Chapter 23 Page 1

Decision Analysis

Consider a situation in which several decisions are to made sequentially. To add a degree of complication, suppose that the results of some of the decisions are not deterministic but rather are governed by known probability distributions. A typical illustration, as described later in the chapter, might concern a repairperson searching for the source of a failure in a computer. She has available a number of tests to isolate the source but before any are performed she only knows the probability distribution of their results rather than their exact outcomes. Moreover, the tests and the repair activity all involve costs. The problem is to determine a conditional strategy for the repair process that minimize the total expected cost as a function of the observed results. The conditional strategy is in the form of an optimal policy that specifies a course of action, such as which test to run and whether to abandon the testing activity, at each decision point that might arise during the failure isolation and repair process.

This is a problem perfect for decision analysis, the subject of this chapter. Similar decision problems naturally arise in parlor games, construction projects, and formation of battle strategies, to name a few. It makes some difference whether or not the forces affecting the chance events are competitive. In the competitive case, the procedures of game theory, discussed in the next chapter, are appropriate. If the forces are benign, then the methods of decision analysis are available. Here we will call the forces "nature" and assume that nature does not operate in a competitive or vindictive fashion, but rather according to probabilistic laws that can be estimated by experimentation or expert judgment. Given enough data, decision analysis allows one to determine an optimal policy.

23.1 Prototype Example

Decision analysis first requires a modeling step and then a solution step. As usual, the modeling step is critical and will be illustrated with an example that encompasses most of the interesting characteristics of a decision analysis problem.

An economically strapped student has just brought his failed computer to a local repair shop. The technician, being a friend of the student, would like to fix it with the least expenditure of money. She guesses that one of four parts has probably failed. They are A, B, C, and D, with the cost of replacement \$100, \$200, \$30, and \$80, respectively. We use the letters A through D to represent the events that the failure is caused by the corresponding part, and the letter E to be the event that something other than the four parts is causing the problem.

Three tests, X, Y and Z, are available to help locate the failed part and can be performed at a cost of \$50, \$70, and \$80, respectively. If E is the result or if the technician decides to abandon testing, the motherboard can be replaced at a cost of \$500. This is guaranteed to fix the problem.

Before any testing is done we estimate from past experience the probability that each of the five events is causing the failure. These are called the *prior* probabilities. The events are mutually exclusive and their probabilities sum to 1.

$$
P(A) = 0.15
$$
, $P(B) = 0.2$, $P(C) = 0.3$, $P(D) = 0.25$, $P(E) = 0.1$

Decision Node

When the technician first gets the machine she can repair it by replacing the motherboard at a cost of \$500 or use test X to get a better idea what's wrong. This represents the first decision in the process, whether to perform the test. In Fig. 1 we begin to construct the *decision tree*. The decision tree is a graphical description of a sequential decision process and constitutes the major part of the model.

The first decision is indicated by the rectangular node labeled 1. The symbol *D1* interior to the node identifies the decision. Two branches leave the node indicating the two possible decisions available at this point, either replace the motherboard or perform the X test. The labels on the branches, 1 and 2, correspond to these two possibilities. The cost associated with the test (50) is shown adjacent to the branch entering node 3.

Figure 1. The first decision

Terminal Node

On the figure the decision to replace the motherboard leads to a *terminal node*, shown as a black circle, labeled node 2. At terminal nodes the process stops, and the cost associated with the terminal state can be evaluated. The number adjacent to the node (500) is the cost associated with reaching this node, that is, the cost of replacing the motherboard.

Chance Node

The decision to perform test X leads to node 3. This is a *chance node* because the result of the test is uncertain. This kind of node is shown as a white circle.

Test X can indicate one of three possibilities: x_1 (failure probably due to A or B), x_2 (failure probably due to C or D), or x_3 (an indication to replace the motherboard). Figures 2 shows the three possible events associated with the experiment as branches leaving the chance node. The numbers on those branches are the probabilities of the three test results: $P(x_1)$, $P(x_2)$ and $P(x_3)$. The events are mutually exclusive and constitute the entire range of possibilities, so their probabilities sum to 1. The branches terminate at nodes that either represent additional decisions or terminal events.

Figure 2. The chance node showing the possible results of test X

Conditional Probabilities

Often in cases of tests, we are not given the probabilities of the various outcomes directly. Rather, we must estimate them using controlled experiments from which we can approximate the conditional probabilities of the results given the cause of the failure.

