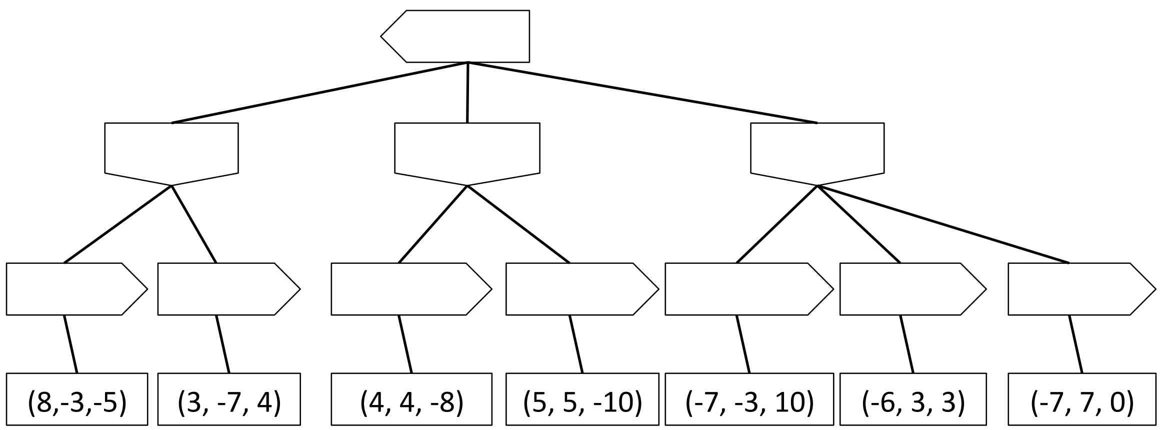
Q1.	3-Player Games	/8
Q2.	HMMs and Particle Filtering	/12
Q3.	Occupy Cal	/12
Q4.	Instantiated Elimination	/14
Q5.	Bayes' Net Representation	/12
Q6.	Exponential Utilities	/14
Q7.	Bayes Net CSPs	/9
Q8.	Q-Learning Strikes Back	/8
Q9.	Probability and Bayes Nets	/11
	Total	/100

Q1. [8 pts] 3-Player Games

(a) [4 pts] **A 3-Player Game.**

Consider the 3-player game shown below. The player going first (top of the tree) is the Left player, the player going second is the Middle player, and the player going last is the Right player, optimizing the left, middle and right utility value respectively. Fill in the values at all nodes. Note all players *maximize* their respective utility value shown in the tree.



(b) [4 pts] **Pruning for Zero-Sum 3-Player Game.** The same game tree is shown again below.

Now assume that we have the knowledge that the sum of the utilities of all 3 players is always zero. Under this assumption is any pruning possible similar to $\alpha - \beta$ pruning? If so mark the pruning on the tree above. If not, briefly explain why not below.

Q2. [12 pts] HMMs and Particle Filtering

Consider a Markov Model with a binary state X (i.e., X_t is either 0 or 1). The transition probabilities are given as follows:

X_t	X_{t+1}	$P(X_{t+1} \mid X_t)$
0	0	0.9
0	1	0.1
1	0	0.5
1	1	0.5

(a) [2 pts] The prior belief distribution over the initial state X_0 is uniform, i.e., $P(X_0 = 0) = P(X_0 = 1) = 0.5$. After one timestep, what is the new belief distribution, $P(X_1)$?

X_1	$P(X_1)$
0	
1	

(b) [2 pts] Now, we incorporate sensor readings. The sensor model is parameterized by a number $\beta \in [0, 1]$:

X_t	E_t	$P(E_t \mid X_t)$
0	0	β
0	1	$(1 - \beta)$
1	0	$(1 - \beta)$
1	1	β

(c) [2 pts] At $t = 1$, we get the first sensor reading, $E_1 = 0$. Use your answer from part (a) to compute $P(X_1 = 0 \mid E_1 = 0)$. Leave your answer in terms of β.

(d) [2 pts] For what range of values of β will a sensor reading $E_1 = 0$ increase our belief that $X_1 = 0$? That is, what is the range of β for which $P(X_1 = 0 \mid E_1 = 0) > P(X_1 = 0)$?

