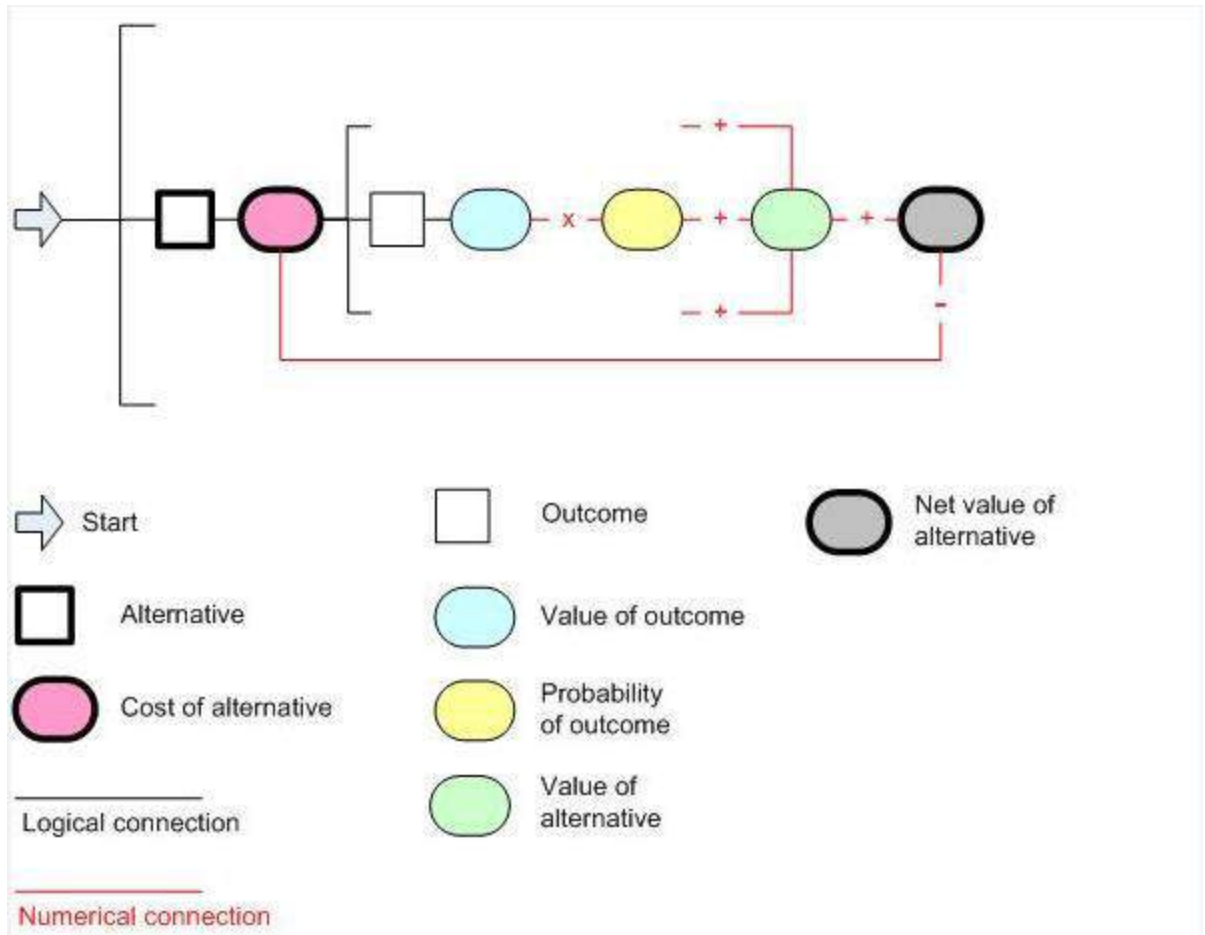**Note:  The ACTUAL case is on the last page of the readings below.**