Say, for the example, we are given $P(x_i/Y)$, the probability of test result *xj* given that event *Y* has caused the failure. Here, *Y* is one of our supposed causes indicated by events *A* through E. These are called the conditional probabilities and are given in Table 1.

What we need for the decision tree are the unconditional probabilities of the results of the test: $P(x_1)$, $P(x_2)$, and $P(x_3)$. These are also called the *marginal probabilities.* The marginal probability for test result x_i is computed by adding up the probabilities of the joint events whose union is the event x_i . Each joint probability term is a product of the known conditional probability and a prior probability.

$$
P(x_i \quad Y) = P(x_i | Y)P(Y).
$$

For x_1 :

$$
P(x_1) = P(x_1 \quad A) + P(x_1 \quad B) + P(x_1 \quad C) + P(x_1 \quad D) + P(x_1 \quad E).
$$

The required marginal probability is computed from the known quantities.

$$
P(x_1) = P(x_1|A)P(A) + \dots + P(x_1|E)P(E)
$$

= (0.6)(0.15) + (0.7)(0.2) + (0.15)(0.3) + (0.05)(0.25) + (0.1)(0.1)
= 0.2975.

Similarly, $P(x_2) = 0.4775$ and $P(x_3) = 0.225$. These probabilities are seen leaving the chance node in Fig. 2.

A Second Round of Decisions

If test X indicates x_1 , the technician can either use test Y or replace the motherboard. If the test indicates x_2 , she can use test Z or replace the motherboard. If the test indicates x_3 , she must replace the motherboard.

Figure 3 shows the details associated with a second round of decisions. If test X indicates x_1 , the technician can either continue the fault isolation process with test Y or immediately replace the motherboard. This is shown as decision $D2$. The cost of test Y is \$70. If test X indicates x_2 she can continue the isolation with test Z or replace the motherboard. This is decision $D3$. The cost of test Z is \$80. In both cases, the motherboard replacement costs \$500.

Figure 3. The second round of tests

Posterior Probabilities

We assume for the example that test Y can accurately identify the cause of failure if it is due to components A or B. If either of these indications are observed, the faulty component is repaired. Every other cause (C, D, or E) is grouped into a third category. If the test does not indicate a failure of A or B, the motherboard is replaced.

Similarly test Z accurately identifies the problem if it is due to C or D. If the test does not indicate C or D, the fault must be in A, B, or E. Rather than continue testing, the motherboard is replaced. The results of the tests together with the appropriate probabilities are shown in Fig. 4.

Figure 4. Results of the tests together with posterior probabilities

The probability of an event *Y* yielding the result A is the probability that event A is true given that test X returns x_1 *and* test Y indicates A. Because test Y is perfectly reliable at identifying cause A, the probability of the test indicating A is just $P(A | x_1)$. This is called the posterior probability and can be computed with Bayes' law. In general, the posterior probability of event *Y* given that test *X* indicates x_i is:

$$
P(Y \mid x_i) = \frac{P(x_i \mid Y)}{P(x_i)}.
$$

For events A and x_1 this is

$$
P(A \mid x_1) = \frac{P(x_1 \cdot A)}{P(x_1)} = \frac{(0.6)(0.15)}{0.2975} = 0.3025.
$$

The complete set of posterior probabilities is shown in Table 2.

We see in Fig. 4, $P(A | x_1)$ and $P(B | x_1)$ on the branches leaving the chance node *Y* and terminating at the events for replacing components A and B, respectively. The probability for the remaining branch is the sum of the remaining probabilities in the row for x_1 . Similarly, we see $P(C | x_1)$ and $P(D | x_1)$ on the branches leaving the chance node *Z* and terminating at the events for replacing components C and D, respectively. The probability for the remaining branch is the sum of the remaining probabilities in the row for x_2 . Because no test is performed if the outcome is x_3 , the numbers in the last row of Table 2 of the posterior probabilities are not used for this example.

Decision Tree

The *decision tree* is the combination of components that we have described and is the graphical description of a sequential process. The complete tree for the technician's decision problem is shown in Fig. 5.