(e) [2 pts] Unfortunately, the sensor breaks after just one reading, and we receive no further sensor information. Compute $P(X_\infty \mid E_1 = 0)$, the stationary distribution *very many* timesteps from now.

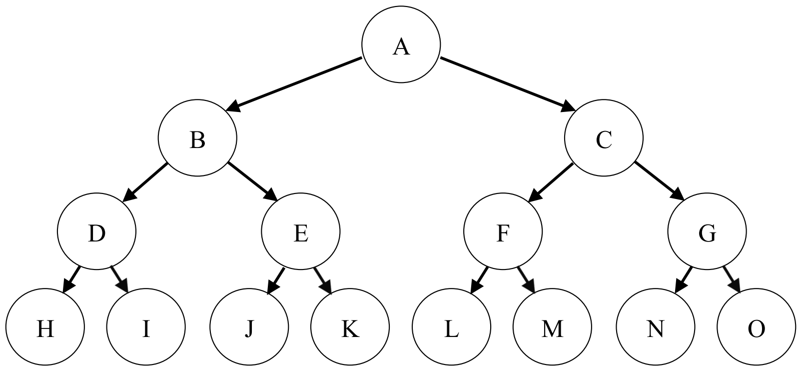
X_∞	$P(X_\infty \mid E_1 = 0)$
0	
1	

(f) [2 pts] How would your answer to part (d) change if we never received the sensor reading E_1, i.e. what is $P(X_\infty)$ given no sensor information?

X_∞	$P(X_\infty)$
0	
1	

Q3. [12 pts] Occupy Cal

(a) [3 pts] You are at Occupy Cal, and the leaders of the protest are deciding whether or not to march on California Hall. The decision is made centrally and communicated to the occupiers via the "human microphone"; that is, those who hear the information repeat it so that it propagates outward from the center. This scenario is modeled by the following Bayes net:



A	$P(A)$
$+m$	0.5
$-m$	0.5

$\pi(X)$	X	$P(X\|\pi(X))$
$+m$	$+m$	0.9
$+m$	$-m$	0.1
$-m$	$+m$	0.1
$-m$	$-m$	0.9

Each random variable represents whether a given group of protestors hears instructions to march ($+m$) or not ($-m$). The decision is made at A, and both outcomes are equally likely. The protestors at each node relay what they hear to their two child nodes, but due to the noise, there is some chance that the information will be misheard. Each node except A takes the same value as its parent with probability 0.9, and the opposite value with probability 0.1, as in the conditional probability tables shown.

(b) [3 pts] Compute the probability that node A sent the order to march ($A = +m$) given that both B and C receive the order to march ($B = +m$, $C = +m$).

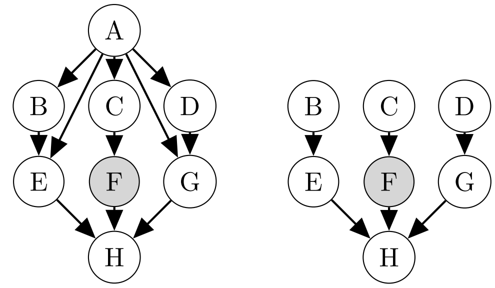
(c) [3 pts] Compute the probability that D receives the order $+m$ given that A sent the order $+m$.

You are at node D, and you know what orders have been heard at node D. Given your orders, you may either decide to march (*march*) or stay put (*stay*). (Note that these actions are distinct from the orders $+m$ or $-m$ that you hear and pass on. The variables in the Bayes net and their conditional distributions still behave exactly as above.) If you decide to take the action corresponding to the decision that was actually made at A (not necessarily corresponding to your orders!), you receive a reward of $+1$, but if you take the opposite action, you receive a reward of -1.

(d) [3 pts] Given that you have received the order $+m$, what is the expected utility of your optimal action? (Hint: your answer to part (b) may come in handy.)

Q4. [14 pts] Instantiated Elimination

(a) Difficulty of Elimination. Consider answering $P(H \mid +f)$ by variable elimination in the Bayes' nets N and N'.
Elimination order is alphabetical.
All variables are binary $+/-$.
Factor size is the number of unobserved variables in a factor made during elimination.



N **N′**

(i) [2 pts] What is the size of the largest factor made during variable elimination for N?

____

(ii) [2 pts] What is the size of the largest factor made during variable elimination for N'?

