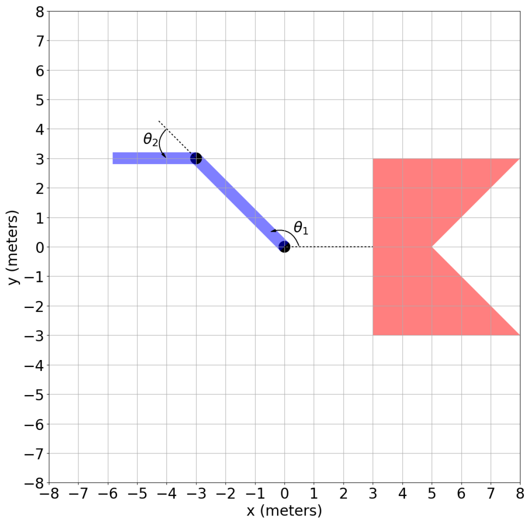
Exam 2 - Practice Problems

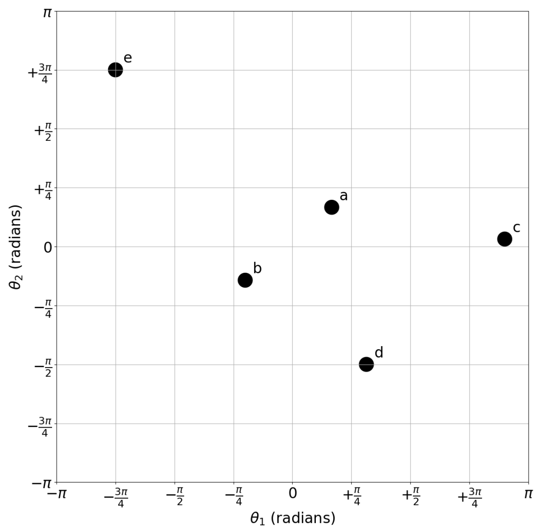
CS 182 - Artificial Intelligence

Note: we have pulled many of these problems from previous Berkeley exams which can be found at: http://ai.berkeley.edu/exams.html. We cannot promise that all of those exams cover relevant material but will note that there are additional questions of these types that are applicable for certain parts of this review.

1 Robot Motion Planning

Problem 1: We have a two link robotic arm with rotational joints one fixed at (0,0) and the other fixed to the end of the first link. Both links are $3\sqrt{2}$ meters long and the 0 angle is defined by the dashed line in the left side of the picture below (from the x-axis for θ_1 and aligned with the first link for θ_2. There is one obstacle, the large red polygon. We have decided to build a PRM to plan collision free paths (where a collision is defined as any part of the robot touching an obstacle). We samples 5 points in configuration space labeled a to e in the configuration space diagram on the right side of the picture below. Which points can we keep as milestone for our PRM?
Hint: $45^o = \frac{\pi}{4}$ *radians and 45-45-90 triangles have side lengths 1-1-*$\sqrt{2}$*)*





Problem 2: When would you consider using trajectory optimization over RRT?

Problem 3: Consider the AdaptiveStep-RRT (AS-RRT) algorithm (the blue lines indicate additional steps taken in AS-RRT compared to standard RRT):

- Algorithm AS-RRT (input: start state s_0, goal s_G, initial tree $T = s_0$, max extend distance d_{max}, min extend distance d_{min}, step size change δ)
- For $i = 1, \ldots, N$:
- Sample states $s \in \mathcal{S}$ until s is collision-free
- Find closest state $s_c \in T$
- Set $d = d_{max}$
- While($d > d_{min}$):
- Extend s_c toward s with distance d resulting in state s'
- If isCollisionFreePath(s_c, s'):
- Add s' to T
- Break
- Else:
- $d = d - \delta$
- Return T

a) Assume that we are using a k-d tree where insert and nearest searches take $O(log|T|)$ and that sample, extend, and isCollisionFreePath are $O(1)$ operations, what is the worst case time coplexity for one iteration of AdaptiveStep-RRT?

b) Do you think that the assumption that isCollisionFreePath is always an $O(1)$ operation is reasonable? Why or why not?

