EXERCISES 15

MAKING SIMPLE DECISIONS

15.1 Combining Beliefs and Desires under Uncertainty

Exercise 15.ALMG

(Adapted from David Heckerman.) This exercise concerns the **Almanac Game**, which is used by decision analysts to calibrate numeric estimation. For each of the questions that follow, give your best guess of the answer, that is, a number that you think is as likely to be too high as it is to be too low. Also give your guess at a 25th percentile estimate, that is, a number that you think has a 25% chance of being too high, and a 75% chance of being too low. Do the same for the 75th percentile. (Thus, you should give three estimates in all—low, median, and high—for each question.)

a. Number of passengers who flew between New York and Los Angeles in 1989.
b. Population of Warsaw in 1992.
c. Year in which Coronado discovered the Mississippi River.
d. Number of votes received by Jimmy Carter in the 1976 presidential election.
e. Age of the oldest living tree, as of 2002.
f. Height of the Hoover Dam in feet.
g. Number of eggs produced in Oregon in 1985.
h. Number of Buddhists in the world in 1992.
i. Number of deaths due to AIDS in the United States in 1981.
j. Number of U.S. patents granted in 1901.

The correct answers appear after the last exercise of this chapter. From the point of view of decision analysis, the interesting thing is not how close your median guesses came to the real answers, but rather how often the real answer came within your 25% and 75% bounds. If it was about half the time, then your bounds are accurate. But if you're like most people, you will be more sure of yourself than you should be, and fewer than half the answers will fall within the bounds. With practice, you can calibrate yourself to give realistic bounds, and thus be more useful in supplying information for decision making. Try this second set of questions and see if there is any improvement:

a. Year of birth of Zsa Zsa Gabor.
b. Maximum distance from Mars to the sun in miles.
c. Value in dollars of exports of wheat from the United States in 1992.
d. Tons handled by the port of Honolulu in 1991.

e. Annual salary in dollars of the governor of California in 1993.
f. Population of San Diego in 1990.
g. Year in which Roger Williams founded Providence, Rhode Island.
h. Height of Mt. Kilimanjaro in feet.
i. Length of the Brooklyn Bridge in feet.
j. Number of deaths due to automobile accidents in the United States in 1992.

It is interesting to create a histogram of accuracy on this task for the students in the class. It is also interesting to record how many times each student comes within, say, 10% of the right answer. Then you get a profile of each student: this one is an accurate guesser but overly cautious about bounds, etc.

Exercise 15.STPT

In 1713, Nicolas Bernoulli stated a puzzle, now called the St. Petersburg paradox, which works as follows. You have the opportunity to play a game in which a fair coin is tossed repeatedly until it comes up heads. If the first heads appears on the nth toss, you win 2^n dollars.

a. Show that the expected monetary value of this game is infinite.
b. How much would you, personally, pay to play the game?
c. Nicolas's cousin Daniel Bernoulli resolved the apparent paradox in 1738 by suggesting that the utility of money is measured on a logarithmic scale (i.e., $U(S_n) = a \log_2 n + b$, where S_n is the state of having \$$n$). What is the expected utility of the game under this assumption?
d. What is the maximum amount that it would be rational to pay to play the game, assuming that one's initial wealth is \$$k$?

a. The probability that the first heads appears on the nth toss is 2^{-n}, so

$$EMV(L) = \sum_{n=1}^{\infty} 2^{-n} \cdot 2^n = \sum_{n=1}^{\infty} 1 = \infty$$

b. Typical answers range between \$4 and \$100.

c. Assume initial wealth (after paying c to play the game) of $\$(k - c)$; then

$$U(L) = \sum_{n=1}^{\infty} 2^{-n} \cdot (a \log_2(k - c + 2^n) + b)$$

Assume $k - c = \$0$ for simplicity. Then

$$\begin{aligned} U(L) &= \sum_{n=1}^{\infty} 2^{-n} \cdot (a \log_2(2^n) + b) \\ &= \sum_{n=1}^{\infty} 2^{-n} \cdot an + b \\ &= 2a + b \end{aligned}$$

d. The maximum amount c is given by the solution of

$$a \log_2 k + b = \sum_{n=1}^{\infty} 2^{-n} \cdot (a \log_2(k - c + 2^n) + b)$$

For our simple case, we have

$$a \log_2 c + b = 2a + b$$

or $c = \$4$.

15.2 Utility Functions

Exercise 15.USED
Chris considers four used cars before buying the one with maximum expected utility. Pat considers ten cars and does the same. All other things being equal, which one is more likely to have the better car? Which is more likely to be disappointed with their car's quality? By how much (in terms of standard deviations of expected quality)?

Pat is more likely to have a better car than Chris because she has more information with which to choose. She is more likely to be disappointed, however, if she takes the expected utility of the best car at face value. Using the results of exercise 15.11, we can compute the expected disappointment to be about 1.54 times the standard deviation by numerical integration.

Exercise 15.ASSU
Assess your own utility for different incremental amounts of money by running a series of preference tests between some definite amount M_1 and a lottery $[p, M_2; (1-p), 0]$. Choose different values of M_1 and M_2, and vary p until you are indifferent between the two choices. Plot the resulting utility function.

This is an interesting exercise to do in class. Choose $M_1 = \$100$, M_2 = \$100, \$1000, \$10000, \$1000000. Ask for a show of hands of those preferring the lottery at different values of p. Students will almost always display risk aversion, but there may be a wide spread in its onset. A curve can be plotted for the class by finding the smallest p yielding a majority vote for the lottery.

Exercise 15.ASSE

Write a computer program to automate the process in Exercise 15..ASSU. Try your program out on several people of different net worth and political outlook. Comment on the consistency of your results, both for an individual and across individuals.

The program itself is pretty trivial. But note that there are some studies showing you get better answers if you ask subjects to move a slider to indicate a proportion, rather than asking for a probability number. So having a graphical user interface is an advantage. The main point of the exercise is to examine the data, expose inconsistent behavior on the part of the subjects, and see how people vary in their choices.

Exercise 15.SURC

The Surprise Candy Company makes candy in two flavors: 75% are strawberry flavor and 25% are anchovy flavor. Each new piece of candy starts out with a round shape; as it moves along the production line, a machine randomly selects a certain percentage to be trimmed into a square; then, each piece is wrapped in a wrapper whose color is chosen randomly to be red or brown. 70% of the strawberry candies are round and 70% have a red wrapper, while 90% of the anchovy candies are square and 90% have a brown wrapper. All candies are sold individually in sealed, identical, black boxes.

Now you, the customer, have just bought a Surprise candy at the store but have not yet opened the box. Consider the three Bayes nets in Figure 15.1.

a. Which network(s) can correctly represent $\mathbf{P}(Flavor, Wrapper, Shape)$?

b. Which network is the best representation for this problem?

c. Does network (i) assert that $\mathbf{P}(Wrapper|Shape) = \mathbf{P}(Wrapper)$?

d. What is the probability that your candy has a red wrapper?

e. In the box is a round candy with a red wrapper. What is the probability that its flavor is strawberry?

f. A unwrapped strawberry candy is worth s on the open market and an unwrapped anchovy candy is worth a. Write an expression for the value of an unopened candy box.

g. A new law prohibits trading of unwrapped candies, but it is still legal to trade wrapped candies (out of the box). Is an unopened candy box now worth more than less than, or the same as before?

a. Networks (ii) and (iii) can represent this network but not (i).

