CMSI 3300 – Classwork 4

Instructions:
This worksheet will not only provide you with practice problems for your upcoming exam, but will add to your deep understanding of the mechanics of many probabilistic reasoning systems, managing priorities through utility theory, and curiosity through VPI.

- Provide answers to each of the following questions and write your responses in the blanks. If you are expected to show your work in arriving at a particular solution, space will be provided for you.
- Place the names of your group members below:

Group Members:

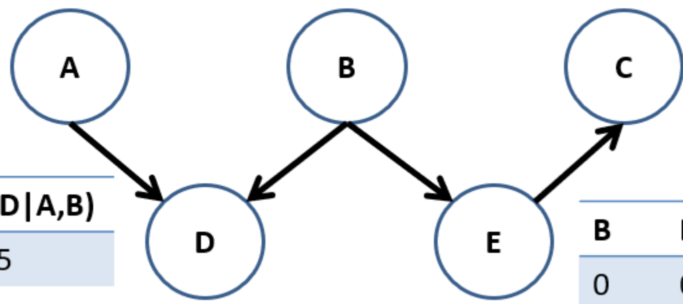
1. Julian Mazzier
2. Cameron Scolari
3. ______________________________

Problem 1 – Bayesian Network Exact Inference

Consider the following Bayesian Network and use it to answer the questions that follow.

B	P(B)
0	0.6

A	P(A)
0	0.8



E	C	P(C|E)
0	0	0.4
1	0	0.2

A	B	D	P(D|A,B)
0	0	0	0.5
0	1	0	0.7
1	0	0	0.8
1	1	0	0.1

B	E	P(E|B)
0	0	0.9
1	0	0.4

While examining Exact Inference in Bayesian Networks, we saw some methods for *simplifying* queries and the resulting computations that can yield large performance improvements when implemented. E.g., variables whose CPTs never affect the query outcome can be ignored.

That rule: *To computer $P(Q|e)$, you require ONLY the CPTs of variables that are NOT ancestors of EITHER the Query Q NOR the evidence e. i.e., any hidden variable not meeting this condition are marginalized during step 2 of enumeration inference.*

For example, to answer the query $P(A|B = b, D = d)$, we do not need the CPTs for C nor E because they are not ancestors of A, B, or D in the network. A brief proof using the normalizing constant $\alpha = \frac{1}{P(e)}$:

CPTs Used: ✓ A ✓ B C ✓ D E

Justification:

$$\begin{aligned}
P(A|b,d) &= \alpha \sum_e \sum_c P(A,b,c,d,e) \\
&= \alpha \sum_e \sum_c P(A)P(d|A,b)P(b)P(e|b)P(c|e) \\
&= \alpha\, P(A)P(d|A,b)P(b) \sum_e P(e|b) \sum_c P(c|e) \\
&= \alpha P(A)P(d|A,b)P(b)
\end{aligned}$$

[!] Note how the above is more computationally desirable – no loops / sums required!

1. Using the Bayesian Network and enumeration inference simplification on the previous page, find the solution to the query:

$$P(A = 0|B = 1, D = 1)$$

1.1. Label variables using the notation Q = queries, e = evidence, and Y = hidden vars.

Q = { $A=0$ }	e = { $B=1, D=1$ }	Y = { C, E }

1.2. Find $P(Q, e) = \sum_y P(Q, e, y)$, substituting into this formula the variable labels in 1.1.

$$P(A=0, B=1, D=1, C=c, E=e) = \sum_e \sum_c P(A)P(d|A,B)P(B)P(C|E)P(E|B)$$
$$= P(A)P(d|A,B)P(B)\sum_e P(E|B) \sum_c P(C|E)$$

Since B is given, A is independent from E and C

$$P(A=0, B=1, D=1) = P(A)P(d|A,B)P(B) = P(A=0)P(D=1|A=0,B=1)P(B=1)$$
$$= 0.8 \cdot 0.3 \cdot 0.4 = 0.096$$

