Q2. [10 pts] Bayes Net and Sampling

(a) [2 pts] A, B, C are discrete random variables. Given $A \perp\!\!\!\perp B|C$, which of the following equations must hold?

[A] $P(A|B,C)P(B|A,C) = P(A,B|C)$

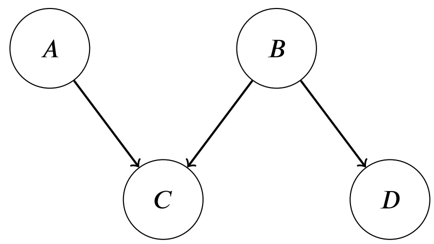
[B] $P(A,B,C) = P(A)P(B)P(A,B|C)$

[C] $P(A|C) = \frac{P(A)P(C|A)}{P(C)}$

[D] $P(A,B|C) = P(A,B)$

(E) None of the above.

Consider the following Bayes Net involving binary random variables A, B, C, D. The relevant probability tables are given.



$P(A)$	
$A = 0$	0.5
$A = 1$	0.5

$P(B)$	
$B = 0$	0.5
$B = 1$	0.5

A	B	C	$P(C\|A,B)$
0	0	0	0.6
0	0	1	0.4
0	1	0	0.4
0	1	1	0.6
1	0	0	0.8
1	0	1	?
1	1	1	0.6
1	1	0	0.4

B	D	$P(D\|B)$
0	0	0.6
0	1	0.4
1	0	0.4
1	1	0.6

(b) **(i)** [1 pt] Calculate $P(C = 1|A = 1, B = 0)$ (the ? entry in the table).

(ii) [1 pt] Calculate $P(A = 1|B = 0, D = 1)$.

(c) Instead of calculating the exact quantity, suppose we want to estimate $P(C = 1|D = 1)$ using different sampling methods.

(i) [1 pt] In this subpart we use rejection sampling. Which of the following is a valid topological order and is most efficient for rejection sampling to estimate $P(C = 1|D = 1)$?

(A) A, B, C, D

(B) B, A, D, C

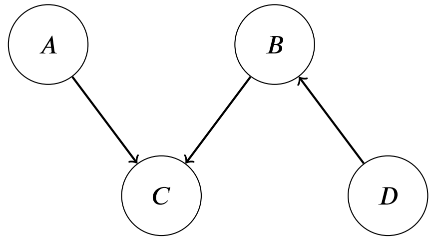
(C) B, D, A, C

(D) D, C, B, A

(ii) [1 pt] In this subpart we use likelihood weighting. What is the weight of the sample $(A = 0, B = 0, C = 0, D = 1)$?

(iii) [2 pts] In this subpart we use Gibbs sampling. We initialize $A = 0, B = 0, C = 0, D = 1$, and choose to re-sample A. What is the probability that we still get $A = 0$ after re-sampling?

(d) [2 pts] We reverse the arrow between B and D to create a new Bayes Net (shown below).



Which of the following statements are true?

(A) The set of joint distributions $P(A, B, C, D)$ that can be modeled by the two Bayes nets are the same.

(B) The set of joint distributions that can be modeled by the old Bayes net is a subset of the set of joint distributions that can be modeled by the new Bayes net.

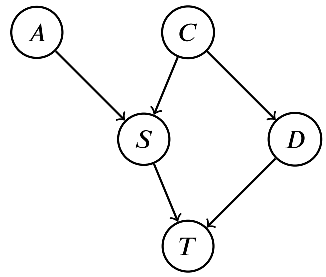
(C) The set of joint distributions that can be modeled by the new Bayes net is a subset of the set of joint distributions that can be modeled by the old Bayes net.

(D) None of the above.

Q2. [13 pts] Bayesian Networks: Across the Spider-Verse

Miles is curious about the probability that he arrives to school on time. The factors involved with him arriving to school on time can be represented by the following Bayes Net (assume that each variable is a binary variable):

- A: Sets alarm
- S: Over sleeps
- D: Dad is late
- C: Fighting crime last night
- T: Arrives at school on time



(a) [7 pts] Miles wants to calculate $P(C|T)$ using variable elimination. Assume he eliminates variables in alphabetical order (A, D, S).

(i) [1 pt] What factors does he have available at the start?

(ii) [1 pt] First, he eliminates A, and get the new factor

$f_1($ $) =$

Write out the remaining factors

(iii) [1 pt] Then, he eliminates D, and get the new factor

$f_2($ $) =$

Write out the remaining factors

(iv) [1 pt] Then, he eliminates S, and get the new factor

$f_3($ $) =$

Write out the remaining factors

(v) [1 pt] Finally, join any remaining factors to calculate

$f_4($ $) =$

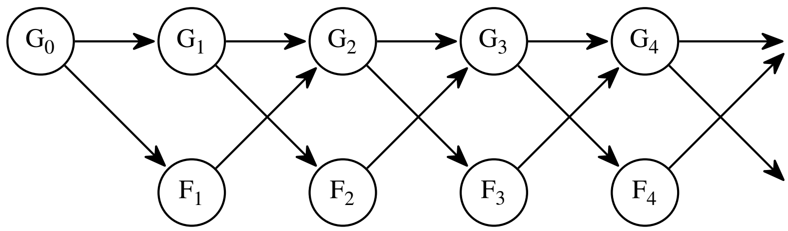
(vi) [1 pt] How can he use this to calculate $P(C = +c|T = -t)$? Your answer should be in terms of f_4.

