KOÇ UNIVERSITY
College of Engineering

COMP 341: INTRODUCTION TO ARTIFICIAL INTELLIGENCE

Midterm 2

Fall 2022, 26/12/2022
Duration: 120 Minutes

Name: Solutions

ID: 00110001 00110000 00110000

- This exam contains 10 pages including this cover page and 7 questions. Check to see if any pages are missing. Put your initials on the top of every page, in case the pages become separated.
- By submitting this exam, you **agree** to fully comply with Koç University Student Code of Conduct, and accept any punishment in case of failure to comply. If you do not submit this exam on time, you will receive 0 credits.
- The exam is **closed book** and **closed notes**. You are **not** allowed to use any additional material including any electronic equipment such as computers and mobile phones.
- You are expected to be able provide clear and concise answers. Gibberish will not receive any credit. Your answers need to be readable by a human, illegible writing is not gradable.
- Read each question carefully and make sure to follow instructions. The written instructions take precedence over an answer that the instructor might give you during the exam, unless the instructor makes a class wide announcement.
- Do not write in the table below.

Question:	1	2	3	4	5	6	7	Total
Points:	24	10	12	12	10	12	20	100
Score:								

1. (24 points) True or False :

   True ___ There are about 2^{16} rows in a joint distribution of 16 boolean variables.

   False ___ If X and Y are conditionally independent given Z, then $P(Z|X,Y) = P(Z|X)P(Z|Y)$.

   False ___ Bayesian Network Graphs can have any topology.

   True ___ Conditional independence allows us to use only a subset of factors during variable elimination.

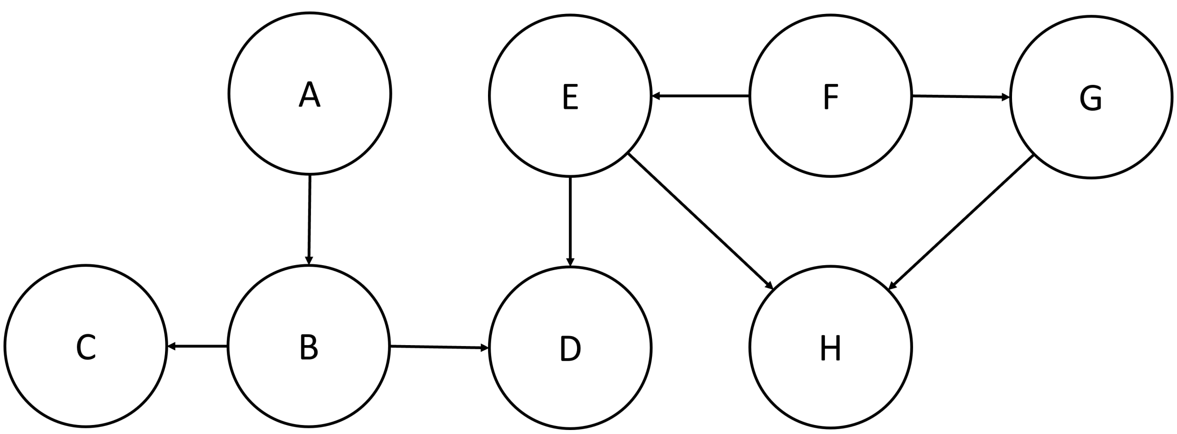
   False ___ In Gibbs sampling, only dependent variables influence their parents.

   True ___ In decision networks, action nodes can be parents to probability nodes.

   True ___ Most Markov chains will converge on a constant probability distribution while forward simulating, independent of their start distribution.

   False ___ Emission probabilities encode the conditional probability of the state given the observations.

2. (10 points) Given the Bayesian Network below, answer the questions about conditional independence



   False ___ (2 points) A and C are guaranteed to be independent.

   True ___ (2 points) A and F are guaranteed to be independent.

   False ___ (3 points) A and F are guaranteed to be conditionally independent given D.

   True ___ (3 points) A and H are guaranteed to be conditionally independent given D and E.

