Decision Trees

Decision trees are utilized to address discrete decision problems. In other words, the problem is to determine the best of several possible decisions. What makes the choice difficult is that the outcome in the aftermath of the decision is uncertain. The decision alternatives are called actions and denoted A_i. Each action is associated with a cost, $c(A_i)$. The possible outcomes are denoted θ_j and each outcome is associated with a cost, $c(\theta_j|A_i)$, and a probability of occurrence, $P(\theta_j)$. Before drawing the decision tree it is necessary to:

1. Enumerate the decision alternatives, A_i
2. Compute the cost of each decision alternative, $c(A_i)$
3. Enumerate the possible outcomes, θ_j
4. Assess the cost of each outcome, $c(\theta_j|A_i)$; often organized in a "payoff table" that often includes $c(A_i)$
5. Compute the probability of each outcome, $P(\theta_j)$

The objective of the decision tree analysis is to identify the action with lowest expected cost. Or conversely, if the cost values are translated into utility values the objective is to identify the action with highest expected utility.

Decision Tree

The setup of a decision tree to determine the expected cost of each action is shown in Figure 1. The starting point is the left-most decision fork, which is drawn as a rectangle. Each action branch ends at a chance fork, which is drawn as a circle. As illustrated in Figure 1, the expected cost of each action branch is

$$E\left[c\middle|A_i\right] = c(A_i) + \sum_{j=1}^{J} c(\theta_j|A_i) \cdot P(\theta_j) \tag{1}$$

where J is the total number of outcomes. The action with the lowest expected cost, or equivalently the one with highest expected utility, is the optimal decision. However, two other decision strategies exist: 1) Minimize the maximum cost, which may be selected by a risk-averse decision maker in one-off situations; 2) Maximize the possible benefit, which may be selected by a gambling-inclined decision maker.