(ii) is fully connected, so it can represent any joint distribution.

(iii) follows the generative story given in the problem: the flavor is determined (presumably) by which machine the candy is made by, then the shape is randomly cut, and the wrapper randomly chosen, the latter choice independently of the former.

(i) cannot represent this, as this network implies that the wrapper color and shape are marginally independent, which is not so: a round candy is likely to be strawberry, which is in turn likely to be wrapped in red, whilst conversely a square candy is likely to be anchovy which is likely to be wrapped in brown.

b. Unlike (ii), (iii) is a polytree, which simplifies inference. Its edges also follow the causal direction, so probabilities will be easier to elicit. Indeed, the problem statement has already given them.

c. Yes, because Wrapper and Shape are d-separated.

d. Once we know the Flavor we know the probability its wrapper will be red or brown. So we marginalize Flavor out:

$$\begin{aligned}\mathbf{P}(Wrapper = red) &= \sum_f \mathbf{P}(Wrapper = red, Flavor = f) \\ &= \sum_f \mathbf{P}(Flavor = f)\mathbf{P}(Wrapper = red|Flavor = f) \\ &= 0.75 \times 0.7 + 0.25 \times 0.1 \\ &= 0.55\end{aligned}$$

e. We apply Bayes theorem, by first computing the joint probabilities

$$\begin{aligned}&\mathbf{P}(Flavor = strawberry, Shape = round, Wrapper = red) \\ &\quad = \mathbf{P}(Flavor = strawberry) \times \mathbf{P}(Shape = round|Flavor = strawberry) \\ &\qquad \times \mathbf{P}(Wrapper = red|Flavor = strawberry) \\ &\quad = 0.75 \times 0.7 \times 0.7 \\ &\quad = 0.3675 \\ &\mathbf{P}(Flavor = anchovy, Shape = round, Wrapper = red) \\ &\quad = \mathbf{P}(Flavor = anchovy) \times \mathbf{P}(Shape = round|Flavor = anchovy) \\ &\qquad \times \mathbf{P}(Wrapper = red|Flavor = anchovy) \\ &\quad = 0.25 \times 0.1 \times 0.1 \\ &\quad = 0.0025\end{aligned}$$

Normalizing these probabilities yields that it is strawberry with probability $0.3675/(0.3675+0.0025) \approx 0.9932$.

f. Its value is the probability that you have a strawberry upon unwrapping times the value of a strawberry, plus the probability that you have a anchovy upon unwrapping times the value of an anchovy or

$$0.75s + 0.25a\ .$$

g. The value is the same, by the axiom of decomposability.

Exercise 15.ALLP

Prove that the judgments $B \succ A$ and $C \succ D$ in the Allais paradox (page 528) violate the axiom of substitutability.

First observe that $C \sim [0.25, A; 0.75, \$0]$ and $D \sim [0.25, B; 0.75\$0]$. This follows from the axiom of decomposability. But by substitutability this means that the preference ordering between the lotteries A and B must be the same as that between C and D.

Exercise 15.ALLQ

Consider the Allais paradox described on page 528: an agent who prefers B over A (taking the sure thing), and C over D (taking the higher EMV) is not acting rationally, according to utility theory. Do you think this indicates a problem for the agent, a problem for the theory, or no problem at all? Explain.

As mentioned in the text, agents whose preferences violate expected utility theory demonstrate irrational behavior, that is they can be made either to accept a bet that is a guaranteed loss for them (the case of violating transitivity is given in the text), or reject a bet that is a guaranteed win for them. This indicates a problem for the agent.

Exercise 15.LOTT

Tickets to a lottery cost \$1. There are two possible prizes: a \$10 payoff with probability 1/50, and a \$1,000,000 payoff with probability 1/2,000,000. What is the expected monetary value of a lottery ticket? When (if ever) is it rational to buy a ticket? Be precise—show an equation involving utilities. You may assume current wealth of $\$k$ and that $U(S_k) = 0$. You may also assume that $U(S_{k+10}) = 10 \times U(S_{k+1})$, but you may not make any assumptions about $U(S_{k+1,000,000})$. Sociological studies show that people with lower income buy a disproportionate number of lottery tickets. Do you think this is because they are worse decision makers or because they have a different utility function? Consider the value of contemplating the possibility of winning the lottery versus the value of contemplating becoming an action hero while watching an adventure movie.

The expected monetary value of the lottery L is

$$EMV(L) = \frac{1}{50} \times \$10 + \frac{1}{2000000} \times \$1000000 = \$0.70$$

Although $\$0.70 < \$1$, it is not *necessarily* irrational to buy the ticket. First we will consider just the utilities of the monetary outcomes, ignoring the utility of actually playing the lottery game. Using $U(S_{k+n})$ to represent the utility to the agent of having n dollars more than the current state, and assuming that utility is linear for small values of money (i.e., $U(S_{k+n}) \approx n(U(S_{k+1}) - U(S_k))$ for $-10 \le n \le 10$), the utility of the lottery is:

$$\begin{aligned} U(L) &= \frac{1}{50} U(S_{k+10}) + \frac{1}{2,000,000} U(S_{k+1,000,000}) \\ &\approx \frac{1}{5} U(S_{k+1}) + \frac{1}{2,000,000} U(S_{k+1,000,000}) \end{aligned}$$

This is more than $U(S_{k+1})$ when $U(S_{k+1,000,000}) > 1,600,000 U(\$1)$. Thus, for a purchase to be rational (when only money is considered), the agent must be quite risk-seeking. This would be unusual for low-income individuals, for whom the price of a ticket is non-trivial. It is possible that some buyers do not internalize the magnitude of the very low probability of winning—to imagine an event is to assign it a "non-trivial" probability, in effect. Apparently, these buyers are better at internalizing the large magnitude of the prize. Such buyers are clearly acting irrationally.

Some people may feel their current situation is intolerable, that is, $U(S_k) \approx U(S_{k\pm1}) \approx u_\perp$. Therefore the situation of having one dollar more or less would be equally intolerable, and it would be rational to gamble on a high payoff, even if one that has low probability.

Gamblers also derive pleasure from the excitement of the lottery and the temporary possession of at least a non-zero chance of wealth. So we should add to the utility of playing the lottery the term t to represent the thrill of participation. Seen this way, the lottery is just another form of entertainment, and buying a lottery ticket is no more irrational than buying a movie ticket. Either way, you pay your money, you get a small thrill t, and (most likely) you walk away empty-handed. (Note that it could be argued that doing this kind of decision-theoretic computation decreases the value of t. It is not clear if this is a good thing or a bad thing.)

Exercise 15.MICM

How much is a micromort worth to you? Devise a protocol to determine this. Ask questions based both on paying to avoid risk and being paid to accept risk.

The protocol would be to ask a series of questions of the form "which would you prefer" involving a monetary gain (or loss) versus an increase (or decrease) in a risk of death. For example, "would you pay \$100 for a helmet that would eliminate completely the one-in-a-million chance of death from a bicycle accident."

Exercise 15.KMAX

Let continuous variables $X_1, \ldots, X_k$ be independently distributed according to the same probability density function $f(x)$. Prove that the density function for $\max\{X_1, \ldots, X_k\}$ is given by $kf(x)(F(x))^{k-1}$, where F is the cumulative distribution for f.

