Discussion 9 Solutions

Decision Networks

- **Chance nodes** - Chance nodes in a decision network behave identically to Bayes' nets. Each outcome in a chance node has an associated probability, which can be determined by running inference on the underlying Bayes' net it belongs to. We'll represent these with ovals.
- **Action nodes** - Action nodes are nodes that we have complete control over; they're nodes representing a choice between any of a number of actions which we have the power to choose from. We'll represent action nodes with rectangles.
- **Utility nodes** - Utility nodes are children of some combination of action and chance nodes. They output a utility based on the values taken on by their parents, and are represented as diamonds in our decision networks.

The **expected utility** of taking an action $A = a$ given evidence $E = e$ and n chance nodes is computed with the following formula:

$$EU(A = a|E = e) = \sum_{x_1,...,x_n} P(X_1 = x_1, ..., X_n = x_n|E = e)U(A = a, X_1 = x_1, ..., X_n = x_n)$$

where each x_i represents a value that the i^{th} chance node can take on. The **maximum expected utility** is the expected utility of the action that has the highest expected utility:

$$MEU(E = e) = \max_a EU(A = a|E = e).$$

Value of Perfect Information

Value of perfect information (VPI) quantifies the amount an agent's maximum expected utility is expected to increase if it were to observe some new evidence. Usually observing new evidence comes at a cost. If we observed some new evidence $E' = e'$ before acting, the maximum expected utility of our action at that point would become

$$MEU(E = e, E' = e') = \max_a \sum_x P(X = x|E = e, E' = e')U(X = x, A = a).$$

However, note that *we don't know what new evidence we'll get*. Because we don't know what what new evidence e' we'll get, we must represent it as a random variable E'. We will compute the expected value of the maximum expected utility:

$$MEU(E = e, E') = \sum_{e'} P(E' = e'|E = e)MEU(E = e, E' = e').$$

Observing a new evidence variable yields a different MEU with probabilities corresponding to the probabilities of observing each value for the evidence variable, and so by computing $MEU(E = e, E')$ as above, we compute what we expect our new MEU will be if we choose to observe new evidence. The VPI is the expected maximum expected utility if we were to observe the new evidence, minus the maximum expected utility if we were not to observe the new evidence:

$$VPI(E'|E = e) = MEU(E = e, E') - MEU(E = e).$$

Properties of VPI

The value of perfect information has several very important properties, namely:

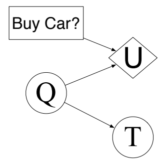
- **Nonnegativity.** $\forall E', e \;\; VPI(E'|E = e) \geq 0$
  Observing new information always allows you to make a *more informed* decision, and so your maximum expected utility can only increase (or stay the same if the information is irrelevant for the decision you must make).

- **Nonadditivity.** $VPI(E_j, E_k|E = e) \neq VPI(E_j|E = e) + VPI(E_k|E = e)$ in general.
  This is probably the trickiest of the three properties to understand intuitively. It's true because generally observing some new evidence E_j might change how much we care about E_k; therefore we can't simply add the VPI of observing E_j to the VPI of observing E_k to get the VPI of observing both of them. Rather, the VPI of observing two new evidence variables is equivalent to observing one, incorporating it into our current evidence, then observing the other. This is encapsulated by the order-independence property of VPI, described more below.

- **Order-independence.** $VPI(E_j, E_k|E = e) = VPI(E_j|E = e) + VPI(E_k|E = e, E_j) = VPI(E_k|E = e) + VPI(E_j|E = e, E_k)$
  Observing multiple new evidences yields the same gain in maximum expected utility regardless of the order of observation. This should be a fairly straightforward assumption - because we don't actually take any action until after observing any new evidence variables, it doesn't actually matter whether we observe the new evidence variables together or in some arbitrary sequential order.