- ○
- **Decision Trees**
- Decision trees are a common topic in the curriculum.  Searching the Web for "decision tree" will produce hundreds of hits, most of them on the websites of undergraduate business programs.  "Real" references, in the form of books and book chapters, are harder to identify, since the topic tends to be immersed in massive tracts that span the entire range of decision making (but see the Background Info page for two of them, listed as Additional Sources.)
- Decision trees are presented in many different ways, with different degrees of elaboration. The following is the "Trident decision tree model."  It may appear unnecessarily complex, and for simple trees, it definitely is.  But as the decision itself become more complex, the multiple features of the model become more useful.
- The various parts of the model are shown below.  At first glance, it may all look a bit confusing.  Don't worry. The confusion will go away as we examine a series of examples.

Start

Alternative

Cost of alternative

Logical connection

Numerical connection

Outcome

Value of outcome

Probability of outcome

Value of alternative

Net value of alternative

We start on the left, with a decision that must be made. Each decision has one or more alternatives. Each alternative has a cost.

Each alternative has one or more possible outcomes. Associated with each outcome is a value (also called a "payoff"), which is the benefit obtained if that particular outcome is realized. Also associated with each outcome is a probability, which ranges between 1.00 (if the outcome is certain) and 0.00 (it's impossible). If there's more than one possible outcome, then the sum of the probabilities must equal 1.00. (That's because there's always an outcome of some kind.)

The expected value of each outcome is its value multiplied by its probability. The value of the alternative is the sum of the expected values of all the outcomes.

The endpoint for the evaluation of each alternative is the net value, which is the expected value of the alternative, minus its cost.

The calculations are repeated for each alternative. The alternative yielding the greatest net value (either greatest gain or smallest loss) is the decision maker's preferred choice.

As an additional feature, the diagram shows two different types of connectors. **Logical connections** are in black, **numerical connections** are in red. For example, writing down an alternative logically implies the existence of a cost associated with that alternative. However, the mere existence of an alternative does not, in itself, determine the *amount* of that cost. For that reason, the line connecting the alternative and its cost is black   On the other hand, the alternative cost is needed to calculate the net value of the alternative, on the far right; for that reason, the line connecting these two entities is red.

Let's look at some worked-out examples.

*Type I.   The first "decision" isn't really a decision.  There's only one alternative, and it's forced upon the decision maker.*

A father, under pressure from his children, "chooses" to buy an AKC Springer spaniel at a cost of $1000.  During its reproductive lifetime, however, the dog whelps eight puppies, which are sold for an average of $1200 each ($9600 total).  What was the "value" of that "choice?"

Cost of alternative: $1000

Value of outcome:  $9600

Probability of outcome (not shown):  1.00.   (Since it happened, it's probability is 100%!)

Diagram (x, + and – signs removed, to reduce the clutter)
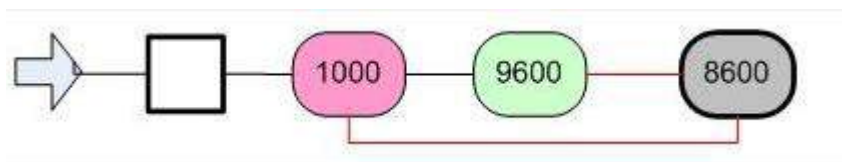
Net value: $8600

*Type II. There are two or more alternatives. The outcomes are known with certainty (probability 1.00) for each, as are the costs and expected values.*

A professional photographer has been offered two contracts, and only has the time to take one of them. Both contracts would require him to lease special equipment. Contract A, which would run for one year, pays $10,000, but requires the lease of a SteadiCam for $3,000. Contract B, which would run for two years, pays $3,500 per year (total value $7,000), but required the lease of a an HD three-dimensional still camera for $800 per year (total cost $1,600). What's the best choice?