Figure 5. Complete decision tree

23.2 The General Decision Tree Model

 The decision tree is made up of nodes and branches. The decision nodes are shown as squares, chance nodes as white circles, and terminal nodes as black circles. Branches, the lines connecting the nodes, represent specific decisions and outcomes of chance events. We draw the tree from left to right. Branches leave nodes going to the right and enter nodes coming from the left. In the numbering convention used, every node has an unique index which is smaller than any of its successor nodes. Node 1, which precedes all other nodes, is called the root node. We now formalize and generalize some of the definitions described in for the example.

Decision Node

The decision node represents a choice that must be made in the sequential process. The capital letter in bold italics within the square node is the name of the set of possible decisions. The branches that leave the node describe the individual decisions in the set. In the figure, the branches are numbered, 1, 2, ... The associated decision is identified with a lowercase letter with the number used as an index. For instance, the decision nodes in the example problem and the corresponding decision sets are

$$
DI = \{d_{11}, d_{21}, d_{31}\}, D2 = \{d_{12}, d_{22}\} \text{ and } D3 = \{d_{13}, d_{23}\}
$$

where DI is the decision whether to use test X , $D2$ is the decision to use test Y, and *D3* is the decision to use test Z.

A number in parentheses adjacent to a branch shows the cost associated with the decision. For the example, the numbers are the costs of performing the tests. In general, we use *C* (*j*) as the cost associated with decision *j* .

Chance Node

The circular nodes represent a chance event that may occur during the sequential process. The upper case letter in bold italics within the circle is the name of the set of possible outcomes of the event. The branches that leave the node describe the individual outcomes in the set. The outcome associated with a branch is identified with a lowercase letter, with the number adjacent to the branch used as an index. For instance, the chance nodes and corresponding outcome sets in Fig. 5 are

$$
\mathbf{X} = \{x_1, x_2, x_3\}, \mathbf{Y} = \{y_1, y_2, y_3\} \text{ and } \mathbf{Z} = \{z_1, z_2, z_3\}.
$$

In the example, the outcomes are the results of the tests.

A number must be associated with each branch leaving a chance node that gives the conditional probability of that particular outcome, given all the decisions and chance outcomes on the path from the root node to the chance node. In general, we show these probabilities in parentheses adjacent to the outcome branch. Because of space limitations in Fig. 5, the probabilities are shown to the right of the terminal nodes.

A number can also be associated with a branch leaving a chance node showing the cost associated with that outcome. Again we use the notation $C(j)$ to specify this cost of outcome j . If this number is not zero, it appears as the second number in the parentheses adjacent to an outcome

branch. For the example, only the outcomes leading to terminal nodes have these costs. They are shown in the parentheses to the right of the terminal nodes.

Conditional Probabilities

Every branch leaving a chance node must be assigned a number which is the conditional probability of that particular outcome given all the decisions and chance outcomes on the path from the root node to the chance node. This probability is written as $\overline{P}(i | \text{Path})$, or the probability of outcome *j*, given the path to the chance node. In the example, this should be apparent for the branches leaving nodes *Y* and *Z*, where the probabilities are conditioned on the results of test *X*.

We use the notation above to indicate this probability with "Path" designating the intersection (the symbol means intersection) of all decisions and chance outcomes on the path from the root node. For instance, for outcome y_1 leaving node 6 in Fig. 5, we can write the conditional probability as

 $P(y_1 | d_{12} \quad x_1 \quad d_{22}) = P(y_1 | x_1) = 0.303.$

The probability of the result *y* 1 depends on the fact that we performed tests *X* and *Y* as well as on the outcome of test X. The probabilities are the posterior probabilities described above.

Decision Function

The decision function specifies the optimal course of action to be taken at each decision node. In Fig. 5, the decision nodes are identified by the indices 1,4, and 5. A decision function must specify a decision for each of these nodes. The criterion for optimality is that all decisions should be made to minimize the expected loss at every decision node.

23.3 Solving the Decision Tree

A solution to the problem is a specification of the optimal decision at each decision node. We call this the *optimal policy*. In general, it is given by $d^*(j)$, where *j* is the index of a decision node and $d^*(j)$ is an index of one of the decision branches leaving node *j*. We derive the optimal policy by assuming that all decisions should be made to minimize the expected loss at each decision node. This is called the *Bayes' criterion*.

