Table 1: For instructor's use

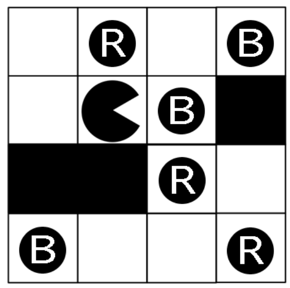
Question	Points Scored	Possible Points
1		12
2		8
3		8
4		6
5		12
6		8
7		8

This exam is closed book. You are allowed 2 sheets of notes (4 pages front and back). You may use any format for your notes that you like. Please explain all of your answers fully to receive full credit.

Here is some extra space. **Show all of your work on the questions!** If you need more paper just ask. Good luck!!

Question 1. **Pacman Search**

You are going to design a search problem to help Pacman get the nourishment he needs. In this game, food pellets are either Red or Blue. Pacman needs *at least one* food pellet of each color to stay healthy. The game is over when he has eaten one of each type of pellet (but he can eat more than one). Pacman has four actions: moving up, down, left, or right, and he does not have a "stay" action. There are K red pellets and K blue pellets, and the dimensions of the board are N by M. Answer the following questions:



$K = 3$, $N = 4$, $M = 4$

(a) (2 pts) Design the space space for this search problem by defining the *minimum* number of variables that are needed. For each variable you define, state its domain (i.e. the range of values it takes on).

We need two variables to describe the location of Pacman, one boolean variable showing whether Pacmac already ate a red pellet, and another boolean variable for the blue pellets. Note that we do not need to track the number of pellets of each color that are eaten. Formally: $(x \in [1, \ldots, N], y \in [1, \ldots, M], eaten_R \in \{T, F\}, eaten_B \in \{T, F\})$

(b) (2 pts) What is the maximum size of the state space (i.e. how many different state assignments are possible) for an arbitrary problem (i.e. not the specific maze shown above)?

There are at most $N \times M$ possible locations for pacman and 4 possible assignments to the boolean variables so the size of the state space is at most $4 \times N \times M$.

(c) (1 pts) What is the maximum branching factor of the search problem?

Each state has at most four successors corresponding to the four possible actions. The branching factor is at most 4.

(d) (1 pts) Assuming Pacman starts the game in position (x, y), what is the initial state?

(x, y, F, F). The two boolean state variables are both false.

(e) (2 pts) Specify a goal test for the search problem.

$(eaten_R == T) \&\& (eaten_B == T)$

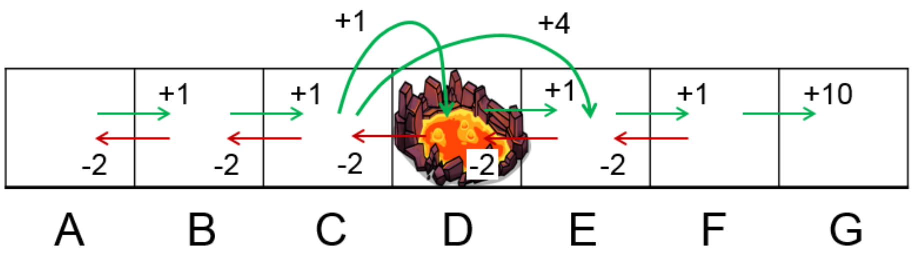
(f) (4 pts) For each of the following search heuristics, indicate (yes/no) whether or not it is *admissible*.

Heuristic	Admissible?
The number of pellets remaining	1. No
The smallest Manhattan distance to any remaining pellet	2. Yes
The maximum Manhattan distance between any two remaining pellets	3. No
The minimum Manhattan distance between any two remaining pellets of opposite colors	4. No

1. Inadmissible because Pacman only cares about eating one pellet of each color. Adding extra pellets to the problem does not increase the optimal cost but it does increase this heuristic.
2. Admissible since Pacman needs to eat at least one extra pellet to reach the goal from any non-goal state and the cost of moving to that extra pellet is greater than of equal to the heuristic. Needs to be defined to equal zero at goal states.
3. Inadmissible. Adding extra pellets to the problem does not increase the optimal cost but it does increase this heuristic.
4. Inadmissible for the states where Pacman has already eaten one pellet.