Summary of data:

|  | Contract A | Contract B |
|---|---|---|
| Value | 10,000 | 7,000 |
| Cost | 3,000 | 1,600 |
| Net Value | 7,000 | 5,400 |

Diagram (numbers in thousands):



Decision: Take contract A, which has the highest net value
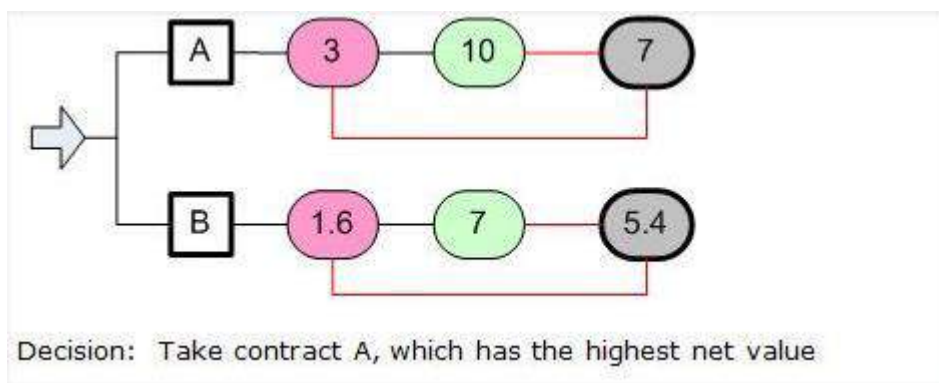
*Type III. There is only one alternative, but that alternative has several possible outcomes. Each outcome has a probability that it will occur. The list of outcomes must consist of all possible outcomes, and the sum of the probabilities must be 1.00. (100%)*

A father needs to buy a puppy for his children (there's no alternative). The usual price for an AKC Springer Spaniel is $2500, but a breeder offers him a puppy -- breeder's choice -- from a litter due to be whelped in one month, at a discount price of $500 cash. The father asks the reason for the discount.

"Well, genetic testing has determined that the sire has a congenital heart condition, and there's a 50% chance a puppy will have it. There's no test for it until the dog is an adult. The condition may shorten the lifespan. A dog that has it shouldn't be bred, and is only worth $200 has a family pet. If a dog doesn't have it, then it's worth more; a breeding male, $2000; a breeding female, $3000. "

"So let me understand," the buyer said. "If I buy the puppy right now for $500, and it's born male without the heart condition, I can turn around and sell it immediately for $2000? And if it's a female, for $3000? Why don't you just wait yourself, and see how the litter turns out?"

"Because I'm risk-adverse," the breeder says. "And on top of that, I need $500 cash today, to pay my kid's orthodontist."

*Flashback to stats: The probability of male vs. female is 0.50. The probability of healthy vs. defective heart is 0.50. Since the outcomes are unrelated, the joint probabilities are 0.25.*

What's the net value of buying the dog? Summary of data

| Alternatives | | Outcomes | | | | Value of alternatives | |
|---|---|---|---|---|---|---|---|
| Name | Cost | Name | Value (V) | Prob (P) | VxP | Total (sum of VxP) | Net |
| Buy dog | 500 | Defect (M or F) | 200 | 0.50 | 100 | 1350 | 850 |
| | | Healthy (M) | 2000 | 0.25 | 500 | | |
| | | Healthy (F) | 3000 | 0.25 | 750 | | |
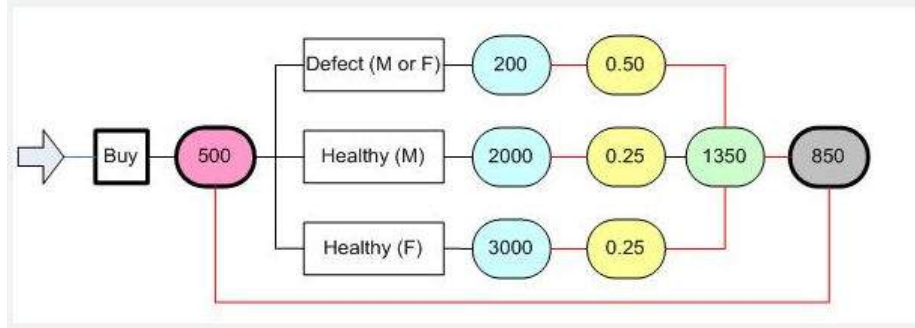


Since the net value of the deal is positive ($850), the buyer should snatch up the deal.  And also ask the breeder if he has anything else he'd like to sell!