The *rollback* procedure is generally used to solve the decision tree. We work backwards starting from the nodes farthest to the right sequentially evaluating the expected loss for each chance node and the minimum expected loss at each decision node. For the latter, we must evaluate the expected loss for each possible decision and select the one yielding the minimum. In the following, we present the general procedure using the example to illustrate the computations.

The Expected Loss at a Chance Node

Let M represent a general chance node with index i , and assume that there are *K* outcomes leaving the node $(m_1, m_2,...,m_K)$. Let $t(m_k)$ be the index of the node terminating the branch representing outcome m_k , $C(m_k)$ be the cost for the outcome, and $P(m_k | \text{Path to } M)$ be the conditional probability associated with the outcome. Then the expected loss at chance node *i* is *f*(*i*), where

$$
f(i) = \frac{K}{k} P(m_k | \text{Path to } M)(C(m_k) + f(t(m_k))). \qquad (1)
$$

In the example, both the chance nodes *Y* and *Z* are at the extreme right of the decision tree. The expected cost given that test Y is performed is $f(6)$, and the expected cost given that test \overline{Z} is performed is $f(\overline{7})$. Computing, we find

$$
f(6) = E [\text{loss at } Y] = \frac{3}{P(y_k | x_1)(C(y_k) + f(t(m_k)))}
$$

\n
$$
k=1
$$

\n
$$
= (0.303)(100) + (0.471)(200) + (0.226)(500) = 237.8
$$

\n
$$
f(7) = E [\text{loss at } Z] = \frac{3}{P(z_k | x_2)(C(z_k) + f(t(m_k)))}
$$

\n
$$
k=1
$$

\n
$$
= (0.346)(30) + (0.445)(80) + (0.209)(500) = 150.7
$$

The Minimum Expected Loss at a Decision Node

Let D represent a general decision node with index j , and assume that there are *J* decisions leaving the node, d_1 , d_2 ,..., d_J . Let $t(d_j)$ be the index of the node terminating the branch representing decision d_j and $C(d_j)$ be the cost for the decision. Then the minimum expected loss at the decision node is

$$
f(j) = \text{Minimum } \{ C(d_j) + f(t(d_j)) \mid d_j \quad J \}. \tag{2}
$$

The optimal decision is the index that obtains the minimum, i.e., $d^*(j)$.

For the example, we illustrate the computations for decision nodes *D2* and *D3* (nodes 4 and 5).

$$
f(4) = \text{Minimum } \{C(d_j) + f(t(d_j)) \mid d_j \quad D2 \}
$$

= \text{Minimum } \{ 500, 70 + 237.8 \} = 307.8; $d^*(4)=2$

$$
f(5) = \text{Minimum } \{ C(d_j) + f(t(d_j)) \mid d_j \quad D3 \}
$$

= \text{Minimum } \{ 500, 80 + 104.7 \} = 184.7; $d^*(5) = 2$

In both cases the optimal decision is to test.

Algorithm

The procedures indicated by Eqs. (1) and (2) suggest a simple algorithm to determine the optimal course of action at each decision node, and hence the optimal solution for the problem. It is similar to a dynamic programming algorithm in that it starts at the end of the tree and works backwards toward the root. It is a one-pass algorithm since it visits each node only once.

- *Step* 1. Choose the node with the largest index. Let this be node *i* . (Recall that the nodes are numbered so that every node has an index smaller than any of its successor.)
- *Step* 2. If node *i* is a terminal node go to step 3.
	- If node *i* is a chance node, use Eq. (1) to evaluate its expected loss and go to Step 3.
	- If node *i* is a decision node, use Eq. (2) to evaluate its minimum expected loss. Record the optimal decision as part of the optimal decision function and go to Step 3.
- *Step* 3*.* Decrease *i* by 1. If *i* is greater than 0, go to Step 2. Otherwise, stop with the optimal decision function.