3. (12 points) You have done 1000 experiments and measured 3 boolean variables, namely M1, M2 and X. The table below summarizes the number of times you have observed a certain combination.

M1	**M2**	**X**	**Counts**
+	+	+	94
+	+	-	96
+	-	+	367
+	-	-	33
-	+	+	9
-	+	-	121
-	-	+	130
-	-	-	150

Answer the questions below.

(a) (3 points) Estimate $P(M1 = +)$

$P(M1 = +) = \frac{94+96+367+33}{1000} = 0.59$

(b) (3 points) Estimate $P(X = +)$

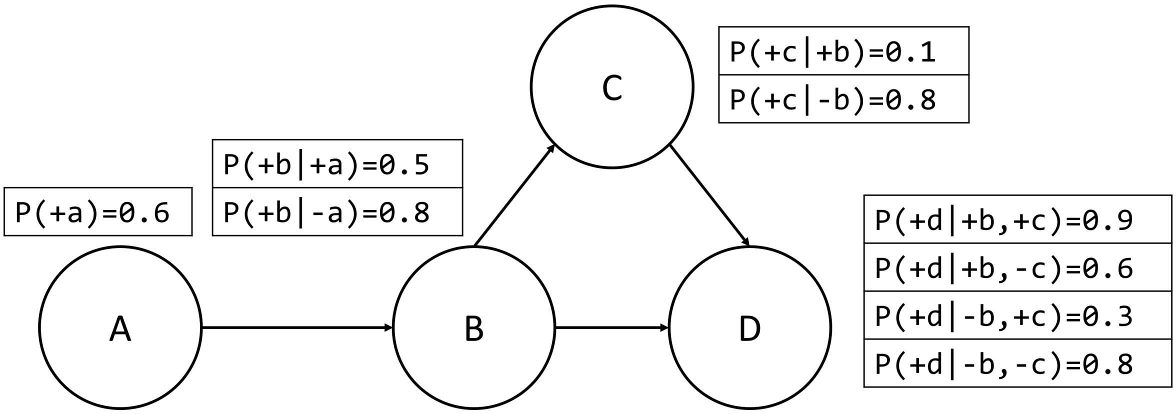
$P(X = +) = \frac{94+367+9+130}{1000} = 0.6$

(c) (6 points) Fill in the probability table below for $P(M1 = +|X)$. Show your work in the empty spaces to get partial points.

X	**M1**	$P(M1 = +\|X)$
+	+	$\frac{94+367}{1000}/0.6 = 0.461/0.6 = 461/600 \approx 0.7683$
-	+	$\frac{96+33}{1000}/0.4 = 0.129/0.4 = 129/400 \approx 0.3225$ ($\frac{121+150}{1000}/0.4 = 271/400$ if M1=- is used)

For this, we need to first calculate $P(M1 = +, X)$, then divide the relevant entries with $P(X)$. Note that $P(X = +) = 0.6, P(X = -) = 0.4$.

4. (12 points) You are given the Bayesian Network Below. All the variables are binary and True versions are denoted by a plus sign followed by a lower case letter. The CPTs are given for only the True values. You subtract this from 1 to get the probabilities for the False values. E.g. $P(-a) = 0.4, P(+c|-b) = 0.8, P(-d|+b,-c) = 0.4$. Answer the following questions and show your work whereever applicable.



(a) (2 points) What is the joint distribution, $P(A, B, C, D)$?
$P(A, B, C, D) = P(A)P(B|A)P(C|B)P(D|B, C)$

(b) (2 points) Calculate $P(+a, -b, +c, -d)$. You do not have to evaluate the final expression.
$P(+a, -b, +c, -d) = 0.6 \cdot (1 - 0.5) \cdot 0.8 \cdot (1 - 0.3) = 0.6 \cdot 0.5 \cdot 0.8 \cdot 0.7 = 0.168$

