CS 188
Fall 2021

Introduction to
Artificial Intelligence

Challenge Q6 HW8

- **Due:** Tuesday 11/9 at 10:59pm.
- **Policy:** Can be solved in groups (acknowledge collaborators) but must be submitted individually.
- **Make sure to show all your work and justify your answers**.
- **Note:** This is a typical exam-level question. On the exam, you would be under time pressure, and have to complete this question on your own. We strongly encourage you to first try this on your own to help you understand where you currently stand. Then feel free to have some discussion about the question with other students and/or staff, before independently writing up your solution.
- Your submission on Gradescope should be a PDF that matches this template. Each page of the PDF should align with the corresponding page of the template (page 1 has name/collaborators, question begins on page 2.). **Do not reorder, split, combine, or add extra pages**. The intention is that you print out the template, write on the page in pen/pencil, and then scan or take pictures of the pages to make your submission. You may also fill out this template digitally (e.g. using a tablet.)

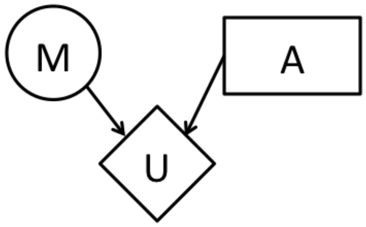
First name	NAMEERA
Last name	FAISAL AKHTAR
SID	3034244256
Collaborators	SARA IMAM

For staff use only:

Q6.	Decision Networks and HMMs	/29

Q6. [29 pts] Decision Networks and HMMs

Assume that you want to watch a film M that can either be great $+m$ or pretty bad $-m$. You can either watch the film in a theater or at home by renting it. This is controlled by your choice A. Consider the following decision network and tables:



M	P(M)
+m	0.5
-m	0.5

M	A	U(M,A)
+m	*theater*	100
-m	*theater*	10
+m	*rent*	80
-m	*rent*	40

Figure 1: Decision network and tables.

6.1) (5 pts) Compute the following quantities. EU and MEU stand for expected and maximum expected utility respectively.

$EU(theater) =$ $P(+m)\,U(+m, \text{theater}) + P(-m)\,U(-m, \text{theater}) = (0.5)(100) + (0.5)(10) = 50 + 5 \boxed{= 55}$

$EU(rent) =$ $P(+m)\,U(+m, \text{rent}) + P(-m)\,U(-m, \text{rent}) = (0.5)(80) + (0.5)(40) = 40 + 20 \boxed{= 60}$

$MEU(\emptyset) =$ $\max\{EU(\text{theater}), EU(\text{rent})\} = \max\{55, 60\} \boxed{= 60}$

$\text{argmax}_A EU(A) =$ $60 > 55 \therefore \boxed{\text{rent}}$

6.2) (5 pts) You would like obtain more information about whether the film is good or not. For that, we introduce another variable F which designates the "fullness" (how sold-out the tickets are) in the theaters. This variable is affected by another variable S which designates possible Covid-19 restrictions. The prior of M and the utilities are the same as before. Assuming that both F and S are binary, consider the following network and tables:

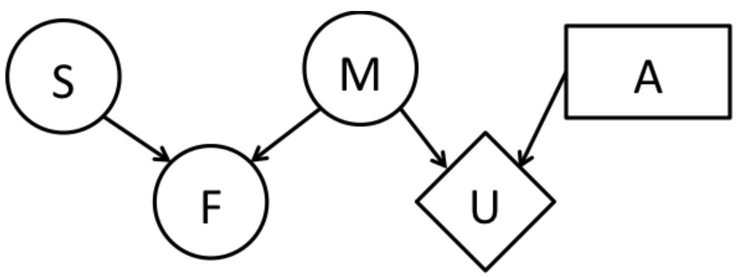


Figure 2: Decision network.

S	M	F	$P(F\|S,M)$
+s	+m	+f	0.6
+s	+m	-f	0.4
+s	-m	+f	0.0
+s	-m	-f	1.0

S	M	F	$P(F\|S,M)$
-s	+m	+f	1.0
-s	+m	-f	0.0
-s	-m	+f	0.3
-s	-m	-f	0.7

S	$P(S)$
+s	0.2
-s	0.8

Figure 3: Tables.

We want to figure out the value of revealing the Covid-19 restrictions S. Compute the values of the following quantities.

$EU(theater|+s) =$ EU(theater) $\boxed{= 55}$

$EU(rent|+s) =$ EU(rent) $\boxed{= 60}$

$MEU(\{+s\}) =$ max{EU(theater | +s), EU(rent | +s)} = max{55, 60} $\boxed{= 60}$

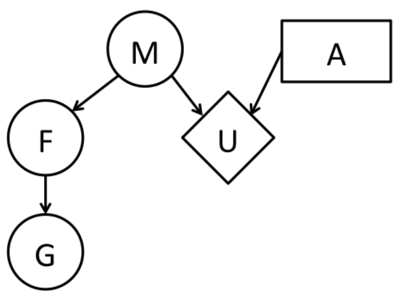
Optimal action for $+s$ = $\boxed{\text{rent}}$ 60 > 55.