2 Probability Review

Note: I wouldn't expect a question dedicated to this on the exam but you need to be comfortable with joint vs. marginal vs. conditional probabilites and how to move between them when working with HMMs and Bayes' Nets.

(a) [2 pts] Fill in the circles of **all** expressions that are equivalent to $\mathbf{P(A,B \mid C)}$, **given no independence assumptions**:

○ $\frac{P(C|A)\ P(A|B)\ P(B)}{P(C)}$

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

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

○ $\frac{P(B,C|A)\ P(A)}{P(B,C)}$

○ $\frac{P(A|C)\ P(B,C)}{P(C)}$

○ $P(A \mid B, C)\ P(B \mid C)$

○ None of the above.

(b) [2 pts] Fill in the circles of **all** expressions that are equivalent to $\mathbf{P(A,B \mid C)}$, **given that** $\mathbf{A \perp\!\!\!\perp B \mid C}$:

○ $\frac{P(C|A)\ P(A|B)\ P(B)}{P(C)}$

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

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

○ $\frac{P(B,C|A)\ P(A)}{P(B,C)}$

○ $\frac{P(A|C)\ P(B,C)}{P(C)}$

○ $P(A \mid B, C)\ P(B \mid C)$

○ None of the above.

(c) [2 pts] Fill in the circles of **all** expressions that are equivalent to $\mathbf{P(A \mid B, C)}$, **given no independence assumptions**:

○ $\frac{P(C|A)\ P(A|B)\ P(B)}{P(C)}$

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

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

○ $\frac{P(B,C|A)\ P(A)}{P(B,C)}$

○ $\frac{P(A|C)\ P(B,C)}{P(C)}$

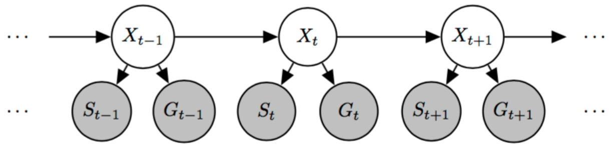
○ $P(A \mid B, C)\ P(B \mid C)$

○ None of the above.

3 HMMs

Problem 1: Transportation researchers are trying to improve traffic in the city but, in order to do that, they first need to estimate the location of each of the cars in the city. They need our help to model this problem as an inference problem of an HMM. For this question, assume that only one car is being modeled.

a) The structure of this modified HMM is given below, which includes X, the location of the car; S, the noisy location of the car from the signal strength at a nearby cell phone tower; and G, the noisy location of the car from GPS.



We want to perform filtering with this HMM. That is, we want to compute the belief $P(x_t|s_{1:t}, g_{1:t})$, the probability of a state x_t given all past and current observations.

The **dynamics update** expression has the following form:

$$P(x_t|s_{1:t-1}, g_{1:t-1}) = \underline{\quad\textbf{(i)}\quad}\ \underline{\quad\textbf{(ii)}\quad}\ \underline{\quad\textbf{(iii)}\quad}\ P(x_{t-1}|s_{1:t-1}, g_{1:t-1}).$$

Complete the expression by choosing the option that fills in each blank.

(i) [1 pt] ○ $P(s_{1:t}, g_{1:t})$ ○ $P(s_{1:t-1}, g_{1:t-1})$ ○ $P(s_{1:t})P(g_{1:t})$ ○ $P(s_{1:t-1})P(g_{1:t-1})$ ○ 1

(ii) [1 pt] ○ $\sum_{x_t}$ ○ $\sum_{x_{t-1}}$ ○ $\max_{x_{t-1}}$ ○ $\max_{x_t}$ ○ 1

(iii) [1 pt] ○ $P(x_{t-2}, x_{t-1})$ ○ $P(x_{t-1} \mid x_{t-2})$ ○ $P(x_{t-1}, x_t)$ ○ $P(x_t|x_{t-1})$ ○ 1

The **observation update** expression has the following form:

$$P(x_t|s_{1:t}, g_{1:t}) = \underline{\quad\textbf{(iv)}\quad}\ \underline{\quad\textbf{(v)}\quad}\ \underline{\quad\textbf{(vi)}\quad}\ P(x_t|s_{1:t-1}, g_{1:t-1}).$$

Complete the expression by choosing the option that fills in each blank.

