Chapter 2 Decision Analysis Primer

Spreadsheet software has made it possible to apply the speed and precision of a personal computer to the basic tasks of business analysis. Taking advantage of the computer's accuracy and speed, analysts often attempt to calculate the effects of scenario upon scenario, producing reports in such volume that decision makers begin to long for the day of pencil and paper and problems of too little, rather than too much, information.

The fundamental difficulty with spreadsheet analysis is that decision makers, whether they are business managers, government regulators, engineers, economists, attorneys, or physicians, generally face problems which require more than computational ability: what is the probability that a particular R&D project will be successful: what is the optimal price to bid in a competition to supply disk drives to a large computer manufacturer; what is the settlement value of recently commenced, treble-damage antitrust litigation; should a promising, but risky, treatment for AIDS be cleared for use before clinical testing is complete?

For these types of decisions, there are many factors to be analyzed, but in the end, the decision maker must make subjective judgments about the likelihood of particular scenarios, and make decisions based not only on costs and benefits, but also on assessments of risk.

The analysis of these problems is seriously complicated by uncertainty because, invariably, the decision maker lacks control over the consequences of one or more of the scenarios under consideration.

The appropriate way to deal with these problems is through decision analysis – a structured methodology which first puts the uncertainties into perspective and then takes them into account in the decision making process.

Decision analysis rests on the concept of *expected value*. For example, if 25% of the money wagered in a lottery is paid out in prizes, then every dollar you spend on a lottery ticket has an expected return of twenty-five cents. This does not mean that you, as the holder of a $1 ticket, will win (or even have the possibility of winning) exactly twenty-five cents when the lottery results are announced. Rather, a majority of lottery tickets will return nothing, while others will return substantially more than a dollar.

The assignment (while the outcome of the lottery remains uncertain) of a twenty-five cent "expected value" to every $1 lottery ticket indicates (1) that the ratio of all money spent on lottery tickets to all money paid as lottery winnings is estimated to be four to one; and (2) every ticket provides the same chance of winning. Determining the expected value of an event is a cornerstone of decision analysis.

There are other important considerations when structuring decisions. For example, apples must always be compared with other apples, not with any other kind of fruit. If you need to compare apples with peaches, you must convert them both to a common unit of measurement, such as dollars.

Two methodologies are available in DATA 3.0 for modeling a decision analysis problem. They have distinctly different means of visually representing the problem.

The first is through a *decision tree*, a branching structure in which each branch represents an event that may take place in the future. The second is through an *influence diagram*, in which each node represents a different factor that influences the outcome, and arcs between the nodes reflect the ways in which one factor influences another.

While DATA's influence diagrams can be extremely helpful for simplifying and presenting complex decisions, they must be converted into decision trees in order to be analyzed. Therefore, the next section of this chapter explains the structure and analysis of decision trees, and the final section describes the design of influence diagrams as an alternative method of generating an analyzable decision tree.

Decision Trees

The design of a decision tree is subject to a few guidelines:

1. *Time flows from left to right.* Decision trees are horizontal structures which proceed from the "root" of the tree (at the far left) to the branches (at the right). Each successive branch represents an event or decision as it occurs in time.

2. *All outcomes must be represented.* Each final outcome must be represented as an endpoint on the right side of the tree.

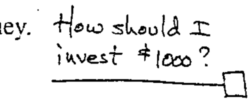
3. *Several types of "nodes" may be used.* In general, a *node* represents a decision, an uncertain event, or an outcome. Each branch of the tree has an associated node located at the right hand end of the branch. A *terminal* node (triangle) is used to denote a final outcome: the end of a path, often referred to as a *scenario*. A *decision* node (square) is used to indicate a decision facing the decision maker. A *chance* node (circle) is used to represent an uncertain or risky event. All of the nodes at the right edge of the tree must be terminal nodes.

4. *Branches emanating from a decision node represent the options.* All available choices must be represented, and none may overlap.

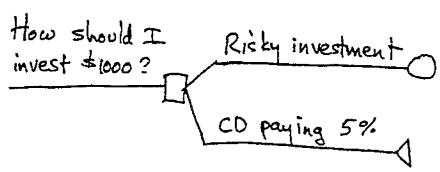
5. *Branches emanating from a chance node represent the possible outcomes of the event.* All possible outcomes must be represented, and none may overlap.

In accordance with these guidelines, let's design a decision tree that represents the investment problem posed at the end of Chapter 1.

