UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

CS440/ECE448 Artificial Intelligence

Conflict Exam 2

Spring 2022

Week of April 4, 2022

Your Name: ______________________________

Your NetID: ______________________________

Instructions

- Please write your NetID on the top of every page.
- This is a CLOSED BOOK exam. You will be permitted to bring one 8.5x11 page of handwritten notes (front & back).
- Calculators are not permitted. You need not simplify explicit numerical expressions.

Possibly Useful Formulas

Search:

Admissible Heuristic: $h(n) \leq d(n)$

Consistent Heuristic: $h(m) - h(n) \leq d(m) - d(n)$ if $d(m) - d(n) \geq 0$

Belief Propagation:

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

Expectation Maximization:

$$P(B=b|A=a) \leftarrow \frac{E\left[\#\text{ times } A=a, B=b\right]}{E\left[\#\text{ times } A=a\right]}$$
$$E\left[\#\text{ times } A=a, B=b\right] = \sum_t P\left(A_t=a, B_t=b|\text{observations on day } t\right)$$

HMM:

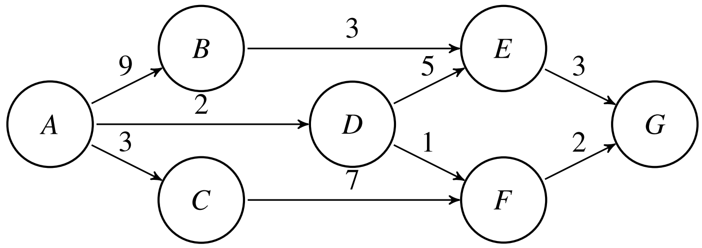
$$P(Y_1, X_1, \ldots, Y_T, X_T) = \prod_{t=1}^{T} P(Y_t|Y_{t-1})P(X_t|Y_t)$$

Viterbi Algorithm:

$$e_{i,j,t} = a_{j,i}b_{j,k}$$
$$v_{i,1} = \pi_i b_{i,x_1}$$
$$v_{j,t} = \max_i v_{i,t-1}e_{i,j,t}$$

Question 1 ***(7 points)***

Consider the following search graph. The starting state is A, the goal state is G, and the cost of each possible action is shown on the corresponding edge:



Suppose that the states shown have the following heuristics:

	A	B	C	D	E	F	G
$h(n)$:	4	3	1	1	0	2	0

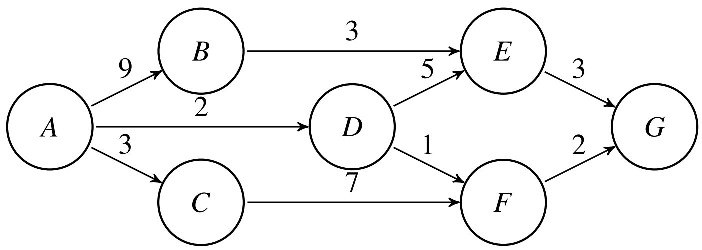
A* search (with repetitions avoided using an explored set) is applied to this graph to find the shortest path.

- What states are expanded, and
- what is the shortest path?

Solution: Expanded states: A,D,C,F,G
Shortest path: A,D,F,G

Question 2 ***(7 points)***

Consider the following search graph. The starting state is A, the goal state is G, and the cost of each possible action is shown on the corresponding edge:



You are considering trying to implement A* search over this graph. You are trying to decide whether to use $h_1(n)$ or $h_2(n)$ as the heuristic, where $h_1(n)$ and $h_2(n)$ are as given in the following table:

	A	B	C	D	E	F	G
$h_1(n)$:	4	5	8	2	2	1	0
$h_2(n)$:	4	5	8	3	2	1	0

(a) Suppose you are using an explored set to prevent repeated states, i.e., you will only expand a state if you have never expanded it before. Which of these two heuristics is better to use, and why?

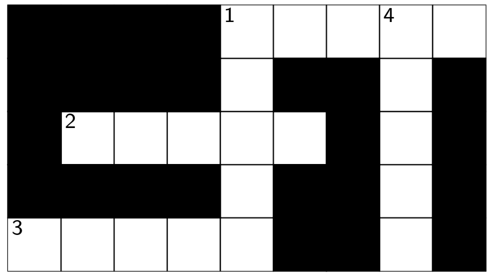
Solution: Use $h_1[n]$, because it is consistent.

(b) Your friend convinces you to use an explored dict, i.e., you will only expand a state if you have never expanded it before with a smaller $g(n)$. Which of these two heuristics is better to use, and why?

Solution: Use $h_2(n)$, because it dominates $h_1(n)$.

Question 3 ***(7 points)***

Consider the following crossword puzzle:



Imagine solving this puzzle as a constraint satisfaction problem. There are five variables:

- "1 across" is the five boxes horizontally, starting at the number 1.
- "1 down" is the five boxes vertically, starting at the number 1.
- "2 across" is the five boxes horizontally, starting at the number 2.
- "3 across" is the five boxes horizontally, starting at the number 3.
- "4 down" is the five boxes vertically, starting at the number 4.

Your goal is to fill each of those five variables with exactly one of the following five "words:" "EDBDC," "AACCC," "DDAAA," "CEEEA," or "ACAAC."