(iv) [1 pt] ○ $P(s_t, g_t|s_{1:t-1}, g_{1:t-1})$ ○ $P(s_{1:t-1}, g_{1:t-1}|s_t, g_t)$ ○ $P(s_t|s_{1:t-1})P(g_t|g_{1:t-1})$
○ $P(s_{1:t-1}|s_t)P(g_{1:t-1}|g_t)$ ○ $\frac{1}{P(s_t, g_t|s_{1:t-1}, g_{1:t-1})}$ ○ $\frac{1}{P(s_{1:t-1}, g_{1:t-1}|s_t, g_t)}$
○ $\frac{1}{P(s_t|s_{1:t-1})P(g_t|g_{1:t-1})}$ ○ $\frac{1}{P(s_{1:t-1}|s_t)P(g_{1:t-1}|g_t)}$ ○ 1

(v) [1 pt] ○ $\sum_{x_{t-1}}$ ○ $\sum_{x_t}$ ○ $\max_{x_{t-1}}$ ○ $\max_{x_t}$ ○ 1

(vi) [1 pt] ○ $P(s_{t-1}|x_{t-1})P(g_{t-1}|x_{t-1})$ ○ $P(x_t, s_t)P(x_t, g_t)$ ○ $P(x_t, s_t, g_t)$
○ $P(x_{t-1}, s_{t-1})P(x_{t-1}, g_{t-1})$ ○ $P(x_{t-1}, s_{t-1}, g_{t-1})$ ○ $P(x_t|s_t)P(x_t|g_t)$
○ $P(x_{t-1}|s_{t-1})P(x_{t-1}|g_{t-1})$ ○ $P(s_t|x_t)P(g_t|x_t)$ ○ 1

b) It turns out that if the car moves too fast, the quality of the cell phone signal decreases. Thus, the signal-dependent location S_t not only depends on the current state X_t but it also depends on the previous state X_{t-1}. Thus, we modify our original HMM for a new more accurate one, which is given below. Again, we want to compute the belief $P(xt|s_{1:t}, g_{1:t})$. In this part we consider an update that combines the dynamics and observation update in a single update.

$$P(x_t|s_{1:t}, g_{1:t}) = \underline{\quad\textbf{(i)}\quad}\ \underline{\quad\textbf{(ii)}\quad}\ \underline{\quad\textbf{(iii)}\quad}\ \underline{\quad\textbf{(iv)}\quad}\ P(x_{t-1}|s_{1:t-1}, g_{1:t-1}).$$

Complete the **forward update** expression by choosing the option that fills in each blank.

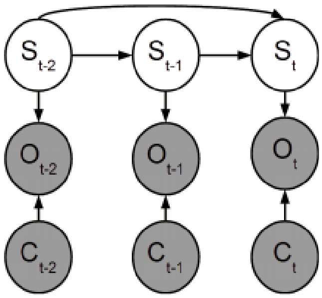
(i) [1 pt] ○ $P(s_t, g_t|s_{1:t-1}, g_{1:t-1})$ ○ $P(s_{1:t-1}, g_{1:t-1}|s_t, g_t)$ ○ $P(s_t|s_{1:t-1})P(g_t|g_{1:t-1})$
○ $\frac{1}{P(s_t, g_t|s_{1:t-1}, g_{1:t-1})}$ ○ $\frac{1}{P(s_{1:t-1}, g_{1:t-1}|s_t, g_t)}$ ○ $P(s_{1:t-1}|s_t)P(g_{1:t-1}|g_t)$
○ $\frac{1}{P(s_t|s_{1:t-1})P(g_t|g_{1:t-1})}$ ○ $\frac{1}{P(s_{1:t-1}|s_t)P(g_{1:t-1}|g_t)}$ ○ 1