Question 2. American Ninja MDP

You need to solve the following MDP to prepare for your upcoming appearance on the hit TV show *American Ninja Warrior*. Your objective is to start at state A, race down the track, jump the lava pit in state D, and reach the finish line (terminal state G). Actions are *Right*, *Left*, and *Jump*, but *Jump* can only be used in state C (and *Right* cannot be used there). Rewards for each state transition are shown in the figure. The discount is $\gamma = 1$.



All actions are *deterministic* except for *Jump*, which succeeds half the time, landing in state E, and fails half the time, landing in D. In summary, the action model is:

Right: Deterministically move to the right.
Left: Deterministically move to the left.
Jump: Stochastically jump to the right. This action is available for square C only.
$T(C, \textit{Jump}, E) = 0.5$ (jump succeeds)
$T(C, \textit{Jump}, D) = 0.5$ (jump fails)

(a) (2 pts) For the policy π of always moving forward (i.e., using actions *Right* or *Jump*), compute $V^\pi(C)$.

Since there is no discounting, the value is the sum of the rewards. If the jump fails, reward to go from C is $1 + 1 + 1 + 10 = 13$. If the jump succeeds it is $4 + 1 + 10 = 15$. Since both are equally likely, $V^\pi(C) = (13 + 15)/2 = 14$.

(b) (2 pts) Perform two iterations of value iteration and fill in the table below. All values are initialized to zero.

$V^2(B)$	3.5
$Q^2(B, \textit{Right})$	3.5
$Q^2(B, \textit{Left})$	-1

SOLUTION:

The update rule for reward on transition is:

$$V^{i+1}(s) = \max_a \sum_{s'} P(s'|s,a)\left[R(s,a,s') + \gamma V^i(s')\right]$$

For state C this gives:

$$V^{i+1}(C) = \max_{J,L} \begin{cases} J : 0.5(1 + V^i(D)) + 0.5(4 + V^i(E)) \\ L : -2 + V^i(B) \end{cases}$$

For the other states, each possible action gives its reward plus the value of the successor state. The following table shows the values for each iteration.

	A	B	C	D	E	F	G
V^1	1	1	2.5	1	1	10	0
V^2	2	3.5	3.5	2	11	10	0

The expression for the Q values is:

$$Q(s,a) = \sum_{s'} P(s'|s,a)\left[R(s,a,s') + \gamma V(s')\right],$$

which implies that $V^i(s) = \max_a Q^i(s,a)$. For state B:

$$V^2(B) = \max_{R,L} \begin{cases} R : 1 + 2.5 = Q^2(B,R) = 3.5 \\ L : -2 + 1 = Q^2(B,L) = -1 \end{cases} = 3.5.$$

(c) (4 pts) You decide to use Q-Learning to obtain the optimal policy. After some number of iterations of Q-Learning, the Q table has the values given in the row "Initial" below. Apply the Q-Learning update rule to update the Q-values according to the four transitions in the episode given below. Q-values that are unaffected by a particular transition can be left blank. Use a learning rate α of 0.5. *Be sure to use the initial Q-values provided in the top row.*

Episode Data for Q-Learning

s	a	r	s	a	r	s	a	r	s	a	r	s
C	*Jump*	$+4$	E	*Right*	$+1$	F	*Left*	-2	E	*Right*	$+1$	F

	Q(C, *Left*)	Q(C, *Jump*)	Q(E, *Left*)	Q(E, *Right*)	Q(F, *Left*)	Q(F, *Right*)
Initial	-1	1	0	2	0	-2
Transition 1		3.5				
Transition 2				1.5		
Transition 3					-0.25	
Transition 4				1.125		

SOLUTION:

The Q-learning update rule for transition i is

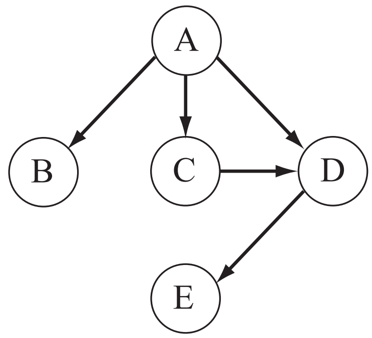
$$Q^i(s,a) = (1-\alpha)Q^{i-1}(s,a) + \alpha\left[R(s,a,s') + \max_{a'} \gamma Q^{i-1}(s',a')\right].$$

Details of the calculations:

$$\begin{aligned} 3.5 &= 0.5(1) + 0.5(4+2) \\ 1.5 &= 0.5(2) + 0.5(1+0) \\ -0.25 &= 0 + 0.5(-2+1.5) \\ 1.125 &= 0.5(1.5) + 0.5(1-0.25) \end{aligned}$$

A state-action is only updated when a transition is made from it. $Q(C, \textit{Left})$, $Q(E, \textit{Left})$, and $Q(F, \textit{Right})$ state-actions are never experienced and so these values are never updated.