(a) Suppose that none of the variables have been filled yet. Which variable should you try to fill first, and why?

> **Solution:** 1 down should be filled first. Explanation: either say "because it constrains the largest number of other variables" or say "MCV."

(b) Suppose that you have decided to try filling the variable "3 across" first, and you need to decide whether to try the value "EDBDC" or the value "DDAAA." Which value would be better to try, and why?

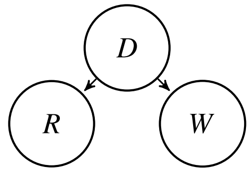
> **Solution:** Try the value "EDBDC." Either say: "EDBDC" leaves two options for "1 down," while "DDAAA" leaves only one option for that variable, or just say: LCV.

(c) Suppose that variable "4 down" has been filled with the word "CEEEA." Which variable should you try next, and why?

> **Solution:** 1 across should be filled next, because there is only one option that can go there (LRV).

Question 4 ***(7 points)***

Consider the following Bayesian network. All variables are binary:



Suppose that the model parameters are as follows:

$$P(D=T)=0.3$$
$$P(R=T|D)=\begin{cases}0.8 & D=F\\ 0.1 & D=T\end{cases}$$
$$P(W=T|D)=\begin{cases}0.1 & D=F\\ 0.6 & D=T\end{cases}$$

What is $P(D=T|W=T)$?

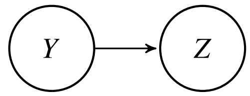
Solution:

$$P(D=T|W=T)=\frac{\sum_{r=1}^{2}P(D=T,W=T)}{\sum_{d=F}^{T}P(D=d,W=T)}$$
$$=\frac{(0.3)(0.6)}{(0.3)(0.6)+(0.7)(0.1)}$$

If you wish, you can simplify the last formula to $\frac{18}{25}=0.72$, but it's not required.

Question 5 ***(7 points)***

Consider the following Bayesian network, showing the relationship between two binary variables Y and Z:



Suppose that you've been given the following initial estimates of the model parameters, where a, b, and c are some arbitrary constants:

$$P(Y=T)=a, \quad P(Z=T|Y=F)=b, \quad P(Z=T|Y=T)=c$$

You are now trying to re-estimate the values of these model parameters. You have observed the values of the variables on five consecutive days, but on the fifth day, the value of Z was unobserved (labeled with a "?" in the table below:

	Day 1	Day 2	Day 3	Day 4	Day 5
Value of Y:	T	F	T	F	F
Value of Z:	F	F	F	T	?

For this table, in terms of the current model parameters a, b, and c, what is the expected number of days on which $Y=F$ and $Z=T$?

Solution:

$$\begin{aligned} E\left[\#\text{ days } Y=F, Z=T\right] &= 1+P(Z=T|Y=F) \\ &= 1+b \end{aligned}$$

Question 6 ***(7 points)***

Suppose that you have a hidden Markov model in which the state variables (Y_t) and observation variables (X_t) are binary. The initial-state probability is $P(Y_1 = T) = 0.6$, and the other model parameters are as follows:

Y_{t-1}	$P(Y_t = T \mid Y_{t-1})$
F	0.2
T	0.3

Y_t	$P(X_t = T \mid Y_t)$
F	0.2
T	0.9

What is $P(Y_1 = T \mid X_1 = T, X_2 = T)$?

Solution:

$P(Y_1 = T \mid X_1 = T, X_2 = T)$

$$= \frac{P(Y_1 = T, X_1 = T, Y_2 = T, X_2 = T) + P(Y_1 = T, X_1 = T, Y_2 = F, X_2 = T)}{P(Y_1 = T, X_1 = T, Y_2 = T, X_2 = T) + P(Y_1 = T, X_1 = T, Y_2 = F, X_2 = T) + P(Y_1 = F, X_1 = T, Y_2 = T, X_2 = T) + P(Y}$$

$$= \frac{(0.6)(0.9)(0.3)(0.9) + (0.6)(0.9)(0.7)(0.2)}{(0.6)(0.9)(0.3)(0.9) + (0.6)(0.9)(0.7)(0.2) + (0.4)(0.2)(0.2)(0.9) + (0.4)(0.2)(0.8)(0.2)}$$

Question 7 ***(7 points)***

Suppose that you have a hidden Markov model in which the state variables (Y_t) and observation variables (X_t) are binary. The initial-state probability is $P(Y_1 = T) = 0.6$, and the other model parameters are as follows:

Y_{t-1}	$P(Y_t = F \| Y_{t-1})$
F	0.0
T	1.0

Y_t	$P(X_t = T \| Y_t)$
F	0.9
T	0.3

Suppose that the observations are $X_1 = F, X_2 = T$. What is the most likely sequence of values Y_1, Y_2?

Solution: In this problem, the sequences (F,F) and (T,T) are impossible. The other two sequences have the probabilities:

$$P(Y_1 = T, X_1 = F, Y_2 = F, X_2 = T) = (0.6)(0.7)(1)(0.9)$$
$$P(Y_1 = F, X_1 = F, Y_2 = T, X_2 = T) = (0.4)(0.1)(1)(0.3)$$

The sequence (T,F) has higher probability.