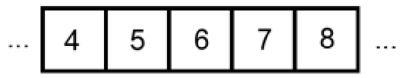
(ii) [1 pt] ○ $\sum_{x_{t-1}}$ ○ $\sum_{x_t}$ ○ $\max_{x_{t-1}}$ ○ $\max_{x_t}$ ○ 1

(iii) [1 pt] ○ $P(x_{t-2}, x_{t-1}, s_{t-1})P(x_{t-1}, g_{t-1})$ ○ $P(x_{t-1}, x_t, s_t)P(x_t, g_t)$ ○ $P(s_{t-1}, g_{t-1}|x_{t-1})$
○ $P(s_{t-1}|x_{t-2}, x_{t-1})P(g_{t-1}|x_{t-1})$ ○ $P(s_t|x_{t-1}, x_t)P(g_t|x_t)$ ○ $P(s_t, g_t|x_t)$
○ $P(x_{t-2}, x_{t-1}|s_{t-1})P(x_{t-1}|g_{t-1})$ ○ $P(x_{t-1}, x_t|s_t)P(x_t|g_t)$ ○ 1
○ $P(x_{t-2}, x_{t-1}, s_{t-1}, g_{t-1})$ ○ $P(x_{t-1}, x_t, s_t, g_t)$

(iv) [1 pt] ○ $P(x_{t-1}, x_t)$ ○ $P(x_t|x_{t-1})$ ○ $P(x_{t-2}, x_{t-1})$ ○ $P(x_{t-1}|x_{t-2})$ ○ 1

Problem 2: Pacman is trying to hunt a ghost in an infinite hallway with positions labeled as in the picture below. He's become more technologically savvy, and decided to locate find the ghosts actual position, S_t, using some sensors he set up. From the sensors, Pacman can find, at each time step, a noisy reading of the ghost's location, O_t. However, just as Pacman has gained technology, so has the ghost. It is able to cloak itself at each time step, given by C_t, adding extra noise to Pacman's sensor readings.





Pacman has generated an error model, given in the table below, for the sensor depending on whether the ghost is cloaked or not. Pacman has also generated a dynamics model, given in the table below, that takes into account the position of the ghost at the two previous timesteps.

Dynamics model:

$P(S_t|S_{t-1}, S_{t-2}) = F(D_1, D_2)$
$D_1 = |S_t - S_{t-1}|$
$D_2 = |S_t - S_{t-2}|$

D_1	D_2	$F(D_1, D_2)$
0	0	0.7
0	1	0.2
0	2	0
1	0	0.3
1	1	0.3
1	2	0.5

Observation model:

$P(O_t|S_t, C_t) = E(C_t, D)$
$D = |O_t - S_t|$

C	D	E(C, D)
+	0	0.4
+	1	0.2
+	2	0.1
-	0	0.6
-	1	0.2
-	2	0

a) Assume that you currently have the following two particles: $(S_6 = 7, S_7 = 8)$ and $(S_6 = 6, S_7 = 6)$. Compute the weights for each particle given the observations:

$$C_6 = +, C_7 = -, O_6 = 5, O_7 = 8$$

b) Assume that Pacman can no longer see whether the ghost is cloaked or not, but assumes that it will be cloaked at each timestep with probability 0.5. Compute the weights for each particle given the observations $O_6 = 5, O_7 = 8$:

c) To prevent error propagation, assume that after weighting the particles and resampling, one of the particles you end up with is $(S_6 = 6, S_7 = 7)$.

(a) What is the probability that after passing this particle through the dynamics model it becomes $(S_7 = 6, S_8 = 6)$?

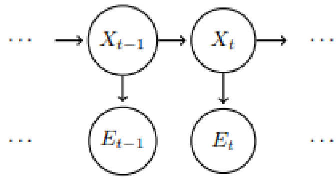
(b) What is the probability the particle becomes $(S_7 = 7, S_8 = 8)$?

d) To again decouple this part from previous parts, assume that you have the following three particles with the specified weights: $(S_7 = 5, S_8 = 6) : .1$, $(S_7 = 7, S_8 = 6) : .25$, $(S_7 = 7, S_8 = 7) : .3$. What is Pacman's belief for the ghost's position at time t = 8?