Question 3. Bayesian Network Structure
Consider the following Bayesian network model.



(a) (2 pts) Identify which of the following are topological orders for this Bayes net by *circling the topological orders.*

A, B, D, C, E $\underline{A, C, D, B, E}$ $\underline{A, B, C, D, E}$ A, D, E, B, C

(b) (2 pts) Write the factorized form of $P(A, B, C, D, E)$ using one of the topological orders above from part (b).

$$P(A, B, C, D, E) = P(A)P(B|A)P(C|A)P(D|A, C)P(E|D)$$

(c) (2 pts) Suppose that A, B, and C are binary random variables, and variables D and E have three possible values. How many free parameters does the joint distribution $P(A, B, C, D, E)$ have? In other words, how many unique numbers need to be specified in order to completely describe the joint density?

In the order of the factorization above, the sum of parameters per table is:

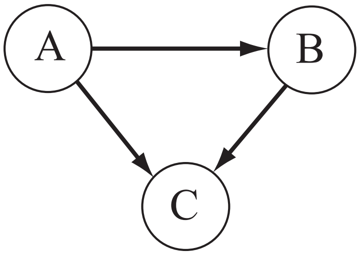
$params = 1 + 2 + 2 + 8 + 6 = 19$

(d) (2 pts) Identify which of the following statements of independency are true for this Bayes net by *circling the true statements.*

$A \perp D \mid C$ $\underline{A \perp E \mid C, D}$ $\underline{D \perp B \mid A}$ $D \perp B \mid C$

Question 4. Inference in Bayes Nets

(a) (4 pts) Consider the following Bayes net model for binary variables A, B, C with the conditional probability tables provided. Note also that $P(A = 0) = 0.9$.



$P(B\|A)$	$B = 0$	$B = 1$
$A = 0$	0.3	0.7
$A = 1$	0.6	0.4

A	B	$P(C = 0\|A, B)$	$P(C = 1\|A, B)$
0	0	0.1	0.9
0	1	0.2	0.8
1	0	0.1	0.9
1	1	0.2	0.8

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

SOLUTION:

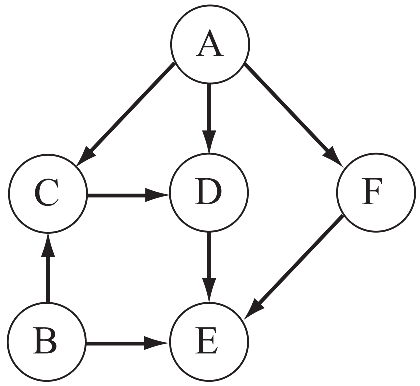
$$P(B, C|A = 1) = \frac{P(B, C, A = 1)}{P(A = 1)} = \frac{P(B|A = 1)P(C|B, A = 1)P(A = 1)}{P(A = 1)}$$

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

$$= \begin{array}{c|cc} & C = 0 & C = 1 \\ \hline B = 0 & (0.6)(0.1) & (0.6)(0.9) \\ B = 1 & (0.4)(0.2) & (0.4)(0.8) \end{array}$$

$$= \begin{array}{c|cc} & C = 0 & C = 1 \\ \hline B = 0 & 0.06 & 0.54 \\ B = 1 & 0.08 & 0.32 \end{array}$$

(b) (2 pts) Now consider the following Bayes net. Use *Variable Elimination* to eliminate the variable F followed by B. In each step, identify the sum that will be computed and the factors that will be generated.