First observe that the cumulative distribution function for $\max\{X_1, \ldots, X_k\}$ is $(F(x))^k$ since

$$\begin{aligned} P(\max\{X_1, \ldots, X_k\} \le x) &= P(X_1 \le x, \ldots, X_k \le x) \\ &= P(X_1 \le x) \ldots P(X_k \le x) \\ &= F(x)^k \end{aligned}$$

where the second to last step follows by independence. The result follows as the probability density function is the derivative of the cumulative distribution function.

Exercise 15.EXPU

Economists often make use of an exponential utility function for money: $U(x) = -e^{-x/R}$, where R is a positive constant representing an individual's risk tolerance. Risk tolerance reflects how likely an individual is to accept a lottery with a particular expected monetary value (EMV) versus some certain payoff. As R (which is measured in the same units as x) becomes larger, the individual becomes less risk-averse.

a. Assume Mary has an exponential utility function with $R = \$500$. Mary is given the choice between receiving \$500 with certainty (probability 1) or participating in a lottery which has a 60% probability of winning \$5000 and a 40% probability of winning nothing. Assuming Marry acts rationally, which option would she choose? Show how you derived your answer.

b. Consider the choice between receiving \$100 with certainty (probability 1) or participating in a lottery which has a 50% probability of winning \$500 and a 50% probability of winning nothing. Approximate the value of R (to 3 significant digits) in an exponential utility function that would cause an individual to be indifferent to these two alternatives. (You might find it helpful to write a short program to help you solve this problem.)

a. Getting \$400 for sure has expected utility

$$-e^{-400/400} = -1/e \approx -0.3679$$

while the getting \$5000 with probability 0.6 and \$0 otherwise has expected utility

$$0.6 - e^{-5000/400} + 0.5 - e^{-0/400} = -(0.6e^{-12.5} + 0.5) \approx -0.5000$$

so one would prefer the sure bet.

b. We want to find R such that

$$e^{-100/R} = 0.5e^{-500/R} + 0.5$$

Solving this numerically, we find $R = 152$ up to 3sf.

Exercise 15.FGAM

Alex is given the choice between two games. In Game 1, a fair coin is flipped and if it comes up heads, Alex receives \$100. If the coin comes up tails, Alex receives nothing. In Game 2, a fair coin is flipped t wice. Each time the coin comes up heads, Alex receives \$50, and Alex receives nothing for each coin flip that comes up t ails. Assuming that Alex has a monotonically increasing utility function for money in the range [\$0, \$100], show mathematically that if Alex prefers Game 2 to Game 1, then Alex is risk averse (at least with respect to this range of monetary amounts).

Since Alex prefers Game 2 to Game 1, we have

$$0.5U(\$0) + 0.5U(\$100) < U(\$50).$$

Since $\$50 = 0.5\$0 + 0.5\$100$ is the expected monetary value of the Game 1 lottery, Alex is risk averse.

Exercise 15.188-UTIL

Consider the following lotteries:

- $L_1 = [1, 1]$.
- $L_2 = [0.5, 2; 0.5, 0]$.
- $L_3 = [1, 2]$.

Four students have expressed their preferences over these lotteries as follows:

- Adam is indifferent between lottery L_2 and lottery L_1.
- Becky prefers lottery L_1 to lottery L_2.
- Charles is indifferent between lottery L_3 and lottery L_2.
- Diana prefers lottery L_2 to lottery L_1.

Match each student with a utility function that is consistent with their stated preferences. Each student has a different utility function.

a. Which student has the utility function of $U(x) = x^2$

b. Which student has the utility function of $U(x) = x$

c. Which student has the utility function of $U(x) = \sqrt{x}$

d. Which student has the utility function of none of the ones above

a. Diana, since $0.5 \times 0^2 + 0.5 \times 2^2 > 1^2$

b. Adam, since $0.5 \times 0 + 0.5 \times 2 = 1$

c. Becky, since $0.5 \times \sqrt{0} + 0.5 \times \sqrt{2} < \sqrt{1}$

d. Charles, since none of the three utility functions satisfies $0.5 \times U(0) + 0.5 \times U(2) = U(2)$

Exercise 15.188-UTPR

True or False: Assume Agent 1 has a utility function U_1 and Agent 2 has a utility function U_2. If $U_1 = k_1 U_2 + k_2$ with $k_1 > 0, k_2 > 0$ then Agent 1 and Agent 2 have the same preferences.

True. For any $a, b : U_2(a) > U_2(b)$ equivalent to $k_1 U_2(a) > k_1 U_2(b)$ since $k_1 > 0$. Then we have $k_1 U_2(a) > k_1 U_2(b)$ equivalent to $k_1 U_2(a) + k_2 > k_1 U_2(b) + k_2$ for any k_2

Exercise 15.188-PELL

Origin:

`sp16_midterm_utilities`

PacLad and PacLass are arguing about the value of eating certain numbers of pellets. Neither knows their exact utility functions, but it is known that they are both rational and that PacLad prefers eating more pellets to eating fewer pellets. For any n, let E_n be the event of eating n pellets. So for PacLad, if $m \geq n$, then $E_m \succeq E_n$. For any n and any $k < n$, let $L_{n\pm k}$ refer to a lottery between E_{n-k} and E_{n+k}, each with probability $\frac{1}{2}$.
Reminder: For events A and B, $A \sim B$ denotes that the agent is indifferent between A and B, while $A \succ B$ denotes that A is preferred to B.

a. Are the following statements guaranteed to be true?
   (i) Under PacLad's preferences, for any n, k, $L_{n\pm k} \sim E_n$.
   (ii) Under PacLad's preferences, for any k, if $m \geq n$, then $L_{m\pm k} \succeq L_{n\pm k}$
   (iii) Under PacLad's preferences, for any k, l, if $m \geq n$, then $L_{m\pm k} \succeq L_{n\pm l}$.

b. To decouple from the previous part, suppose we are given now that under PacLad's preferences, for any n, k, $L_{n\pm k} \sim E_n$. Suppose PacLad's utility function in terms of the number of pellets eaten is U_1. For each of the following, suppose PacLass's utility function, U_2, is defined as given in terms of U_1. Choose **all** statements which are guaranteed to be true of PacLass's preferences under each definition. You should assume that all utilities are positive (greater than 0).
   (i) $U_2(n) = aU_1(n) + b$ for some positive integers a, b
      (A) $E_4 \succeq E_3$
      (B) $L_{4\pm1} \sim L_{4\pm2}$
      (C) $L_{4\pm1} \succ E_4$
   (ii) $U_2(n) = \frac{1}{U_1(n)}$
      (A) $E_4 \succeq E_3$
      (B) $L_{4\pm1} \sim L_{4\pm2}$
      (C) $L_{4\pm1} \succ E_4$

a. (i) False
All we know is that PacLad's utility is an increasing function of the number of pellets. One utility function consistent with this is $U(E_n) = 2^n$. Then the expected utility of $L_{2\pm1}$ is $\frac{1}{2}U(E_1) + \frac{1}{2}U(E_3) = \frac{1}{2}(2+8) = 5$. Since $U(E_2) = 2^2 = 4$, $L_{2\pm1} \succ E_2$. The only class of utility functions that give the guarantee that this claim is true is linear utility functions. This is a mathematical way of writing the PacLad is risk-neutral; but this is not given as an assumption in the problem. 2^n is a good counterexample because it is a risk-*seeking* utility function. A risk-avoiding utility function would have worked just as well.