Position	$P(S_8)$
$S_8 = 5$	
$S_8 = 6$	
$S_8 = 7$	
$S_8 = 8$	

Problem 3: Below is a full derivation of the forward algorithm updates for Hidden Markov Models. As seen in lecture, we used $e1 : t$ to denote all the evidence variables $e1, e2, \ldots, et$. For reference, the Bayes net corresponding to the usual Hidden Markov Model is shown on the right side of the derivation below.

$$
\begin{aligned}
P(x_t|e_{1:t}) &\propto P(x_t, e_{1:t}) && (1)\\
&= \sum_{x_{t-1}} P(x_{t-1}, x_t, e_{1:t}) && (2)\\
&= \sum_{x_{t-1}} P(x_{t-1}, x_t, e_{1:t-1}, e_t) && (3)\\
&= \sum_{x_{t-1}} P(e_t \mid x_{t-1}, x_t, e_{1:t-1}) P(x_{t-1}, x_t, e_{1:t-1}) && (4)\\
&= \sum_{x_{t-1}} P(e_t \mid x_t) P(x_{t-1}, x_t, e_{1:t-1}) && (5)\\
&= \sum_{x_{t-1}} P(e_t \mid x_t) P(x_t \mid x_{t-1}, e_{1:t-1}) P(x_{t-1}, e_{1:t-1}) && (6)\\
&= \sum_{x_{t-1}} P(e_t \mid x_t) P(x_t \mid x_{t-1}) P(x_{t-1}, e_{1:t-1}) && (7)\\
&= P(e_t \mid x_t) \sum_{x_{t-1}} P(x_t \mid x_{t-1}) P(x_{t-1}, e_{1:t-1}) && (8)
\end{aligned}
$$



(a) [2 pts] The following assumption(s) are needed to justify going from step (4) to step (5): (select all that apply)

- ○ $E_t \perp\!\!\!\perp X_{t-1} \mid X_t$
- ○ $E_t \perp\!\!\!\perp E_k \mid X_t$ for all $1 \leq k \leq t-1$
- ○ $E_t \perp\!\!\!\perp E_k$ for all $1 \leq k \leq t-1$
- ○ $E_t \perp\!\!\!\perp E_{t+1} \mid X_t$
- ○ $E_t \perp\!\!\!\perp E_k \mid X_{t-1}$ for all $1 \leq k \leq t-1$
- ○ $X_t \perp\!\!\!\perp E_{t+1} \mid X_{t+1}$
- ○ $X_t \perp\!\!\!\perp E_k \mid X_{t-1}$ for all $1 \leq k \leq t-1$
- ○ none

(b) [2 pts] The following assumption(s) are needed to justify going from step (5) to step (6): (select all that apply)

- ○ $E_t \perp\!\!\!\perp X_{t-1} \mid X_t$
- ○ $E_t \perp\!\!\!\perp E_k \mid X_t$ for all $1 \leq k \leq t-1$
- ○ $E_t \perp\!\!\!\perp E_k$ for all $1 \leq k \leq t-1$
- ○ $E_t \perp\!\!\!\perp E_{t+1} \mid X_t$
- ○ $E_t \perp\!\!\!\perp E_k \mid X_{t-1}$ for all $1 \leq k \leq t-1$
- ○ $X_t \perp\!\!\!\perp E_{t+1} \mid X_{t+1}$
- ○ $X_t \perp\!\!\!\perp E_k \mid X_{t-1}$ for all $1 \leq k \leq t-1$
- ○ none

(c) [2 pts] The following assumption(s) are needed to justify going from step (6) to step (7): (select all that apply)