Example

Applying the algorithm to the decision tree in Fig. 5, we get the results listed in Table 3. Only decision and chance nodes are shown because no action is required at terminal nodes.

| Node | Name | Type | Expected loss | Optimal decision |
|------|-----------------|----------|------------------|-----------------------|
| | Z | Chance | 150.7 | |
| 6 | Y | Chance | 237.8 | |
| 5 | $\overline{D}3$ | Decision | 184.7 | \mathcal{D}_{\cdot} |
| 4 | $\overline{D2}$ | Decision | 307.8 | 2 |
| 3 | X | Chance | 292.3 | |
| | D1 | Decision | 342.3 | |

Table 3. Algorithmic Results for Decision Tree in Figure 5

The optimal policy is

$$
d^*(1) = 2 \qquad \text{do test X}
$$

d * $(4) = 2$ if the result of test X is x_1 do test Y

d * $(5) = 2$ if the result of test X is x_2 do test Z

The expected cost of the repair process is \$342.30. This is the minimum expected cost solution.

23.4 Discussion and Assessment

The unique feature of decision trees is that they allow management to view the logical order of a sequence of decisions. They afford a clear graphical presentation of the various alternative courses of action and their possible consequences. By using decision trees, management can also examine the impact of a series of decisions (over many periods) on the objectives of the organization. Such models reduce abstract thinking to a rational visual pattern of cause and effect. When costs and benefits are associated with each branch and probabilities are estimated for each possible outcome, analysis of the flow network can clarify choices and risks.

On the down side, the methodology has several weaknesses that should not be overlooked. A basic limitation of its representational properties is that only small and relatively simple models can be shown at the level of detail that makes trees so descriptive. Every variable added expands the tree's size multiplicatively. Although this problem can be overcome to some extent by generalizing the diagram, significant information may be lost in doing so. This loss is particularly acute if the problem structure is highly dependent or asymmetric.

Regarding the computational properties of trees, for simple problems in which the endpoints are precalculated or assessed directly, the rollback procedure is very efficient. However, for problems that require a roll-forward procedure, the classic tree-based algorithm has a fundamental drawback: it is essentially an enumeration technique. That is, every path through the tree is traversed in order to solve the problem and generate the full range of outputs. This feature raises the "curse of dimensionality" common to many stochastic models: for every variable added, the computational requirements increase multiplicatively. This implies that the number of chance variables that can be included in the model tends to be small. There is also a strong incentive to simplify the value model, since it is recalculated at the end of each path through the tree.

Nevertheless, the enumeration property of tree-based algorithms in theory can be reduced dramatically by taking advantage of certain structural properties of a problem. Two such properties are referred to as "asymmetry" and "coalescence." For more discussion and some practical aspects of implementation, consult Call and Miller (1990).

23.5 Exercises

1. A bank manager wants to develop a decision policy to assist the bank's credit officers. Consider a situation involving a \$20,000 loan. The following information has been gathered from historical records regarding the percentage of customers in various income categories who seek such a loan.

There are several steps in the decision process for each customer. The officer first decides whether to reject the loan application, accept the application, or call for a credit check. Whenever a customer is rejected there is a loss of good will, which can be reflected as a cost that depends on the class of customer. When the application is accepted, there is a probability that the customer will default and the loan principal lost. If the customer does not default, the gain is the interest income from the loan. All this information is shown in the following table.

The cost of a credit check is \$300. The results of the check are "good risk" or "bad risk." The good risk customer has a probability of default 0.05 less than the entries in the table above. The bad risk customer has a probability of default 0.05 more than the entries in the table above. After the credit check, the officer must accept or reject the loan application. When a customer is rejected, assume that the alternative investment for the \$20,000 has a \$1500 return. Show the decision tree for this situation and find the optimal decision policy.

2. You own a stretch of land in western Texas at which there is the possibility of discovering oil. Your options are as follows.

*a*1 , drill for oil yourself

*a*2 , lease the site to someone else to drill

*a*3 , lease the site but maintain an interest in the results

There are four possible outcomes regarding the success of the well.

*q*1 , 600,000 barrel well

*q*2, 400,000 barrel well

*q*3 , 100,000 barrel well

 q_4 , dry hole.

Your profit (\$1000) depends on which development option you choose and the results of the drilling, as shown in the following table.

Based on information about the site, you have estimated the prior probabilities of the four results to be

$$
P(q_1) = 0.1
$$
, $P(q_2) = 0.15$, $P(q_3) = 0.25$, $P(q_4) = 0.5$.

- a. Based on the prior probabilities, which development option should you choose?
- b. Rather than making a firm decision among the three options, you now have a fourth option of running a seismic test to get more information on the likelihood of success. From your experience in drilling in this area, you have estimated the conditional probabilities of well success given the test result. These values are denoted by $P(x_i | q_j)$, and appear in the following table.