(ii) True
The expected utility of $L_{m\pm k}$ is $\frac{1}{2}U(E_{m-k}) + \frac{1}{2}U(E_{m+k})$, and that of $L_{n\pm k}$ is $\frac{1}{2}U(E_{n-k}) + \frac{1}{2}U(E_{n+k})$. Since $m - k \geq n - k$, $E_{m-k} \succeq E_{n-k}$, so $U(E_{m-k}) \geq U(E_{n-k})$. Similarly, since $m + k \geq n + k$, $E_{m+k} \succeq E_{n+k}$, so $U(E_{m+k}) \geq U(E_{n+k})$. Thus $\frac{1}{2}U(E_{m-k}) + \frac{1}{2}U(E_{m+k}) \geq \frac{1}{2}U(E_{n-k}) + \frac{1}{2}U(E_{n+k})$ and there-

fore $L_{m\pm k} \succeq L_{n\pm k}$.

(iii) False

Consider again the utility function $U(E_n) = 2^n$. It is a risk-seeking utility function as mentioned in part (i), so we should expect that if this were PacLad's utility function, he would prefer a lottery with higher variance (i.e. a higher k value). So for a counterexample, we look to $L_{3\pm1}$ and $L_{3\pm2}$ (i.e. $m = n = 3$, $k = 1$, $l = 2$). The expected utility of $L_{3\pm1}$ is $\frac{1}{2}U(E_2) + \frac{1}{2}U(E_4) = \frac{1}{2}(4 + 16) = 10$. The expected utility of $L_{3\pm2}$ is $\frac{1}{2}U(E_1) + \frac{1}{2}U(E_5) = \frac{1}{2}(2 + 32) = 17 > 10$. Thus $L_{n\pm l} \succ L_{m\pm k}$. Once again, this is a statement that would only be true for a risk-neutral utility function. A risk-avoiding utility function could also have been used for a counterexample.

b. (i) A, B

The guarantee that under PacLad's preferences for any n, k, $L_{n\pm k} \sim E_n$ means that PacLad is risk-neutral and therefore his utility function is linear. An affine transformation, as this $aU_1(n) + b$ is called, of a linear function is still a linear function, so we have that PacLass's utility function is also linear and thus she is also risk-neutral. Therefore she is indifferent to the variance of lotteries with the same expectation (first option) and she does *not* prefer a lottery to deterministically being given the expectation of that lottery (**not** third option). Since a is positive, U_2 is also an increasing function (second option).

(ii) C

Since U_1 is an increasing function, U_2 is decreasing, and thus the preferences over deterministic outcomes are flipped (**not** second option).

The expected utility of $L_{4\pm1}$ is $\frac{1}{2}(U_2(3) + U_2(5)) = \frac{1}{2}\left(\frac{1}{U_1(3)} + \frac{1}{U_1(5)}\right)$. We know that U_1 is linear, so write $U_1(n) = an + b$ for some a, b. Then substituting this into this expression for $\mathbb{E}[U_2(L_{4\pm1})]$ and simplifying algebraically yields $\frac{1}{2}\left(\frac{8a+2b}{15a^2+8ab+b^2}\right) = \frac{4a+b}{15a^2+8ab+b^2}$. By the same computation for $L_{4\pm2}$, we get $\mathbb{E}[U_2(L_{4\pm2})] = \frac{4a+b}{12a^2+8ab+b^2}$. Since we only know that U_1 is increasing and linear, the only constraint on a and b is that a is positive. So let $a = 1$, $b = 0$. Then $\mathbb{E}[U_2(L_{4\pm2})] = \frac{1}{3} > \frac{4}{15} = \mathbb{E}[U_2(L_{4\pm1})]$ and thus $L_{4\pm2} \succ L_{4\pm1}$ (**not** first option). Similarly, for this U_1, $U_2(4) = \frac{1}{U_1(4)} = \frac{1}{4} < \frac{1}{3} = \mathbb{E}[U_2(L_{4\pm2})]$ and thus $L_{4\pm1} \succ E_4$ (third option).

What follows is a more general argument that could have been used to answer this question if particular numbers were not specified.

In order to determine PacLass's attitude toward risk, we take the second derivative of U_2 with respect to n. By the chain rule, $\frac{\mathrm{d}U_2(n)}{\mathrm{d}n} = \frac{\mathrm{d}U_2(n)}{\mathrm{d}U_1(n)} \cdot \frac{\mathrm{d}U_1(n)}{\mathrm{d}n}$. Since U_1 is an increasing linear function of n, $\frac{\mathrm{d}U_1(n)}{\mathrm{d}n}$ is some positive constant a, so $\frac{\mathrm{d}U_2(n)}{\mathrm{d}n} = a\frac{\mathrm{d}U_2(n)}{\mathrm{d}U_1(n)} = -a\frac{1}{(U_1(n))^2}$. Taking the derivative with respect to n again and using the chain rule yields $\frac{\mathrm{d}^2U_2(n)}{\mathrm{d}n^2} = \frac{\mathrm{d}}{\mathrm{d}U_1(n)}\left(-a\frac{1}{(U_1(n))^2}\right) \cdot \frac{\mathrm{d}U_1(n)}{\mathrm{d}n} = \frac{1}{2}a^2\frac{1}{(U_1(n))^3}$. U_1 is always positive, so this is a positive number and thus the second derivative of PacLass's utility function is everywhere positive. This means the utility function is strictly convex (equivalently "concave up"), and thus all secant lines on the plot of the curve lie above the curve itself.

In general, strictly convex utility functions are risk-seeking. To see this, consider $L_{n\pm k}$ and E_n. The expected utility of $L_{n\pm k}$ is $\frac{1}{2}U_2(n-k) + \frac{1}{2}U_2(n+k)$, which corresponds to the midpoint of the secant line drawn between the points $(n-k, U_2(n-k))$ and $(n+k, U_2(n+k))$, which both lie on the curve. That point is $(n, \mathbb{E}[U(L_{n\pm k})]) = (n, \frac{1}{2}U_2(n-k) + \frac{1}{2}U_2(n+k))$. The utility of E_n is $U(n)$, which lies on the curve at the point $(n, U_2(n))$. Since U_2 is strictly convex, the secant line lies above the curve, so we must have $\mathbb{E}[U_2(L_{n\pm k})] > U(n)$.

With that proof that PacLass is risk-seeking, we can address the remaining two options: she is not indifferent to the variance of a lottery (**not** the first option), and she prefers the lottery over the deterministic outcome (the third option).

Exercise 15.188-INSU

PacBaby just found a \$100 bill—it is the only thing she owns. Ghosts are nice enough not to kill PacBaby, but when they find PacBaby they will steal all her money. The probability of the ghosts finding PacBaby is 20%. PacBaby's utility function is $U(x) = \log(1+x)$ (this is the natural logarithm, i.e., $\log e^x = x$), where x is the total monetary value she owns. When PacBaby gets to keep the \$100 (ghosts don't find her) her utility is $U(\$100) = \log(101)$. When PacBaby loses the \$100 (per the ghosts taking it from her) her utility is $U(\$0) = \log(1+0) = 0$.

a. What is the expected utility for PacBaby?

b. Pacgressive offers theft insurance: if PacBaby pays an insurance premium of \$30, then they will reimburse PacBaby \$70 if the ghosts steal all her money (after paying \$30 in insurance, she would only have \$70 left). What is the expected utility for PacBaby if she takes insurance? For PacBaby to maximize her expected utility should she take this insurance?

c. In the above scenario, what is the expected *monetary* value of selling the insurance from Pacgressive's point of view?

a. $0.8 \times \log(101) + 0.2 \times \log(1) = 0.8 \times \log(101) + 0.2 \times 0 \approx 3.6921$

b. When taking insurance, PacBaby's expected utility equals $0.8\log(1+70) + 0.2\log(1+70) = \log(71) \approx 4.2627$. Yes, PacBaby should take the insurance.

c. The expected monetary value equals $0.8 \times 30 + 0.2 \times (-40) = 16$.