$P(C = +c|T = -t) =$

(vii) [1 pt] Order factors f_1, f_2, and f_3 in increasing order of size. $\quad < \quad < \quad$

Q6. [10 pts] Dynamic Bayes' Nets: Book Club

Each week t ($t \geq 0$), the members of a book club decide on the genre of the book G_t to read that week: romance ($+g_t$) or science fiction ($-g_t$). Each week's choice is influenced by the genre chosen in the previous week. In addition, after week t ($t \geq 0$), they collect feedback (F_{t+1}) from the members, which can be overall positive ($+f_{t+1}$) or negative ($-f_{t+1}$), and that influences the choice at week $t+2$. The situation is shown in the Bayes' net below:



(a) [1 pt] Recall that the joint distribution up to time T for the "standard" HMM model with state X_t and evidence E_t is given by $P(X_{0:T}, E_{1:T}) = P(X_0) \prod_{t=1}^{T} P(X_t \mid X_{t-1}) P(E_t \mid X_t)$.

Write an expression for the joint distribution in the model shown above.

$P(F_{1:T}, G_{0:T}) =$

(b) [2 pts] Which of the following Markov assumptions are implied by the network structure, assuming $t \geq 2$?

- ☐ $P(G_t \mid G_{0:t-1}, F_{1:t-1}) = P(G_t \mid F_{t-1})$
- ☐ $P(G_t \mid G_{0:t-1}, F_{1:t-1}) = P(G_t \mid G_{t-1})$
- ☐ $P(G_t, F_t \mid G_{0:t-1}, F_{1:t-1}) = P(G_t, F_t \mid G_{t-1})$
- ☐ $P(G_t, F_t \mid G_{0:t-1}, F_{1:t-1}) = P(G_t, F_t \mid G_{t-1}, F_{t-1})$
- ☐ $P(F_t \mid G_{0:t-1}, F_{1:t-1}) = P(F_t \mid F_{t-1})$
- ☐ $P(F_t \mid G_{0:t-1}, F_{1:t-1}) = P(F_t \mid G_{t-1}, F_{t-2})$
- ○ None of the above

The conditional probability tables for G_t and F_t are shown below, where a, b, c, d, p, q are constants between 0 and 1.

F_t	G_{t-1}	$P(F_t \mid G_{t-1})$
+	+	p
−	+	$1-p$
+	−	q
−	−	$1-q$

G_t	F_{t-1}	G_{t-1}	$P(G_t \mid F_{t-1}, G_{t-1})$
+	+	+	a
−	+	+	$1-a$
+	+	−	b
−	+	−	$1-b$
+	−	+	c
−	−	+	$1-c$
+	−	−	d
−	−	−	$1-d$

Let's consider this model as a Markov chain with state variables (F_t, G_t).

(c) [2 pts] Write out the transition probabilities below.

Your answer should be an expression, possibly in terms of a, b, c, d, p, and q.

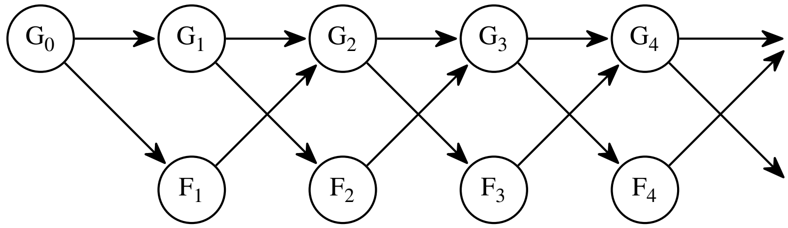
$x_1 = P(+f_t, +g_t \mid +f_{t-1}, +g_{t-1}) =$

$x_2 = P(+f_t, +g_t \mid +f_{t-1}, -g_{t-1}) =$

$x_3 = P(+f_t, +g_t \mid -f_{t-1}, +g_{t-1}) =$

$x_4 = P(+f_t, +g_t \mid -f_{t-1}, -g_{t-1}) =$

The Bayes' net, repeated for your convenience:



(d) [2 pts] As t goes to infinity, the stationary distribution of this Markov chain is shown below, where y_1, y_2, y_3, y_4 are constants between 0 and 1.

$$\begin{aligned} y_1 &= P(+f_\infty, +g_\infty) \\ y_2 &= P(+f_\infty, -g_\infty) \\ y_3 &= P(-f_\infty, +g_\infty) \\ y_4 &= P(-f_\infty, -g_\infty) \end{aligned}$$

Write an equation that must be true if this is a stationary distribution.

Your answer should be an equation, possibly in terms of x_1, x_2, x_3, x_4 (from the previous part), y_1, y_2, y_3, y_4.

(e) [3 pts] Select all true statements.

☐ There are values of a, b, c, d such that the feedback model defined by p, q is irrelevant to the stationary distribution of the genre G.

☐ Even if $P(F_t, G_t)$ reaches a unique stationary distribution, it is possible that the marginal probabilities $P(F_t)$ and $P(G_t)$ do not reach unique stationary distributions.

☐ For any values of a, b, c, d, p, q, there is at least one stationary distribution.

○ None of the above