Schoolworkout Maths

Conditional Probability

Recap: Independent events.

If events A and B are *independent*, then the probability of B happening does not depend upon whether A has happened or not. Therefore

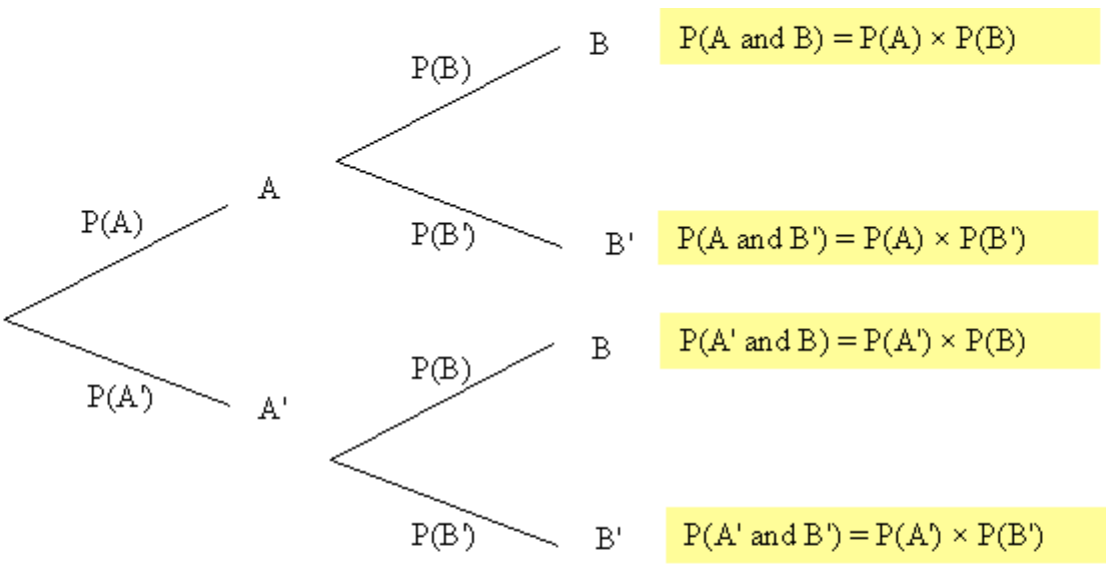
$$P(B \mid A) = P(B \mid A') = P(B)$$

Probability of B given that A has occurred

and

$$P(A \text{ and } B) = P(A) \times P(B).$$

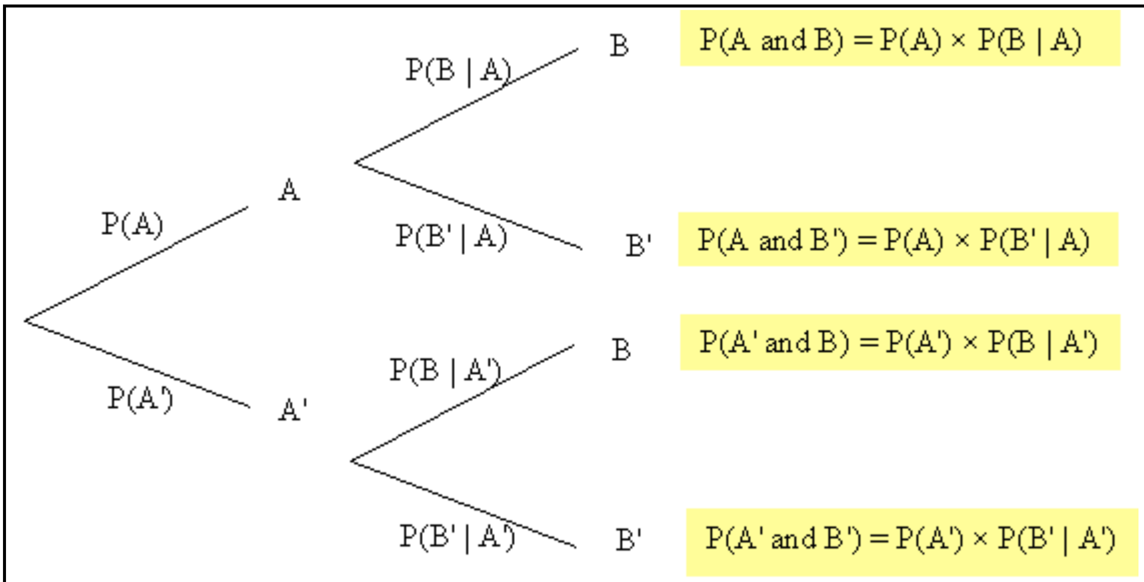
A tree diagram can be drawn:



Dependent events:

If A and B are *dependent* events, the probability of B happening will depend upon whether A has happened or not.