Q1. Decision Networks and VPI

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 $Q = +q$) or in bad shape (of bad quality Q=¬q), 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 buyer's estimate is that c has 70% chance of being in good shape. The Decision Network is shown below.



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

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

2. 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:

$$P(T = \text{pass}|Q = +q) = 0.9$$
$$P(T = \text{pass}|Q = \neg q) = 0.2$$

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.

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

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

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

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

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

$$VPI(T) = (\sum_t P(T = t)MEU(T = t)) - MEU(\phi)$$
$$= 0.69 \cdot 437 + 0.31 \cdot 0 - 290 \approx 11.53$$

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

Q2. VPI

You are the latest contestant on Monty Hall's game show, which has undergone a few changes over the years.
In the game, there are n closed doors: behind one door is a car ($U(car) = 1000$), while the other $n - 1$ doors each have a goat behind them ($U(goat) = 10$). You are permitted to open exactly one door and claim the prize behind it.

You begin by choosing a door uniformly at random.

(a) What is your expected utility?

Answer: $(1000 * \frac{1}{n} + 10 * \frac{n-1}{n})$ or $(10 + 990 * \frac{1}{n})$

We can calculate the expected utility via the usual formula of expectation, or we can note that there is a guaranteed utility of 10, with a small probability of a bonus utility. The latter is a bit simpler, so the answers to the following parts use this form.

(b) After you choose a door but before you open it, Monty offers to open k other doors, each of which are guaranteed to have a goat behind it.
If you accept this offer, should you keep your original choice of a door, or switch to a new door?

$EU(keep)$: $10 + 990 * \frac{1}{n}$

$EU(switch)$: $10 + 990 * \frac{(n-1)}{n*(n-k-1)}$

Action that achieves MEU: switch

The expected utility if we keep must be the same as the answer from the previous part: the probability that we have a winning door has not changed at all, since we have gotten no meaningful information.
In order to win a car by switching, we must have chosen a goat door previously (probability $\frac{n-1}{n}$) and then switch to the car door (probability $\frac{1}{n-k-1}$).
Since $n - 1 > n - k - 1$ for positive k, switching gets a larger expected utility.

(c) What is the value of the information that Monty is offering you?

Answer: $990 * \frac{1}{n} * \frac{k}{n-k-1}$

The formula for VPI is $MEU(e) - MEU(\emptyset)$. Thus, we want the difference between $EU(switch)$ (the optimal action if Monty opens the doors) and our expected utility from part (a).
(It is true that $EU(keep)$ happens to have the same numeric expression as in part (a), but this fact is not meaningful in answering this part.)

(d) Monty is changing his offer!
After you choose your initial door, you are given the offer to choose any other door and open this second door. If you do, after you see what is inside the other door, you may switch your initial choice (to the newly opened door) or keep your initial choice.
What is the value of this new offer?

Answer: $\frac{990}{n}$

Intuitively, if we take this offer, it is as if we just chose two doors in the beginning, and we win if either door has the car behind it. Unlike in the previous parts, if the new door has a goat behind it, it is not more optimal to switch doors. Mathematically, letting D_i be the event that door i has the car, we can calculate this as $P(D_2 \cup D_1) = P(D_1) + P(D_2) = 1/n + 1/n = 2/n$, to see that $MEU(\text{offer}) = 10 + 990 * \frac{2}{n}$. Subtracting the expected utility without taking the offer, we are left with $990 * \frac{1}{n}$.

(e) Monty is generalizing his offer: you can pay $\$d^3$ to open d doors as in the previous part. (Assume that $U(\$x) = x$.) You may now switch your choice to any of the open doors (or keep your initial choice). What is the largest value of d for which it would be rational to accept the offer?

Answer: $d = \sqrt{\frac{990}{n}}$

It is a key insight (whether intuitive or determined mathematically) that the answer to the previous part is constant for each successive door we open. Thus, the value of opening d doors is just $d * 990 * \frac{1}{n}$. Setting this equal to d^3, we can solve for d.