# Section Handout 9 V2 Solutions

## Sampling

Suppose we want to evaluate $P(Q|E)$ where $Q$ are the query variables and $E$ are the evidence variables.

**Prior Sampling**: Draw samples from the Bayes net by sampling the parents and then sampling the children given the parents. $P(Q|E) \approx \frac{count(QandE)}{count(E)}$.

**Rejection Sampling**: Like prior sampling, but ignore all samples that are inconsistent with the evidence.

**Likelihood Weighting**: Fix the evidence variables, and weight each sample by the probability of the evidence variables given their parents.

**Gibbs Sampling**:

1. Fix evidence.

2. Initialize other variables randomly

3. Repeat:

   (a) Choose non-evidence variable $X$.

   (b) Resample $X$ from $P(X|markovblanket(X))$

## 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.

# HMMs

State variables $W_t$ and observation (evidence) variables ($O_t$), which are supposed to be shaded below. Transition model $P(W_{t+1}|W_t)$. Sensor model $P(O_t|W_t)$. The joint distribution of the HMM can be factorized as

$$P(W_1, ..., W_T, O_1, ....O_T) = P(W_1) \prod_{t=1}^{T-1} P(W_{t+1}|W_t) \prod_{t=1}^{T} P(O_t|W_t) \tag{1}$$
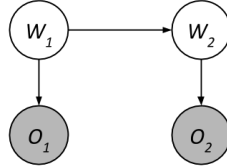


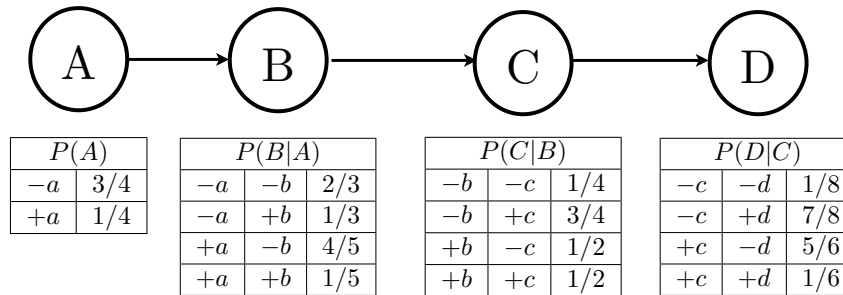Define the following belief distribution

- $B(W_t) = P(W_t|O_1, ..., O_t)$: Belief about state $W_t$ given all the observations up until (and including) timestep $t$.

- $B'(W_t) = P(W_t|O_1, ..., O_{t-1})$: Belief about state $W_t$ given all the observations up until (but not including) timestep $t$.

# Forward Algorithm

- *Prediction update*: $B'(W_{t+1}) = \sum_{w_t} P(W_{t+1}|w_t)B(w_t)$
- *Observation update*: $B(W_{t+1}) \propto P(O_{t+1}|W_{t+1})B'(W_{t+1})$

# Q1. Bayes' Nets Sampling

Assume the following Bayes' net, and the corresponding distributions over the variables in the Bayes' net:



| P(A) | |
|---|---|
| $-a$ | 3/4 |
| $+a$ | 1/4 |

| P(B\|A) | | |
|---|---|---|
| $-a$ | $-b$ | 2/3 |
| $-a$ | $+b$ | 1/3 |
| $+a$ | $-b$ | 4/5 |
| $+a$ | $+b$ | 1/5 |

| P(C\|B) | | |
|---|---|---|
| $-b$ | $-c$ | 1/4 |
| $-b$ | $+c$ | 3/4 |
| $+b$ | $-c$ | 1/2 |
| $+b$ | $+c$ | 1/2 |

| P(D\|C) | | |
|---|---|---|
| $-c$ | $-d$ | 1/8 |
| $-c$ | $+d$ | 7/8 |
| $+c$ | $-d$ | 5/6 |
| $+c$ | $+d$ | 1/6 |

**(a)** You are given the following samples:

$$
\begin{array}{cccc}
+a & +b & -c & -d \\
+a & -b & +c & -d \\
-a & +b & +c & -d \\
-a & -b & +c & -d
\end{array}
\qquad
\begin{array}{cccc}
+a & -b & -c & +d \\
+a & +b & +c & -d \\
-a & +b & -c & +d \\
-a & -b & +c & -d
\end{array}
$$

**(i)** Assume that these samples came from performing Prior Sampling, and calculate the sample estimate of $P(+c)$.

5/8

**(ii)** Now we will estimate $P(+c \mid +a, -d)$. Above, clearly cross out the samples that would **not** be used when doing Rejection Sampling for this task, and write down the sample estimate of $P(+c \mid +a, -d)$ below.