(c) (8 points) Calculate $P(+b| + d)$ using exact inference (variable elimination). You do not have to evaluate the final expression. If you are out of time, write your variable elimination plan for partial points.
Evidence variable is $D = +d$, hidden variables are A, C, and the starting factors are:
$f_1 : P(A), f_2 : P(B|A), f_3 : P(C|B), f_4 : P(+d|B, C)$
Step 1: Join $f_1(A)$ and $f_2(A, B)$ to get $f_5(A, B)$:

A	B	$f_5(A,B) = f_1(A)f_2(A,B)$
+a	+b	$0.6 \cdot 0.5 = 0.3$
+a	-b	$0.6 \cdot 0.5 = 0.3$
-a	+b	$0.4 \cdot 0.8 = 0.32$
-a	-b	$0.4 \cdot 0.2 = 0.08$

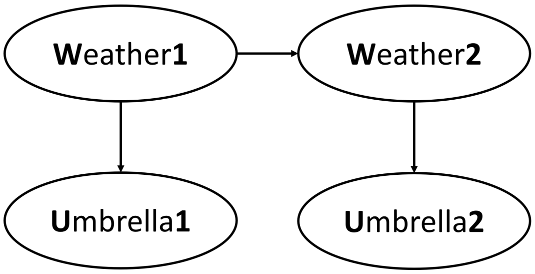
Step 2: Sum out A from $f_5(A, B)$ to get $f_6(B)$:
$f_6(+b) = 0.3 + 0.32 = 0.62$ and $f_6(-b) = 0.3 + 0.08 = 0.38$
Step 3: Join $f_3(C, B)$ and $f_4(B, C, +d)$ to get $f_7(B, C, +d)$

B	C	$f_7(B,C,+d) = f_3(C,B)f_4(B,C,+d)$
+b	+c	$0.1 \cdot 0.9 = 0.09$
+b	-c	$0.9 \cdot 0.6 = 0.54$
-b	+c	$0.8 \cdot 0.3 = 0.24$
-b	-c	$0.2 \cdot 0.8 = 0.16$

Step 4: Sum out C from $f_7(B, C, +d)$ to get $f_8(B, +d)$:
$f_8(+b, +d) = 0.9 + 0.54 = 0.63$ and $f_8(-b, +d) = 0.24 + 0.16 = 0.4$
Step 5: Join $f_6(B)$ and $f_8(B, +d)$ to get $f_9(B, +d)$:
$f_9(+b, +d) = 0.63 \cdot 0.62 = 0.3906$ and $f_9(-b, +d) = 0.4 \cdot 0.38 = 0.152$

Step 6: Normalize on B to get the desired result:
$P(+b|+d) = \frac{0.63 \cdot 0.62}{0.63 \cdot 0.62 + 0.4 \cdot 0.38} \approx 0.7199$

5. (10 points) Recall the Sunny/Rainy and Umbrella HMM from the lectures. We are going to perform an analysis for two days and as a result decide to treat everything as a static BN. You can find the network below. $W1$ and $W2$ correspond to weather at days 1 and 2. Their values, S means sunny and R means rainy. The $U1$ and $U1$ correspond to whether we see an umbreallа or not at days 1 and 2. T means an umbrella was observed and F means an umbrella was not observed.



W1	P(W1)
S	0.6
R	0.4

W1	W2	P(W2|W1)
R	R	0.7
R	S	0.3
S	R	0.5
S	S	0.5

Wi	Ui	P(Ui|Wi)
S	T	0.9
S	F	0.1
R	T	0.2
R	F	0.8

(a) (2 points) Suppose we produce the following 10 samples of $W1, U1, W2, U2$ from this BN. You want to estimate $P(W2|U1 = F, U2 = T)$. Cross off samples above which are rejected by rejection sampling for this:

$\cancel{(R,T,R,T)}$, $\cancel{(R,T,R,T)}$, $\cancel{(S,T,S,F)}$, $\cancel{(S,F,S,F)}$, (S,F,R,T),
$\cancel{(R,T,R,F)}$, $\cancel{(S,F,S,F)}$, $\cancel{(S,F,S,F)}$, (S,F,S,T), $\cancel{(R,T,S,F)}$

0.25 per cross

(b) (4 points) Calculate the weights of the following samples given the evidence $U1 = F$ and $U2 = T$ for likelihood weighing.