We therefore have to introduce conditional probabilities in the tree diagram (as shown below):



The rule for combining probabilities for dependent events is

$$P(A \text{ and } B) = P(A) \times P(B \mid A)$$

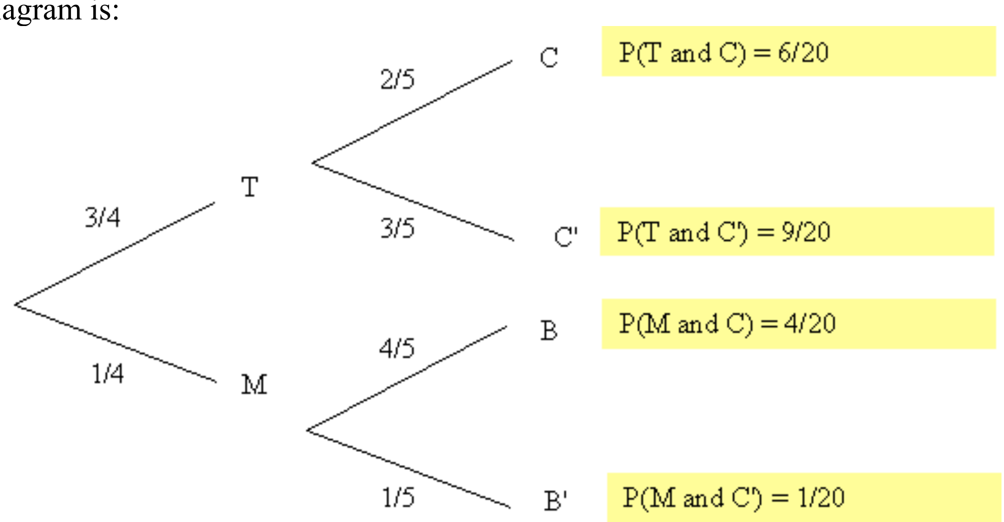
This is equivalent to saying

$$P(B \mid A) = \frac{P(A \text{ and } B)}{P(A)}$$

Example:
Every morning I buy either The Times or The Mail. The probability that I buy The Times is ¾ and the probability that I buy The Mail is ¼. If I buy The Times, the probability that I complete the crossword is $\frac{2}{5}$, whereas if I buy The Mail the probability that I complete the crossword is $\frac{4}{5}$.

a) Find the probability that I complete the crossword on any particular day.
b) If I have completed the crossword, find the probability that I bought The Mail.

The tree diagram is:



Notation: T = buy The Times M = buy The Mail C = complete crossword

a) From the tree diagram,

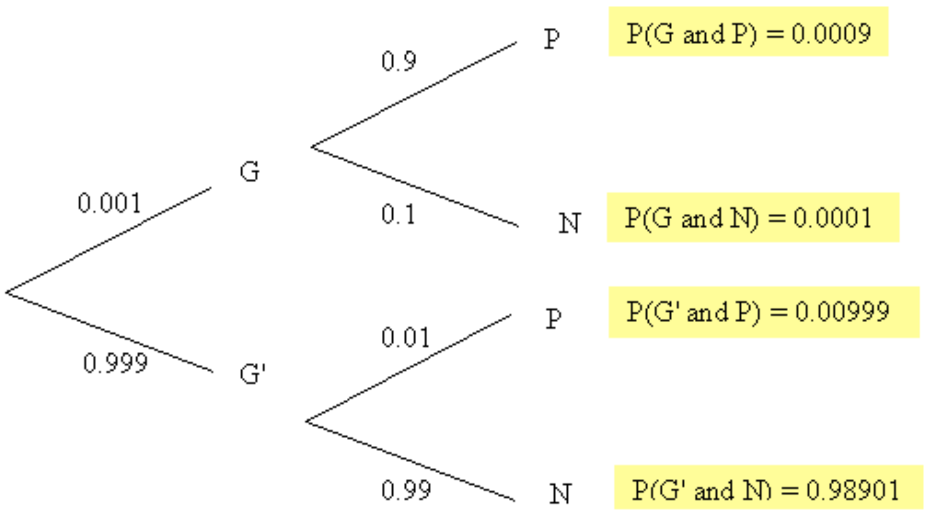
$$\text{P(complete crossword)} = \text{P(T and C)} + \text{P(M and C)} = \frac{6}{20} + \frac{4}{20} = \frac{1}{2}$$

b) The probability that I bought The Mail given that I completed the crossword is given by

$$\text{P}(M \mid C) = \frac{\text{P(M and C)}}{\text{P(C)}} = \frac{4/20}{1/2} = \frac{2}{5}$$

