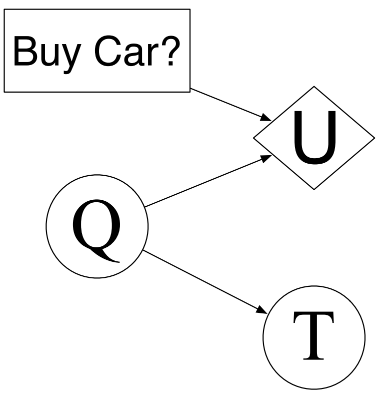
Used Car Purchase

A used car buyer can decide to carry out various tests with various costs (e.g., kick the tires, take the car to a qualified mechanic) and then, depending on the outcome of the tests, decide which car to buy. We will assume that the buyer is deciding whether to buy car c and that there is time to carry out at most one test which costs \$50 and which can help to figure out the quality of the car. A car can be in good shape (of good quality $Q = good$) or in bad shape (of bad quality $Q = bad$), and the test might help to indicate what shape the car is in. There are only two outcomes for the test T: pass (T=pass) or fail (T=fail). Car c costs \$1,500, and its market value is \$2,000 if it is in good shape; if not, \$700 in repairs will be needed to make it in good shape. The buyers estimate is that c has 70% chance of being in good shape.

1. Draw the decision network that represents this problem. (The only action is whether to buy or not the car).



2. Calculate the expected net gain from buying car c, given no test.

$$\begin{aligned} EU(\text{buy}) &= P(Q = good) \cdot U(good, \text{buy}) + P(Q = bad) \cdot U(bad, buy) \\ &= .7 \cdot 500 + 0.3 \cdot -200 = 290 \end{aligned}$$

3. Tests can be described by the probability that the car will pass or fail the test given that the car is in good or bad shape. We have the following information:

$$P(T = \text{pass}|Q = good) = 0.9$$
$$P(T = \text{pass}|Q = bad) = 0.2$$

Calculate the probability that the car will pass (or fail) its test, and then the probability that it is in good (or bad) shape given each possible test outcome.

$$\begin{aligned}
P(T = \text{pass}) &= \sum_q P(T = \text{pass}, Q = q) \\
&= P(T = \text{pass}|Q = good)P(Q = good) + P(T = \text{pass}|Q = bad)P(Q = bad) \\
&= 0.69 \\
P(T = fail) &= 0.31 \\
P(Q = good|T = \text{pass}) &= \frac{P(T = \text{pass}|Q = good)P(Q = good)}{P(T = \text{pass})} \\
&= \frac{0.9 \cdot 0.7}{0.69} = \frac{21}{23} \approx 0.91 \\
P(Q = good|T = \text{fail}) &= \frac{P(T = \text{fail}|Q = good)P(Q = good)}{P(T = \text{fail})} \\
&= \frac{0.1 \cdot 0.7}{0.31} = \frac{7}{31} \approx 0.22
\end{aligned}$$

4. Calculate the optimal decisions given either a pass or a fail, and their expected utilities.

$$\begin{aligned}
EU(\text{buy}|T = pass) &= P(Q = good|T = \text{pass})U(good, \text{buy}) + P(Q = bad|T = \text{pass})U(bad, \text{buy}) \\
&\approx 0.91 \cdot 500 + 0.09 \cdot (-200) \approx 437 \\
EU(\text{buy}|T = \text{fail}) &= P(Q = good|T = \text{fail})U(good, \text{buy}) + P(Q = bad|T = \text{fail})U(bad, \text{buy}) \\
&\approx 0.22 \cdot 500 + 0.78 \cdot (-200) = -46 \\
EU(\neg\text{buy}|T = pass) &= 0 \\
EU(\neg\text{buy}|T = fail) &= 0
\end{aligned}$$

Therefore: $MEU(T = \text{pass}) = 437$ (with buy) and $MEU(T = \text{fail}) = 0$ (using $\neg$buy)

5. Calculate the value of (perfect) information of the test. Should the buyer pay for a test?

$$\begin{aligned}
VPI(T) &= (\sum_t P(T = t)MEU(T = t)) - MEU(\phi) \\
&= 0.69 \cdot 437 + 0.31 \cdot 0 - 290 \approx 11.53
\end{aligned}$$

You shouldn't pay for it, since the cost is \$50.