Sample	**Weight**
(S,F,S,T)	$0.1 \cdot 0.9 = 0.09$
(S,F,R,T)	$0.1 \cdot 0.2 = 0.02$
(R,F,S,T)	$0.8 \cdot 0.9 = 0.72$
(R,F,R,T)	$0.8 \cdot 0.2 = 0.16$

(c) (4 points) Estimate $P(W2|U1 = F, U2 = T)$, given the following samples using likelihood weighing. You do not have to evaluate the final expression.

$(S,F,R,T), (R,F,R,T), (S,F,R,T), (S,F,S,T), (S,F,S,T), (R,F,S,T)$

Sample numbers: $(S,F,R,T) : 2, (R,F,R,T) : 1, (S,F,S,T) : 2, (R,F,S,T) : 1$

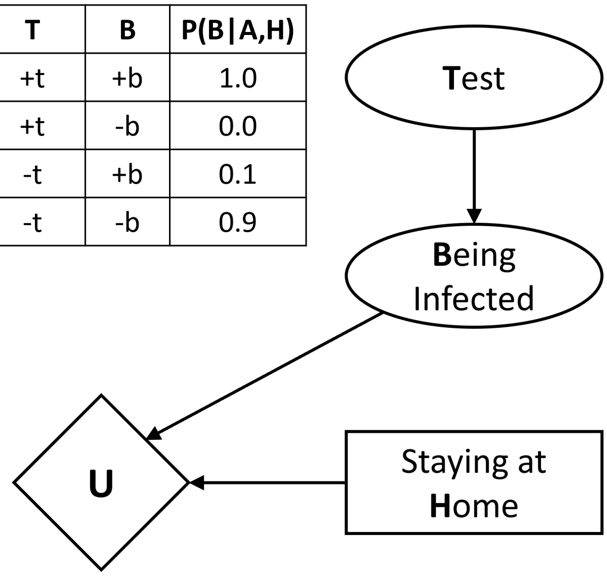
$$P(W2 = S|U1 = F, U2 = T) = \frac{2 \cdot 0.09 + 0.72}{2 \cdot 0.02 + 1 \cdot 0.16 + 2 \cdot 0.09 + 1 \cdot 0.72} = \frac{0.9}{1.1} = 9/11 \approx 0.818$$

$$P(W2 = R|U1 = F, U2 = T) = \frac{2 \cdot 0.02 + 1 \cdot 16}{2 \cdot 0.02 + 1 \cdot 0.16 + 2 \cdot 0.09 + 1 \cdot 0.72} = \frac{0.2}{1.1} = 2/11 \approx 0.182$$

6. (12 points) You are given the below decision network. It is about deciding to go out or staying at home depending on whether you are sick or not (being infected). There is a new test that was developed. If the test is positive, you are guaranteed to be sick. If it is negative, you are highly likely not sick. This is modelled in the Conditional Probability Tables (note that this is not a causal structure!).

T	P(T)
+t	0.2
-t	0.8

T	B	P(B|A,H)
+t	+b	1.0
+t	-b	0.0
-t	+b	0.1
-t	-b	0.9



B	H	U
+b	stay	0
+b	go out	-100
-b	stay	0
-b	go out	20

(a) (6 points) Without any evidence, what is the expected utility of each action? What is the best action? Note that you need to perform a short inference step.

We need to calculate $P(B)$ to be able to calculate the expected utilities. We do this by joining $P(T)$ and $P(B|T)$ to get $P(B,T)$ and summing out T:

T	B	$P(T,B)$
+t	+b	$1.0 \cdot 0.2 = 0.2$
+t	-b	$0.0 \cdot 0.2 = 0.0$
-t	+b	$0.1 \cdot 0.8 = 0.08$
-t	-b	$0.9 \cdot 0.8 = 0.72$

From the table, $P(B=+b) = 0.2+0.08 = 0.28$ and $P(B=-b) = 0.0+0.72 = 0.72$. Then we can calculate the expected utilities as follows:

$EU(stay) = P(+b)U(+b,s) + P(-b)U(-b,s) = 0.28 \cdot 0 + 0.72 \cdot 0 = 0$
$EU(go\ \ out) = P(+b)U(+b,g) + P(-b)U(-b,g) = 0.28 \cdot -100 + 0.72 \cdot 20 = -13.6$
Then the maximum expected utility is $MEU() = 0$ with *stay* being the optimal action.