*Type IV :  Multiple alternatives, multiple outcomes per alternative. For each alternative, the list of possible outcomes must include all possible outcomes. For each alternative, the sum of the probabilities associated with the outcomes must be 1.00 (100%).*

Mr. Entre is interested in selling his business.  He has two possible buyers, A and B.  Both of them would require some capital improvements before they buy;  either updating the store fronts, or updating the IT system.  Mr. Entre is inclined to do one or the other, but not both.

The store upgrade would cost $6M, the IT upgrade $3.5M.   If the store is upgraded, there's a 20% chance that A would buy, paying $9M.  There's a 50% chance that B would buy, paying $8M.

If the IT system were upgraded, there's a 40% chance that A would buy, for $8M.  There's a 30% chance that B would buy, for $6M.

The third alternative is to do nothing, and hold onto his company. What should Entre do?
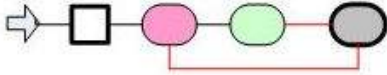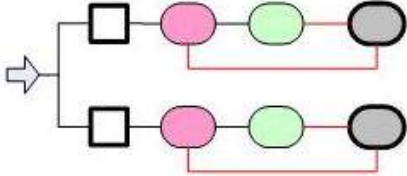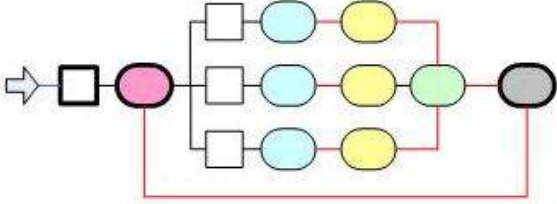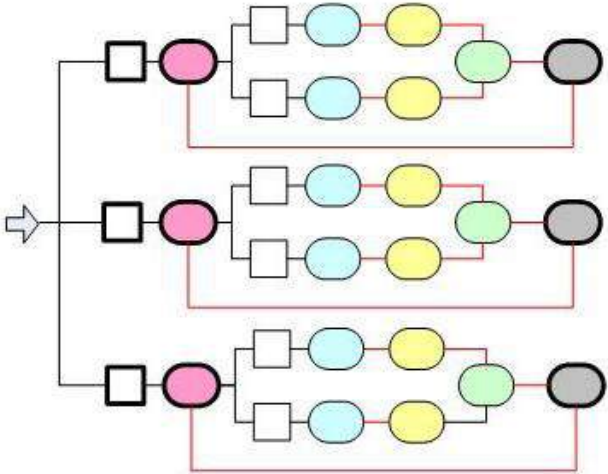
Summary of data (Costs/values in millions of $):

| Alternatives | | Outcomes | | | | Value of alternatives | |
|---|---|---|---|---|---|---|---|
| Name | Cost | Name | Value (V) | Prob (P) | VxP | Total (sum of VxP) | Net |
| Update store | 6 | Buyer A | 9 | .2 | 1.8 | 5.8 | -0.2 |
| | | Buyer B | 8 | .5 | 4.0 | | |
| | | No sale | 0 | .3 | 0 | | |
| Update IT | 3.5 | Buyer A | 8 | .4 | 3.2 | 5.0 | 1.5 |
| | | Buyer B | 6 | .3 | 1.8 | | |
| | | No sale | 0 | .3 | 0 | | |
| Do nothing | 0 | Buyer A | 0 | | | 0 | 0 |
| | | Buyer B | 0 | | | | |
| | | No sale | 0 | | | | |