____

Variable elimination in N can take a lot of work! If only A were observed...

(b) Instantiation Sets. To simplify variable elimination in N, let's pick an *instantiation set* to pretend to observe, and then do variable elimination with these additional instantiations.

Consider the original query $P(H \mid +f)$, but let A be the instantiation set so $A = a$ is observed. Now the query is H with observations $F = +f, A = a$.

(i) [2 pts] What is the size of the largest factor made during variable elimination with the $A = a$ instantiation? ______________________

(ii) [1 pt] Given a Bayes' net over n binary variables with k variables chosen for the instantiation set, how many instantiations of the set are there?

____

(c) Inference by Instantiation. Let's answer $P(H \mid +f)$ by variable elimination with the instantiations of A.

(i) [2 pts] What quantity does variable elimination for $P(H \mid +f)$ with the $A = +a$ instantiation compute *without normalization*? That is, which choices are equal to the entries of the last factor made by elimination?

○ $P(H \mid +f)$ ○ $P(H, +a, +f)$ ○ $P(H, +f \mid +a)$
○ $P(H \mid +a)$ ○ $P(H, +a \mid +f)$ ○ $P(H \mid +a, +f)$

(ii) [2 pts] Let $I_+(H) = F(H, +a, +f)$ and $I_-(H) = F(H, -a, +f)$ be the last factors made by variable elimination with instantiations $A = +a$ and $A = -a$. Which choices are equal to $p(+h \mid +f)$?

○ $I_+(+h) \cdot p(+a) \cdot I_-(+h) \cdot p(-a)$
○ $I_+(+h) \cdot p(+a) + I_-(+h) \cdot p(-a)$
○ $I_+(+h) + I_-(+h)$
○ $\frac{I_+(+h) \cdot p(+a) \cdot I_-(+h) \cdot p(-a)}{\sum_h I_+(h) \cdot p(+a) \cdot I_-(h) \cdot p(-a)}$
○ $\frac{I_+(+h) \cdot p(+a) + I_-(+h) \cdot p(-a)}{\sum_h I_+(h) \cdot p(+a) + I_-(h) \cdot p(-a)}$
○ $\frac{I_+(+h) + I_-(+h)}{\sum_h I_+(h) + I_-(h)}$

(d) [3 pts] **Complexity of Instantiation.** What is the time complexity of instantiated elimination? Let n = number of variables, k = instantiation set size, f = size of the largest factor made by elimination without instantiation, and i = size of the largest factor made by elimination with instantiation. Mark the tightest bound. Variable elimination without instantiation is $O(n \exp(f))$.

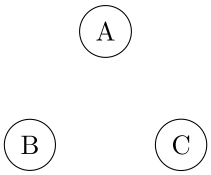
○ $O(n \exp(k))$ ○ $O(n \exp(i))$ ○ $O(n \exp(i + k))$
○ $O(n \exp(f))$ ○ $O(n \exp(f - k))$ ○ $O(n \exp(i/f))$

Q5. [12 pts] Bayes' Net Representation

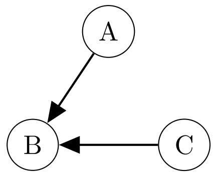
(a) [4 pts] Consider the joint probability table on the right.

Clearly fill in all circles corresponding to BNs that can correctly represent the distribution on the right. If no such BNs are given, clearly select *None of the above.*

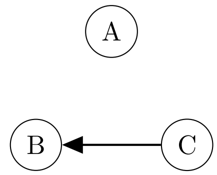
A	**B**	**C**	$\mathbf{P(A, B, C)}$
0	0	0	.15
0	0	1	.1
0	1	0	0
0	1	1	.25
1	0	0	.15
1	0	1	.1
1	1	0	0
1	1	1	.25