(b) (6 points) What is the value of getting the test?

The value of gettig the test is $MEU(T) - MEU()$. To calculate $MEU(T)$, we need to calculate the expected utility of both the outcome cases and sum them up with the outcome probabilities.

$EU(+t, stay) = P(+b|+t)U(+b,s) + P(-b|+t)U(-b,s) = 1.0 \cdot 0 + 0 \cdot 0 = 0$
$EU(+t, goout) = P(+b|+t)U(+b,g) + P(-b|+t)U(-b,g) = 1.0 \cdot -100 + 0 \cdot 20 = -100$
$\Rightarrow MEU(+t) = 0$ with the *stay* action being optimal

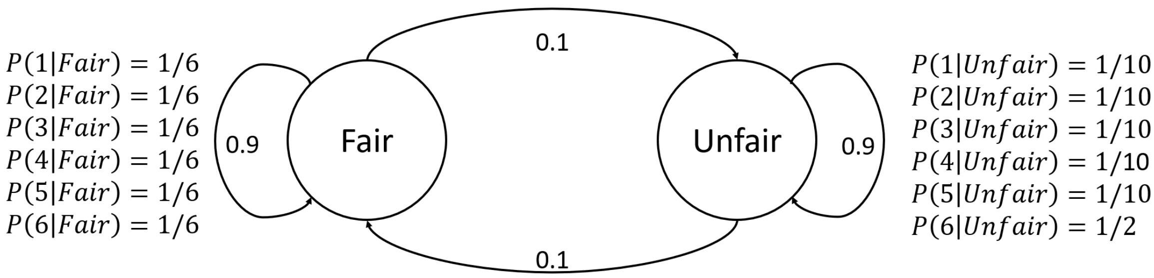
$EU(-t, stay) = P(+b|-t)U(+b, s) + P(-b|-t)U(-b, s) = 0.1 \cdot 0 + 0.9 \cdot 0 = 0$
$EU(-t, go\ \ out) = P(+b|-t)U(+b, g) + P(-b|-t)U(-b, g) = 0.1 \cdot -100 + 0.9 \cdot 20 = 8$
$\Rightarrow MEU(-t) = 8$ with the *go out* action being optimal

Then $MEU(T) = P(+t)MEU(+t) + P(-t)MEU(-t) = 0.2 \cdot 0 + 0.8 \cdot 8 = 6.4$
$\Rightarrow VPI(T) = MEU(T) - MEU() = 6.4 - 0 = 6.4$

7. (20 points) It is your last day in Las Vegas for work. You are done with all your work and have some time to kill before your shuttle takes you to the airport. You see a simple dice game and decide to play even though you know gambling is bad. The rules of the game are as follows:

- You bet some money
- You roll a 6 sided die
- Casino player rolls a 6 sided die
- If your roll is higher, you get double your money. If not, you lose the money you bet

You suddenly hear two people talking about the casino switching back and forth between a fair die and an unfair die about every 10 rolls and that the unfair die has a 50% chance of coming up 6, while chances for the all other numbers are uniform. Luckily, you are taking the AI class to figure out what is going on. You formulate the problem in your head as a Hidden Markov Model with the following structure:



As you play you see the following sequence of die rolls $\{1, 1, 6, 2, 6, 3, 4, 3, 6, 2, 6, 6\}$. You have no idea which die the casino player started with and thus assume both are equally likely.

