

## Decision Trees

### Decision Theory and Decision Trees

Decision trees are excellent tools for making financial or number based decisions where a lot of complex information needs to be taken into account. They provide an effective structure in which alternative decisions, and the implications of taking those decisions, can be laid down and evaluated. They also help you to form an accurate, balanced picture of the risks and rewards that can result from a particular choice.

### How to Draw a Decision Tree

You start a decision tree with a decision that needs to be made. This decision is represented by a small square towards the left of a large piece of paper. From this box draw out lines towards the right for each possible solution, and write that solution along the line. Keep the lines apart as far as possible so that you can expand your thoughts.

At the end of each solution line, consider the results. If the result of taking that decision is uncertain, draw a small circle. If the result is another decision that needs to be made, draw another square. Squares represent decisions; circles represent uncertainty or random factors. Write the decision or factor to be considered above the square or circle. If you have completed the solution at the end of the line, just leave it blank.

Starting from the new decision squares on your diagram, draw out lines representing the options that could be taken. From the circles draw out lines representing possible outcomes. Again mark a brief note on the line saying what it means. Keep on doing this until you have drawn down as many of the possible outcomes and decisions as you can leading on from your original decision.

An example of the sort of thing you will end up with is shown below in Illustration 1.11:

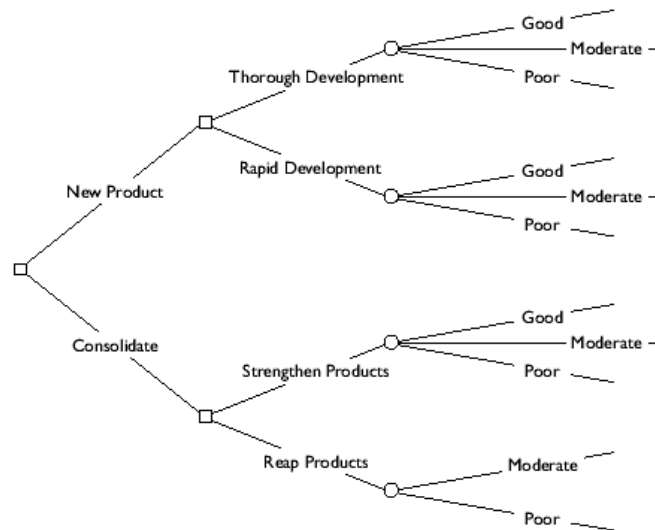


Illustration 1.11. Example decision tree.

Once you have done this, review your tree diagram. Challenge each square and circle to see if there are any solutions or outcomes you have not considered. If there are, draw them in. If necessary, redraft your tree if parts of it are too congested or untidy.

You should now have a good understanding of the range of possible outcomes.

**Starting to Evaluate Your Decision Tree**

Now you are ready to evaluate the decision tree. This is where you can calculate the decision that has the greatest worth to you. Start by assigning a currency or numeric value to each possible outcome – how much you think it would be worth to you.

Next look at each circle (representing an uncertainty point) and estimate the probability of each outcome. If you use percentages, the total must come to 100% at each circle. If you use fractions, these must add up to 1. If you have data on past events you may be able to make rigorous estimates of the probabilities. Otherwise write down your best guess.

This will give you a tree like the one below in Illustration 1.12:

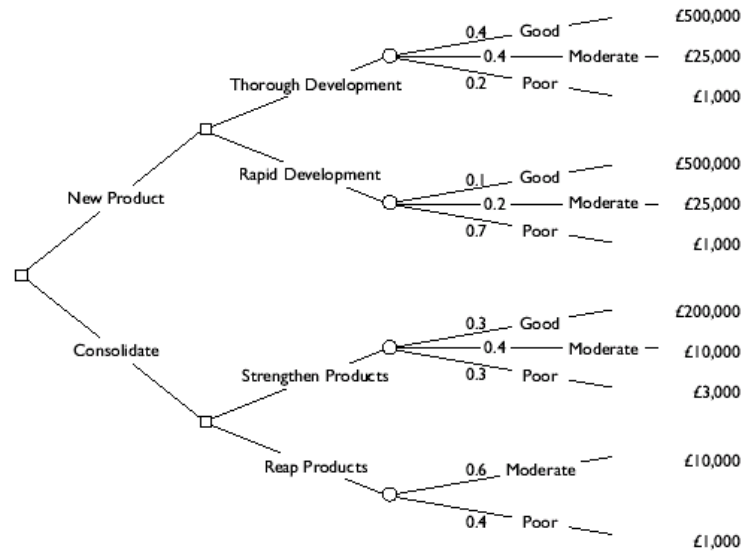


Illustration 1.12. Example decision tree 2.

### Calculating Tree Values

Once you have worked out the value of the outcomes and have assessed the probability of the outcomes of uncertainty, it is time to start calculating the values that will help you make your decision.

We start on the right hand side of the decision tree and work back towards the left. As we complete a set of calculations on a node (decision square or uncertainty circle), all we need to do is to record the result. All the calculations that lead to that result can be ignored from now on – effectively that branch of the tree can be discarded. This is called ‘pruning the tree’.

### Calculating The Value of Uncertain Outcome Nodes

Where we are assessing the value of an uncertain outcome (circles on the diagram), we do this by multiplying the value of the

outcomes by their probability, and noting the result. The total value of that node of the tree is gained by adding these together.

In the example above, the value for 'new product, thorough development' is:

0.4 (probability good outcome)	X	£500,000 (value)	=	£200,000
0.4 (probability moderate outcome)	X	£25,000 (value)	=	£10,000
0.2 (probability poor outcome)	X	£1,000 (value)	=	£200
				£210,200

This is shown across our example tree in Illustration 1.13:

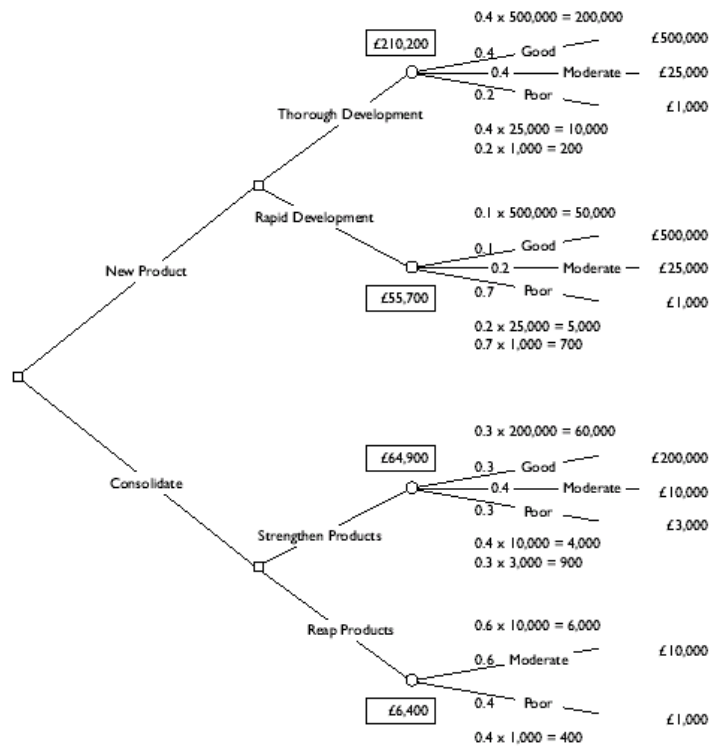


Illustration 1.13. Example decision tree 3.

Note that the values calculated for each node are shown in the boxes.

### Calculating The Value of Decision Nodes

When you are evaluating a decision node, write down the cost of each option along each decision line. Then subtract the cost from the value of that outcome that you have already calculated. This will give you a value, which represents the benefit of that decision.

Sunk costs, amounts already spent, do not count for this analysis.

When you have calculated the benefit of each decision, select the decision that has the largest benefit, and take that as the decision made and the value of that node.

