Question 1

You have three baskets of fruit: the first one contains two apples, the second one contains two oranges, and the third one contains one apple and one orange. Assume that a basket is selected randomly and that a piece of fruit is picked randomly from that basket. Let B be the random variable corresponding to the basket number selected (B can have as value 1, 2 or 3) and let F be the random variable corresponding to the type of fruit picked (F can have as value $apple$ or $orange$).

1. What is the distribution $P(B)$ and what are the conditional distributions $P(F|B)$? Write your answers in tabular form.

B	**P(B)**
1	1/3
2	1/3
3	1/3

F	**B**	**P(F|B)**
Apple	1	1
Orange	1	0
Apple	2	0
Orange	2	1
Apple	3	1/2
Orange	3	1/2

2. What is the joint probability of selecting the first basket and picking an apple, i.e. what is $P(B = 1, F = apple)$?

$$P(B = 1, F = apple) = P(F = apple|B = 1)P(B = 1)$$

$$= \frac{1}{1} * \frac{1}{3}$$

$$= \frac{1}{3}$$

3. If we observe that the picked fruit is an apple, what is the conditional probability that the chosen basket is the basket containing two apples, i.e. what is $P(B = 1|F = apple)$?

$$P(A \mid B) = \frac{P(B \mid A)\, P(A)}{P(B)},$$

$$P(B = 1|F = apple) = \frac{P(F=apple|B=1)P(B=1)}{P(F=apple)}$$

$$= \frac{P(F=apple|B=1)P(B=1)}{\sum_B P(F=apple|B)P(B)}$$

$$= \frac{\frac{1}{3}}{\frac{1}{3}*\frac{1}{1}+\frac{1}{3}*0+\frac{1}{3}*\frac{1}{2}}$$

$$= \frac{2}{3}$$

Question 2

A CS student may or may not show up in class (C) on a given day. If they do not show up ($C = \neg c$), there two possible explanations. It could be that the student was abducted by aliens ($A = a$). It could also be that the student has the stomach bug ($B = b$) that is going around. The possible outcomes are listed in a joint probability table $P(A,B,C)$ in Table 1:

a	b	c	0.00
a	b	$\neg c$	0.01
a	$\neg b$	c	0.00
a	$\neg b$	$\neg c$	0.02
$\neg a$	b	c	0.01
$\neg a$	b	$\neg c$	0.04
$\neg a$	$\neg b$	c	0.90
$\neg a$	$\neg b$	$\neg c$	0.02

Table 1: $P(A,B,C)$

1. What is the distribution $P(A,B)$? Your answer should be in the form of a table.

A	**B**	**P(A,B)**
a	b	$\sum_C P(A=a, B=b, C) = 0.00 + 0.01 = 0.01$
a	$\neg b$	$\sum_C P(A=a, B=\neg b, C) = 0.00 + 0.02 = 0.02$
$\neg a$	b	$\sum_C P(A=\neg a, B=b, C) = 0.01 + 0.04 = 0.05$
$\neg a$	$\neg b$	$\sum_C P(A=\neg a, B=\neg b, C) = 0.90 + 0.02 = 0.92$

2. Are A and B independent? Justify your answer using the actual probabilities computed in part 1.

$$\mathrm{P}\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} \mathrm{P}(A_i).$$

$$P(A=a) = \sum_B P(A=a, B) = 0.03$$

$$P(B=b) = \sum_A P(B=b, A) = 0.06$$

$$P(A=a)P(B=b) = 0.0018 \neq P(A=a, B=b)$$

$$\therefore A \text{ and } B \text{ are not independent}$$

3. We would like to reason about whether or not a certain student has the bug. What is the marginal distribution over B given no evidence (sometimes this is called the prior distribution)? Is it most likely that the student does or does not have the bug?