Example 2:
0.1% of the population carry a particular faulty gene.
A test exists for detecting whether an individual is a carrier of the gene.
In people who actually carry the gene, the test provides a positive result with probability 0.9.
In people who don't carry the gene, the test provides a positive result with probability 0.01.
If someone gives a positive result when tested, find the probability that they actually are a carrier of the gene.

Let G = person carries gene P = test is positive for gene N = test is negative for gene
The tree diagram then looks as follows:



We want to find $\text{P}(G \mid P) = \dfrac{\text{P(G and P)}}{\text{P(P)}}$

However, P(P) = P(G and P) + P(G' and P) = 0.0009 + 0.00999 = 0.01089

Therefore, $\text{P}(G \mid P) = \dfrac{0.0009}{0.01089} = 0.0826$ (to 3 significant figures)
So there is a very low chance of actually having the gene even if the test says that you have it.

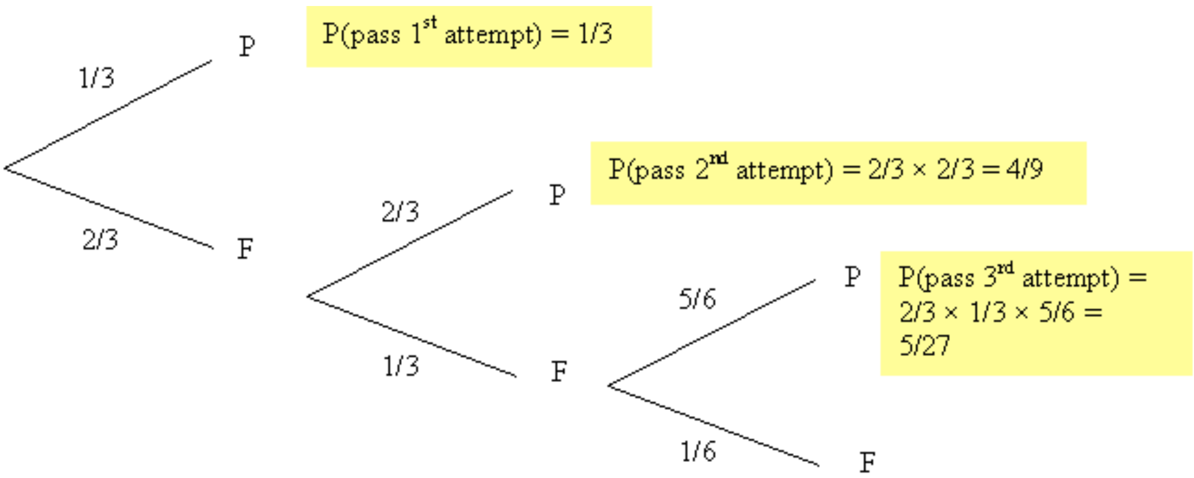
Note: This example highlights the difficulty of detecting rare conditions or diseases.

Past Examination Question: (OCR)
Students have to pass a test before they are allowed to work in a laboratory. Students do not retake the test once they have passed it. For a randomly chosen student, the probability of passing the test at the first attempt is 1/3. On any subsequent attempt, the probability of failing is half the probability of failing on the previous attempt. By drawing a tree diagram or otherwise,

a) show that the probability of a student passing the test in 3 attempts of fewer is 26/27,

b) find the conditional probability that a student passed at the first attempt, given that the student passed in 3 attempts or fewer.

Solution:
Let P = pass test F = fail test



a) So P(pass in 3 attempts or fewer) = 1/3 + 4/9 + 5/27 = 26/27 as required

b) Let X stand for when a student passes the test.

We need $P(X = 1 \mid X \leq 3) = \dfrac{P(X = 1 \text{ and } X \leq 3)}{P(X \leq 3)}$.

But $P(X = 1 \text{ and } X \leq 3)$ is the same as just $P(X = 1)$.

Therefore $P(X = 1 \mid X \leq 3) = \dfrac{1/3}{26/27} = \dfrac{9}{26}$