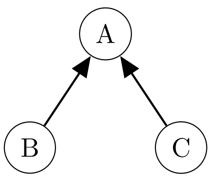
○ $\mathbf{G_1}$



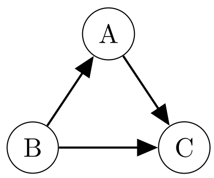
○ $\mathbf{G_2}$



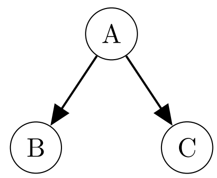
○ $\mathbf{G_3}$



○ $\mathbf{G_4}$



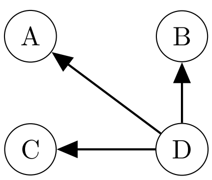
○ $\mathbf{G_5}$



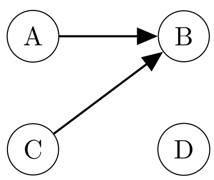
○ $\mathbf{G_6}$

○ **None of the above.**

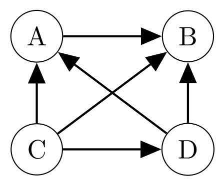
(b) [4 pts] You are working with a distribution over A, B, C, D that can be fully represented by just three probability tables: $P(A \mid D)$, $P(C \mid B)$, and $P(B, D)$. Clearly fill in the circles of those BNs that can correctly represent this distribution. If no such BNs are given, clearly select *None of the above.*



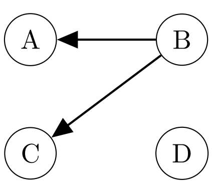
○ $\mathbf{G_1}$



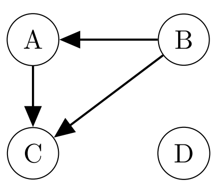
○ $\mathbf{G_2}$



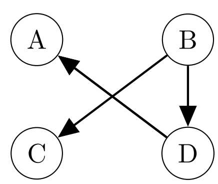
○ $\mathbf{G_3}$



○ $\mathbf{G_4}$



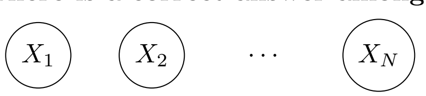
○ $\mathbf{G_5}$



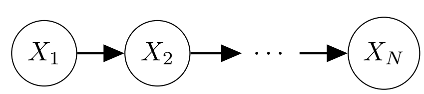
○ $\mathbf{G_6}$

○ **None of the above.**

(c) [4 pts] We are dealing with two probability distributions over N variables, where each variable can take on exactly d values. The distributions are represented by the two Bayes' Nets shown below. If S is the amount of storage required for the CPTs for $X_2, \ldots, X_N$ in D_1, how much storage is required for the CPTs for $X_2, \ldots, X_N$ in D_2? **There is a correct answer among the options.**



$\mathbf{D_1}$



$\mathbf{D_2}$

○ S ○ 2^S ○ Sd

○ S^2 ○ S^d ○ $S + 2^d$

Q6. [14 pts] Exponential Utilities

(a) The ghosts offer Pacman a deal: upon rolling a fair 6-sided die, they will give Pacman a reward equal to the number shown on the die minus a fee x, so he could win $1-x, 2-x, 3-x, 4-x, 5-x$ or $6-x$ with equal probability. Pacman can also refuse to play the game, getting 0 as a reward.

(i) [2 pts] Assume Pacman's utility is $U(r) = r$. Pacman should accept to play the game if and only if:

○ $x \leq 7/6$ ○ $x \leq 7/2$ ○ $x \leq 21/2$ ○ $x \leq 21$

(ii) [2 pts] Assume Pacman's utility is $U'(r) = 2^r$. Pacman should accept to play the game if and only if:

○ $x \leq \log_2(7/2)$ ○ $x \leq \log_2(20)$ ○ $x \leq \log_2(21)$ ○ $x \leq 21$

(b) For the following question assume that the ghosts have set the price of the game at $x = 4$. A fortune-teller is able to accurately predict whether the die roll will be even $(2, 4, 6)$ or odd $(1, 3, 5)$.

(i) [3 pts] Assume Pacman's utility is $U(r) = r$. The VPI (value of perfect information) of the prediction is:

○ 0 ○ $\frac{1}{16}$ ○ $\frac{7}{8}$ ○ 1 ○ $\frac{7}{4}$

(ii) [3 pts] Assume Pacman's utility is $U'(r) = 2^r$. The VPI of the prediction is:

○ 0 ○ $\frac{1}{16}$ ○ $\frac{7}{8}$ ○ 1 ○ $\frac{7}{4}$

(c) [4 pts] For simplicity the following question concerns only Markov Decision Processes (MDPs) with no discounting ($\gamma = 1$) and parameters set up such that the total reward is always finite. Let J be the total reward obtained in the MDP:

$$J = \sum_{t=1}^{\infty} r(S_t, A_t, S_{t+1}).$$

The utility we've been using implicitly for MDPs with no discounting is $U(J) = J$. The value function $V(s)$ is equal to the maximum expected utility $E[U(J)] = E[J]$ if the start state is s, and it obeys the Bellman equation seen below:

$$V^*(s) = \max_a \sum_{s'} T(s, a, s')(r(s, a, s') + V^*(s')).$$

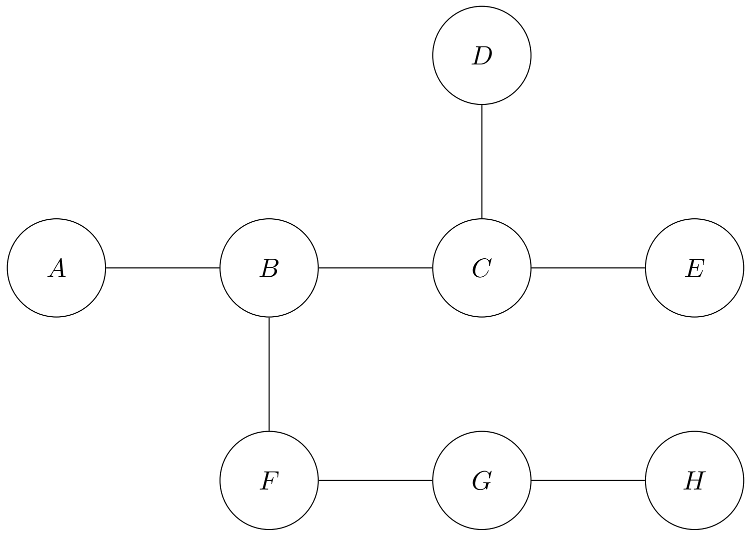
Now consider using the exponential utility $U'(J) = 2^J$ for MDPs. Write down the corresponding Bellman equation for $W^*(s)$, the maximum expected exponential utility $E[U'(J)] = E[2^J]$ if the start state is s.

$$W^*(s) = \max_a \underline{\hspace{6cm}} W^*(s')$$

Q7. [9 pts] Bayes Net CSPs

(a) For the following Bayes' Net structures that are missing a direction on their edges, assign a direction to each edge such that the Bayes' Net structure implies the requested conditional independences and such that the Bayes' Net structure does not imply the conditional independences requested not to be true. Keep in mind that Bayes' Nets cannot have directed cycles.

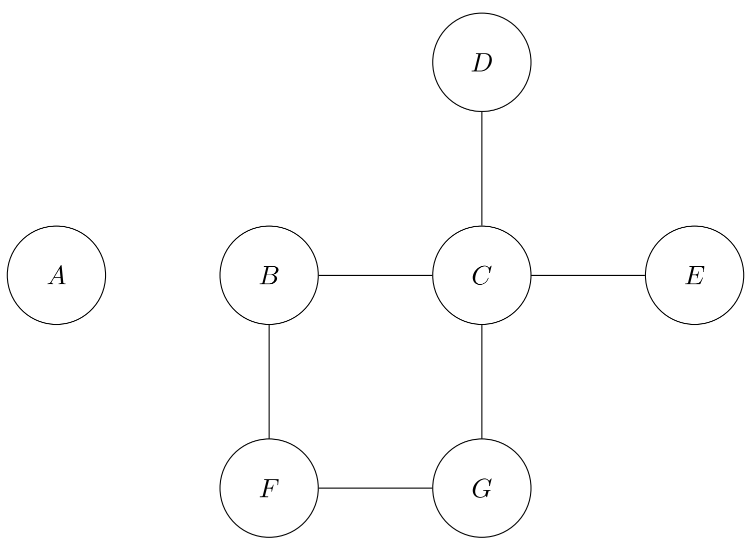
(i) [2 pts]



Constraints:

- $D \perp\!\!\!\perp G$
- not $D \perp\!\!\!\perp A$
- $D \perp\!\!\!\perp E$
- $H \perp\!\!\!\perp F$

(ii) [2 pts]