B	**P(B)**
b	$\sum_A P(B = b, A) = 0.01 + 0.05 = 0.06$
$\neg b$	$\sum_A P(B = \neg b, A) = 0.02 + 0.92 = 0.94$

4. If we observe that the student does not come to class ($C = \neg c$), what is the conditional distribution $P(B|C = \neg c)$ (sometimes this is called a posterior distribution)? Is it most likely that the student did or did not have the bug given the evidence? Does it make intuitive sense how the numbers have shifted from part 3 given the new evidence that has been observed?

$$\begin{aligned} P(C = \neg c) &= \textstyle\sum_{A,B} P(C = \neg c, A, B) \\ &= 0.01 + 0.02 + 0.04 + 0.02 \\ &= 0.09 \end{aligned}$$

B	$P(B\vert C = \neg c)$
b	$P(B = b\vert C = \neg c) = \frac{\sum_A P(B = b, C = \neg c, A)}{P(C = \neg c)} = \frac{0.01 + 0.04}{0.09} = \frac{5}{9}$
$\neg b$	$P(B = \neg b\vert C = \neg c) = \frac{\sum_A P(B = \neg b, C = \neg c, A)}{P(C = \neg c)} = \frac{0.02 + 0.02}{0.09} = \frac{4}{9}$

5. If we further discover that the student has been abducted by aliens ($A = a$), what is the new posterior distribution over B, i.e. $P(B|C = \neg c, A = a)$? Is it most likely that the student did or did not have the bug given all the evidence? Does it make intuitive sense how the numbers have shifted from part 4 given the new evidence that has been observed? The general phenomenon where the discovery that one possible cause is true decreases belief in other possible causes is called explaining-away.

$$\begin{aligned} P(A = a, C = \neg c) &= \textstyle\sum_B P(A = a, C = \neg c, B) \\ &= 0.01 + 0.02 \\ &= 0.03 \end{aligned}$$

B	$P(B\vert A = a, C = \neg c)$
b	$P(B = b\vert A = a, C = \neg c) = \frac{P(B = b, A = a, C \neg c) =}{P(A = a, C = \neg c)} = \frac{0.01}{0.03} = \frac{1}{3}$
$\neg b$	$P(B = \neg b\vert A = a, C = \neg c) = \frac{P(B = \neg b, A = a, C \neg c) =}{P(A = a, C = \neg c)} = \frac{0.02}{0.03} = \frac{2}{3}$

6. Are A and B conditionally independent given C? Justify your answer using specific probabilities (hint: you should already have sufficient probabilities computed).

A and B conditionally independent given C means:

$$A \perp B \iff P(A,B|C) = P(B|C)P(A|C)$$

And

$$A \perp B \iff P(B|C) = P(B|A,C)$$

Does $P(B|A,C) = P(B|C)$?

$$P(B = b|A = a, C = \neg c) = \frac{1}{3}$$

$$P(B = b|C = \neg c) = \frac{5}{9}$$

$$\therefore P(B = b|A = a, C = \neg c) \neq P(B = b|C = \neg c)$$

$$\therefore A \text{ and } B \text{ are not conditionally independent given } C$$

Extra:

We also covered the rule (which may be useful for your assignment):

$$\frac{P(A,B|C)}{P(A|C)} = P(B|A,C)$$

I can't find the name, but you can still use it if you find it will help. I've included the proof below:

$P(A,B,C) = P(A,B|C)P(C)$ (1)

$P(A,B,C) = P(B|A,C)P(A,C) = P(B|A,C)P(A|C)P(C)$ (2)

$\frac{P(A,B,C)}{P(A,B,C)} = 1$ (3)

Substituting (1) and (2) into (3)

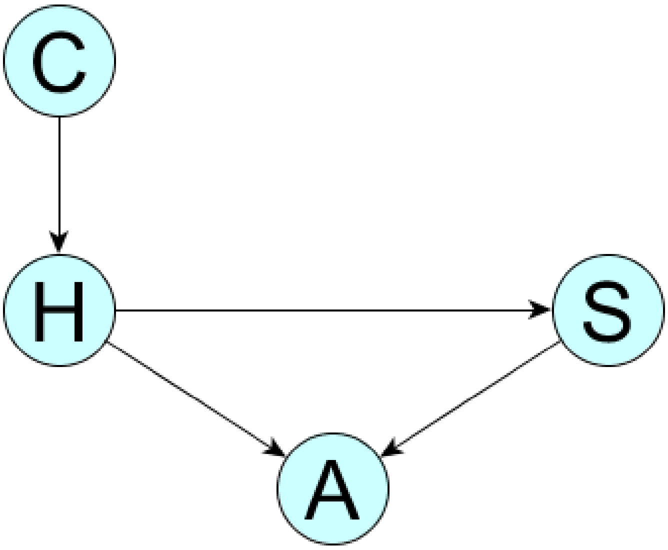
$$\frac{P(A,B|C)\cancel{P(C)}}{P(B|A,C)P(A|C)\cancel{P(C)}} = 1$$

