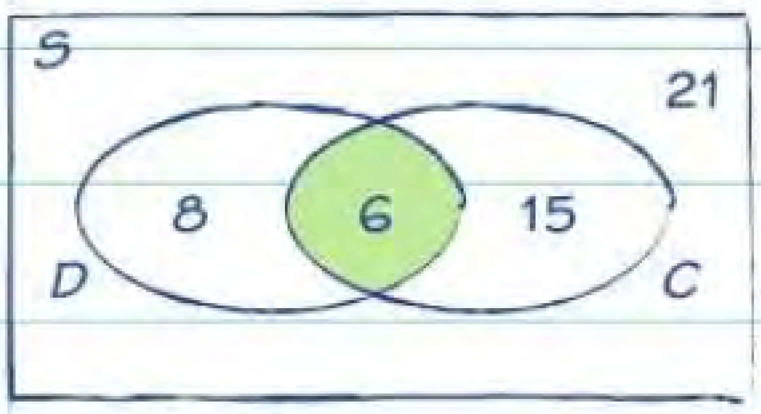
4 A **tree diagram** can be used to show the outcomes of two (or more) events happening in succession.

Drawing Venn diagrams

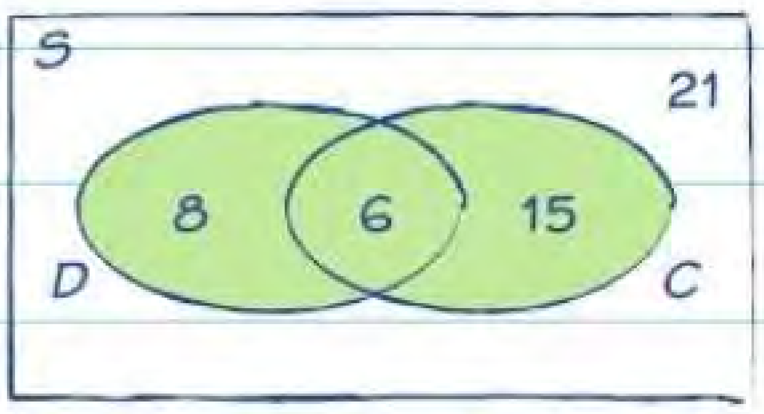
You can use a Venn diagram to represent different **events** in a **sample space**.

This Venn diagram shows the results when 50 people were surveyed about whether they owned a dog (D) or a cat (C). The rectangle represents the whole sample space (S), and each event is represented by an oval.

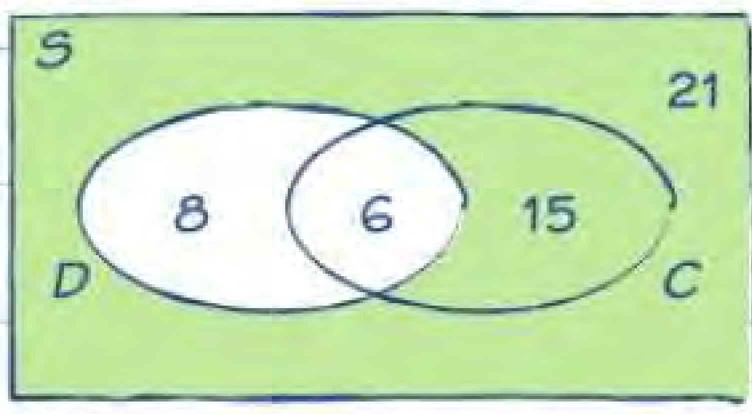




6 people owned a dog **and** a cat. You can write this event as 'D and C'.



8 + 6 + 15 = 29 people owned a dog **or** a cat.
You can write this event as 'D or C'.