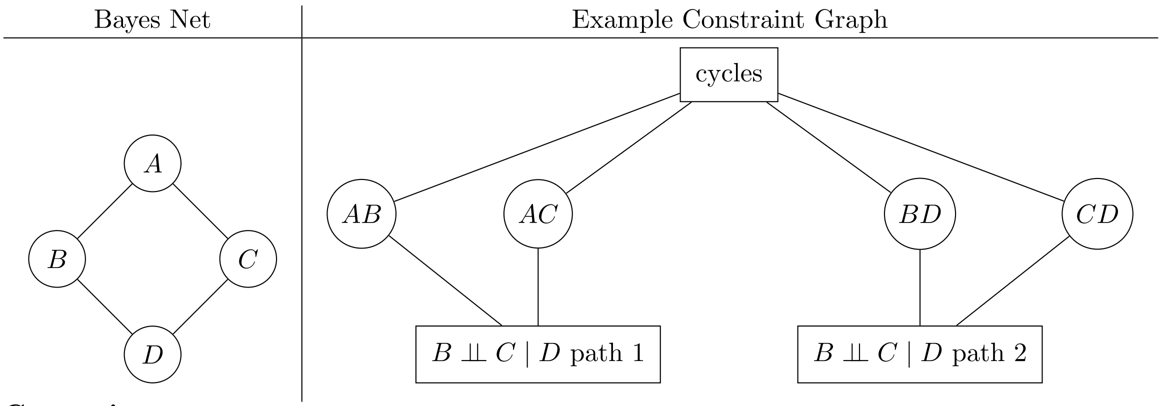


Constraints:

- $D \perp\!\!\!\perp F$
- not $D \perp\!\!\!\perp G$
- $D \perp\!\!\!\perp E$
- Bayes Net has no directed cycles

(b) For each of the following Bayes Nets and sets of constraints draw a constraint graph for the CSP. Remember that the constraint graph for a CSP with non-binary constraints, i.e., constraints that involve more than two variables, is drawn as a rectangle with the constraint connected to a node for each variable that participates in that constraint. A simple example is given below.

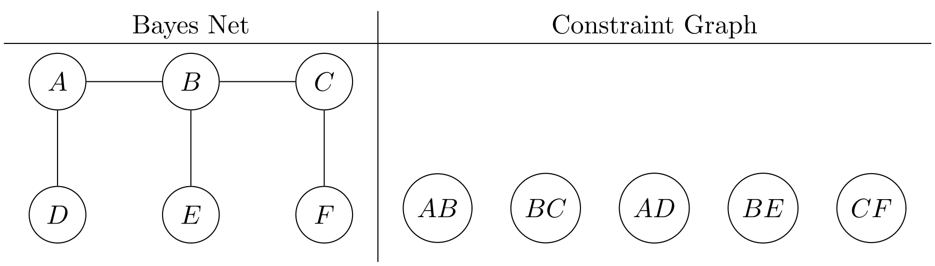
Note: As shown in the example below, if a constraint can be broken up into multiple constraints, do so.



Constraints:

- $B \perp\!\!\!\perp C \mid D$
- No directed cycles

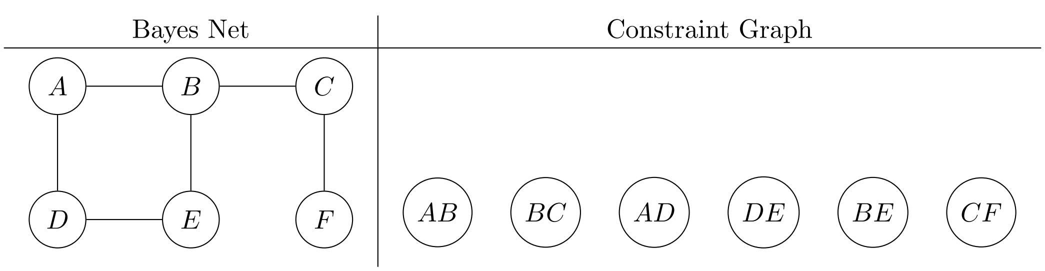
(i) [2 pts]



Constraints:

- $A \perp\!\!\!\perp F \mid E$
- not $D \perp\!\!\!\perp C$

(ii) [3 pts]