15.3 Multiattribute Utility Functions

Exercise 15.PREI

Show that if X_1 and X_2 are preferentially independent of X_3, and X_2 and X_3 are preferentially independent of X_1, then X_3 and X_1 are preferentially independent of X_2.

The complete proof is given by Keeney and Raiffa (1976).

15.4 Decision Networks

Exercise 15.APID

This exercise completes the analysis of the airport-siting problem in Figure 15.6.

a. Provide reasonable variable domains, probabilities, and utilities for the network, assuming that there are three possible sites.

b. Solve the decision problem.

c. What happens if changes in technology mean that each aircraft generates half the noise?

d. What if noise avoidance becomes three times more important?

e. Calculate the VPI for $AirTraffic$, $Litigation$, and $Construction$ in your model.

This exercise can be solved using an influence diagram package such as IDEAL. The specific values are not especially important. Notice how the tedium of encoding all the entries in the utility table cries out for a system that allows the additive, multiplicative, and other forms sanctioned by MAUT.

One of the key aspects of the fully explicit representation in Figure 15.5 is its amenability to change. By doing this exercise as well as Exercise 15.9, students will augment their appreciation of the flexibility afforded by declarative representations, which can otherwise seem tedious.

a. For this part, one could use symbolic values (high, medium, low) for all the variables and not worry too much about the exact probability values, or one could use actual numerical ranges and try to assess the probabilities based on some knowledge of the domain. Even with three-valued variables, the cost CPT has 54 entries.

b. This part almost certainly should be done using a software package.

c. If each aircraft generates half as much noise, we need to adjust the entries in the $Noise$ CPT.

d. If the noise attribute becomes three times more important, the utility table entries must all be altered. If an appropriate (e.g., additive) representation is available, then one would only need to adjust the appropriate constants to reflect the change.

e. This part should be done using a software package. Some packages may offer VPI calculation already. Alternatively, one can invoke the decision-making package repeatedly to do all the what-if calculations of best actions and their utilities, as required in the VPI formula. Finally, one can write general-purpose VPI code as an add-on to a decision-making package.

Exercise 15.ARPT

Repeat Exercise 15.APID, using the action-utility representation shown in Figure 15.7.

The information associated with the utility node in Figure 15.6 is an action-value table, and can be constructed simply by averaging out the $Deaths$, $Noise$, and $Cost$ nodes in Figure 15.5. As explained in the text , modifications to aircraft noise levels or to the importance of noise do not result in simple changes to the action-value table. Probably the easiest way to do it is to go back to the original table in Figure 15.5. The exercise therefore illustrates the tradeoffs involved in using compiled representations.

Exercise 15.ARPS

For either of the airport-siting diagrams from Exercises **??** and **??**, to which conditional probability table entry is the utility most sensitive, given the available evidence?

The answer to this exercise depends on the probability values chosen by the student.

Exercise 15.BACO

Modify and extend the Bayesian network code in the code repository to provide for creation and evaluation of decision networks and the calculation of information value.

This is relatively straightforward in the AIMA2e code release. We need to add node types for action nodes and utility nodes; we need to be able to run standard Bayes net inference on the network given fixed actions, in order to compute the posterior expected utility; and we need to write an "outer loop" that can try all possible actions to find the best. Given this, adding VPI calculation is straightforward, as described in the answer to Exercise 15.8.

Exercise 15.TXBK

Consider a student who has the choice to buy or not buy a textbook for a course. We'll model this as a decision problem with one Boolean decision node, B, indicating whether the agent chooses to buy the book, and two Boolean chance nodes, M, indicating whether the student has mastered the material in the book, and P, indicating whether the student passes the course. Of course, there is also a utility node, U. A certain student, Sam, has an additive utility function: 0 for not buying the book and -\$100 for buying it; and \$2000 for passing the course and 0 for not passing. Sam's conditional probability estimates are as follows:

$$\begin{array}{ll} P(p|b,m) = 0.9 & P(m|b) = 0.9 \\ P(p|b,\neg m) = 0.5 & P(m|\neg b) = 0.7 \\ P(p|\neg b,m) = 0.8 & \\ P(p|\neg b,\neg m) = 0.3 & \end{array}$$

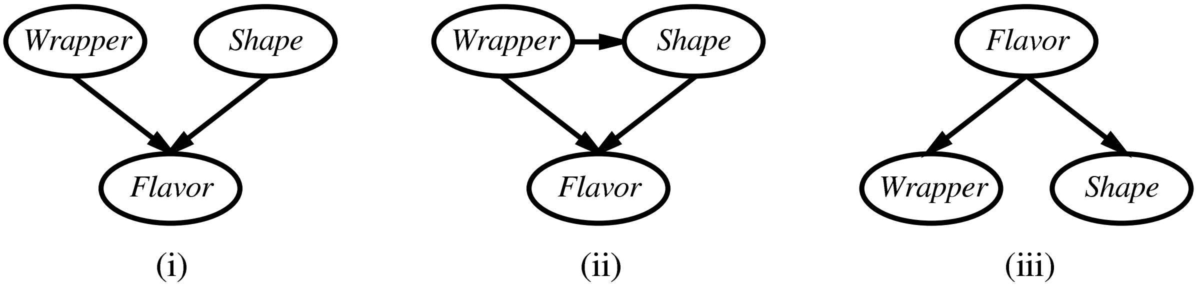
You might think that P would be independent of B given M, But this course has an open-book final—so having the book helps.

a. Draw the decision network for this problem.

b. Compute the expected utility of buying the book and of not buying it.

c. What should Sam do?

a. See Figure S15.2.



Three proposed Bayes nets for the Surprise Candy problem, **Exercise 15.**SURC.

b. For each of $B = b$ and $B = \neg b$, we compute $P(p|B)$ and thus $P(\neg p|B)$ by marginalizing out M, then use this to compute the expected utility.

$$\begin{aligned}
P(p|b) &= \sum_m P(p|b,m)P(m|b) \\
&= 0.9 \times 0.9 + 0.5 \times 0.1 \\
&= 0.86 \\
P(p|\neg b) &= \sum_m P(p|\neg b,m)P(m|\neg b) \\
&= 0.8 \times 0.7 + 0.3 \times 0.3 \\
&= 0.65
\end{aligned}$$

The expected utilities are thus:

$$\begin{aligned}
EU[b] &= \sum_p P(p|b)U(p,b) \\
&= 0.86(2000 - 100) + 0.14(-100) \\
&= 1620 \\
EU[\neg b] &= \sum_p P(p|\neg b)U(p,\neg b) \\
&= 0.65 \times 2000 + 0.14 \times 0 \\
&= 1300
\end{aligned}$$

c. Buy the book, Sam.