2/3

**(b)** Using Likelihood Weighting Sampling to estimate $P(-a \mid +b, -d)$, the following samples were obtained. Fill in the weight of each sample in the corresponding row.

**Sample** **Weight**

$-a \quad +b \quad +c \quad -d \quad$ $P(+b \mid -a)P(-d \mid +c) = 1/3 * 5/6 = 5/18 = 0.277$

$+a \quad +b \quad +c \quad -d \quad$ $P(+b \mid +a)P(-d \mid +c) = 1/5 * 5/6 = 5/30 = 1/6 = 0.17$

$+a \quad +b \quad -c \quad -d \quad$ $P(+b \mid +a)P(-d \mid -c) = 1/5 * 1/8 = 1/40 = 0.025$

$-a \quad +b \quad -c \quad -d \quad$ $P(+b \mid -a)P(-d \mid -c) = 1/3 * 1/8 = 1/24 = 0.042$

**(c)** From the weighted samples in the previous question, estimate $P(-a \mid +b, -d)$.

$\frac{5/18+1/24}{5/18+5/30+1/40+1/24} = 0.625$

**(d)** Which query is better suited for likelihood weighting, $P(D \mid A)$ or $P(A \mid D)$? Justify your answer in one sentence.

$P(D \mid A)$ is better suited for likelihood weighting sampling, because likelihood weighting conditions only on upstream evidence.

**(e)** Recall that during Gibbs Sampling, samples are generated through an iterative process.

Assume that the only evidence that is available is $A = +a$. Clearly fill in the circle(s) of the sequence(s) below that could have been generated by Gibbs Sampling.

Sequence 1

|   | | | | |
|---|---|---|---|---|
| 1 : | $+a$ | $-b$ | $-c$ | $+d$ |
| 2 : | $+a$ | $-b$ | $-c$ | $+d$ |
| 3 : | $+a$ | $-b$ | $+c$ | $+d$ |

Sequence 2

|   | | | | |
|---|---|---|---|---|
| 1 : | $+a$ | $-b$ | $-c$ | $+d$ |
| 2 : | $+a$ | $-b$ | $-c$ | $-d$ |
| 3 : | $-a$ | $-b$ | $-c$ | $+d$ |

Sequence 3

|   | | | | |
|---|---|---|---|---|
| 1 : | $+a$ | $-b$ | $-c$ | $+d$ |
| 2 : | $+a$ | $-b$ | $-c$ | $-d$ |
| 3 : | $+a$ | $+b$ | $-c$ | $-d$ |

Sequence 4

|   | | | | |
|---|---|---|---|---|
| 1 : | $+a$ | $-b$ | $-c$ | $+d$ |
| 2 : | $+a$ | $-b$ | $-c$ | $-d$ |
| 3 : | $+a$ | $+b$ | $-c$ | $+d$ |

<span style="color:red">Gibbs sampling updates one variable at a time and never changes the evidence.</span>

<span style="color:red">The first and third sequences have at most one variable change per row, and hence could have been generated from Gibbs sampling. In sequence 2, the evidence variable is changed. In sequence 4, the second and third samples have both $B$ and $D$ changing.</span>

# 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 buyer's estimate is that $c$ has 70% chance of being in good shape. The Decision Network is shown below.
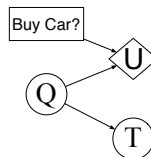


**(a)** Calculate the expected net gain from buying car c, given no test.

<span style="color:red">$$EU(\text{buy}) = P(Q = +q) \cdot U(+q, \text{buy}) + P(Q = \neg q) \cdot U(\neg q, buy)$$
$$= .7 \cdot 500 + 0.3 \cdot -200 = 290$$</span>

