CS-271P, Intro to A.I. — Final Exam — Winter Quarter, 2018

YOUR NAME: ____

YOUR ID: __________ ID TO RIGHT: __________ ROW: ________ SEAT: ________

Please turn off all cell phones now.

The exam will begin on the next page. Please, do not turn the page until told.

When told to begin, check first to ensure that your copy has all the pages, as numbered 1-14 in the bottom-right corner of each page. We will supply a new exam for any copy problems.

The exam is closed-notes, closed-book. No calculators, cell phones, electronics.

Clear your desk except for pen, pencil, eraser, & water bottle. Put backpacks under your seat. Please do not detach the provided scratch paper from the exam.

After you first stand up from your seat, your exam is over and must be turned in immediately. You may turn in your Final exam early and leave class when you are finished.

This page summarizes the points for each question, so you can plan your time.

1. (10 pts total, 2 pts each) PROBABILITY.
2. (10 pts total) DECISION TREES.
3. (12 pts total, 3 pts each) CROSS-VALIDATION.
4. (14 pts total, 2 pts each) Bayesian Networks.
~~**5. (8 pts total, 2 pts each) Probability.**~~
6. (8 pts total, 4 pts each) Bayesian Networks.
7. (4 pts total, 2 pts each) Naive Bayes Classifier.
8. (10 points total, 2 pts each) Constraint Satisfaction Problems
9. (10 pts total, -1 for each error, but not negative) Alpha-Beta Pruning.
10. (4 pts total, 1 pt each) Task Environment.
11. (10 pts total, 1/2 pt each) Search Properties.

Question 5 is canceled on the grounds that it was confusing to non-native English speakers who do not know what it means to flip a coin. Some thought that it meant simply to turn the coin over, in place. Everyone gets it right, regardless of your answer.

The Exam is printed on both sides to save trees! Work both sides of each page!

1. (10 pts total, 2 pts each) PROBABILITY. Using the joint distribution below for the Boolean variables X, Y, and Z, calculate the following probabilities. Your final answer should be a number in the interval [0,1]. **Show your work. Correct answer + no work = no credit. Correct answer + correct work = full credit.**

X	Y	Z	P(X, Y, Z)
t	t	t	0.16
t	t	f	0.09
t	f	t	0.03
t	f	f	0.02
f	t	t	0.15
f	t	f	0.25
f	f	t	0.20
f	f	f	0.10

1.a. (2 pts) P(Y=t) = P(X=t, Y=t, Z=t) + P(X=t, Y=t, Z=f) + P(X=f, Y=t, Z=t) + P(X=f, Y=t, Z=f)

= 0.16 + 0.09 + 0.15 + 0.25

= 0.65

1.b. (2 pts) P(NOT(X=t)) = 1 − P(X=t)

= 1 − {P(X=t, Y=t, Z=t) + P(X=t, Y=t, Z=f) + P(X=t, Y=f, Z=t) + P(X=t, Y=f, Z=f)}

= 1 - {0.16 + 0.9 + 0.03 + 0.02) = 0.7

Also OK:

P(NOT(X=t)) = P(X=f) = P(X=f, Y=t, Z=t) + P(X=f, Y=t, Z=f) + P(X=f, Y=f, Z=t) + P(X=f, Y=f, Z=f)
= 0.15 + 0.25 + 0.20 + 0.10 = 0.7

1.c. (2 pts) P(X=t, Z=t) = P(X=t AND Z=t) = P(X=t, Y=t, Z=t) + P(X=t, Y=f, Z=t)

= 0.16 + 0.03

= 0.19

1.d. (2 pts) P(Y=t | X=t) = P(X=t AND Y=t) / P(X=t)

= {P(X=t, Y=t, Z=t) + P(X=t, Y=t, Z=f)}
/ {P(X=t, Y=t, Z=t) + P(X=t, Y=t, Z=f) + P(X=t, Y=f, Z=t) + P(X=t, Y=f, Z=f)}

= (0.16 + 0.09) / (0.16 + 0.09 + 0.03 + 0.02) = 0.25 / 0.3 = 5/6 = 0.8333...