$$\frac{P(A,B|C)}{P(A|C)} = (B|A,C)$$

Question 3

Consider the following network, in which a mouse is reasoning about the behaviour of a cat. The mouse really wants to know whether the cat will attack (A), which depends on whether the cat is hungry (H) and whether the cat is sleepy (S). The mouse can observe two things, whether the cat is sleepy (S) and whether the cat has a collar (C). The cat is more often sleepy (S) when it's either full (f) or starved (v) than when it is peckish (p) and the collar (C) tends to indicate that the cat is not starved. Note that entries are omitted, such as $P(C = \neg c)$, when their complements are given.

1. Draw the Bayesian network corresponding to the above joint probability distribution on C, H, S and A.



2. If the conditional probability tables for the network are as follows,

   compute the following probabilities:

P(C)	
C	P(C)
c	0.4

P(H|C)		
H	C	P
f	c	0.70
v	c	0.10
p	c	0.20
f	¬c	0.20
v	¬c	0.50
p	¬c	0.30

P(S|H)		
S	H	P
s	f	0.90
s	v	0.60
s	p	0.30

P (A|H,S)			
A	H	S	P
a	f	s	0.01
a	f	¬s	0.10
a	v	s	0.40
a	v	¬s	0.90
a	p	s	0.20
a	p	¬s	0.70

(a) $P(A = \text{a}, C = c, S = s, H = f)$

$= P(C = c)P(H = f|C = c)P(S = s|H = f)P(A = a|H = f, S = s)$

$= 0.4 * 0.7 * 0.9 * 0.01$

$= 0.00252$

(b) $P(A = a, C = c, S = s)$

$$= \sum_{H} P(C = c)P(H|C = c)P(S = s|H)P(A = a|H, S = s)$$

$$= P(C = c) \sum_{H} P(H|C = c)P(S = s|H)P(A = a|H, S = s)$$

$$= 0.4 * (0.7 * 0.9 * 0.01 + 0.1 * 0.6 * 0.4 + 0.2 * 0.3 * 0.2)$$

$$= 0.01692$$

(c) $P(C = c, S = s)$

$$= \sum_{H} P(C = c)P(H|C = c)P(S = s|H)$$

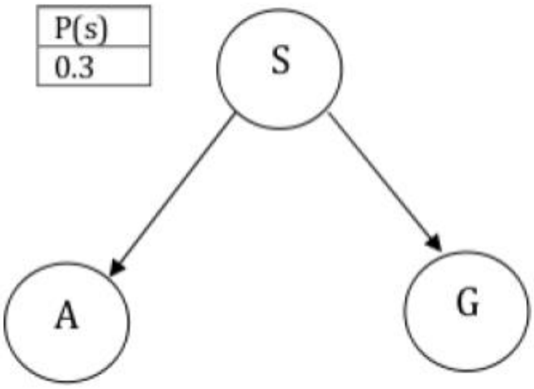
$$= P(C = c) \sum_{H} P(H|C = c)P(S = s|H)$$

$$= 0.4 * (0.7 * 0.9 + 0.1 * 0.6 + 0.2 * 0.3)$$

$$= 0.3$$

Question 4

A CS student in the AI class notices that people who drive 4WDs vehicles (S) consume large amounts of gas (G) and are involved in more accidents than the national average (A). They have constructed the Bayesian network in Figure 1 (here t implies "true" and f implies "false"):



P(s)
0.3

S	P(a)
t	0.6
f	0.3

S	P(g)
t	0.7
f	0.2

Figure 1: 4WD Bayesian network

1. Compute $P(A)$ (this is the probability distribution of A).

$$
\begin{aligned}
P(A = a) &= \textstyle\sum_S P(A = a, S) \\
&= \sum_S P(A = a|S)P(S) \\
&= P(A = a|S = s)P(S = s) + P(A = a|S = \neg s)P(S = \neg s) \\
&= 0.3 * 0.6 + 0.7 * 0.3 \\
&= 0.39 \\
P(A = \neg a) &= 1 - 0.39 \\
&= 0.61
\end{aligned}
$$

2. Using conditional independence, compute $P(\neg g, \mathrm{a}|s)$ and $P(\neg g, \mathrm{a}|\neg s)$. Then use Bayes' rule to compute $P(s|\neg g, \mathrm{a})$.