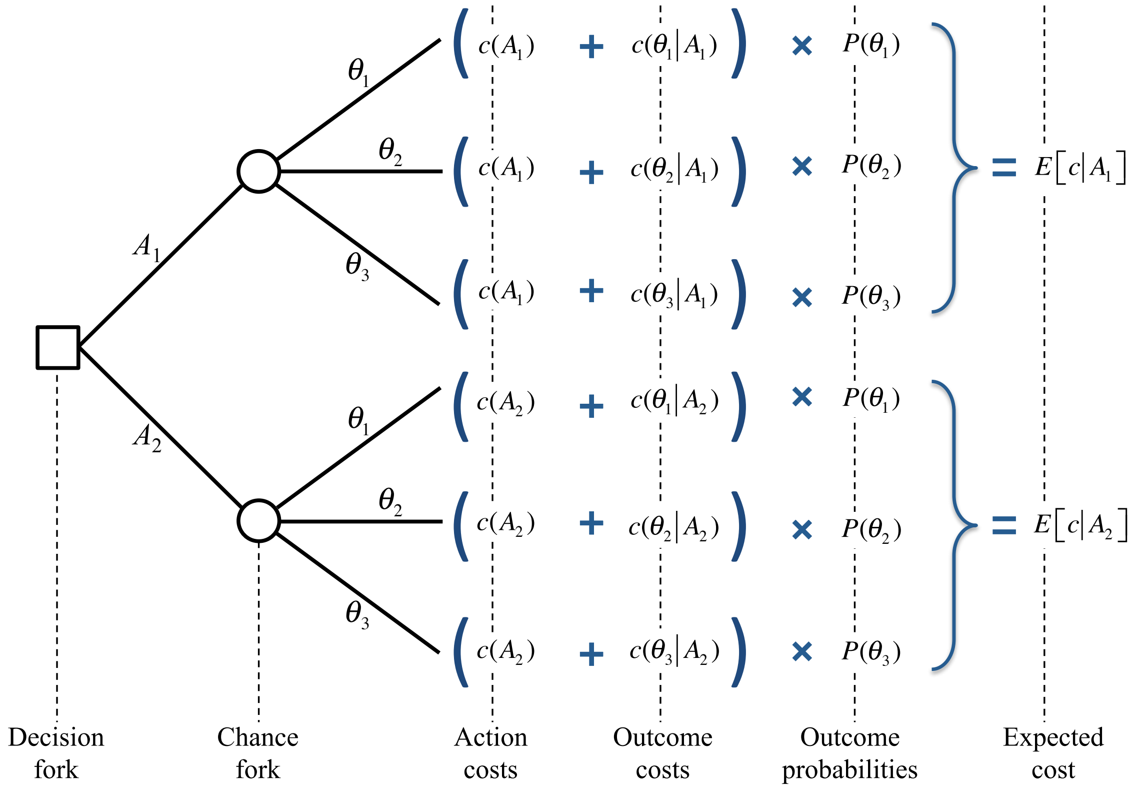

Figure 1: Decision tree.

Terminal Analysis

Terminal analysis extends the decision tree in circumstances when new information becomes available. The information is not conclusive; otherwise the decision would be easy. Rather, the information is associated with some uncertainty. The typical example is test data obtained with an imperfect device, from which the probability of various outcomes is provided as "sample likelihoods," i.e., in the conditional form

$$P\left[I_k|\theta_j\right] \tag{2}$$

where I_k is the indicator value from the test device and θ_j is the real state. For each j, i.e., for each possible real state, the following condition must be satisfied:

$$\sum_{k=1}^{K} P\left[I_k|\theta_j\right] = 1 \tag{3}$$

where K is the number of possible test indicator values. The key step in terminal analysis is to use the test probabilities in Eq. (2) to update the outcome probabilities in the decision tree. I.e., the objective is to compute the probabilities P[$\theta_j|I_k$]. Notice that the outcome probabilities are dependent on the indicator value from the test device; for each indicator value there will be a unique decision tree, as shown in Figure 2. The new indicator-dependent outcome probabilities are computed by Bayes' rule:

$$P\left[\theta_j \middle| I_k\right] = \frac{P\left[I_k \middle| \theta_j\right]}{P\left[I_k\right]} P\left[\theta_j\right] \tag{4}$$

For each indicator branch in Figure 2 a basic decision tree is drawn, with outcome probabilities from Eq. (4). That is, for each indicator value there will be an optimal decision that should be made if that indicator value is observed.

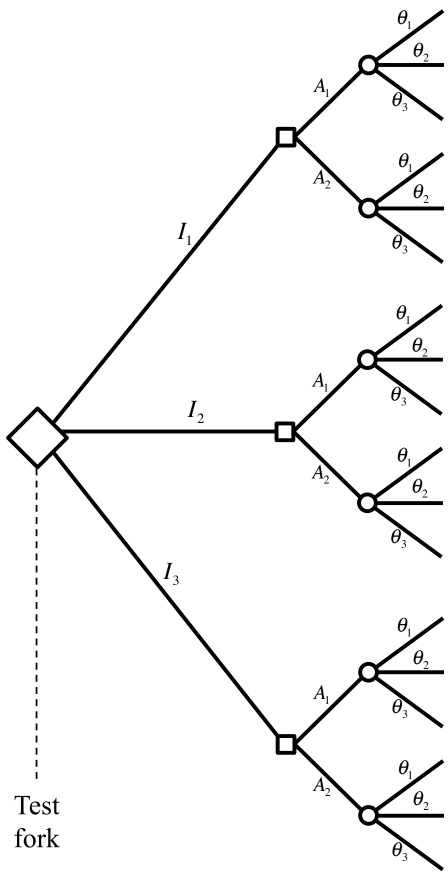

Figure 2: Decision tree for terminal analysis.