- ○ $E_t \perp\!\!\!\perp X_{t-1} \mid X_t$
- ○ $E_t \perp\!\!\!\perp E_k \mid X_t$ for all $1 \leq k \leq t-1$
- ○ $E_t \perp\!\!\!\perp E_k$ for all $1 \leq k \leq t-1$
- ○ $E_t \perp\!\!\!\perp E_{t+1} \mid X_t$
- ○ $E_t \perp\!\!\!\perp E_k \mid X_{t-1}$ for all $1 \leq k \leq t-1$
- ○ $X_t \perp\!\!\!\perp E_{t+1} \mid X_{t+1}$
- ○ $X_t \perp\!\!\!\perp E_k \mid X_{t-1}$ for all $1 \leq k \leq t-1$
- ○ none

Problem 4: 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) 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	

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

1.

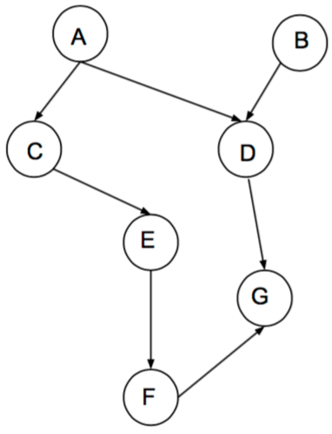
X_t	E_t	$P(E_t \mid X_t)$
0	0	β
0	1	$(1 - \beta)$
1	0	$(1 - \beta)$
1	1	β

b) 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.
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 $P(X_1 = 0 \mid E_1 = 0) > P(X_1 = 0)$?

4 Bayes' Nets

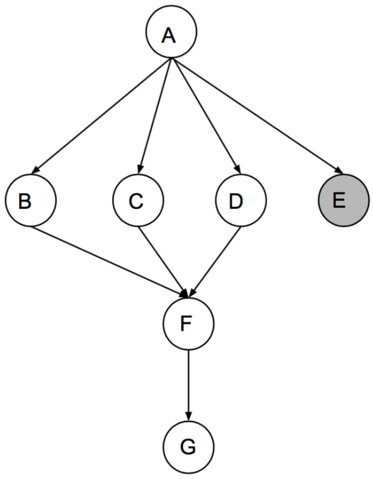
Problem 1: Consider a Bayes Net with the following graph:



Which of the following are guaranteed to be true without making any additional conditional independence assumptions, other than those implied by the graph? (Cirlce all true statements)

- $P(A|C, E) = P(A|C)$
- $P(A, E|G) = P(A|G) \cdot P(E|G)$
- $P(A|B, G) = P(A|G)$
- $P(A|B = b) = P(A)$
- $P(A, B|F) = P(A|F) \cdot P(B|F)$
- $P(E, G|D) = P(E|D) \cdot P(G|D)$

Problem 2: Now consider a Bayes Net with the following graph:



The factors associated with the Bayes Net are $P(A), P(B|A), P(C|A), P(D|A), P(E|A), P(F|B, C, D)$ and $P(G|F)$. We will consider variable elimination to answer the query $P(G| + e)$.

a) Suppose the first variable we eliminate is A. Provide an expression for the resulting factor as a function of the original factors.

b) Instead suppose the first variable we eliminate is B. Provide an expression for the resulting factor as a function of the original factors.

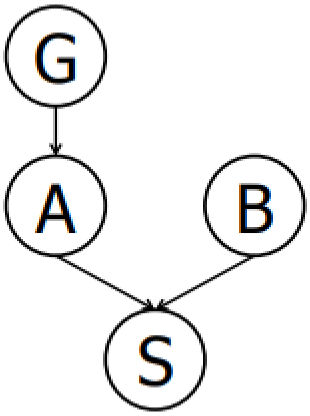
c) Instead suppose the first variable we eliminate is F. Provide an expression for the resulting factor as a function of the original factors.

d) Instead suppose we eliminated the variables A, B, C, D, F in that order, and the single remaining factor is $f(+e, G)$. How do we obtain $P(G|+e)$ from this remaining factor? (Your answer should be in the form of an equation.)

