Decision-Making under Uncertainty

Connecticut Electronics, in its use of gold-plating in a semiconductor memory module, has had adhesion problems caused by irregularities in electric flow during the plating process. Based on historical data, they know that approximately 70% of the time the current is fairly uniform, in which case the batch of modules they produce (one thousand at a time) will consist of roughly 90% good modules and only 10% defectives. The other 30% of the time, when the current is somewhat irregular, only 60% of the modules are good and 40% are defective.

Unfortunately, there is no way the engineers at CE can determine the uniformity of the current flow. However, they have several alternative ways to handle each batch. One alternative is to send the batch directly to the assembly operation and hope for the best; their records show that when they do this they incur subsequent delay and adjustment costs of about $1000 for each 90%-good batch, and of about $4000 for each 60%-good batch. Another alternative is to replate the entire batch; this ensures that the batch will be sufficiently free from defects that no delay or adjustment costs will occur, but the replating costs $2000.

1. What should the company do? What is the (expected) cost per batch if they follow this strategy?

   Testing a module is an expensive proposition. (A 4096-bit unit has more than 16 million interconnections.) Nevertheless, CE is thinking of testing one unit from each batch. The test will definitely establish whether that (single) unit is good or not.

2. How should the outcome of such a test influence the replating decision? What is the expected value of the information gained from the test? (That is, how much should CE be willing to pay, in order to conduct the test?

A unit from a particular batch is tested, and proves to be good.

3. What is the probability that a second unit from the same batch will also test out to be good?

4. How much should CE be willing to pay to test a second unit, after the first test is positive (i.e., the first item tested is good)? (Note: It is not immediately obvious that the outcome of a second test would affect their decision; if it wouldn't, then they shouldn't be willing to pay anything.)

Decision Trees

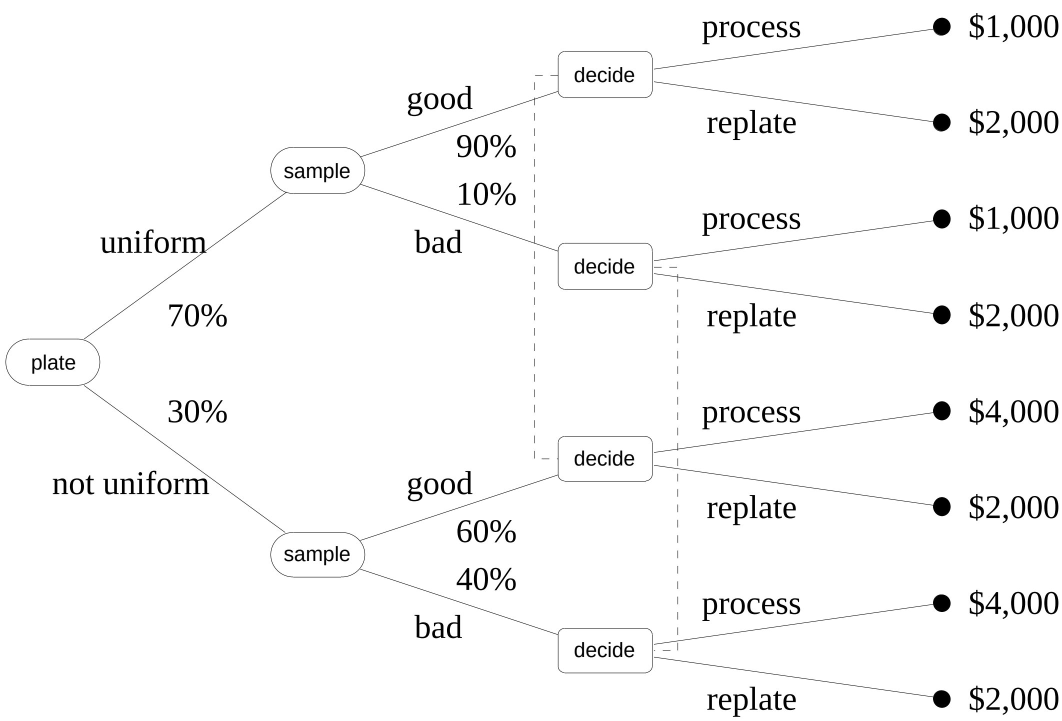
Often, managers must make decisions in settings where relevant information about the state of the world is not available at the time the decisions must be made (either because the information is not directly observable, or because it will not be revealed until some time in the future), but indirect information *is* available. For example, oil drilling decisions are based on the results of prior drilling experience and geological survey data, while the truly-relevant state of the world — the nature of the petroleum deposit — is not known. Marketing managers must decide upon promotional campaigns on the basis of the results of market surveys, without knowing the true state of overall demand. Commercial loan officers must make credit decisions on the basis of balance-sheet information and prior defaults, without knowing whether an applicant is currently credit-worthy. Insurance underwriters must set premiums without knowing the precise risks associated with prospective purchasers. R&D projects must be decided upon before it is known for certain that they will be successful. Boards of directors must select CEO's without knowing precisely how well they will perform. Physicians must make diagnoses and decide upon courses of treatment before knowing the precise nature of a patient's state of health. Firms must decide whether to settle claims out of court, before knowing how a judge or jury would rule.