Here's the outline of the decision tree. By now, the reader should be able to fill in the data. The "Do nothing" alternative, which has zero values from left to right, has been omitted.
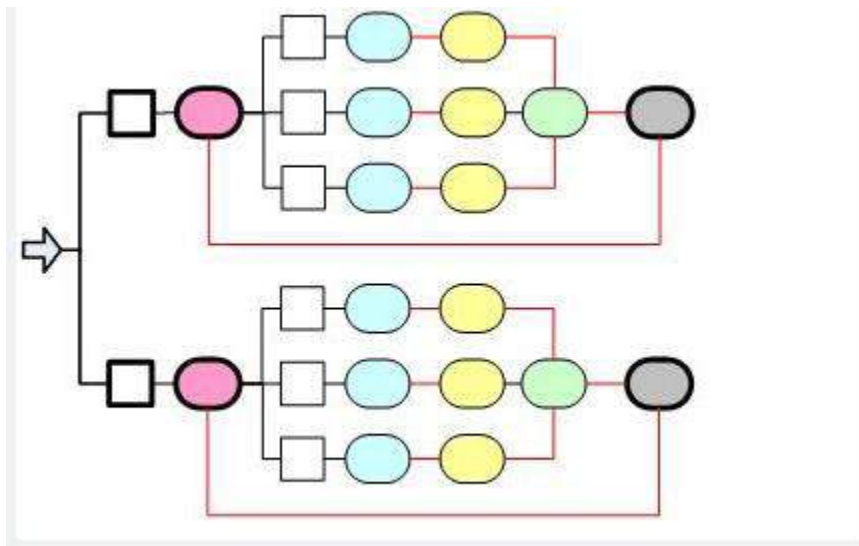
As seen above, the preferred alternative is to update the IT and hope for a sale.

Summary of types:

| | |
|---|---|
| **I:**<br>One alternative, one outcome (100% probable) per alternative |  |
| **II:**<br>Multiple alternatives, one outcome (100% probable) per alternative |  |
| **III. One alternative, multiple outcomes (total 100% probable) per alternative** |  |

| | |
|---|---|
| **IV. Multiple alternatives, multiple outcomes (total 100% probable) per alternative** |  |

There are also mixed types. One type combines alternatives having probabilistic outcomes with those having known outcomes. The diagram below illustrates this type. The upper alternative is probabilistic, the bottom alternative is known.
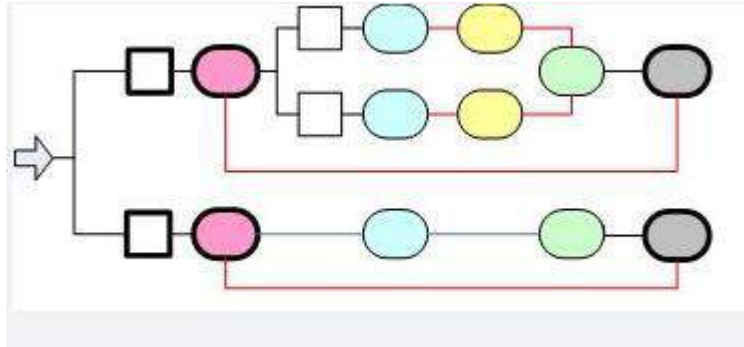


This is a straightforward combination of Types II and IV, above, and we won't bother with an example.