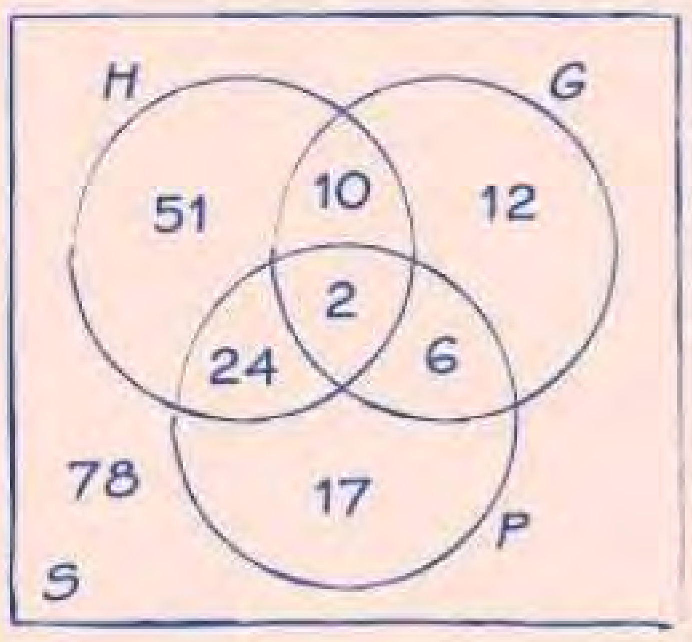
15 + 21 = 36 people **did not** own a dog.
You can write this event as 'not D'.

Worked example

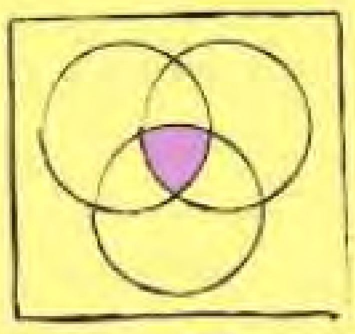
200 students were surveyed about whether they are taking history, German or physics A-Level:

87 take history
30 take German
49 take physics
12 take history and German
26 take history and physics
8 take German and physics
2 take all three subjects.

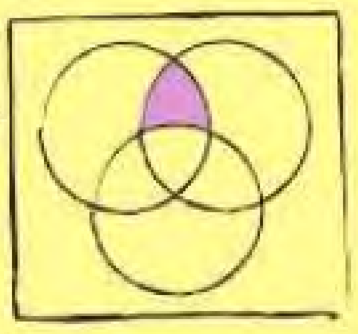
Draw a Venn diagram to represent this information. **(5 marks)**



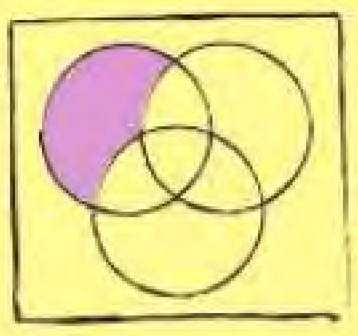
H and *C* and *P*
You can fill in the centre of the Venn diagram first. 2 students take all three subjects.



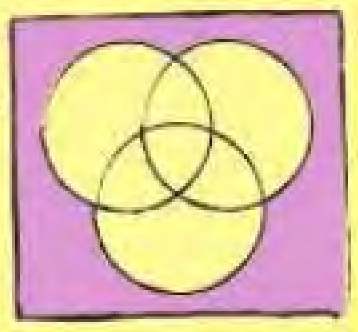
H and *C* but not *P*
The 12 students who take history and German **include** the 2 students who take all three subjects. So 12 − 2 = 10 students take history and German but **not** physics.



H but not *C* and not *P*
The 87 students who take history **include** the 10 + 2 + 24 = 36 students who take at least one other subject. So 87 − 36 = 51 take **only** history.



None of *H*, *C* or *P*
51 + 12 + 17 + 10 + 24 + 6 + 2 = 122 students take at least one subject. So 200 − 122 = 78 take **none** of the three subjects.



Now try this

An Indian restaurant recorded the bread choices at 100 tables. 56 tables ordered naan, 22 tables ordered roti and 40 ordered paratha. There were 8 tables that ordered naan and roti, 19 that ordered naan and paratha, and 15 that ordered roti and paratha. 6 tables ordered all three types of bread.