6. The value of the information in this problem depends greatly on the prior probability $P(Q = good)$. What do you think happens to the VPI as you vary $P(Q = good)$? What happens when $P(Q = good)$ approaches 1? Approaches 0? Approaches 0.5?

Intuitively, the more unsure you are of something, the more value you stand to gain by getting "perfect information" related to it. Thus, we expect that near 1 and near 0, there is little value in the information, but somewhere in the middle there is higher value. Utilities mess with this intuition: if you have an action that is low risk (not buying) and one that is high risk (buying), the midpoint might shift. Thus, as you might see in the next part, the critical range where getting the test is useful lies somewhere between 0.18 and 0.52.

MDPs

MDPs can formulated with a reward function R(s), R(s, a) that depends on the action taken or R(s, a, s') that depends on the action taken and outcome state.

1. Write down the Bellman equations for R(s)

   $U(s) = R(s) + \gamma max_a[\sum_{s'} T(s,a,s')U(s')]$

2. Write down the Bellman equations for R(s, a)

   $U(s) = max_a[R(s,a) + \gamma \sum_{s'} T(s,a,s')U(s')]$

3. Write down the Bellman equations for R(s, a, s')

   $U(s) = max_a \sum_{s'} T(s,a,s')[R(s,a,s') + \gamma U(s')]$

4. Show how am MDP with a reward function R(s, a, s') can be transformed into a different MDP with reward function R(s, a), such that optimal policies in the new MDP correspond exactly to optimal policies in the original MDP.
   One solution is to define a 'pre-state' pre(s, a, s') for every s, a, s' such that executing a in s leads not to s' but to pre(s, a, s'). From the pre-state there is only one action b that always leads to s'. Let the new MDP have transition T', reward R', and discount γ'.
   $T'(s,a,pre(s,a,s')) = T(s,a,s')$
   $T'(pre(s,a,s'),b,s') = 1$
   $R'(s,a) = 0$
   $R'(pre(s,a,s'),b) = \gamma^{-\frac{1}{2}}R(s,a,s')$
   $\gamma' = \gamma^{\frac{1}{2}}$

   Then, using pre as shorthand for pre(s, a, s'):

   $U(s) = max_a[R'(s,a) + \gamma^{\frac{1}{2}} \sum_{pre} T'(s,a,pre)(max_b[R'(pre,b) + \gamma^{\frac{1}{2}} \sum_{s'} T'(pre,b,s') * U(s'))]]$
   $U(s) = max_a[\sum_{s'} T(s,a,s')(R(s,a,s') + \gamma U(s')]$

5. Now do the same to convert MDPs with R(s, a) into MDPs with R(s).
   Similar to part (c), create a state post(s, a) for every s, a such that
   $T'(s,a,post(s,a,s')) = 1$
   $T'(post(s,a,s'),b,s') = T(s,a,s')$
   $R'(s) = 0$
   $R'(post(s,a,s')) = \gamma^{-\frac{1}{2}}R(s,a)$
   $\gamma' = \gamma^{\frac{1}{2}}$

   Then, using post as shorthand for post(s, a, s'):

   $U(s) = R'(s) + \gamma^{\frac{1}{2}} max_a[\sum_{post} T'(s,a,post)(R'(post) + \gamma^{\frac{1}{2}} max_b[\sum_{s'} T'(post,b,s')U(s'))]]$
   $U(s) = max_a[R(s,a) + \gamma \sum_{s'} T(s,a,s')U(s')]$