In the second mixed type, the choice of one alternative forces the consideration of one or more additional alternatives. The value of the first alternative is included in the calculation of the sub sequent alternatives, as indicated by the green lines in the figure below.
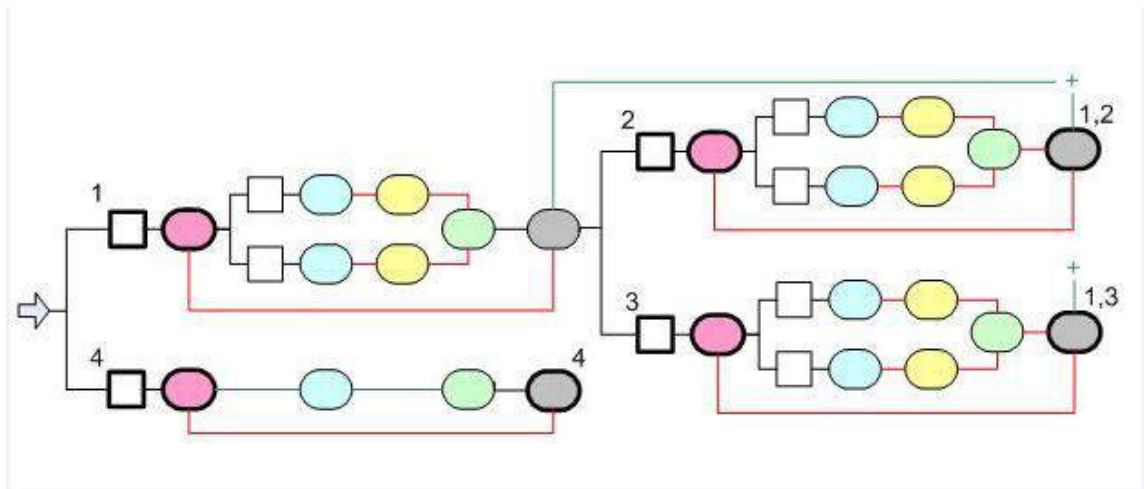
In step 1, the decision maker must choose between alternatives 1 and 4. (The numbering is arbitrary).If alternative 1 is selected, then it has a net value; but the decision-maker is forced into an additional choice between alternatives 2 and 3. Each has a net value, which is calculated as before, but now includes the net value of alternative 1. The decision maker then chooses between the net values of alternative 1 followed by 2 (1,2), alternative 1 followed by 3 (1,3), or the net value of alternative 4.

Here's an example that corresponds to the model.

Mr. Digit, a company CIO, decides he must upgrade its database software. His geeks offer him the choice between DB1 and DBX. (All dollar amounts are lease prices, per year.)

DB1 is a safe, established app costing $1M. The expected and assured productivity gain from DB1 would be $2M.

DBX is a new, state-of-the-art app costing $1.5M. The productivity gain is less certain, depending upon how rapidly the IT people trained. When forced to give estimates, the geeks estimated a 30% probability of a $4M gain, and a 70% probability of a $3M gain.

The geeks go on to tell Digit that while DBX is great, its vendor may not survive an ongoing market shakeout. If Digit chooses DBX, then he must consider buying a third party support contract, at a cost of $10,000 ($0.01M). The cost of a catastrophic crash is expected to be $0.5M, but the support contract would cover it; the contract would, in effect, have a value of $0.5M. Without the contract, the company would be stuck with the $0.5M bill (a "value" of -$0.05). The geeks estimate a 10% probability of a catastrophic crash.

What should Mr. Digit do?

Data summary (Costs and values in millions of $):

| Alternatives | | Outcomes | | | | Value of alternatives | | |
|---|---|---|---|---|---|---|---|---|
| Name | Cost | Name | Value (V) | Prob (P) | VxP | Total (Σ VxP) | Net | |
| DB1 | 1 | Gain | 2 | 1.00 | 2 | 1 | 1 | |
| DBX | 1.5 | Low gain | 3 | 0.70 | 2.1 | 3.3 | 1.8 (subtotal) | |
| | | High gain | 4 | 0.30 | 1.2 | | | |

| Name | Cost | Name | Value (V) | Prob (P) | VxP | Total (Σ VxP) | Net | Net + subtotal (= 1.8) |
|---|---|---|---|---|---|---|---|---|
| DBX + Svc | 0.01 | Crash | 0.5 | 0.10 | 0.05 | 0.05 | 0.04 | 1.84 |
| | | No crash | 0 | 0.90 | 0 | | | |
| DBX + no Svc | 0.00 | Crash | -0.5 | 0.10 | -0.05 | -0.05 | -0.05 | 1.75 |
| | | No crash | 0 | 0.90 | | | | |

There are three net-value outcomes, highlighted above.  Buy DB1, buy DBX with a service contract, or buy DBX without a service contract.  The best choice of alternative for Digit (+1.84M) is to buy DBX with a contract.