Problem 3: Suppose that a patient can have a symptom (S) that can be caused by two different diseases (A and B). It is known that the variation of gene G plays a big role in the manifestation of disease A. The Bayes' Net and corresponding conditional probability tables for this situation are shown below. For each part, you may leave your answer as an arithmetic expression.

$\mathbb{P}(G)$	
$+g$	0.1
$-g$	0.9

$\mathbb{P}(A\|G)$		
$+g$	$+a$	1.0
$+g$	$-a$	0.0
$-g$	$+a$	0.1
$-g$	$-a$	0.9



$\mathbb{P}(B)$	
$+b$	0.4
$-b$	0.6

$\mathbb{P}(S\|A,B)$			
$+a$	$+b$	$+s$	1.0
$+a$	$+b$	$-s$	0.0
$+a$	$-b$	$+s$	0.9
$+a$	$-b$	$-s$	0.1
$-a$	$+b$	$+s$	0.8
$-a$	$+b$	$-s$	0.2
$-a$	$-b$	$+s$	0.1
$-a$	$-b$	$-s$	0.9

a) Compute the following entry from the joint distribution: $P(+g, +a, +b, +s) =$

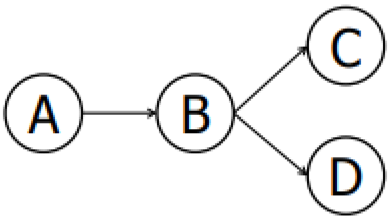
b) What is the probability that a patient has disease A?

c) What is the probability that a patient has disease A given that they have disease B?

d) What is the probability that a patient has the disease carrying gene variation G given that they have disease A?

e) What is the probability that a patient has the disease carrying gene variation G given that they have disease B?

Problem 4: Assume the following Bayes net, and the corresponding distributions over the variables in the Bayes net:



A	$\mathbb{P}(A)$
$+a$	$1/5$
$-a$	$4/5$

B	C	$\mathbb{P}(C\|B)$
$+b$	$+c$	$1/4$
$+b$	$-c$	$3/4$
$-b$	$+c$	$2/5$
$-b$	$-c$	$3/5$

A	B	$\mathbb{P}(B\|A)$
$+a$	$+b$	$1/5$
$+a$	$-b$	$4/5$
$-a$	$+b$	$1/2$
$-a$	$-b$	$1/2$

B	D	$\mathbb{P}(D\|B)$
$+b$	$+d$	$1/2$
$+b$	$-d$	$1/2$
$-b$	$+d$	$4/5$
$-b$	$-d$	$1/5$

a) Your task is now to estimate $P(+b|-a,-c,-d)$ using rejection sampling. Below are some samples that have been produced by prior sampling (that is, the rejection stage in rejection sampling hasn't happened yet). Cross out the samples that would be rejected by rejection sampling:

$-a$	$-b$	$+c$	$+d$	$-a$	$-b$	$-c$	$-d$
$+a$	$-b$	$-c$	$+d$	$-a$	$+b$	$+c$	$+d$
$-a$	$-b$	$+c$	$-d$	$+a$	$-b$	$-c$	$-d$

b) Using those samples, what value would you estimate for P(+b| − a, −c, −d) using rejection sampling?

c) Using the following three samples (which were generated using likelihood weighting), estimate $P(+b|-a,-c,-d)$ using likelihood weighting, or state why it cannot be computed.