Pre-posterior Analysis

Now consider the problem of deciding whether to purchase test information. Two situations arise; the information may be perfect or associated with uncertainty. Consider first the case of perfect test information. The expected cost in a situation where the test removes all uncertainty about the outcome, albeit excluding the cost of the test itself, is

$$E\left[c_{\text{test}}\right] = \sum_{j=1}^{J} \left[\left(c(A_{oj}) + c(\theta_j \middle| A_{oj}) \right) \cdot P(\theta_j) \right] \tag{5}$$

where J is the number of possible outcomes (remember, the test device is perfect) and A_{oj} is the obvious decision once θ_j is the known outcome. Eq. (5) states that the expected cost in a perfect-test situation is the sum of the probability of each possible outcome multiplied by the certain cost of the action that is obvious once that outcome is known.

To determine whether it is cost-effective to purchase the test, the expected cost in Eq. (5) is compared with the expected cost of the optimal decision in the classical decision tree in Figure 1. The expected cost in Eq. (5) is lower and the discrepancy is the saving associated with having perfect test information. It is expected to be cost-effective to purchase the test if the purchase-cost does not exceed this saving.

Another type of pre-posterior analysis arises when deciding whether to pay for imperfect information. This analysis has three steps: First, probabilities for the possible test results (indicator values) are computed by the rule of total probability:

$$P\left[I_k\right] = \sum_{i=1}^{J} P\left[I_k \middle| \theta_j\right] \cdot P\left[\theta_j\right] \tag{6}$$

Each of these probabilities, P[I_k], is multiplied by the expected cost of the optimal decision that follows the test result I_k, which is available from a terminal analysis shown in Figure 2. The complete pre-posterior decision tree for this situation is shown in Figure 3. Summation of these products yields the expected cost in the situation where we have help from an imperfect test device. This value is compared with the expected cost of the optimal decision in the classical decision tree in Figure 1. The discrepancy is the saving associated with performing and imperfect test. It is expected to be cost-effective to purchase the test if the purchase-cost does not exceed this saving.

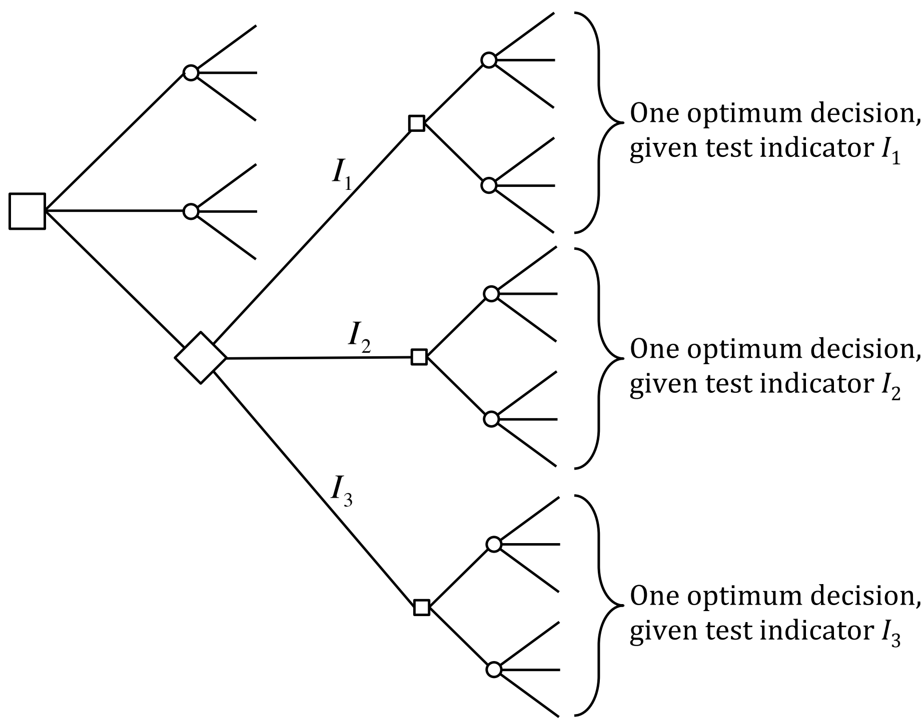

Figure 3: Decision tree for pre-posterior analysis with the option of selecting imperfect test.