**(b)** 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.

$$P(T = \text{pass}) = \sum_q P(T = \text{pass}, Q = q)$$

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

$$= 0.69$$

$$P(T = fail) = 0.31$$

$$P(Q = +q|T = \text{pass}) = \frac{P(T = \text{pass}|Q = +q)P(Q = +q)}{P(T = \text{pass})}$$

$$= \frac{0.9 \cdot 0.7}{0.69} = \frac{21}{23} \approx 0.91$$

$$P(Q = +q|T = \text{fail}) = \frac{P(T = \text{fail}|Q = +q)P(Q = +q)}{P(T = \text{fail})}$$

$$= \frac{0.1 \cdot 0.7}{0.31} = \frac{7}{31} \approx 0.22$$

**(c)** Calculate the optimal decisions given either a pass or a fail, and their expected utilities.

$$EU(\text{buy}|T = pass) = P(Q = +q|T = \text{pass})U(+q, \text{buy}) + P(Q = \neg q|T = \text{pass})U(\neg q, \text{buy})$$

$$\approx 0.91 \cdot 500 + 0.09 \cdot (-200) \approx 437$$

$$EU(\text{buy}|T = \text{fail}) = P(Q = +q|T = \text{fail})U(+q, \text{buy}) + P(Q = \neg q|T = \text{fail})U(\neg q, \text{buy})$$

$$\approx 0.22 \cdot 500 + 0.78 \cdot (-200) = -46$$

$$EU(\neg\text{buy}|T = pass) = 0$$

$$EU(\neg\text{buy}|T = fail) = 0$$

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

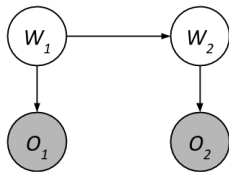**(d)** Calculate the value of (perfect) information of the test. Should the buyer pay for a test?

$$VPI(T) = \left(\sum_t P(T = t)MEU(T = t)\right) - MEU(\phi)$$

$$= 0.69 \cdot 437 + 0.31 \cdot 0 - 290 \approx 11.53$$

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

# 3 HMMs

Consider the following Hidden Markov Model. $O_1$ and $O_2$ are supposed to be shaded.



| $W_1$ | $P(W_1)$ |
|---|---|
| 0 | 0.3 |
| 1 | 0.7 |

| $W_t$ | $W_{t+1}$ | $P(W_{t+1}|W_t)$ |
|---|---|---|
| 0 | 0 | 0.4 |
| 0 | 1 | 0.6 |
| 1 | 0 | 0.8 |
| 1 | 1 | 0.2 |

| $W_t$ | $O_t$ | $P(O_t|W_t)$ |
|---|---|---|
| 0 | a | 0.9 |
| 0 | b | 0.1 |
| 1 | a | 0.5 |
| 1 | b | 0.5 |

Suppose that we observe $O_1 = a$ and $O_2 = b$.
Using the forward algorithm, compute the probability distribution $P(W_2|O_1 = a, O_2 = b)$ one step at a time.

**(a)** Compute $P(W_1, O_1 = a)$.

$P(W_1, O_1 = a) = P(W_1)P(O_1 = a|W_1)$
$P(W_1 = 0, O_1 = a) = (0.3)(0.9) = 0.27$
$P(W_1 = 1, O_1 = a) = (0.7)(0.5) = 0.35$

**(b)** Using the previous calculation, compute $P(W_2, O_1 = a)$.

$P(W_2, O_1 = a) = \sum_{w_1} P(w_1, O_1 = a)P(W_2|w_1)$
$P(W_2 = 0, O_1 = a) = (0.27)(0.4) + (0.35)(0.8) = 0.388$
$P(W_2 = 1, O_1 = a) = (0.27)(0.6) + (0.35)(0.2) = 0.232$

**(c)** Using the previous calculation, compute $P(W_2, O_1 = a, O_2 = b)$.

$P(W_2, O_1 = a, O_2 = b) = P(W_2, O_1 = a)P(O_2 = b|W_2)$
$P(W_2 = 0, O_1 = a, O_2 = b) = (0.388)(0.1) = 0.0388$
$P(W_2 = 1, O_1 = a, O_2 = b) = (0.232)(0.5) = 0.116$

**(d)** Finally, compute $P(W_2|O_1 = a, O_2 = b)$.

Renormalizing the distribution above, we have
$P(W_2 = 0|O_1 = a, O_2 = b) = 0.0388/(0.0388 + 0.116) \approx 0.25$
$P(W_2 = 1|O_1 = a, O_2 = b) = 0.116/(0.0388 + 0.116) \approx 0.75$