$-a$	$-b$	$-c$	$-d$
$-a$	$+b$	$-c$	$-d$
$-a$	$-b$	$-c$	$-d$

5 ML

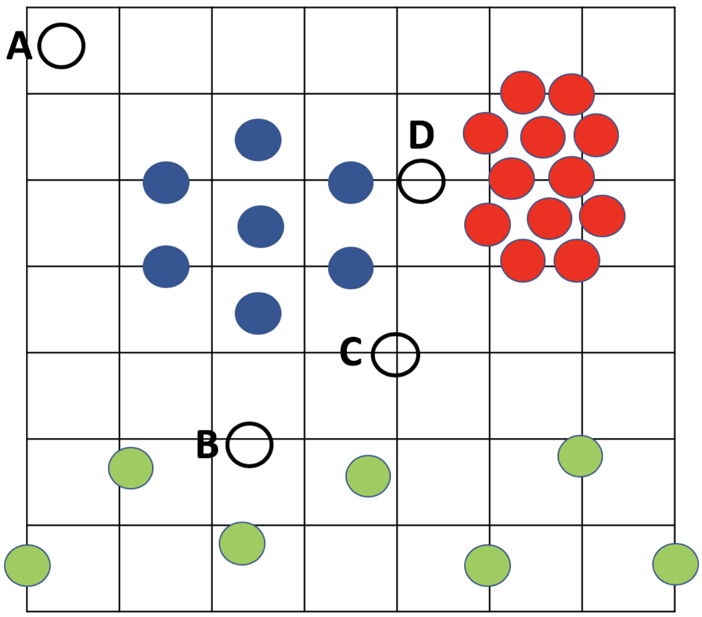
Problem 1:
For each of the following algorithms, describe how to evaluate whether k is too small:

- k-means?

- k-NN?

Problem 2:

Consider the following labeled data, along with four new points A, B, C, and D:



What labels would be assigned to points A, B, C, and D if $k = 1$?

What labels would be assigned to points A, B, C, and D if $k = 4$?

Problem 3:

You are a turnip farmer, and after a few years of tumultuous harvests, you decide to take a quantitative approach to agriculture. You decide to model your annual turnip yield (y) as a linear function of a single variable: the amount of rainfall in the previous year (x). So, after normalization, you have the following data points:

turnip yield (y)	rainfall (x)
24	10
33	14
34	12

Your are looking for a parameter θ, and your model is $h_\theta(x) = \theta x$.

1. Suppose you are using a quadratic loss function ($L(h_\theta(x), y) = (h_\theta(x) - y)^2$), and you can choose θ to be equal to either 0, 1, 2, 3, or 4. Which is the best option for your data?

2. Suppose you are using an L_1 loss function ($L(h_\theta(x), y) = |h_\theta(x) - y|$), and you can choose θ to be equal to either 0, 1, 2, or 3. Which is the best option for your data?

Problem 4:

A few years later, you are a very successful farmer due to your skillful use of regression. Now, you farm not only turnips, but a large variety of vegetables; in fact, you decide which vegetables to plant each year using logistic regression:

$$h_\theta(x) = \frac{1}{1 + e^{-\theta^T x}}$$

Based on the total rainfall, normalized average temperature, and normalized market price (your input variable x), you decide whether the harvest will or will not be profitable (your output variable y, where $profitable = 1$ and $not\ profitable = 0$). Using a training dataset, you calculate the vector of optimal parameters:

$$\theta = (-2, 1, 0.5)^T$$

1. This year, the total rainfall is 11, the normalized average temperature is 15, and the normalized current market price is -10. Does logistic regression predict that the harvest will or will not be profitable?

2. In another year, the total rainfall is 11, the normalized average temperature is 15, and the normalized current market price is 24. Does logistic regression predict that the harvest will or will not be profitable?

6 Game Theory

Problem 1: True or False: Every game with a Nash Equilibrium has a dominant strategy.

Problem 2: This following is the matrix representation for the classic rock-paper-scissor game between a row player and a column player.

	Rock	Paper	Scissor
Rock	(0,0)	(-1,1)	(1,-1)
Paper	(1, -1)	(0,0)	(-1,1)
Scissor	(-1,1)	(1, -1)	(0,0)

What is the Nash Equilibrium of this zero-sum game? Briefly explain why.

Problem 3: This following is the matrix representation for a zero-sum cop-robber game played between a cop (row player) and an robber (column player). The row player wants to catch the column player while the column player wants to avoid being caught.

	Location 1	Location 2
Location 1	(3,-3)	(-2, 2)
Location 2	(-1, 1)	(2, -2)

What is the Nash Equilibrium of this zero-sum game? Briefly explain why.