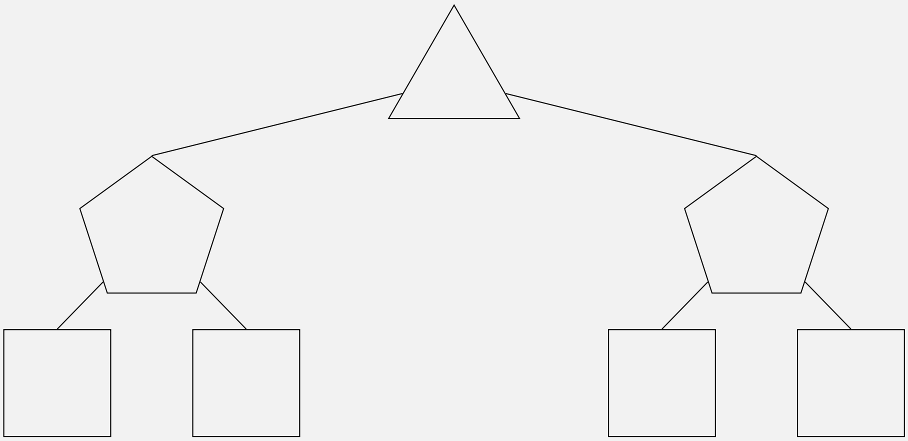
Exercise 15.188-VALE

Origin:

`su19_midterm2_decision_networks`

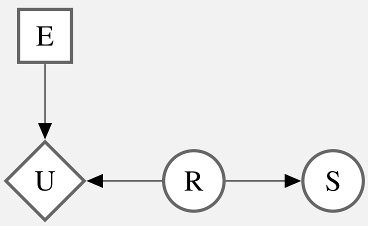
Valerie has just found a cookie on the ground. She is concerned that the cookie contains raisins, which she really dislikes but she still wants to eat the cookie. If she eats the cookie and it contains raisins she will receive a utility of -100 and if the cookie doesn't contain raisins she will receive a utility of 10. If she doesn't eat the cookie she will get 0 utility. The cookie contains raisins with probability 0.1.

a. We want to represent this decision network as an expectimax game tree. Fill in the nodes of the tree below, with the top node representing her maximizing choice.



b. Should Valerie eat the cookie?

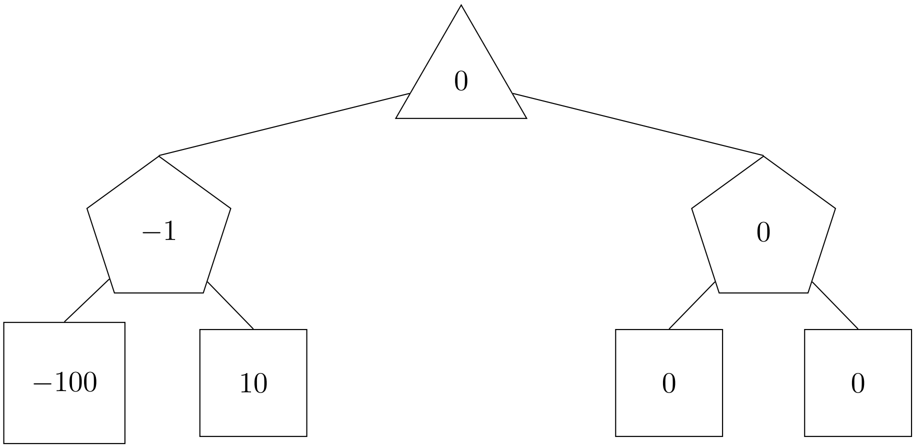
c. Valerie can now smell the cookie to judge whether it has raisins before she eats it. However, since she dislikes raisins she does not have much experience with them and cannot recognize their smell well. As a result she will incorrectly identify raisins when there are no raisins with probability 0.2 and will incorrectly identify no raisins when there are raisins with probability 0.3. This decision network can be represented by the diagram below where E is her choice to eat, U is her utility earned, R is whether the cookie contains raisins, and S is her attempt at smelling.



Valerie has just smelled the cookie and she thinks it doesn't have raisins. Write the probability, X, that the cookie has raisins given that she smelled no raisins as a simplest form fraction or decimal.

d. What is her maximum expected utility, Y given that she smelled no raisins? You can answer in terms of X or as a simplest form fraction or decimal.

e. What is the Value of Perfect Information (VPI) of smelling the cookie? You can answer in terms of X and Y or as a simplest form fraction or decimal.



a.

b. No

c. 0.04
$P(+r|-s) = \frac{P(-s|+r)P(+r)}{P(-s)} = \frac{P(-s|+r)P(+r)}{P(-s|+r)P(+r)+P(-s|-r)P(-r)} = \frac{.3*.1}{.3*.1+.8*.9} = \frac{.03}{.75} = .04$

d. $-100X + 10(1 - X)$ or 5.6
$MEU(-s) = max(MEU(eating|-s), MEU(noteating|-s)) =$
$max(P(+r|-s)*EU(eating, +r)+P(-r|-s)*EU(eating, -r), MEU(noteating)) =$
$max(X*(-100) + (1 - X)*10, 0) =$
$X * 100 + (1 - X) * 10$

e. $0.75 \times Y$ or 4.2
$VPI(S) = MEU(S) - MEU(\{\})$
$MEU(S) = P(-s)MEU(-s) + P(+s)MEU(+s)$
$P(-s) = .75$ from part (c), $MEU(-s) = Y$
$MEU(+s) = 0$ because it was better for her to not eat the raisin without knowing anything, smelling raisins will only make it more likely for the cookie to have raisins and it will still be best for her to not eat and earn a utility of 0. Note this means we do not have to calculate P(+s).
$MEU(\{\}) = 0$
$VPI(S) = .75 * Y + 0 - 0 = .75 * Y$

15.5 The Value of Information

Exercise 15.VPIX

(Adapted from Pearl (1989).) 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_1, that there is time to carry out at most one test, and that t_1 is the test of c_1 and costs \$50.

A car can be in good shape (quality q^+) or bad shape (quality q^-), and the tests might help indicate what shape the car is in. Car c_1 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_1 has a 70% chance of being in good shape.

a. Draw the decision network that represents this problem.

b. Calculate the expected net gain from buying c_1, given no test.

c. 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:
$P(pass(c_1, t_1)|q^+(c_1)) = 0.8$
$P(pass(c_1, t_1)|q^-(c_1)) = 0.35$
Use Bayes' theorem to calculate the probability that the car will pass (or fail) its test and hence the probability that it is in good (or bad) shape given each possible test outcome.

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

e. Calculate the value of information of the test, and derive an optimal conditional plan for the buyer.