All of these problems can be systematically analyzed using *decision trees* — graphical structures which help to organize the relevant data. A decision tree consists of (1) a *root node*, which represents the beginning of the decision process, (2) *branches* from the root to subsequent *nodes*, and branches from those nodes, and so on to the (3) terminal nodes (representing possible outcomes of the decision process). Each non-terminal node "belongs" either to the decision-maker (and indicates a point at which a decision is required), or to Nature (and indicates a point where a chance event occurs). Each terminal node has an attached *payoff* (which represents the value of the outcome to the decision-maker, measured on some scale with respect to which the decision-maker's objective is to maximize his or her expected payoff). The branches coming from each chance node carry *conditional probabilities* (showing the probability of each branch being taken by Nature, given that the chance node has been reached). And, finally, the decision nodes are clustered into *information sets*, or *situations*, which are informationally indistinguishable at the time a decision must be made.

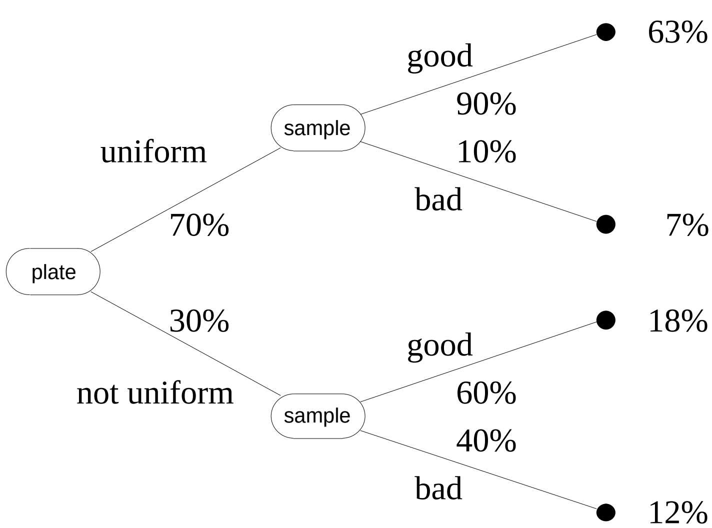
This is all much simpler in practice than it might seem to be in words: Consider the Connecticut Electronics case. The first diagram is a decision tree displaying the problem in *temporal order* — Each move by Nature or the decision-maker is listed before other moves which occur later in time. Notice that a full temporal-order representation of the problem requires the use of nontrivial (multinode) information sets. The last diagram displays the problem in *informational order* — No chance move appears until its outcome becomes observable to the decision-maker.

Connecticut Electronics:

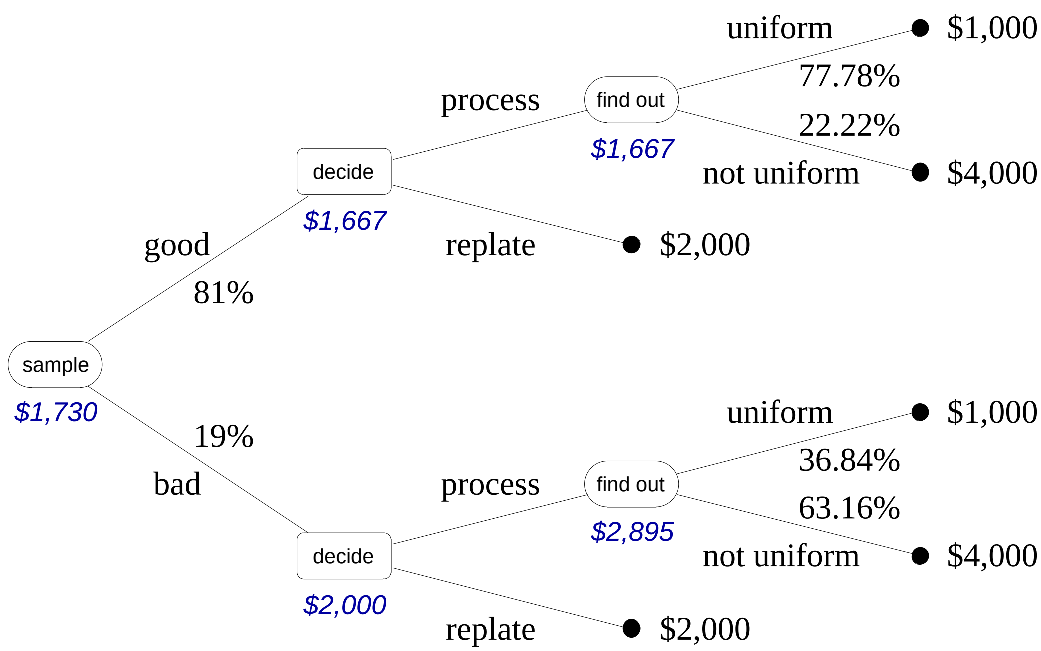
The temporal-order tree



The probability tree



The decision (informational-order) tree



The fundamental principle of decision trees: Typically, the representation of a decision problem in temporal order requires the use of “information sets” which link decision nodes corresponding to the same “situation” (i.e., nodes at which precisely the same information is available at the time an action must be chosen). It is impossible to take systematically different actions at nodes in the same information set.

However, every single-person decision problem can also be represented in informational order, by not listing any moves of “Nature” until they become observable. When written in informational order, decision trees never have multinode information sets. Each decision node is an information set by itself, and the problem can be analyzed by “pruning” the tree from the end.

Analytic simplicity, however, comes at a cost. The conditional probabilities on the branches of an informational-order decision tree typically must be computed using Bayes' Rule, since they are probabilities of prior events having occurred, given subsequent observations. Fortunately, the relevant probability tree is embedded in the temporal-order representation of the problem.

Connecticut Electronics:

(a) Send the batch on to assembly, at an expected (rework) cost of $1900 per batch.

(b) If the tested item is good, send the batch on; if bad, replate the batch: The expected cost per batch is $1730, so it's worth $170 to do the test.

(c) 83.33%

(d) The information gained by testing a second unit, given that the first is "good", is worth $100. After the first unit tests OK, sending the batch on to assembly has an expected rework cost of $1666.67; if a second unit is tested, and the batch is replated if the second unit is "bad", the expected cost per batch (given that the first tested unit was good) is $1566.67.

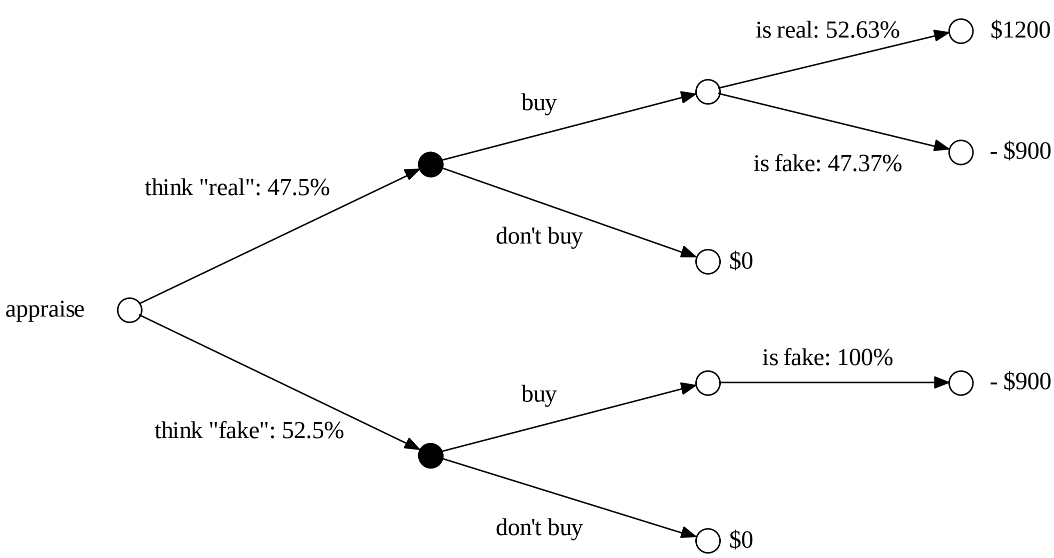
Examples

1. Having recently completed a course in gemology, Mary feels that she never misidentifies real diamonds, but that she is fooled by fakes about 30% of the time. On an Asian holiday, she finds herself in a shop being offered an unset diamond for \$1000. It looks real to her. However, her friends have warned her that 75% of all so-offered stones are fakes. If this *is* a real stone, it will be worth about \$2200 back home; if it's a fake, it will be worth only \$100. Should she purchase the diamond? If she does, what is her expected profit from the deal?

