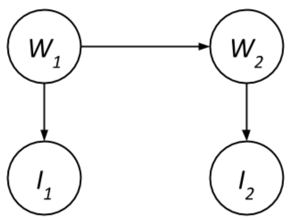
CS188 Fall 2017 Section 9: Sampling and Decision Nets

1 Sampling and Dynamic Bayes Nets

We would like to analyze people's ice cream eating habits on sunny and rainy days Suppose we consider the weather, along with a person's ice-cream eating, over the span of two days. We'll have four random variables: W_1 and W_2 stand for the weather on days 1 and 2, which can either be rainy $\texttt{R}$ or sunny $\texttt{S}$, and the variables I_1 and I_2 represent whether or not the person ate ice cream on days 1 and 2, and take values $\texttt{T}$ (for truly eating ice cream) or $\texttt{F}$. We can model this as the following Bayes Net with these probabilities.



W_1	$P(W_1)$
S	0.6
R	0.4

W_1	W_2	$P(W_2\|W_1)$
S	S	0.7
S	R	0.3
R	S	0.5
R	R	0.5

W	I	$P(I\|W)$
S	T	0.9
S	F	0.1
R	T	0.2
R	F	0.8

Suppose we produce the following samples of (W_1, I_1, W_2, I_2) from the ice-cream model:

~~R, F, R, F~~ ~~R, F, R, F~~ ~~S, F, S, T~~ ~~S, T, S, T~~ S, T, R, F
~~R, F, R, T~~ ~~S, T, S, T~~ ~~S, T, S, T~~ S, T, R, F ~~R, F, S, T~~

1. What is $\widehat{P}(W_2 = \texttt{R})$, the probability that sampling assigns to the event $W_2 = \texttt{R}$?
   Number of samples in which $W_2 = \texttt{R}$: 5. Total number of samples: 10. Answer $5/10 = 0.5$.

2. Cross off samples above which are rejected by rejection sampling if we're computing $P(W_2|I_1 = \texttt{T}, I_2 = \texttt{F})$.

   Rejection sampling seems to be wasting a lot of effort, so we decide to switch to likelihood weighting. Assume we generate the following six samples given the evidence $I_1 = \texttt{T}$ and $I_2 = \texttt{F}$:

$$(W_1, I_1, W_2, I_2) = \Big\{(\texttt{S},\texttt{T},\texttt{R},\texttt{F}), (\texttt{R},\texttt{T},\texttt{R},\texttt{F}), (\texttt{S},\texttt{T},\texttt{R},\texttt{F}), (\texttt{S},\texttt{T},\texttt{S},\texttt{F}), (\texttt{S},\texttt{T},\texttt{S},\texttt{F}), (\texttt{R},\texttt{T},\texttt{S},\texttt{F})\Big\}$$

3. What is the weight of the first sample (S, T, R, F) above?

   The weight given to a sample in likelihood weighting is

$$\prod_{\text{Evidence variables } e} \Pr(e|\text{Parents}(e)).$$

   In this case, the evidence is $I_1 = \texttt{T}, I_2 = \texttt{F}$. The weight of the first sample is therefore

$$w = \Pr(I_1 = \texttt{T}|W_1 = \texttt{S}) \cdot \Pr(I_2 = \texttt{F}|W_2 = \texttt{R}) = 0.9 \cdot 0.8 = 0.72$$

4. Use likelihood weighting to estimate $P(W_2|I_1 = \texttt{T}, I_2 = \texttt{F})$.
   The sample weights are given by

(W_1, I_1, W_2, I_2)	w
S, T, R, F	0.72
R, T, R, F	0.16
S, T, R, F	0.72

(W_1, I_1, W_2, I_2)	w
S, T, S, F	0.09
S, T, S, F	0.09
R, T, S, F	0.02

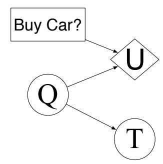
   To compute the probabilities, we thus normalize the weights and find

$$\widehat{P}(W_2 = \texttt{R}|I_1 = \texttt{T}, I_2 = \texttt{F}) = \frac{0.72 + 0.16 + 0.72}{0.72 + 0.16 + 0.72 + 0.09 + 0.09 + 0.02} = 0.889$$

$$\widehat{P}(W_2 = \texttt{S}|I_1 = \texttt{T}, I_2 = \texttt{F}) = 1 - 0.889 = 0.111.$$

2 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 buyers 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 $\neg$buy)

4. 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.