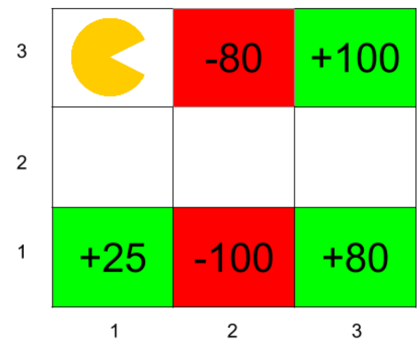


Constraints:

- $A \perp\!\!\!\perp E \mid F$
- $C \perp\!\!\!\perp E$
- No directed cycles

Q8. [8 pts] Q-Learning Strikes Back

Consider the grid-world given below and Pacman who is trying to learn the optimal policy. If an action results in landing into one of the shaded states the corresponding reward is awarded during that transition. All shaded states are terminal states, i.e., the MDP terminates once arrived in a shaded state. The other states have the *North, East, South, West* actions available, which deterministically move Pacman to the corresponding neighboring state (or have Pacman stay in place if the action tries to move out of the grad). Assume the discount factor $\gamma = 0.5$ and the Q-learning rate $\alpha = 0.5$ for all calculations. Pacman starts in state (1, 3).



(a) [2 pts] What is the value of the optimal value function V^* at the following states:

$V^*(3,2) =$ ________ $V^*(2,2) =$ ________ $V^*(1,3) =$ ________

(b) [3 pts] The agent starts from the top left corner and you are given the following episodes from runs of the agent through this grid-world. Each line in an Episode is a tuple containing (s, a, s', r).

Episode 1	Episode 2	Episode 3
(1,3), S, (1,2), 0	(1,3), S, (1,2), 0	(1,3), S, (1,2), 0
(1,2), E, (2,2), 0	(1,2), E, (2,2), 0	(1,2), E, (2,2), 0
(2,2), S, (2,1), -100	(2,2), E, (3,2), 0	(2,2), E, (3,2), 0
	(3,2), N, (3,3), +100	(3,2), S, (3,1), +80

Using Q-Learning updates, what are the following Q-values after the above three episodes:

Q((3,2),N) = ________ Q((1,2),S) = ________ $Q((2,2), E) =$ ________

(c) Consider a feature based representation of the Q-value function:

$$Q_f(s,a) = w_1 f_1(s) + w_2 f_2(s) + w_3 f_3(a)$$

$f_1(s)$: The x coordinate of the state $f_2(s)$: The y coordinate of the state

$$f_3(N) = 1,\ f_3(S) = 2,\ f_3(E) = 3,\ f_3(W) = 4$$

(i) [2 pts] Given that all w_i are initially 0, what are their values after the first episode:

$w_1 =$ ________ $w_2 =$ ________ $w_3 =$ ________

(ii) [1 pt] Assume the weight vector w is equal to $(1, 1, 1)$. What is the action prescribed by the Q-function in state $(2, 2)$?

____

Q9. [11 pts] Probability and Bayes Nets

(a) [2 pts] Suppose $A \perp\!\!\!\perp B$. Determine the missing entries (x, y) of the joint distribution $P(A, B)$, where A and B take values in $\{0, 1\}$.

$$\begin{aligned} P(A = 0, B = 0) &= 0.1 \\ P(A = 0, B = 1) &= 0.3 \\ P(A = 1, B = 0) &= x \\ P(A = 1, B = 1) &= y \end{aligned}$$

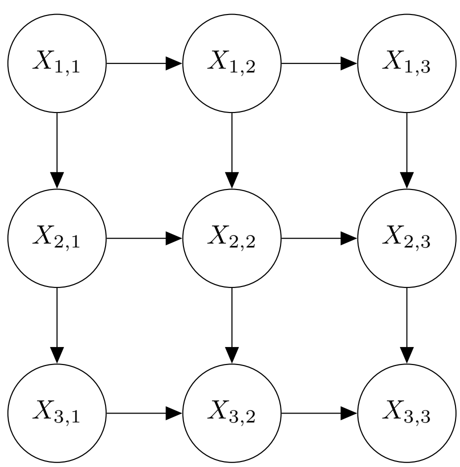
$x =$ ______________, $y =$ ______________

(b) [3 pts] Suppose $B \perp\!\!\!\perp C \mid A$. Determine the missing entries (x, y, z) of the joint distribution $P(A, B, C)$.