When faced with a decision, does the decision-maker actually have to draw out a tree?  That's a personal preference (although it may be a mandatory school exercise).  But when drawing a tree, must one include all the special nomenclature shown above;  i.e., the different boxes, colors, colored lines, and the like?

The answer to that question is definitely NO.  The diagrams are only useful as heuristics; that is, they help us organize our thinking, and make sure we don't leave anything out.  An actual decision will probably be based on a freehand sketch

Here's how the database / support decision, just discussed, first appeared when roughed out by the course developer on an engineering pad, using black and red Sharpies.

*Limitations of decision trees:*

The GIGO aphorism has never been more true.  The decision tree approach is only as useful as the initial data are accurate.

It's safe to assume that Mr. Digit's experts were reluctant to give him fixed-point estimates of the various outcomes.  In the real world, who could say with certainty that there's exactly a 70% chance of earning exactly 3 million dollars?  There would, of course, be a whole range of outcomes, but this particular approach doesn't handle ranges;  only fixed-point approximations.

Such approximations may be useful as a starting point.  After all, some data are better than no data, and having some decision procedure is better than merely guessing.  If nothing else, drawing a decision tree forces one to list and consider all the outcomes, and the degrees to which they're known.  This is always a useful first step.

*Some additional reading:*

As noted at the top of this page, there are many sources on the Web.  One interesting short paper, unfortunately without an illustration,  discusses using a decision tree to make career choices (Kautt, 2010).   Bratvold and Begg (2010)  present decision trees in the context of petroleum exploration and production.  For a more general discussion, see Simon (2000).

Bratvold, R. & Begg, S. (2010).  *Making good decisions* (Chap. 5: Structuring and solving decision problems).  Richardson, TX: Society of Petroleum Engineers [Available in the Trident ProQuest Ebrary].

Kautt, G. (2010).  Decision tree:  Diagramming the options in your professional life.  *Financial Planning* 40 (9) (no pagination).

Simon, J. (2000).  *Developing decision-making skills for business* (Part II:  Introduction to evaluative thinking.)  London:  M.E. Sharpe

## ASSIGNMENT CASE

- Calculate the expected value of an outcome, given its nominal value and probability of occurrence.
- Summarize data bearing on a quantitative decision.
- Represent the decision-making process as a decision tree

You're expected to provide references and citations.  It's recommended you use the sources listed in the Module.  Resist the temptation to browse the Web, looking for "clues."  There's a lot of confusing stuff out there, and much of it is irrelevant.  This Module contains all the information you need.  Study it – use it!

All land within the municipality of Springfield is zoned for either agricultural, residential or commercial use.  You have just purchased a parcel of agricultural land for $20,000 in the expectation that it will be rezoned next year.  But the rezoning is controversial, and will be decided by the full City Council rather than just the Zoning Commission.

According to your inside sources, there's a 30% chance the land will be rezoned for commercial use; in that event, you'll be able to sell the land for $50,000.  But the populists on the Council are pushing for more affordable housing; if they win the vote, which your sources think has a likelihood of 50%, then the land will be rezoned for residential use, and you'll only be able to sell it for $30,000.  Of course, the Greens may win, and the land won't be rezoned at all.  In that event, it will still be worth what you paid for it, but no more than that.

You've just been approached by Mr. Hi Roller.  He's a land speculator like yourself, but he doesn't have your inside sources.  He thinks the land will be rezoned next year, and has offered you $30,000 cash for it, right now.

Which alternative should you choose?  Explain your decision process in detail.

**Please read and heed the hints given in Case 1.** The same general advice is applicable in this Case, in particular:

1. Read the source materials before beginning.
2. The Case is about decision trees, not flipping real estate.
3. Follow standard format. A cover page, a short discussion, references and citations are all required.