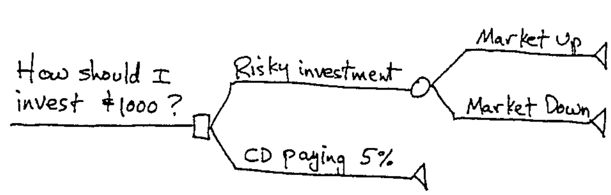
First, you must decide where to invest your money. This decision is represented by a decision node:



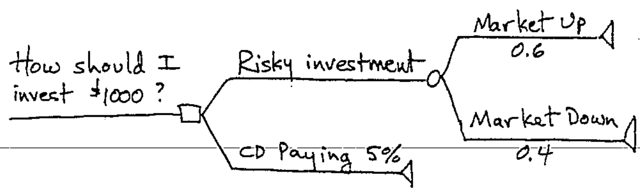
The branches of the decision node must represent all available options. The options are *Risky investment* and *CD paying 5%*. The branch *CD paying 5%* is a final outcome (neither an uncertainty nor a new decision), so it is represented by a triangular terminal node. However, the option *Risky investment* requires consideration of an uncertain event: which way the market will go. So *Risky investment* is followed in your tree by a circular chance node.



The branches of *Risky Investment* must represent all possible outcomes. In this example, the outcomes are *Market up* and *Market down*. They are both final outcomes, so they are represented by terminal nodes.

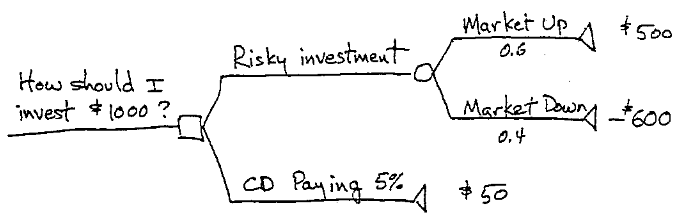


Now the structure of the tree is complete. All that remains is to place the values in the tree. There are two types of values: probabilities and payoffs. Probabilities are assigned to the branches emanating from each chance node, and payoffs are assigned at every terminal node. The



probabilities at the branches emanating from a chance node must sum to 1.0 (100%), as they are *conditional* probabilities. That is, a probability is assigned to a particular chance outcome under the assumption that the events to its left have already occurred. Probabilities are drawn below the branch line of the event they represent.

Payoffs are drawn to the right of the terminal node. Note that a payoff value is assigned to a terminal outcome under the assumption that the outcome is reached; hence, no consideration of probability values is necessary. Here is your complete tree, with all values displayed:



Look at the tree above to be sure you understand it. The decision node *How should I invest $1,000?* is at the root of the tree. You have two choices at that node: *Risky investment* and *CD paying 5%*. The former is a risky event (you can't be sure what will happen to your money), so it is a chance node; the latter involves no uncertainty, so it is represented as a final outcome. The branches of the chance node represent the two possible scenarios as they unfold. Each final outcome has a "payoff" value associated with it.

To calculate a decision tree, one works backward, from right to left. Therefore, calculating a tree is often referred to as "folding back" or "rolling back" the tree. The value of each node is calculated as follows:

- The value of a decision node is equal to the value of its best option.
- The value of a terminal node is equal to the value of its payoff.
- The value of a chance node is equal to its *expected value*, which is found by weighting the values of each of its branches by their respective probabilities.

Applying these rules to the tree, the value of each of the three terminal nodes is already displayed. Now you work leftward from the rightmost branches, to the node *Risky investment.* You can find its expected value by the following calculation:

expected value of *Risky investment* = (500 * 0.6) + (-600 * 0.4)

where 500 is the value (payoff) of *Market up*, and 0.6 is its probability; similarly, -600 is the value (payoff) of *Market down*, and 0.4 is its probability.

Finishing up the calculation, you have

expected value of *Risky investment* = 300 - 240

expected value of *Risky investment* = 60.

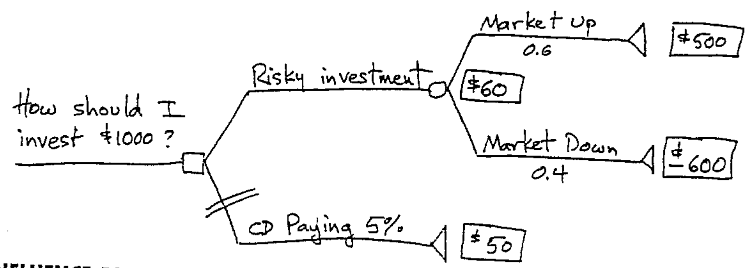
Note that this expression does *not* indicate that if you buy stock you will earn $60, which is clearly untrue, since the model allows for only a gain of $500 or a loss of $600. It simply means that $60 would be your average profit if you were to make the same investment many, many times. This distinction is critical to your understanding of decision analysis.