1.3. Find $P(e) = \sum_q P(Q, e)$.

$$P(B=1, D=1) = P(A=0) \cdot P(D=1|A=0,B=1) \cdot P(B=1) + P(A=1) \cdot P(D=1|A=1,B=1) \cdot P(B=1)$$
$$= 0.8 \cdot 0.3 \cdot 0.4 + 0.2 \cdot 0.9 \cdot 0.4 = 0.096 + 0.072 = 0.168$$

1.4. Solve the original query, $P(Q|e) = P(Q, e)/P(e)$

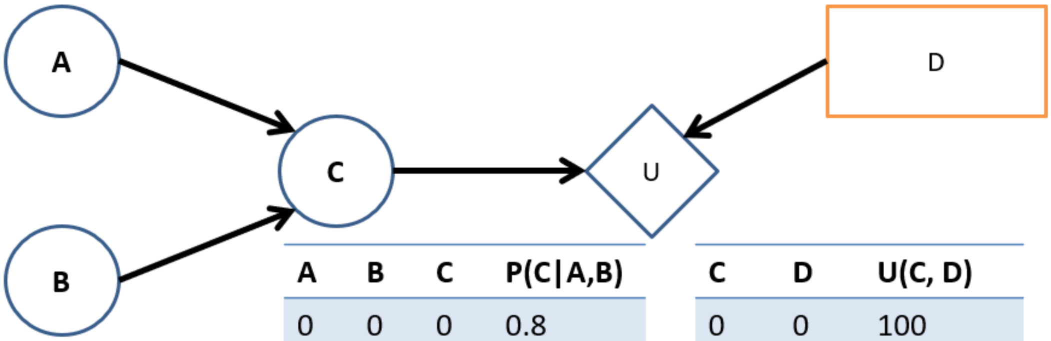
$$\frac{0.096}{0.168} \approx 0.571$$

Problem 2 – Decision Networks & MEU

Use the following Decision Network with chance nodes A, B, C, decision node $D \in \{0,1\}$, and utility node U to answer the questions that follow.

A	P(A)
0	0.2

B	P(B)
0	0.5



A	B	C	P(C|A,B)
0	0	0	0.8
0	1	0	0.6
1	0	0	0.1
1	1	0	0.5

C	D	U(C, D)
0	0	100
0	1	20
1	0	40
1	1	80

2. Find the $MEU(B = 0)$ (i.e., the Maximum Expected Utility with evidence $B = 0$). The individual steps for computing this are outlined for you below.

2.1. Begin by finding $P(S = s|do(D = d), e)\ \forall\, s \in S, d \in D$ using Enumeration Inference for decision nets.

2.1.1. (lol nested problem depth) Label your variables for $P(S = s|do(D = d), e)$

Q = { C }	e = { B = 0 }	D = { D }	Y = { A }

2.1.2. Find $P\big(Q, e|do(D)\big) = \sum_y P\big(Q, e, y|do(D)\big)$ (just the formula, we'll plug n' chug next)

$$P(C, B=0, A=0 \mid D) + P(C, B=0, A=1 \mid D)$$
$$= P(C \mid A=0, B=0, D) \cdot P(A=0 \mid D) \cdot P(B=0 \mid D) + P(C \mid A=1, B=0, D) \cdot P(A=1 \mid D) \cdot P(B=0 \mid D)$$
$$= P(B=0)\left[P(C \mid A=0, B=0) P(A=0) + P(C \mid A=1, B=0) P(A=1)\right]$$