$$
\begin{aligned}
P(\neg g, a|s) &= P(\neg g|s)P(a|s) \\
&= 0.3 * 0.6 \\
&= 0.18 \\
P(\neg g, a|\neg s) &= P(\neg g|\neg s)P(a|\neg s) \\
&= 0.8 * 0.3 \\
&= 0.24 \\
P(s|\neg g, a) &= \frac{P(\neg g, a|s)P(s)}{P(\neg g, a)} \\
&= \frac{P(\neg g, a|s)P(s)}{\sum_S P(\neg g, a, S)} \\
&= \frac{P(\neg g, a|s)P(s)}{\sum_S P(\neg g, a, S)P(S)} \\
&= \frac{P(\neg g, a|s)P(s)}{P(\neg g, a|s)P(s) + P(\neg g, a|\neg s)P(\neg s)} \\
&= \frac{0.18 * 0.3}{0.18 * 0.3 + 0.24 * 0.7} \\
&= \frac{9}{37}
\end{aligned}
$$

3. The enterprising AI student notices that there are two types of people that drive 4WDs, people from the country (C) and people with large families (F). After collecting some statistics, the student arrives at the Bayesian network in Figure 2. Using the chain rule (product rule) compute the probability $P(\neg g, \mathrm{a}, s, c, \neg f)$.

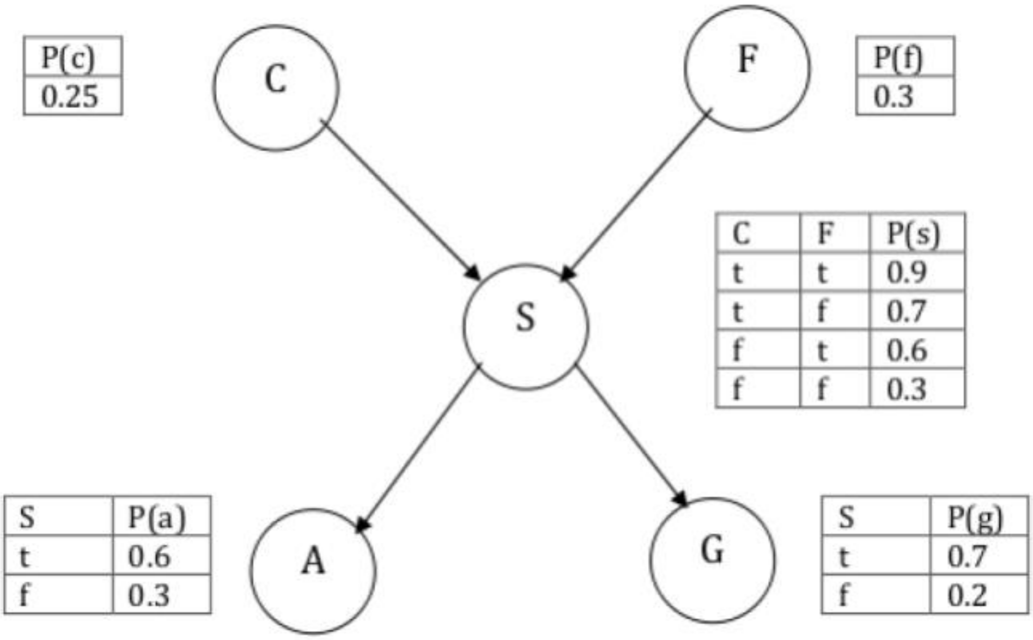


Figure 2: New 4WD Bayesian network

$$
\begin{aligned}
P(\neg g, a, s, c, \neg f) &= P(c)P(\neg f)P(s|c,\neg f)P(a|s)P(\neg g|s) \\
&= 0.25 * 0.7 * 0.7 * 0.6 * 0.3 \\
&= 0.02205
\end{aligned}
$$

4. Using the local semantics implied by the structure of Bayesian Networks,

state whether the following variables from the Bayesian network in Figure 2 are (conditionally) independent.