Solution:

$\Pr(\text{real} \mid \text{thinks real}) = \Pr(\text{thinks real} \mid \text{real})\cdot\Pr(\text{real}) / \Pr(\text{thinks real})$, and
$\Pr(\text{thinks real}) = \Pr(\text{thinks real} \mid \text{real})\cdot\Pr(\text{real}) + \Pr(\text{thinks real} \mid \text{fake})\cdot\Pr(\text{fake})$, so
$\Pr(\text{thinks real}) = 1.0\cdot0.25 + 0.3\cdot0.75 = 0.475$, and
$\Pr(\text{real} \mid \text{thinks real}) = 0.25/0.475 = 10/19$.

So, $E[\text{profit from buying} \mid \text{thinks real}] = E[\text{profit} \mid \text{real, thinks real}]\cdot\Pr(\text{real} \mid \text{thinks real})$
$+ E[\text{profit} \mid \text{fake, thinks real}]\cdot\Pr(\text{fake} \mid \text{thinks real})$
$= 1200\cdot10/19 - 900\cdot9/19 = \$205.26$



$E[\text{ profit }] = 0.475 \cdot \$205.26 + 0.525 \cdot \$0.00 = \$97.50$

Value of information from taking course = [\$97.50] – [\$0.00] = \$97.50.

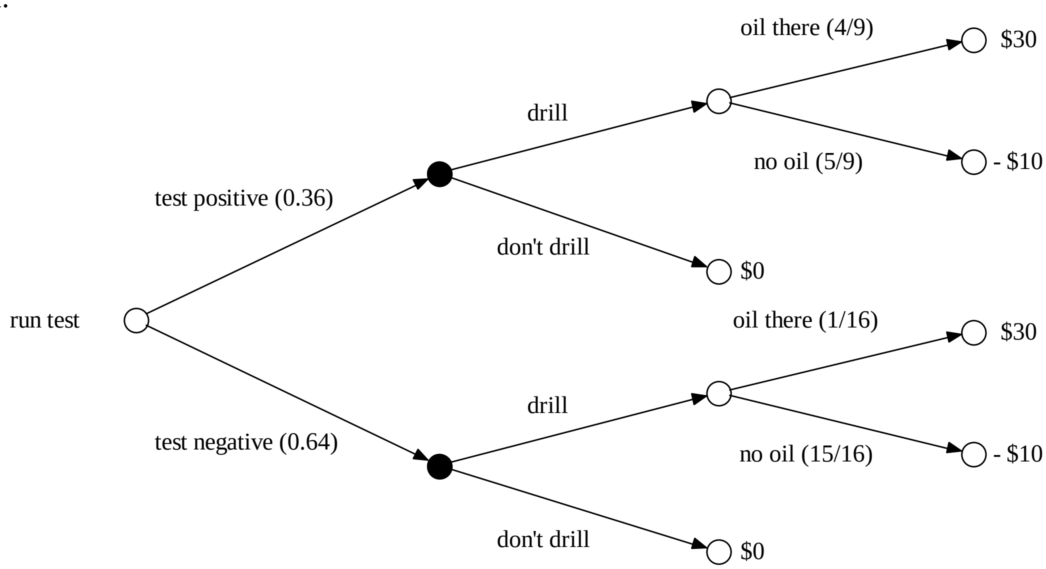
(Without doing the appraisal, she wouldn't buy; with the appraisal, she'll buy if she thinks it's real.)

2. Wilson Drilling has acquired the right to petroleum extraction on a specific tract in the eastern Gulf of Mexico. On the basis of past experience in the region, Wilson thinks there is a 20% chance that an oil deposit is present on the tract. If there *is* oil, the net present value accruing from full-scale drilling is $30 million; if there is no oil, the net present value is —$10 million. Before deciding whether to drill or not, Wilson can conduct a seismic survey.

From previous studies, Wilson knows that oil-bearing tracts yield positive survey results 80% of the time, and dry tracts yield positive survey results 25% of the time.

Given that they *do* run a survey, and the test results are positive, what is Wilson's expected profit if they decide to drill? What is the most Wilson should be willing to pay to run a seismic survey?

Solution:



E[profit | results are positive, drill] = 4/9 · 30 + 5/9 · (-10) = $7.78 million

Value of information from test = [0.36 · $7.78 + 0.64 · $0] – [$0] = $2.8 million

(Without the test, you'd choose not to drill; with the test, you'll drill only if the results are positive.)