This question is a simple exercise in sequential decision making, and helps in making the transition to Chapter 16. It also emphasizes the point that the value of information is computed by examining the *conditional* plan formed by determining the best action for each possible outcome of the test. It may be useful to introduce "decision trees" (as the term is used in the decision analysis literature) to organize the information in this question. (See Pearl (1988), Chapter 5.) Each part of the question analyzes some aspect of the tree. Incidentally, the question assumes that utility and monetary value coincide, and ignores the transaction costs involved in buying and selling.

a. The decision network is shown in Figure S**??**.

b. The expected net gain in dollars is

$$P(q^+)(2000 - 1500) + P(q^-)(2000 - 2200) = 0.7 \times 500 + 0.3 \times -200 = 290$$

c. The question could probably have been stated better: Bayes' theorem is used to compute $P(q^+|Pass)$, etc., whereas conditionalization is sufficient to compute $P(Pass)$.

$$\begin{aligned} P(Pass) &= P(Pass|q^+)P(q^+) + P(Pass|q^-)P(q^-) \\ &= 0.8 \times 0.7 + 0.35 \times 0.3 = 0.665 \end{aligned}$$

Using Bayes' theorem:

$$\begin{aligned} P(q^+|Pass) &= \frac{P(Pass|q^+)P(q^+)}{P(Pass)} = \frac{0.8 \times 0.7}{0.665} \approx 0.8421 \\ P(q^-|Pass) &\approx 1 - 0.8421 = 0.1579 \\ P(q^+|\neg Pass) &= \frac{P(\neg Pass|q^+)P(q^+)}{P(\neg Pass)} = \frac{0.2 \times 0.7}{0.335} \approx 0.4179 \\ P(q^-|\neg Pass) &\approx 1 - 0.4179 = 0.5821 \end{aligned}$$

d. If the car passes the test, the expected value of buying is

$$
\begin{aligned}
& P(q^+|Pass)(2000-1500) + P(q^-|Pass)(2000-2200) \\
= \; & 0.8421 \times 500 + 0.1579 \times -200 = 378.92
\end{aligned}
$$

Thus buying is the best decision given a pass. If the car fails the test, the expected value of buying is

$$
\begin{aligned}
& P(q^+|\neg Pass)(2000-1500) + P(q^-|\neg Pass)(2000-2200) \\
= \; & 0.4179 \times 500 + 0.5821 \times -200 = 92.53
\end{aligned}
$$

Buying is again the best decision.

e. Since the action is the same for both outcomes of the test, the test itself is worthless (if it is the only possible test) and the optimal plan is simply to buy the car without the test. (This is a trivial conditional plan.) For the test to be worthwhile, it would need to be more discriminating in order to reduce the probability $P(q^+|\neg Pass)$. The test would also be worthwhile if the market value of the car were less, or if the cost of repairs were more.

An interesting additional exercise is to prove the general proposition that if α is the best action for all the outcomes of a test then it must be the best action in the absence of the test outcome.

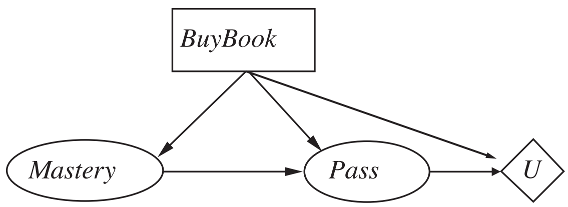


Figure S15.2 A decision network for the book-buying problem.

Exercise 15.NNVP

Recall the definition of *value of information* in Section 15.6.

a. Prove that the value of information is nonnegative and order independent.

b. Explain why it is that some people would prefer not to get some information—for example, not wanting to know the sex of their baby when an ultrasound is done.

c. A function f on sets is **submodular** if, for any element x and any sets A and B such that $A \subseteq B$, adding x to A gives a greater increase in f than adding x to B:

$$A \subseteq B \;\Rightarrow\; (f(A \cup \{x\}) - f(A)) \geq (f(B \cup \{x\}) - f(B)) .$$

Submodularity captures the intuitive notion of *diminishing returns*. Is the value of information, viewed as a function f on sets of possible observations, submodular? Prove this or find a counterexample.

a. Intuitively, the value of information is nonnegative because in the worst case one could simply ignore the information and act as if it was not available. A formal proof therefore begins by showing that this policy results in the same expected utility. The formula for the value of information is

$$VPI_E(E_j) = \left(\sum_k P(E_j = e_{jk}|E) EU(\alpha_{e_{jk}}|E, E_j = e_{jk})\right) - EU(\alpha|E)$$

If the agent does α given the information E_j, its expected utility is

$$\sum_k P(E_j = e_{jk}|E) EU(\alpha|E, E_j = e_{jk}) = EU(\alpha|E)$$

where the equality holds because the LHS is just the conditionalization of the RHS with respect to E_j. By definition,

$$EU(\alpha_{e_{jk}}|E, E_j = e_{jk}) \geq EU(\alpha|E, E_j = e_{jk})$$

hence $VPI_E(E_j) \geq 0$.

b. One explanation is that people are aware of their own irrationality and may want to avoid making a decision on the basis of the extra information. Another might be that the value of information is small compared to the value of surprise—for example, many people prefer not to know in advance what their birthday present is going to be.

c. Value of information is not submodular in general. Suppose that A is the empty set and B is the set $Y = 1$; and suppose that the optimal decision remains unchanged unless both $X = 1$ and $Y = 1$ are observed.

Exercise 15.HUNT

Figure **??** shows a myopic **function** INFORMATION-GATHERING-AGENT(t) **returns** hat chooses the next evidence variable to observe according to the ratio $VPI(E_j)/C(E_j)$, where $C(E_j)$ is the cost of observing E_j. The optimal algorithm for the treasure-hunt problem ranks the prospects according to $P(E_j)/C(E_j)$, where $P(E_j)$ is the probability that the treasure is in location j.

a. Calculate $VPI(E_j)$, assuming the treasure is valued at v.

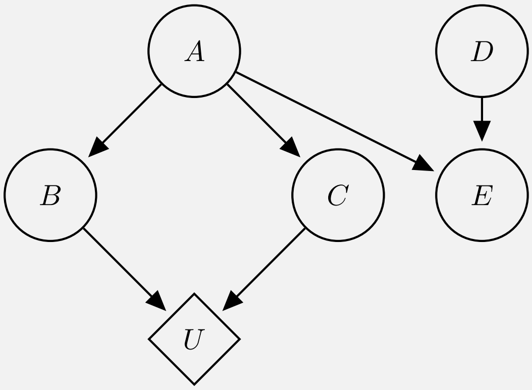
b. Determine if the two algorithms give the same behavior for the treasure-hunt problem.

c. Does the value v matter? Explain how to incorporate it into the treasure-hunt algorithm and give a proof of optimality for the new algorithm.

[[TBC]]

Exercise 15.188-DSEP

Consider a decision network with the following structure, where node U is the utility:



a. In the graph above, how do we know if a node is guaranteed to have $VPI = 0$?

b. Can any node be guaranteed to have $VPI > 0$? Why or why not?

a. Any node which is d-separated from the parents of the utility node is guaranteed to have 0 VPI. In the graph above, $VPI(D) = 0$, $VPI(E|A) = 0$, $VPI(A|B, C) = 0$,

b. No, we can never guarantee any node in the graph to have $VPI > 0$, since here we have no assumptions about the utility function. We could have $U(B, C) = 0$, in which case MEU will always be 0 regardless of the information we have and thus VPI is 0 for all nodes.