Represent these data on a Venn diagram. Answer **(5 marks)**

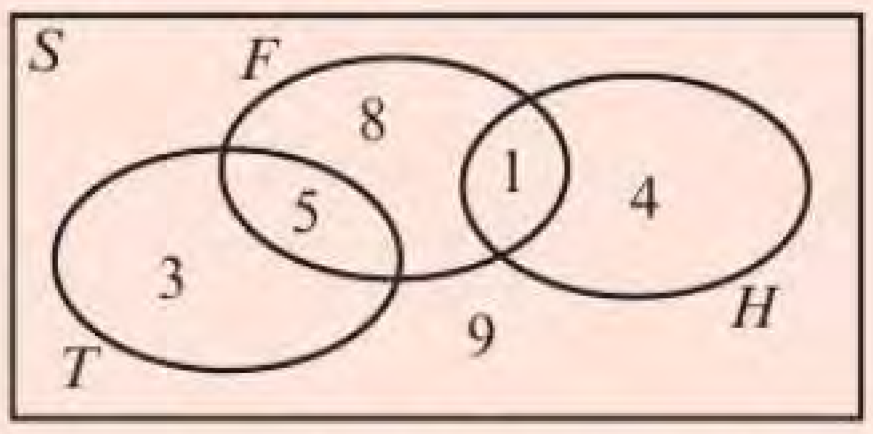
Make sure you label the whole sample space S, and draw three closed, intersecting circles or ovals to represent the three events. Label them N, R and P.

Using Venn diagrams

You can use Venn diagrams to help you answer **probability** questions. Find the region in the Venn diagram which **satisfies the conditions** in the question, add up the total number of **successful outcomes** and divide by the total number of possible outcomes.

Worked example

This Venn diagram shows the numbers of students in a class who play tennis, football and hockey.



One of these students is selected at random.

(a) Show that the probability that the student plays more than one of the sports is $\frac{1}{5}$ **(2 marks)**

$3 + 5 + 8 + 9 + 1 + 4 = 30$

$\frac{5 + 1}{30} = \frac{6}{30} = \frac{1}{5}$

(b) Find the probability that the student plays either football or tennis, or both. **(2 marks)**

$\frac{3 + 8 + 1 + 5}{30} = \frac{17}{30}$

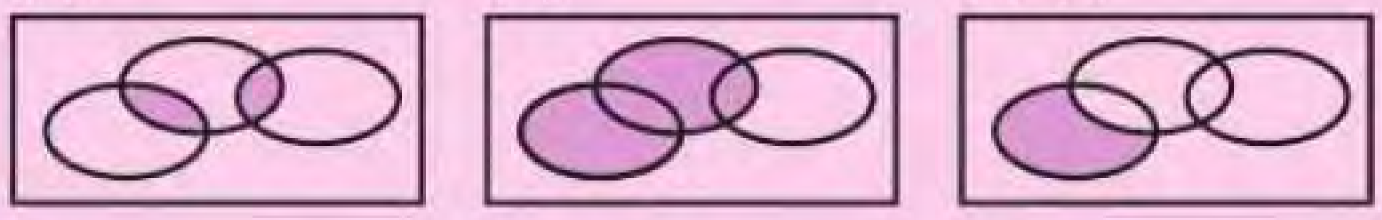
(c) Write down the probability that the student plays
  (i) tennis but not football
  (ii) all three sports. **(2 marks)**

(i) $\frac{3}{30} = \frac{1}{10}$ (ii) 0

Problem solved!

Make sure you add up the outcomes for **all** the regions you are interested in:

(a) (b) (c) (i)



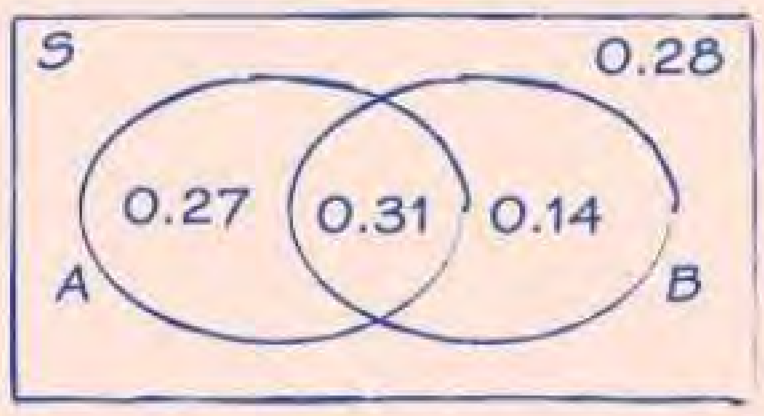
For part (c) (ii), there is no region where all three sports overlap, so no students played all three sports.