Using Bayes' law, compute the posterior probabilities $P(q_j | x_i)$.

- c. The cost of the seismic test is \$20,000. Perform a decision analysis to determine whether or not to do the test. Also determine which of the development options to pursue based on the results of the test. What is the profit associated with the optimal decisions?
- 3. A new oil field has been discovered with different prior probabilities of well success than those identified in Exercise 2. These values are listed below. Using the original conditional probabilities for the results of the seismic test, find the optimal policy.

$$
P(q_1) = 0.20
$$
, $P(q_2) = 0.20$, $P(q_3) = 0.30$, $P(q_4) = 0.30$.

4. Use the original statement of the problem given in Exercise 2 but let the conditional leasing arrangement change. The returns for conditional lease are \$600,000 for a 600,000 barrel well, \$300,000 for a 400,000 barrel well, and no return for a 100,000 barrel well or a dry well. Find the optimal policy.

- 5. The daily demand for a particular type of printed circuit board in an assembly shop can assume one of the following values: 100, 120, or 130 with probabilities 0.2, 0.3, and 0.5. The manager of the shop is thus limiting her alternatives to stocking one of the three levels indicated. If she prepares more boards than are needed in the same day, she must reprocess those remaining at a cost price of 55 cents/board. Assuming that it costs 60 cents to prepare a board for assembly and that each board produces a revenue of \$1.05, find the optimal stocking level by using a decision tree model.
- 6. In Exercise 5, suppose that the owner wishes to consider her decision problem over a 2-day period. Her alternatives for the second day are determined as follows. If the demand in day 1 is equal to the amount stocked, she will continue to order the same quantity on the second day. Otherwise, if the demand exceeds the amount stocked, she will have the options to order higher levels of stock on the second day. Finally, if day 1's demand is less than the amount stocked, she will have the options to order any of the lower levels of stock for the second day. Express the problem as a decision tree and find the optimal solution using the cost data given in Exercise 5.
- 7. Zingtronics Corp. has completed the design of a new graphic-display unit for computer systems and is about to decide on whether it should produce one of the major components internally or subcontract it to another local firm. The advisability of which action to take depends on how the market will respond to the new product. If demand is high, it is worthwhile to make the extra investment for special facilities and equipment needed to product the component internally. For low demand it is preferable to subcontract. The analyst assigned to study the problem has produced the following information on costs (in thousands of dollars) and probability estimates of future demand for the next 5-year period:

- a. Prepare a decision tree that describes the structure of this problem.
- b. Select the best action based on the initial probability estimates for future demand.
- c. Determine the expected cost with perfect information (i.e., knowing future demand exactly).
- 8. Refer to Exercise 7. The management of Zingtronics is planning to hire Dr. Sam Fenichel, an economist, to prepare an economic forecast for the computer industry. The reliability of his forecasts from previous assignments is provided by the following table of conditional probabilities:

- a. Select the best action for Zingtronics if Dr. Fenichel submits a pessimistic forecast for the computer industry.
- b. Prepare a decision tree diagram for the problem with the use of Dr. Fenichel's forecasts.
- c. What is the Bayes strategy for this problem?
- d. Determine the maximum fee that should be paid for the use of Dr. Fenichel's services.
- 9. Allen Konigsberg is an expert in decision support systems and has been hired by a small software engineering firm to help plan their R&D strategy for the next 6 to 12 months. The company wishes to devote up to 3 person-years or roughly \$200,000 to R&D projects. Show how Konigsberg can use a decision tree to structure his analysis. State all your assumptions.

Bibliography

J.H. Call and W.A. Miller, "A Comparison of Approaches and Implementations for Automating Decision Analysis," *Reliability Engineering and System Safety*, Vol. 30, pp. 115-162 (1990).

R.T. Clemen, *Making Hard Decisions*, Second Edition, Duxbury, Belmont, MA, 1996.

S. French, *Decision Theory: An Introduction to the Mathematics of Rationality*, John Wiley & Sons, New York, 1988.

P. Goodwin and G. Wright, *Decision Analysis for Management Judgment*, Second Edition, John Wiley & Sons, New York, 2000.

R.L. Keeney and H. Raiffa, *Decisions with Multiple Objectives*, Cambridge University Press, Cambridge, UK. 1993.