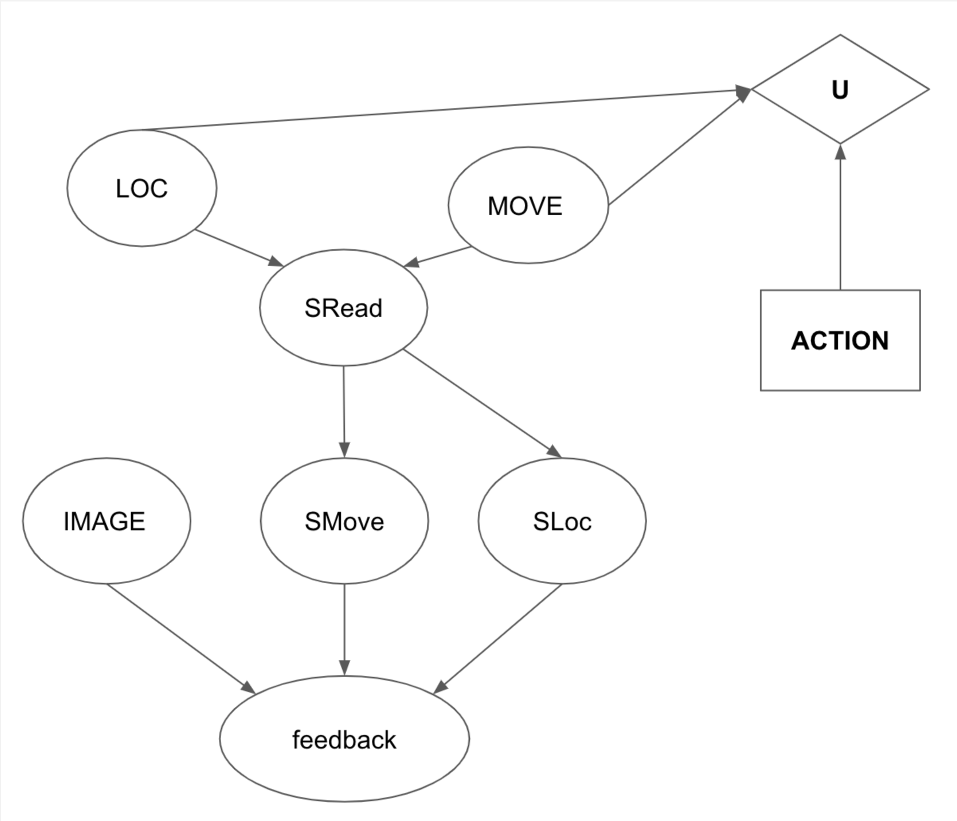
Exercise 15.188-VEHI

Origin:

`fa19_final_lindsay_vpi`

A vehicle is trying to identify the situation of the world around it using a set of sensors located around the vehicle. Each sensor reading is based off of an object's location (LOC) and an object's movement (MOVE). The sensor reading will then produce various values for its predicted location, predicted movement, and image rendered, which is then sent back to the user. The vehicle takes an action, and we assign some utility to the action based on the object's location and movement. Possible actions are MOVE TOWARDS, MOVE AWAY, and STOP.

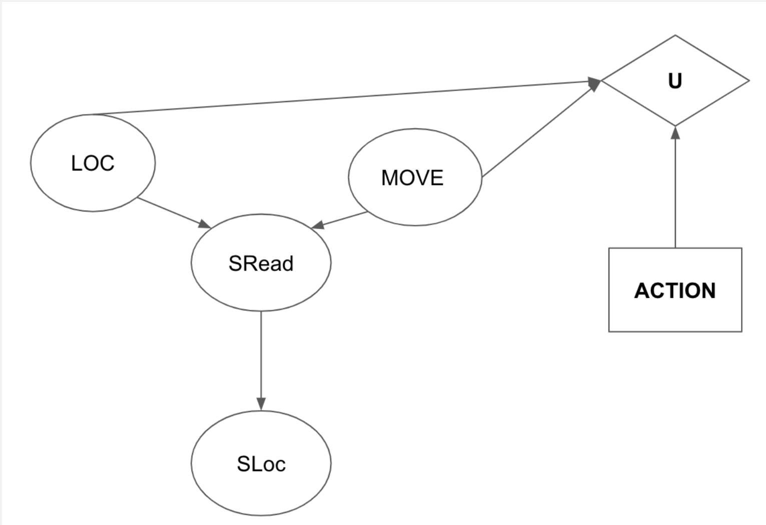
a. For part (a) only, suppose the decision network faced by the vehicle is the following.



(i) Based on the diagram above, is it **possiblely** true that $VPI(SMove, SRead) > VPI(SRead)$?

(ii) Based on the diagram above, is it **necessarily** true that $VPI(SRead) = 0$?

b. For the remainder of this problem, let's assume that your startup has less money, so we use a simpler sensor network. One possible sensor network can be represented as follows.



You have distributions of $P(LOC)$, $P(MOVE)$, $P(SRead|LOC, MOVE)$, $P(SLoc|SRead)$ and utility values $U(ACTION, LOC, MOVE)$. What is equation for determining the expected utility for some ACTION a?

c. Your colleague Bob invented a new sensor to observe values of $SLoc$.

(i) Suppose that your company had no sensors till this point. What is $VPI(SLoc)$ in terms of the given distributions?

(ii) Gaagle, an established company, wants to sell your startup a device that gives you $SRead$. Given that you already have Bob's device (that gives you $SLoc$), what is the maximum amount of money you should pay for Gaagle's device? Suppose you value $1 at 1 utility. Please simplify your answer as much as possible.

a. (i) No
$VPI(SMove, SRead) = VPI(SMove|SRead) + VPI(SRead)$, therefore we can cancel $VPI(SRead)$ from both side, which becomes asking if $VPI(SMove|SRead) > 0$.
And we can see that cannot be true, because after shading in $SRead$, there is no active path connecting $SMove$ and U.

(ii) No
$VPI(SRead)$ could be > 0 because $SRead - MOVE - U$ is an active path between $SRead$ and U

b. $EU(action) = \sum_{loc} P(loc) \sum_{move} P(move)U(action, loc, move)$
We can eliminate $SRead$ and $SLoc$ via marginalization, so they don't need to be included the expression

c. (i) Substituting in the definitions

$$\begin{aligned} VPI(SLoc) &= MEU(SLoc) - MEU(\{\}) \\ &= (\sum_{sloc} P(sloc)\, MEU(SLoc = sloc)) - \max_a EU(a) \end{aligned}$$

(ii) We should pay

$$\begin{aligned} VPI(SRead|SLoc) &= VPI(SRead, SLoc) - VPI(SLoc) \\ &= VPI(SRead) + VPI(SLoc|SRead) - VPI(SLoc) \\ &= VPI(SRead) + 0 - VPI(SLoc) \\ &= VPI(SRead) - VPI(SLoc) \end{aligned}$$

Exercise 15.188-MONT

Origin:

`su16_midterm2_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?

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?

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

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?

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?

a. $(1000 * \frac{1}{n} + 10 * \frac{n-1}{n})$ 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. $EU(keep) = 10 + 990 * \frac{1}{n}$
$EU(switch) = 10 + 990 * \frac{(n-1)}{n*(n-k-1)}$
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. $990 * \frac{1}{n} * \frac{k}{n-k-1}$
The formula for VPI is $MEU(e) - MEU(\{\})$. 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. $\frac{990}{n}$
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. $d = \sqrt{\frac{990}{n}}$
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.

15.6 Unknown Preferences

Exercise 15.OFFS

In the off-switch problem (Section 15.7.2), we have assumed that Harriet acts rationally. Suppose instead that she is **Boltzmann-rational**, i.e., she follows a randomized policy that chooses action x with a softmax probability:

$$\pi(x) = \frac{e^{\beta U_x)}}{\sum_y e^{\beta U_y)}}\,.$$

a. Derive the general condition for Robbie to defer to Harriet, assuming that Robbie's prior for Harriet's utility for the immediate action a is $P(u)$.

b. Determine the minimum value of β such that Robbie defers to Harriet in the example of Figure 15.11.

[[TBC]]