You will need to use problem-solving skills throughout your exam – **be prepared!**

Worked example

For the events A and B, P(A and not B) = 0.27, P(B and not A) = 0.14 and P(A or B) = 0.72

(a) Draw a Venn diagram to illustrate the complete sample space for events A and B. **(3 marks)**



The total of all the probabilities must add up to 1.

(b) Write down the value of P(A) and the value of P(B). **(3 marks)**

P(A) = 0.27 + 0.31 = 0.58

P(B) = 0.31 + 0.14 = 0.45

Now try this

A survey showed that 80% of students in a school own a laptop, and 22% of students own a tablet. 15% of students own neither a laptop nor a tablet.

(a) Draw a Venn diagram to represent this information. Answer **(4 marks)**

A student is chosen at random.

(b) Write down the probability that this student owns a laptop but not a tablet. Answer **(1 mark)**

You can write probabilities as percentages on a Venn diagram. Remember that the percentages in the whole sample space need to add up to 100%.

Independent events

Two events are statistically independent if the outcome of one has **no effect** on the probability of the other.

Determining independence

Two events A and B are independent if and only if:

P(A and B) = P(A) × P(B)

This rule is given in the A-level section of the formulae booklet as:

$$P(A \cap B) = P(A)\,P(B)$$

where the symbol $\cap$ means 'and'.

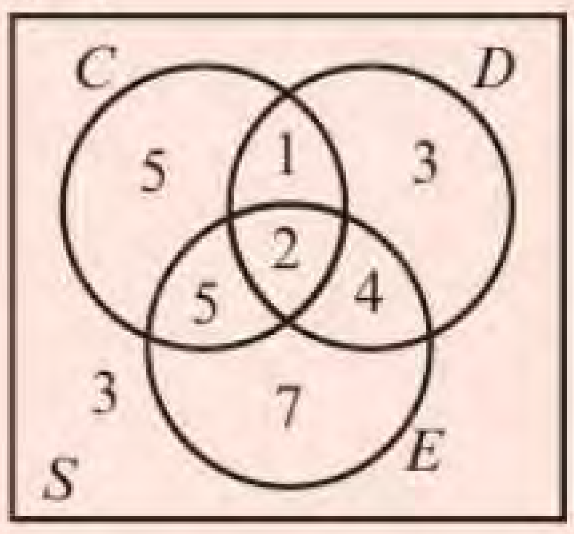
You can use this rule in two ways:

1. You can show that two events are independent by calculating P(A), P(B) and P(A and B) and demonstrating that the relationship is true.
2. If you are told that two events are independent you can use the fact that P(A and B) = P(A) × P(B) to find unknown values.

Remember to count **all** the possible outcomes in each event when determining the probability of that event. Make sure you don't use the rule given above in your working, as you haven't been told that the events are independent.

Worked example

This Venn diagram shows the numbers of students in a class of 30 who watched Coronation Street (C), Doctor Who (D) or EastEnders (E).



Determine whether watching EastEnders and watching Doctor Who are statistically independent. **(3 marks)**