$$\begin{aligned} P(A = 0, B = 0, C = 0) &= 0.01 \\ P(A = 0, B = 0, C = 1) &= 0.02 \\ P(A = 0, B = 1, C = 0) &= 0.03 \\ P(A = 0, B = 1, C = 1) &= x \\ P(A = 1, B = 0, C = 0) &= 0.01 \\ P(A = 1, B = 0, C = 1) &= 0.1 \\ P(A = 1, B = 1, C = 0) &= y \\ P(A = 1, B = 1, C = 1) &= z \end{aligned}$$

$x =$ ______________, $y =$ ______________, $z =$ ______________

(c) [3 pts] For this question consider the Bayes' Net below with 9 variables.



Which random variables are independent of $X_{3,1}$? (Leave blank if the answer is none.)

○ $X_{1,1}$ ○ $X_{1,2}$ ○ $X_{1,3}$ ○ $X_{2,1}$ ○ $X_{2,2}$ ○ $X_{2,3}$ ○ $X_{3,2}$ ○ $X_{3,3}$

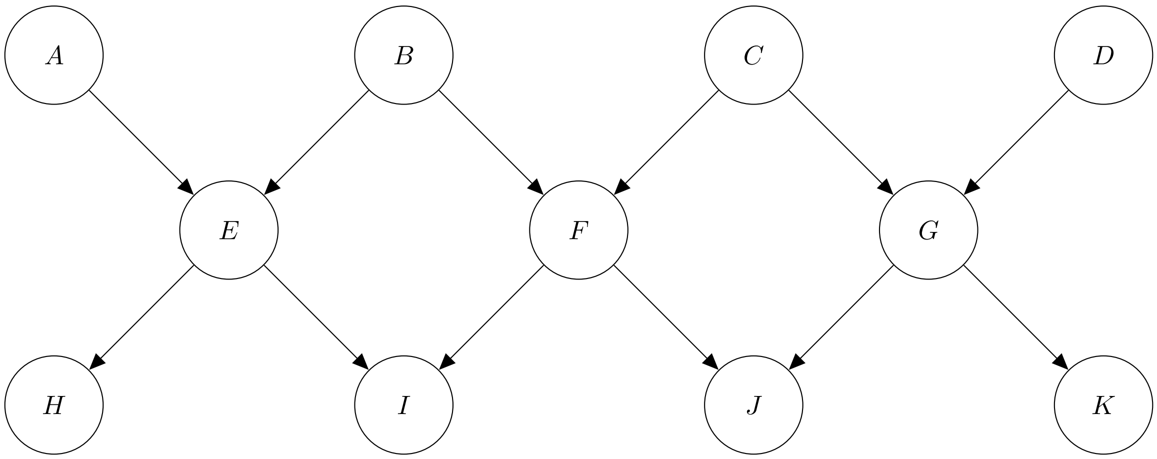
Which random variables are independent of $X_{3,1}$ *given* $X_{1,1}$? (Leave blank if the answer is none.)

○ $X_{1,2}$ ○ $X_{1,3}$ ○ $X_{2,1}$ ○ $X_{2,2}$ ○ $X_{2,3}$ ○ $X_{3,2}$ ○ $X_{3,3}$

Which random variables are independent of $X_{3,1}$ *given* $X_{1,1}$ and $X_{3,3}$? (Leave blank if the answer is none.)

○ $X_{1,2}$ ○ $X_{1,3}$ ○ $X_{2,1}$ ○ $X_{2,2}$ ○ $X_{2,3}$ ○ $X_{3,2}$

For the following questions we will consider the following Bayes' Net:



(d) Consider a run of Gibbs sampling for the query $P(B, C \mid +h, +i, +j)$. The current sample value is $+a, +b, +c, +d, +e, +f, +g, +h, +i, +j, +k$. For each of the following scenarios, write out an expression for the distribution Gibbs sampling would sample from. *Your expression should contain only conditional probabilities available in the network, and your expression should contain a minimal number of such conditional probabilities.*

(i) [1 pt] If A were to be sampled next, the distribution over A to sample from would be:

____

(ii) [1 pt] If F were to be sampled next, the distribution over F to sample from would be:

____

(iii) [1 pt] If K were to be sampled next, the distribution over K to sample from would be:

____