SOLUTION:
Factors in model: $P(A), P(B), P(F|A), P(C|A,B), P(D|A,C), P(E|B,D,F)$

Eliminate F:

$$\phi(A,B,D,E,F) = P(F|A)P(E|B,D,F)$$
$$\psi(A,B,D,E) = \sum_F P(F|A)P(E|B,D,F)$$

Eliminate B:

$$\phi(A,B,C,D,E) = P(B)P(C|A,B)\psi(A,B,D,E)$$
$$\psi(A,C,D,E) = \sum_B P(B)P(C|A,B)\psi(A,B,D,E)$$

Question 5. **Decision Tree Learning**

Sarah organizes a local science fiction book club. She plans to build an agent that will decide whether to buy a book for the club. It will eventually search the internet. But first, she has to learn a decision tree classifier based on books the club has already read. The classifier will learn to identify good books by their attributes:

- Space - the story is not set on Earth
- Time - the story involves someone traveling through time
- Heroine - the protagonist is a female

	Space	Time	Heroine	Good?
1	1	1	0	0
2	1	1	1	1
3	1	0	1	1
4	0	1	1	1
5	1	1	0	0
6	0	0	0	0
7	0	0	1	0
8	0	1	1	1

When determining the root for her decision tree, Susan calculates that Gain(Heroine)=0.54, Gain(Space)=0.00, and Gain(Time)=0.05. She picks **heroine** to be the root.

The questions start on the next page, break any ties between attributes alphabetically.

You may want to use the following table of logarithms:

$$\log_2(1/3) = -1.585$$
$$\log_2(2/3) = -0.585$$
$$\log_2(1/5) = -2.322$$
$$\log_2(2/5) = -1.322$$
$$\log_2(3/5) = -0.737$$
$$\log_2(4/5) = -0.322$$
$$\log_2(3/8) = -1.415$$
$$\log_2(5/8) = -0.678$$

(a) (6 pts) Consider the **TRUE** branch of the heroine node.

Is another attribute necessary for this branch?

$$H(G|H = 1) = B(4/5) = -(0.8)(-0.322) - (0.2)(-2.322) = 0.722,$$

therefore we still have room to improve the entropy, so YES we should split!

If so, calculate Gain(Space) and Gain(Time). Which attribute should go next? (Show your calculations to be eligible for partial credit.)

SOLUTION:

Space: $H(G|S) = 2/5B(1) + 3/5B(2/3) = 0.551$
$B(2/3) = -(0.666)(-0.585) - (0.333)(-1.585) = 0.918$
$Gain(Space) = 0.722 - 0.551 = 0.171$
Time: $H(G|T) = 3/5B(1) + 2/5B(1/2) = 0.4$ **(best)**
$Gain(Time) = 0.722 - 0.4 = 0.322$

Conclusion: Split on Time

(b) (6 pts) Consider the **FALSE** branch of the heroine node.

Is another attribute necessary for this branch?

$$H(G|H=0) = B(0) = 1,$$

therefore we will not have the ability improve this branch. Don't split.

If so, calculate Gain(Space) and Gain(Time). Which attribute should go next? (Show your calculations to be eligible for partial credit.)

Question 6. Dynamic Dean

The AI faculty decide to build an agent to guess the mood of Dean Zvi Galil from hour to hour (so visits to the Dean's office can be carefully timed). After detailed observation, they determine that when the Dean is in a good mood, there is a 75% chance he will be in a good mood an hour later. But when he is in a bad mood, there is a 50% chance he will be in a bad mood an hour later. Unfortunately the Dean's mood can't be observed directly. However, through careful study the faculty have determined that when the Dean is in a good mode, the likelihood that he will send a reminder email about CIOS is 80%, but there is only a 60% chance of sending a CIOS email when he is in a bad mood.

(a) (1 pts) Draw a dynamic Bayesian network model for this problem. Be sure to show at least two time slices. Label the nodes in your graph. Use the random variable M to indicate the Dean's mood. M can take on the values "good" or "bad." Use the random variable C to indicate whether a CIOS email was sent. C can take on the values "true" or "false." Use subscripts on M and C to indicate time slices.

The graph should have the edges $M_{t-1} \rightarrow M_t$, $M_{t-1} \rightarrow C_{t-1}$, and $M_t \rightarrow C_t$.

(b) (2 pts) Fill in the conditional probability tables for the transition model and sensor model elements of the DBN. Be sure the columns and rows are labeled properly in the spaces provided.