$MEU(\{-s\}) =$ max{EU(theater | −s), EU(rent | −s)} = max{55, 60} $\boxed{= 60}$

Optimal action for $-s$ = $\boxed{\text{rent}}$ 60 > 55

$VPI(S) =$ MEU(S) − MEU(∅) = 60 − 60 $\boxed{= 0}$

6.3) (5 pts) Now let's assume that we want to determine the "fullness" of the theaters F without using information about the Covid-19 restrictions but using information about the garbage disposal G outside the theaters. This new variable is also binary and the new decision network and tables are as follows:



G	F	$P(G\|F)$
+g	+f	0.8
-g	+f	0.2
+g	-f	0.3
-g	-f	0.7

M	P(M)
+m	0.5
-m	0.5

F	M	$P(F\|M)$
+f	+m	0.92
-f	+m	0.08
+f	-m	0.24
-f	-m	0.76

M	A	U(M,A)
+m	*theater*	100
-m	*theater*	10
+m	*rent*	80
-m	*rent*	40

Figure 4: Decision network and tables.

We also provide you with the following extra tables that might help you to answer the question.

EU(theater | −g) = P(+m | −g) U(+m, theater) + P(−m | −g) U(−m, theater)
= (0.293)(100) + (0.707)(10)
= 36.37

EU(theater | +g) = P(+m | +g) U(+m, theater) + P(−m | +g) U(−m, theater)
= (0.644)(100) + (0.356)(10)
= 67.96

F	$P(F)$
+f	0.58
-f	0.42

G	$P(G)$
+g	0.59
-g	0.41

M	G	$P(M\|G)$
+m	+g	0.644
-m	+g	0.356
+m	-g	0.293
-m	-g	0.707

G	M	$P(G\|M)$
+g	+m	0.760
-g	+m	0.240
+g	-m	0.420
-g	-m	0.580

M	F	$P(M\|F)$
+m	+f	0.793
-m	+f	0.207
+m	-f	0.095
-m	-f	0.905

EU(rent | −g) = P(+m | −g) U(+m, rent) + P(−m | −g) U(−m, rent)
= (0.293)(80) + (0.707)(40)
= 51.72

Figure 5: Extra tables.

EU(rent | +g) = P(+m | +g) U(+m, rent) + P(−m | +g) U(−m, rent)
= (0.644)(80) + (0.356)(40)
= 65.76

Fill in the following values:

$MEU(+g) =$ max{EU(theater | +g), EU(rent | +g)} = max{67.96, 65.76} $\boxed{= 67.96}$

$MEU(-g) =$ max{EU(theater | −g), EU(rent | −g)} = max{36.37, 51.72} $\boxed{= 51.72}$

$VPI(G) =$ MEU(G) − MEU(∅) = MEU(+g)P(+g) + MEU(−g)P(−g) − MEU(∅)
= (67.96)(0.59) + (51.72)(0.41) − 60
$\boxed{= 1.3016}$

6.4) Now let's reintroduce variable S.

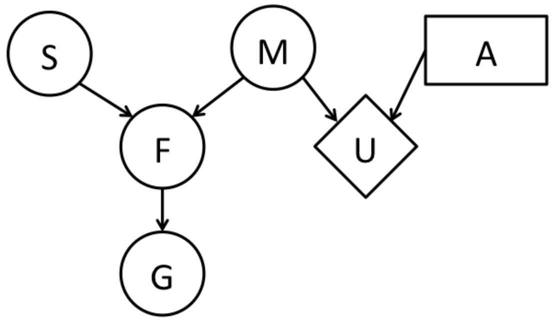


Figure 6: Decision network.

Using the probability tables $P(S)$, $P(M)$, $P(G|F)$ and $P(F|S,M)$ and utility tables from the previous parts, which of the following choices are true for each VPI? Select all that apply. Remember to justify your answers.

(i) (1 pt) $VPI(S)$:

A. $VPI(S) < 0$

$\boxed{\text{B. } VPI(S) = 0}$

C. $VPI(S) > 0$

D. $VPI(S) = VPI(F)$

E. $VPI(S) = VPI(G)$

Here, there is no observed evidence variable.

Since VPI is the expected difference in MEU given the evidence and MEU without the evidence.

No evidence ⇒ zero difference ⇒ VPI(S) = 0

(ii) (1 pt) $VPI(S|G)$ (*To avoid excessive calculations answer this question using only the network structure in Fig. 6 without taking into account the CPTs. Choose the answers that could potentially be true given the network only.*):

A. $VPI(S|G) < 0$

$\boxed{\text{B. } VPI(S|G) = 0}$

$\boxed{\text{C. } VPI(S|G) > 0}$

D. $VPI(S|G) = VPI(F)$

E. $VPI(S|G) = VPI(G)$

→ It is possible that VPI(S|G) = 0 if observing the evidence variable does not improve the MEU.

→ It is possible that VPI(S|G) > 0 if observing the evidence variable improves the MEU.