2.1.3. Using your formula from the previous problem (which should have found that the chance nodes were independent of the decision), find $P(Q, e|do(D))\ \forall\ q \in Q$

$$P(B=0)\left[P(C|A=0,B=0)P(A=0)+P(C|A=1,B=0)P(A=1)\right]$$

$$C=0: 0.5[0.8\cdot 0.2+0.1\cdot 0.8]=0.5[0.16+0.08]=0.5(0.24)=0.12$$

$$C=1: 0.5[0.2\cdot 0.2+0.9\cdot 0.8]=0.5[0.04+0.72]=0.5(0.76)=0.38$$

2.1.4. Using your answer to the previous problem, find $P(e|do(D)) = \sum_q P(Q=q, e|do(D))$

$$P(B=0)=P(B=0,C=0)+P(B=0,C=1)=0.12+0.38=0.5$$

2.1.5. Using your answers to the previous problems, find $P(Q|e, do(D)) = \frac{P(Q,e|do(D))}{P(e|do(D))}$

$$C=0: P(C=0|B=0)=\frac{0.12}{0.5}=0.24$$

$$C=1: P(C=1|B=0)=\frac{0.38}{0.5}=0.76$$

2.1.2. Using your result for $P(S|do(A), e)$ from the previous problem, find $EU(a|e)\ \forall\ a$.

$$D=0: EU(a|e)=EU(D=0|B=0)=P(C=0|B=0)\cdot U(C=0,D=0)+P(C=1|B=0)\cdot U(C=1,D=0)$$
$$=0.24\cdot 100+0.76\cdot 40=24+30.4=54.4$$

$$D=1: EU(a|e)=EU(D=1|B=0)=P(C=0|B=0)\cdot U(C=0,D=1)+P(C=1|B=0)\cdot U(C=1,D=1)$$
$$=0.24\cdot 20+0.76\cdot 80=4.8+60.8=65.6$$

2.1.3. Using your EU calculations above, find $MEU(B=0)$ and $d^* = argmax_d EU(d|B=0)$.

$\max EU(54.4, 65.6)=65.6 \quad d^*=1$

Problem 3 – Value of Perfect Information

Using the network and your answer from the previous problem, we're going to compute the Value of Perfect Information (VPI) of knowing the state of variable A when $B = 0$ is given. Let's do so step-by-step:

3.1. Find $MEU(A = 0, B = 0)$ (hint: $P(S|do(A), e)$ is an immediately answerable query here).

D=0: $EU(a|e) = P(C=0|A=0,B=0) \cdot U(C=0,D=0) + P(C=1|A=0,B=0) \cdot U(C=1,D=0)$
$= 0.8 \cdot 100 + 0.2 \cdot 40 = 80 + 8 = 88$

D=1: $EU(a|e) = P(C=0|A=0,B=0) \cdot U(C=0,D=1) + P(C=1|A=0,B=0) \cdot U(C=1,D=1)$
$= 0.8 \cdot 20 + 0.2 \cdot 80 = 16 + 16 = 32$

$\max EU(88,32) = 88$

3.2. Using your answer to 3.1 and knowledge that $MEU(A = 1, B = 0) = 74$ (freebee!) find $MEU(A, B = 0)$.

A=1: D=0: $P(C=0|A=1,B=0) \cdot U(C=0,D=0) + P(C=1|A=1,B=0) \cdot U(C=1,D=0)$
$= 0.1 \cdot 100 + 0.9 \cdot 40 = 10 + 36 = 46$

D=1: $P(C=0|A=1,B=0) \cdot U(C=0,D=1) + P(C=1|A=1,B=0) \cdot U(C=1,D=1)$
$= 0.1 \cdot 20 + 0.9 \cdot 80 = 2 + 72 = 74$ ✓

$MEU(A,B=0) = P(A=0) \cdot MEU(A=0) + P(A=1) \cdot MEU(A=1) = 0.2 \cdot 88 + 0.8 \cdot 74 = 17.6 + 59.2 = 76.8$

3.3. Compute the $VPI(A|B = 0)$.

$VPI(A|B=0) = MEU(A,B=0) - MEU(B=0) = 76.8 - 65.6 = 11.2$

3.4. If the utility scores represent dollar amounts, what would be a fair price for A when $B = 0$?

from \$0 – \$11.2 so \$11.2