The expected value of an uncertainty is a probabilistic calculation, making it possible to compare one uncertainty with another, or an uncertainty against a certain outcome.

Let's continue to calculate the tree. All that remains is to calculate the value of the root decision node, and to decide which option to take. As indicated above, the expected value of a decision node is equal to the value of its best option. The value of *CD paying 5%* is $50, and the value of *Risky Investment* is $60. By this calculation, the better option is to buy stock, so the value of the node *How should I invest $1,000?* is $60 also.

Again, be sure you understand the meaning of the $60 expected value calculation. It does not mean that if you follow the recommended strategy, you will earn $60. It means that the *expected value* (which is a mathematical construct, not necessarily a possible outcome) of the investment is $60 if you follow the recommended strategy.

At the top of the next page is a version of your tree with these values included. When the tree is rolled back, the value (or expected value) of each node is typically drawn in a box to the right of the node.



INFLUENCE DIAGRAMS

Influence diagrams tend to be simpler on their face than decision trees. While they are less effective at presenting all of the underlying facts at once, they portray more clearly the factors that influence a decision, and how those factors are related.

TIP: See Chapter 20 for more information on the advantages of influence diagrams.

Even in complex problems, where the decision tree is far too large to fit on a single printed page, the associated influence diagram is almost certain to be small enough for simple reproduction and efficient communication. Moreover, influence diagrams can make it easier to undertake certain calculations, such as Bayes' revision and expected value of perfect information, which depend on considerations of influence.

The design of an influence diagram is subject to guidelines:

1. *All factors influencing the decision must be represented.* Each relevant factor or variable influencing the decision should be represented by a node in the diagram.

2. *Several types of "nodes" may be used.* In general, a *node* represents a decision, a variable, or an objective. A *decision* node (square) is used to indicate a decision facing the decision maker. A *chance* node (circle) is used to represent a variable (or event) whose value (or outcome) is uncertain. A *value* node (diamond) denotes a quantity that measures the desirability of any final outcome.

3 *Nodes may be connected by arcs.* An arc drawn from one node to another indicates (a) that the first node influences (conditions) the second node, and/or (b) timing. (See below for more on influence.)

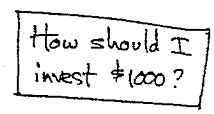
4. *Time flows along the arc lines.* In general, a node which is conditioned on the outcome of another node must occur later in time. There are a few important caveats to this rule which

are discussed in Chapter 20. It is important to note that while arcs generally indicate timing, their primary purpose is the representation of influence.

The "influence" in an influence diagram is, as mentioned above, represented by arcs. If, for instance, the probabilities associated with an uncertain event are different depending on your prior decision, there will be an arc indicating this influence in the diagram. The meaning of arcs and of the influence they represent will be discussed more thoroughly later.

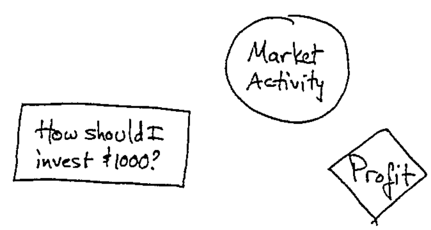
In accordance with these guidelines, let's design an influence diagram that, like the decision tree above, represents the investment problem posed at the end of Chapter 1.

The first step is to draw a node that represents the decision:

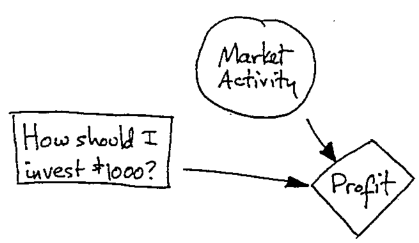


Next, you must consider what uncertainties may influence the final outcome. In this example, there is only one: the activity of the market. Thus, a chance node is added:

The objective by which success is to be measured is the return on the investment, or profit. Thus, a value node, called *Profit*, is added:



Next, we draw in the arcs that denote the influences between nodes. Profit will be affected by your investment decision and by the activity of the market, so two arcs must be drawn:



While the influence diagram is now complete on its face, there is a substantial amount of information missing. For example, the different market states, the probabilities and values associated with those states, and even the specific decision alternatives available to you are not pictured. You will learn to enter these data in the tutorial on influence diagrams in Chapter 4. For the moment, however, it is sufficient to note that the diagram reflects the contours of the decision problem.