Exam 1 Review

CS440/ECE448, Spring 2021

Exam date: Friday, March 5, 1:00pm

Question 1

Imagine a maze with only four possible positions, numbered 1 through 4 in the following diagram. Position 2 is the start position (denoted S in the diagram below), while positions 1, 3, and 4 each contain a goal (denoted as G_1, G_2, and G_3 in the diagram below). Search terminates when the agent finds a path that reaches all three goals, using the smallest possible number of steps.

1 G_1	2 S	3 G_2
	4 G_3	

(a) Define a notation for the state of this agent. How many distinct non-terminal states are there?

(b) Draw a search tree out to a depth of 3 moves, including repeated states. Circle repeated states.

(c) For A* search, one possible heuristic, h_1, is the Manhattan distance from the agent to the nearest goal that has not yet been reached. Prove that h_1 is consistent.

(d) Another possible heuristic is based on the Manhattan distance $M[n, g]$ between two positions, and is given by

$$h_2[n] = M[G_1, G_2] + M[G_2, G_3] + M[G_3, G_1]$$

that is, h_2 is the sum of the Manhattan distances from goal 1 to goal 2, then to goal 3, then back to goal 1. Prove that h_2 is not admissible.

(e) Prove that $h_2[n]$ is dominant to $h_1[n]$.

Question 2

For each type of maze described below, specify the time complexity and space complexity of both breadth-first-search (BFS) and depth-first-search (DFS).

(a) The Albuquerque maze has b possible directions that you can take at each intersection. No path is longer than m steps, where m is finite. There is only one solution, which is known to require exactly d steps, where $d < m$.

(b) The Belmont maze has b possible directions that you can take at each intersection. No path is longer than m steps, where m is finite. All solutions require $d = m$ steps. About half of all available paths are considered solutions to the maze.

(c) The Crazytown maze has b possible directions that you can take at each intersection. The maze is infinite in size, so some paths have infinite length. There is only one solution, which is known to require $d = 25$ steps.

Question 3

Consider the data points in Table 1, representing a set of seven patients with up to three

Sore Throat	Stomachache	Fever	Flu
No	No	No	No
No	No	Yes	Yes
No	Yes	No	No
Yes	No	No	No
Yes	No	Yes	Yes
Yes	Yes	No	Yes
Yes	Yes	Yes	No

Table 1: Symptoms of seven patients, three of whom had the flu.

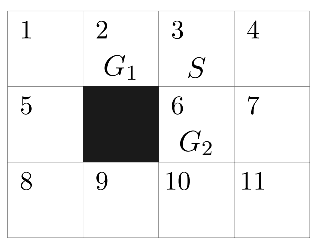
different symptoms. We want to use the Naïve Bayes assumption to diagnose whether a person has the flu based on the symptoms.

(a) Calculate the maximum likelihood conditional probability tables.

(b) If a person has stomachache and fever, but no sore throat, what is the probability of him or her having the flu (according to the conditional probability tables you calculated in part (b))?

Question 4

Consider the following maze. There are 11 possible positions, numbered 1 through 11. The agent starts in the position marked S (position number 3). From any position, there are from one to four possible moves, depending on position: Left, Right, Up, and/or Down. The agent's goal is to find the shortest path that will touch both of the goals (G_1 and G_2).



(a) Define a notation for the state of this agent. How many distinct non-terminal states are there?

(b) Draw a search tree out to a depth of 2 moves, including repeated states. Circle repeated states.

(c) For A* search, one possible heuristic, h_1, is the number of goals not yet reached. Prove that h_1 is consistent.

(d) Another possible heuristic is based on the Manhattan distance $M[n, g]$ between two positions, and is given by

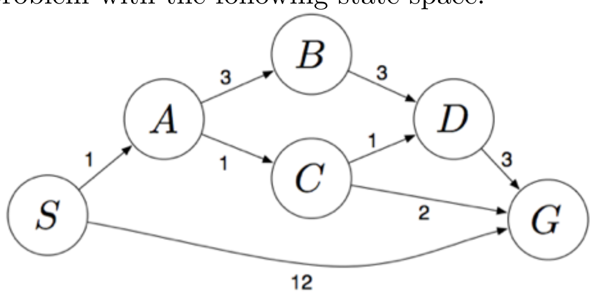
$$h_2[n] = M[n, G_1] + M[G_1, G_2]$$

that is, h_2 is the sum of the Manhattan distance from the current position to G_1, plus the Manhattan distance from G_1 to G_2. Prove that h_2 is not admissible.

(e) Prove that $h_2[n]$ is dominant to $h_1[n]$.