$$P(D) = \frac{1+3+2+4}{30} = \frac{10}{30} = \frac{1}{3}$$

$$P(E) = \frac{5+2+4+7}{30} = \frac{18}{30} = \frac{3}{5}$$

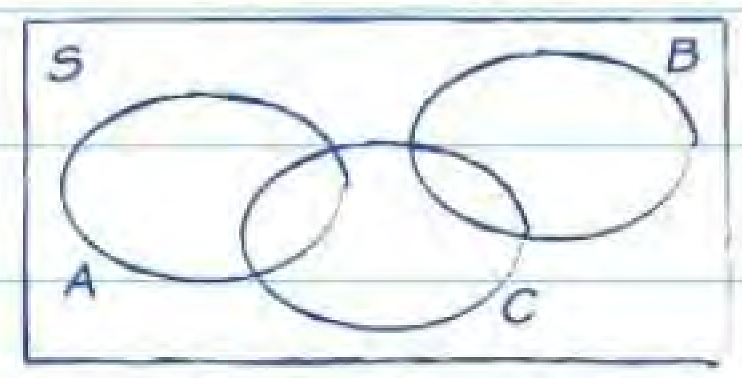
$$P(D \text{ and } E) = \frac{2+4}{30} = \frac{6}{30} = \frac{1}{5}$$

$$P(D) \times P(E) = \frac{1}{\cancel{3}_1} \times \frac{\cancel{3}^1}{5} = \frac{1}{5} = P(D \text{ and } E)$$

So D and E are statistically independent.

Mutually exclusive events

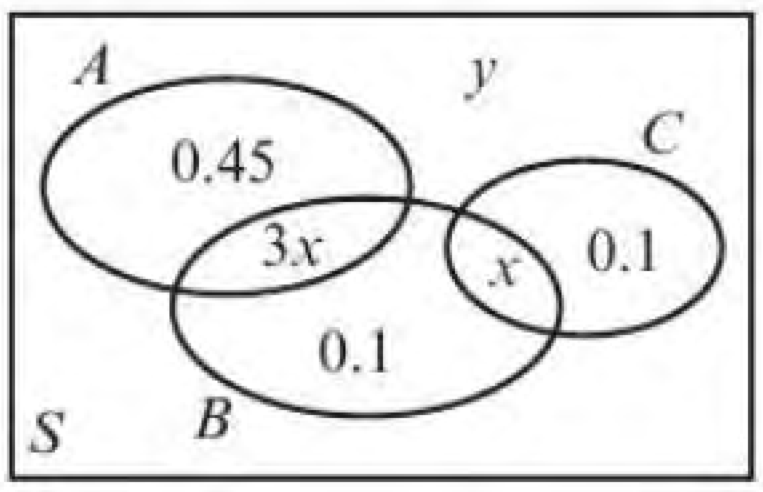
Two events are mutually exclusive if they **cannot both** occur.
On the Venn diagram on the right, the events A and B are mutually exclusive because they **do not overlap**.



For two mutually exclusive events A and B: P(A and B) = 0 and P(A and B) = P(A) + P(B)

Now try this

The Venn diagram shows the probabilities of three events, A, B and C. x and y are unknown probabilities. Answer



(a) Write down two events that are mutually exclusive. **(1 mark)**

Given that events B and C are independent,

(b) show that $400x^2 - 50x + 1 = 0$ Answer **(3 marks)**

(c) find the values of x and y, justifying your choice of solution to the above quadratic equation. Answer **(4 marks)**

Had a look ☐ Nearly there ☐ Nailed it! ☐

Tree diagrams

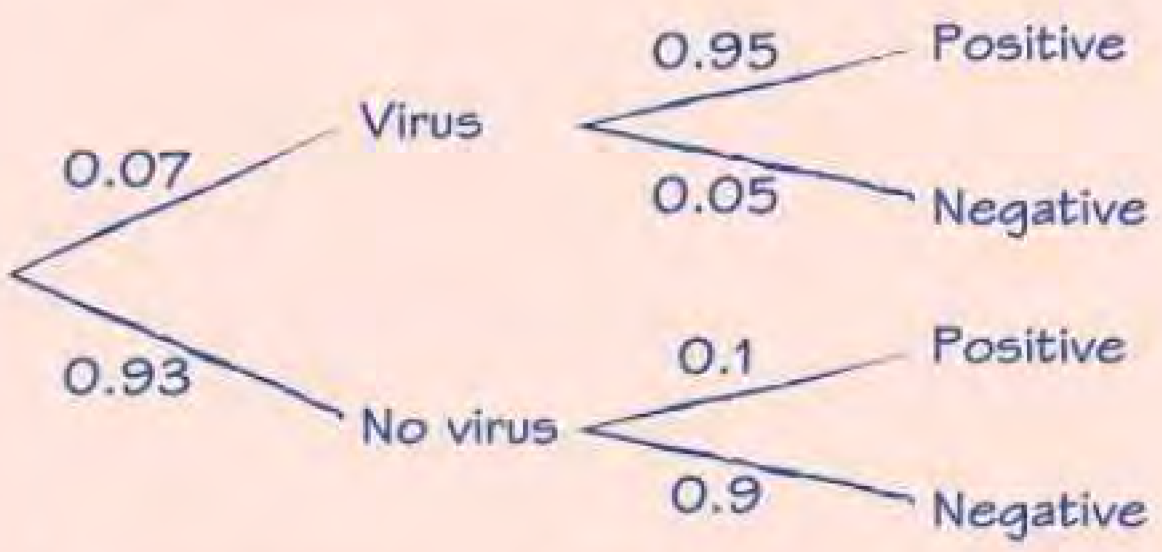
Tree diagrams can be used to solve some probability problems. In your exam, it's usually only a good idea to draw a tree diagram when you are told to do so in the question.