(a) C, G

(b) $F, A|S$

(c) C, F

(d) A, G

(e) $C, F|S$

(f) $C, F|A$

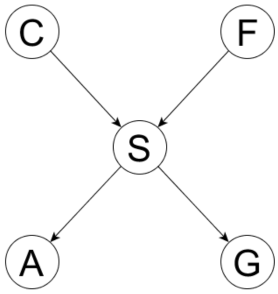
a) No. C gives additional information to S, so it gives additional information to G. This means C and G are not independent.
b) Yes. We know the outcome of S so we cannot receive any additional information about it. Hence, F gives no additional information to S and no additional information to A. This means that F and A are independent given S.
c) Yes. We don't know any information about A, S or G, so C and F are independent.
d) No. A gives some additional information to S, so A gives some additional information to G. This means that A and G are not independent.

Note: We could think of these as A being attending class, S being sunny and G being Go Kart Racing. If we had a student who does not attend class when it's sunny and only goes go kart racing on sunny days, then if we are told the student has not attended class then we know it's more likely to be sunny and that they're more likely to be go kart racing.

e) No. We know S so we know some additional information about C, as well as some additional information about F. However, we also know this additional information about C and F *at the same time.* Because we know information about both at the same time then knowing more information about one of them will tell us more information about the other, this means they're not independent given S.
f) No. A tells us more information about S and this means A tells us information about C and F at the same time. This means they're are not independent given A.

Question 6 (Extra)

Considering the following graph:



1. Write down the factorisation of the joint distribution.

$$P(C,A,S,F,G) = P(C)P(F)P(S|C,F)P(A|S)P(G|S)$$

2. Write down variable elimination for marginal inference to compute P(S).

$P(S)$ $= \sum_{C,A,F,G} P(C,A,S,F,G)$

$= \sum_{C,A,F,G} P(C)P(F)P(S|C,F)P(A|S)P(G|S)$

$= \sum_{C,A,F} P(C)P(F)P(S|C,F)P(A|S)(\sum_G P(G|S) = 1)$, eliminating G

$= \sum_{C,F} P(C)P(F)P(S|C,F)(\sum_A P(A|S) = 1)$, eliminating A

$= \sum_C P(C) \sum_F P(F)P(S|C,F)$

$= \sum_C P(C)m_1(S,C)$, eliminated F

$= m_2(S)$, eliminated C

$$Order\ of\ A\ and\ G\ could\ be\ swapped, order\ of\ F\ and\ C\ can\ also\ be\ swapped$$

3. Write down variable elimination for MAP inference.

$\max_{C,A,S,F,G} P(C,A,S,F,G)$ $= \max_{C,A,S,F,G} P(C)P(F)P(S|C,F)P(A|S)P(G|S)$

$= \max_{C,A,S,F} P(C)P(F)P(S|C,F)P(A|S) \max_G P(G|S)$

$= \max_{C,A,S,F} P(C)P(F)P(S|C,F)P(A|S)m_1(S)$, eliminated G

$= \max_{C,S,F} P(C)P(F)P(S|C,F)[\max_A P(A|S)]m_1(S)$

$= \max_{C,S,F} P(C)P(F)P(S|C,F)m_2(S)m_1(S)$, eliminated A

$= \max_{C,F} P(C)P(F)[\max_S P(S|C,F)m_2(S)m_1(S)]$

$= \max_{C,F} P(C)P(F)m_3(C,F)$, eliminated A

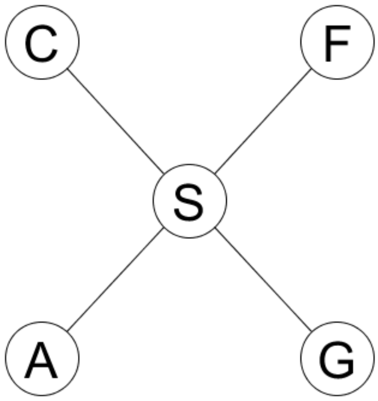
$= \max_C P(C)[\max_F P(F)m_3(C,F)]$

$= \max_C P(C)m_4(C)$, eliminated F

$$Order\ of\ A\ and\ G\ could\ be\ swapped, order\ of\ F\ and\ C\ can\ also\ be\ swapped$$

Question 7 (Extra):

Considering the following graph:



1. Is F independent to A?

No, path though S

2. Is F independent to A, conditioned on S?

Yes, knowing S means that F gives no additional information to S and no additional information to A.

3. What messages are needed to compute message $m_{S \to G}(G)$?

$m_{C \to S}(S)$, $m_{F \to S}(S)$, $m_{A \to S}(S)$