Question 5

Consider the search problem with the following state space:



S denotes the start state, G denotes the goal state, and step costs are written next to each arc. Assume that ties are broken alphabetically (i.e., if there are two states with equal priority on the frontier, the state that comes first alphabetically should be visited first).

(a) What path would BFS return for this problem?

(b) What path would DFS return for this problem?

(c) What path would UCS return for this problem?

(d) Consider the heuristics for this problem shown in the table below.

State	h_1	h_2
S	5	4
A	3	2
B	6	6
C	2	1
D	3	3
G	0	0

i. Is h1 admissible? Is it consistent?

ii. Is h2 admissible? Is it consistent?

Question 6

You're creating sentiment analysis. You have a training corpus with four movie reviews:

Review #	Sentiment	Review
1	+	what a great movie
2	+	I love this film
3	-	what a horrible movie
4	-	I hate this film

Let $Y = 1$ for positive sentiment, $Y = 0$ for negative sentiment.

(a) What's the maximum likelihood estimate of $P(Y = 1)$?

(b) Find maximum likelihood estimates $P(W|Y = 1)$ and $P(W|Y = 0)$ for the ten words $W \in \{\text{what,a,movie,I,this,film,great,love,horrible,hate}\}$.

(c) Use Laplace smoothing, with a smoothing parameter of $k = 1$, to estimate $P(W|Y = 1)$ and $P(W|Y = 0)$ for the ten words $W \in \{\text{what,a,movie,I,this,film,great,love,horrible,hate}\}$.

(d) Using some other method (unknown to you), your professor has estimated the following conditional probability table:

Y	$P(\text{great}\|Y)$	$P(\text{love}\|Y)$	$P(\text{horrible}\|Y)$	$P(\text{hate}\|Y)$
1	0.01	0.01	0.005	0.005
0	0.005	0.005	0.01	0.01

and $P(Y = 1) = 0.5$. All other words (except great, love, horrible, and hate) can be considered out-of-vocabulary, and you can assume that $P(W|Y) = 1$ for all out-of-vocabulary words. Under these assumptions, what is the probability $P(Y = 1|R)$ that the following 14-word review is a positive review?

$R = \{$"I'm horrible fond of this movie, and I hate anyone who insults it."$\}$

Question 7

Consider a Nave Bayes classifier with 100 feature dimensions. The label Y is binary with $P(Y = 0) = P(Y = 1) = 0.5$. All features are binary, and have the same conditional probabilities: $P(X_i = 1|Y = 0) = a$ and $P(X_i = 1|Y = 1) = b$ for $i = 1, \ldots, 100$. Given an item X with alternating feature values $(X_1 = 1, X_2 = 0, X_3 = 1, ..., X_{100} = 0)$, compute $P(Y = 1|X)$.

Question 8

Use the axioms of probability to prove that $P(\neg A) = 1 - P(A)$.

Question 9

Discuss the relative strengths and weaknesses of breadth-first search vs. depth-first search for AI problems.

Question 10

A friend who works in a big city owns two cars, one small and one large. Three-quarters of the time he drives the small car to work, and one-quarter of the time he drives the large car. If he takes the small car, he usually has little trouble parking, and so is at work on time with probability 0.9. If he takes the large car, he is at work on time with probability 0.6. Given that he was on time on a particular morning, what is the probability that he drove the small car?

Question 11

You're trying to determine whether a particular newspaper article is of class $Y = 0$ or $Y = 1$. The prior probability of class $Y = 1$ is $P(Y = 1) = 0.4$. The newspaper is written in a language that only has four words, so that the i^{th} word in the article must be $W_i \in \{0, 1, 2, 3\}$, with probabilities given by:

$$\begin{array}{ll} P(W_i = 0|Y = 0) = 0.3 & P(W_i = 0|Y = 1) = 0.1 \\ P(W_i = 1|Y = 0) = 0.1 & P(W_i = 1|Y = 1) = 0.1 \\ P(W_i = 2|Y = 0) = 0.1 & P(W_i = 2|Y = 1) = 0.3 \end{array}$$

The article is only three words long; it contains the words

$$A = (W_1 = 3, W_2 = 2, W_3 = 0)$$

What is $P(Y = 1, A)$?

Question 12

20% of students at U of I are part of the Greek system. Amongst these students, 10% study engineering. Furthermore, 15% of the entire student body studies engineering. Given that we know that a student studies engineering, what is the probability that the student is not part of the Greek system?

Question 13

Consider the following joint probability distribution:

$$
\begin{aligned}
P(A, B) &= 0.12 \\
P(A, \neg B) &= 0.18 \\
P(\neg A, B) &= 0.28 \\
P(\neg A, \neg B) &= 0.42
\end{aligned}
$$

What are the marginal distributions of A and B? Are A and B independent and why?