(a) (6 points) You want to calculate the likelihood of this observation sequence, given that the casino player used an unfair dice all the way. What is the expression? Show all your work. (You do not need to calculate the final expression, i.e., leave exponentials as they are)

From the text we can infer the prior probabilities as: $P(x_0 = Fair) = P(x_0 = Unfair) = 0.5$.
Let the observation sequence be: $e = \{e_1, \ldots, e_{12}\} = \{1, 1, 6, 2, 6, 3, 4, 3, 6, 2, 6, 6\}$
Let the state sequence be: $x = \{x_1, \ldots, x_{12}\} = \{Unfair, \ldots, Unfair\}$
There is some ambiguity in the text. The desired probability can be interpreted in two ways:

Joint: $P(x, e) = P(x_0) \prod_{i=1}^{12} P(x_i|x_{i-1}) P(e_i|x_i) = 0.5 \cdot (9/10)^{12} \cdot (1/10)^7 \cdot (1/2)^5$

Conditional: $P(e|x) = \prod_{i=1}^{12} P(e_i|x_i) = (1/10)^7 \cdot (1/2)^5$

(b) (14 points) You do not know whether the casino player started with a fair die or an unfair die as before. However, you are a bit suspicious that the casino player used an unfair die for the third roll. Calculate the necessary values and decide whether the casino player used an unfair die or not. (Hint: Uniform at the start, then 3 state changes and 3 observations.)

We want $P(X_3|e_1 = 1, e_2 = 1, e_3 = 6)$. Run the forward algorithm for 3 times to get this:

$$B_t(x) = P(x_t|e_{1:t}) \propto P(e_t|x_t) \sum_{x_{t-1}} P(x_t|x_{t-1})P(x_{t-1}|e_{1:t-1})$$

Time 0:
$P(x_0 = Fair) = B_0(Fair) = 0.5, P(x_0 = Unfair) = B_0(Unfair) = 0.5$

Moving forward $Fair = F$ and $Unfair = U$

Time 1:
$B_1(F) \propto P(e_1 = 1|x_1 = F)\,(P(x_1 = F|x_0 = U)B_0(U) + P(x_1 = F|x_0 = F)B_0(F))$
$= 1/6 \cdot (0.1 \cdot 0.5 + 0.9 \cdot 0.5) = 1/12$
$B_1(U) \propto P(e_1 = 1|x_1 = U)\,(P(x_1 = U|x_0 = U)B_0(U) + P(x_1 = U|x_0 = F)B_0(U))$
$= 1/10 \cdot (0.9 \cdot 0.5 + 0.1 \cdot 0.5) = 1/20$
Normalize to get: $B_1(F) = 5/8, B_1(U) = 3/8$

Time 2:
$B_2(F) \propto P(e_2 = 1|x_2 = F)\,(P(x_2 = F|x_1 = U)B_1(U) + P(x_2 = F|x_1 = F)B_1(F))$
$= 1/6 \cdot (0.1 \cdot 3/8 + 0.9 \cdot 5/8) = 0.1$
$B_2(U) \propto P(e_2 = 1|x_2 = U)\,(P(x_2 = U|x_1 = U)B_1(U) + P(x_2 = U|x_1 = F)B_1(U))$
$= 1/10 \cdot (0.9 \cdot 3/8 + 0.1 \cdot 5/8) = 0.04$
Normalize to get: $B_2(F) = 5/7, B_1(U) = 2/7$

Time 3:
$B_3(F) \propto P(e_3 = 6|x_3 = F)\,(P(x_3 = F|x_2 = U)B_2(U) + P(x_3 = F|x_2 = F)B_2(F))$
$= 1/6 \cdot (0.1 \cdot 2/7 + 0.9 \cdot 5/7) = 47/420$
$B_3(U) \propto P(e_3 = 6|x_3 = U)\,(P(x_3 = U|x_2 = U)B_2(U) + P(x_3 = U|x_2 = F)B_2(U))$
$= 1/2 \cdot (0.9 \cdot 2/7 + 0.1 \cdot 5/7) = 69/420$
Normalize to get: $B_3(F) = 47/116, B_3(U) = 69/116$

At time 3, the probability of die being Unfair is higher than the probability of being Fair. As such, we conclude that the die at time 3 is unfair.