(iii) (1 pt) $VPI(G|F)$:

A. $VPI(G|F) < 0$

$\boxed{\text{B. } VPI(G|F) = 0}$

C. $VPI(G|F) > 0$

D. $VPI(G|F) = VPI(F)$

E. $VPI(G|F) = VPI(G)$

→ VPI(G|F) = 0 because G is independent of the square nodes if F is observed, so observing F doesn't improve the MEU.

(iv) (1 pt) $VPI(G)$:

A. $VPI(G) = 0$

B. $VPI(G) > 0$

C. $VPI(G) > VPI(F)$

D. $VPI(G) < VPI(F)$

E. $VPI(G) = VPI(F)$

We know that VPI(F) > 0 because the evidence is informative.
If VPI(F) > 0, then VPI(G) > 0 because the evidence is also informative here since G depends on F.
However, VPI(G) < VPI(F) because VPI(F) = VPI(G) is only equal if F perfectly provides evidence for G.

Now we change the theme and we switch to a HMM smoothing problem which has state variables X_t and observation variables E_t. The joint distribution is written as follows:

$$P(X_{1:T}, E_{1:T} = e_{1:T}) = P(X_1) \prod_{t=1}^{T-1} P(X_{t+1}|X_t) \prod_{t=1}^{T} P(E_t = e_t|X_t).$$

Recall that the notation $1 : T$ is used to refer to indices $1, \ldots, T$ and $E_{1:T} = e_{1:T}$ stands for $E_1 = e_1, \ldots, E_T = e_T$. The forward algorithm we covered in class can be used to calculate $P(X_t|E_{1:t} = e_{1:t})$ (*filtering*). The *smoothing* problem on the other hand calculates $P(X_t|E_{1:T} = e_{1:T})$ for $1 \leq t < T$, aiming to obtain more accurate estimates given the observed sequence of evidence from $1 : T$. We define for convenience two probability vectors $\alpha(X_t) \coloneqq P(E_{1:t} = e_{1:t}, X_t)$ and $\beta(X_t) \coloneqq P(E_{t+1:T} = e_{t+1:T}|X_t)$.

6.5) (4 pts) Which of the following expressions are equivalent to $P(E_T = e_T|X_{T-1})$? More than one answer might be correct.

A. $\sum_{x_t} P(X_T = x_t|X_{T-1})P(E_T = e_T|X_T = x_t)$

B. $\sum_{x_t} P(X_T = x_t|X_{T-1})P(E_T = e_T|X_T = x_t, X_{T-1})$

C. $P(X_T = x_t|X_{T-1})P(E_T = e_T|X_T = x_t)$

D. $P(X_T = x_t|X_{T-1})P(E_T = e_T|X_T = x_t, X_{T-1})$

6.6) (4 pts) A similar approach to forward recursion for filtering can be used to compute $\alpha(X_1), \ldots, \alpha(X_T)$. We want to derive a *backward* recursion to compute the βs. For each blank (i),...,(iv) choose the correct expression to generate the formula below:

$$P(E_{t+1} = e_{t+1}, \ldots, E_T = e_T|X_t) = \boxed{\text{(i)}}\boxed{\text{(ii)}}\boxed{\text{(iii)}}\boxed{\text{(iv)}}$$

If for a particular location no term is needed, select *None*. Each part (i)-(iv) is worth 1 point.

(i)

A. $\sum_{x_{t-1}}$

B. $\sum_{x_t}$

C. $\sum_{x_{t+1}}$

D. *None*

(ii)

A. $\alpha(X_{t-1} = x_{t-1})$

B. $\alpha(X_t = x_t)$

C. $\alpha(X_{t+1} = x_{t+1})$

D. $\beta(X_{t-1} = x_{t-1})$

E. $\beta(X_{t+1} = x_{t+1})$

F. *None*

(iii)

A. $P(X_t = x_t | X_{t-1})$

B. $P(X_{t+1} = x_{t+1} | X_t)$

C. $P(X_t | X_{t-1} = x_{t-1})$

D. $P(X_{t+1} | X_t = x_t)$

E. *None*

(iv)

A. $P(E_{t-1} = e_{t-1} | X_{t-1})$

B. $P(E_t = e_t | X_t)$

C. $P(E_{t+1} = e_{t+1} | X_{t+1})$

D. $P(E_{t-1} = e_{t-1} | X_{t-1} = x_{t-1})$

E. $P(E_t = e_t | X_t = x_t)$

F. $P(E_{t+1} = e_{t+1} | X_{t+1} = x_{t+1})$

G. *None*

6.7) (2pts) Which of the following expressions are equivalent to $P(X_t = x_t | E_{1:T} = e_{1:T})$? More than one choice might be correct.

A. $\sum_{x'_t} \alpha(X_t = x'_t)\beta(X_t = x'_t)$

B. $\alpha(X_t = x_t)\beta(X_t = x_t)$

C. $\frac{\alpha(X_t=x_t)\beta(X_t=x_t)}{\sum_{x'_t} \alpha(X_t=x'_t)\beta(X_t=x'_t)}$

D. $\frac{\alpha(X_t=x_t)\beta(X_t=x_t)}{\sum_{x'_T} \alpha(X_T=x'_T)}$