Calculation of decision nodes in our example is shown below in illustration 1.14:

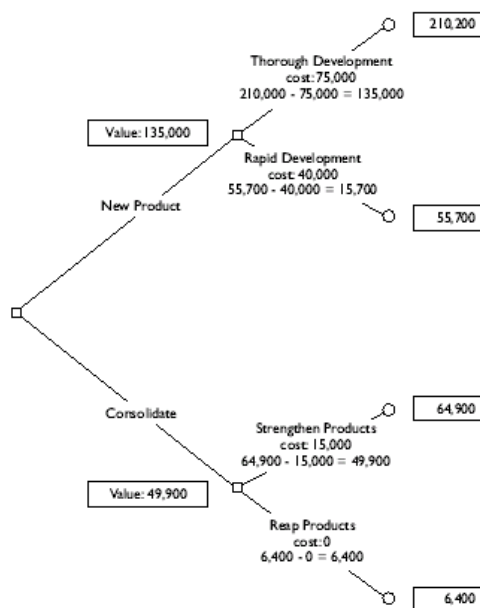


Illustration 1.14. Example decision tree 4.

In this example, the benefit we previously calculated for 'new product, thorough development' was £210,000. This example shows that we calculate the cost of this approach as £75,000. This gives a net benefit of £135,000. The benefit of 'new product, rapid development' was £15,700. On this branch we therefore choose the most valuable option, 'new product, thorough development', and allocate this value to the decision node.

**Result**

By applying this technique we can see that the best option for us may be to develop a new product. What the analysis shows which we might not have appreciated, is that it is worth much more to us to take our time and get the product right than to rush the product to market. In fact, it is better just to improve our existing products than to botch a new product, even though it costs us less.

**Tree Induction Algorithm**

Here is an algorithm for inducing a decision tree from such a collection of classified instances. Originally termed CLS (Concept Learning System) it has been successively enhanced. At the highest level of enhancement the system is known as ID3 – later versions include C4, C4.5 and C5.

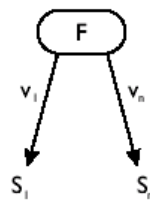


Illustration 1.15. A tree induction algorithm.

- The algorithm operates over a set of training instances,  $C$ .
- If all instances in  $C$  are in class  $P$ , create a node  $P$  and stop, otherwise select a *feature* or *attribute*  $F$  and create a decision node.
- Partition the training instances in  $C$  into subsets according to the values of  $V$ .

- Apply the algorithm recursively to each of the subsets C .

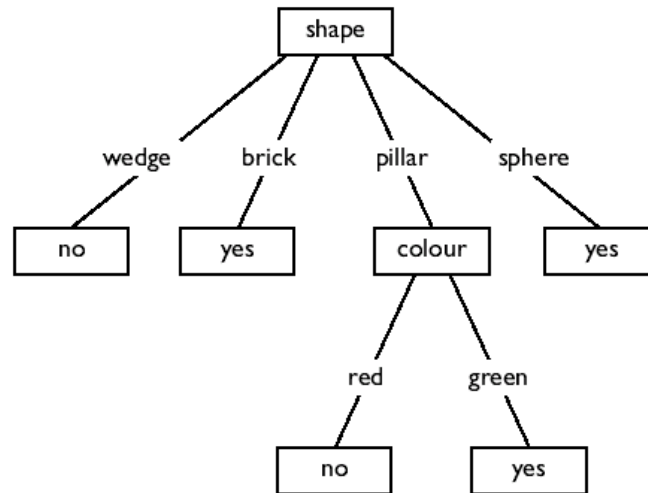


Illustration 1.16. Example output of a tree induction algorithm.

The output of a tree induction algorithm can easily be expressed as a nested if-statement (described later in the unit) :

```
if (shape == wedge)
    return no;
if (shape == brick)
    return yes;
if (shape == pillar)
{
    if (colour == red)
        return no;
    if (colour == green)
        return yes;
}
if (shape == sphere)
    return yes;
```