Figure 1: *

Transition Model:

$P(M_t\|M_{t-1})$	$M_t =$ bad	$M_t =$ good
$M_{t-1} =$ bad	0.5	0.5
$M_{t-1} =$ good	0.25	0.75

Figure 2: *

Measurement Model:

$P(C_t\|M_t)$	$C_t = F$	$C_t = T$
$M_t =$ bad	0.4	0.6
$M_t =$ good	0.2	0.8

(c) (2 pts) The faculty are interested in tracking the Dean's mood carefully during the Final Exam period. The day before Finals start, the faculty have no knowledge of the Dean's mood. In other words, $P(M_0 = \text{good}) = 0.5$. Predict (through forward simulation) the probability of the Dean's mood on the first day of Finals:

$$P(M_1 = \text{good}) = \sum_{M_0} P(M_1 = \text{good}|M_0)P(M_0) = (0.5)(0.5) + (0.75)(0.5) = 0.625$$

Note that this is an example of belief updating with belief function $B(M_0) = P(M_0)$ updated due to passage of time to yield the prediction belief $B'(M_1) = P(M_1)$. The general form of the updated belief for any time t would be $B'(M_t) = P(M_t, C_{1:t-1})$. In this case, evidence has not been introduced yet.

(d) (3 pts) Now suppose that the Dean sends a CIOS reminder email on the first day of Finals, in other words $C_1 = T$. Compute an updated posterior estimate for the Dean's mood M_1.

SOLUTION: The problem is asking for $P(M_1|C_1)$, which we can obtain by updating the prediction belief function $B'(M_1)$ from part (c) with evidence, yielding $B(M_1) = P(M_1, C_1)$, which can then be normalized to obtain $P(M_1|C_1)$. We have:

$$P(M_1|C_1) = \frac{B(M_1)}{P(C_1)} = \frac{P(C_1|M_1)B'(M_1)}{P(C_1)}$$

$$= \frac{1}{P(C_1)} \times \begin{array}{l|c} & P(M_1, C_1) \\ \hline M_1 = \text{bad} & (0.6)(0.375) = 0.225 \\ M_1 = \text{good} & (0.8)(0.625) = 0.5 \end{array}$$

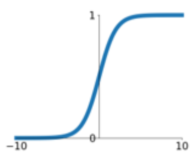
$$= \begin{array}{l|c} & P(M_1|C_1) \\ \hline M_1 = \text{bad} & 0.225/0.725 = 0.31 \\ M_1 = \text{good} & 0.5/0.725 = 0.69 \end{array}$$

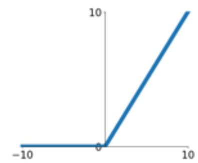
Note that we were able to obtain the probability of evidence by summing the (unnormalized) beliefs. Dividing by this number normalizes the beliefs, resulting in a probability density.

Question 7. Neural Networks

Provide a short answer to each of the questions below.

- **(2 pts)** Name two standard nonlinear functions used in neural network hidden units, and sketch the shape of each function. Be sure to label the axes and identify any relevant values.

  Sigmoid: 

  RELU: 

- **(2 pts)** Name two standard loss functions used in neural network models. For each loss function, give an example of prediction problem for which it would be an appropriate choice.

  Squared Error loss can be used in regression

  Softmax loss can be used in classification

- **(2 pts)** Suppose the input to a convolution layer in a ConvNet is 32x32x3 image. If we apply a set of 10 filters, each with a 7x7x3 receptive field, what are the dimensions of the output activation map? How many parameters need to be learned?

  $Size = (N - F)/stride + 1 = 32 - 7 + 1 = 26 \Rightarrow$ Output map is 26x26x1

  Number of parameters is $10 \times (7 \times 7 \times 3 + 1) = 1,480$

- **(2 pts)** Suppose the input to a 1x1 convolution layer is a feature map that is 56x56 with a depth of 32, and the output map is 56x56x1. How many multiplications and how many additions must be performed in order to compute the output map?

  Convolution kernel is 1x32. At each pixel location, 32 multiplies.
  $Multiplications = 56 \times 56 \times 32 = 100,352$

  31 additions + 1 for bias term = 32 additions/site. $Additions = 100,352$