Worked example

A computer virus is known to infect 7% of all computers. A software company writes a virus checker to determine whether or not a computer is infected. If a computer is infected, the test is positive with probability 0.95

If a computer is not infected, the test is positive with probability 0.1

(a) Draw a tree diagram to represent this information. **(3 marks)**



A computer is selected at random and tested.

(b) Find the probability that the test is positive. **(3 marks)**

$P(\text{Positive}) = 0.07 \times 0.95 + 0.93 \times 0.1$
$= 0.1595$

A sample of 10 computers is selected at random and tested.

(c) Find the probability that more than three of them test positive. **(3 marks)**

Let X = the number of computers that test positive. Then $X \sim B(10, 0.1595)$

$P(X > 3) = 1 - P(X \leqslant 3)$
$= 1 - 0.9392...$
$= 0.06075...$

So the probability is 0.0608 (3 s.f.)

Tree diagram checklist

Make sure that you:

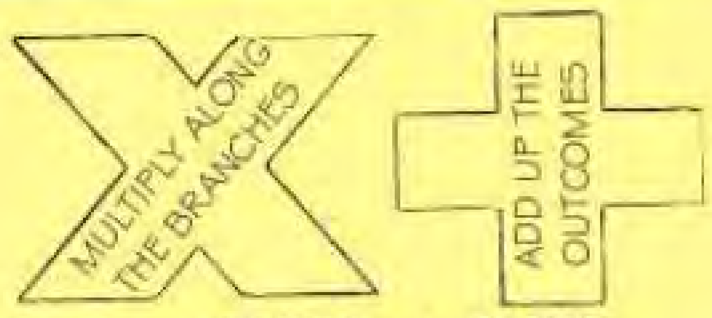
- ✓ write a probability on **every** branch
- ✓ write an outcome at the **end** of every branch.

In your exam you **don't** need to:

- ✗ draw a tree diagram unless it's asked for in the question
- ✗ work out the probabilities of **all** the final outcomes – you will be asked for specific probabilities later in the question.

On each pair of branches the probabilities need to **add up to 1**.
So P(No virus) = 1 − 0.07 = 0.93

On a tree diagram you:



P(Virus and Positive) = 0.07 × 0.95
P(No virus and Positive) = 0.93 × 0.1
Add these together to find P(Positive).

Problem solved!

You are repeating trials with the same probability of success. You can model the number of computers that test positive with a **binomial distribution**. Write down the distribution and use your calculator to find the probability.

There is more about finding binomial probabilities on page 75

You will need to use problem-solving skills throughout your exam – **be prepared!**

Now try this

In a board game, Amy picks a question card. She can pick an easy, medium or hard question, with probabilities 0.54, 0.31 and 0.15 respectively.

The probabilities that she answers each type of question correctly are 0.8, 0.5 and 0.1 respectively.

(a) Draw a tree diagram to represent this information. Answer **(3 marks)**

(b) Amy picks a card at random and answers the question. Work out the probability that she answers the question correctly. **(4 marks)**

Answer