1.e. (2 pts) P(X=t, Z=f | Y=f) = P(X=t AND Y=f AND Z=f) / P(Y=f)

= P(X=t, Y=f, Z=f)
/ {P(X=t, Y=f, Z=t) + P(X=t, Y=f, Z=f) + P(X=f, Y=f, Z=t) + P(X=f, Y=f, Z=f)}

= (0.02) / (0.03 + 0.02 + 0.20 + 0.10) = 0.02 / 0.35 = 0.0517...

****** TURN PAGE OVER AND CONTINUE ON THE OTHER SIDE ******

2. (10 pts total) DECISION TREES. You are a robot that plans to use a decision tree to predict an outcome y using three features, x_1 ,x_2, and x_3. You observe eight training patterns, each of which we represent as [x_1,x_2,x_3] (so, "010" means x_1=0, x_2=1, x_3=0). We observe the labeled training data,

y=0: [001], [010], [010], [110]
y=1: [000], [011], [111], [111]

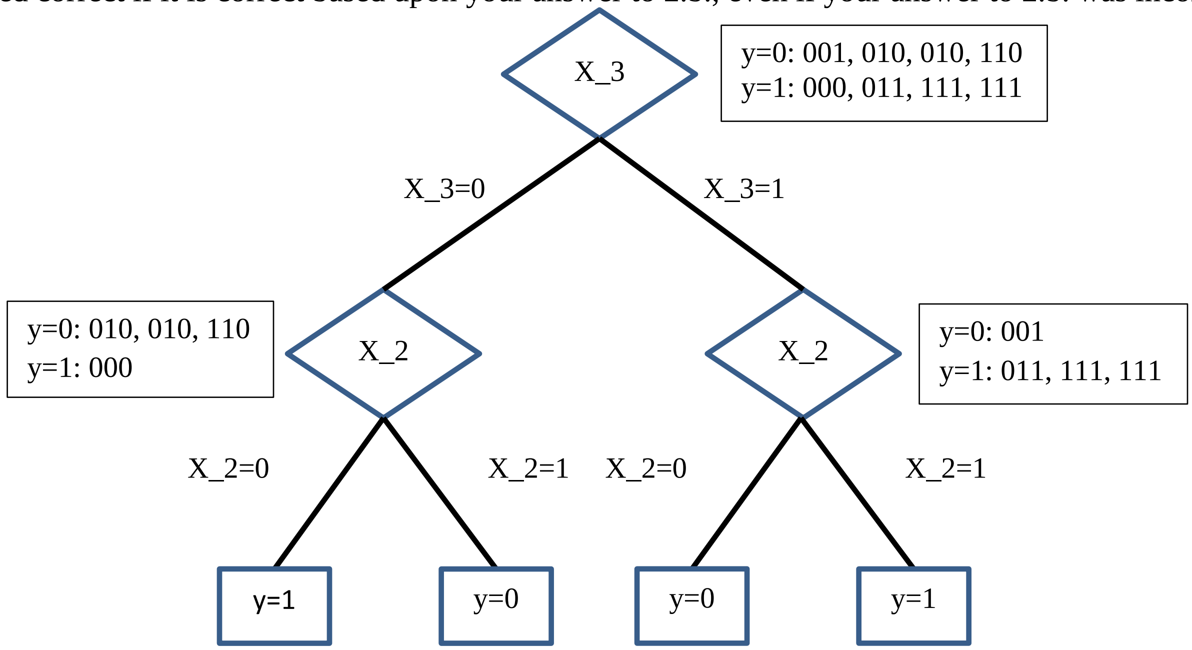
2.a. (3 pts) What is the initial entropy of y before any split? Write one of A, B, C, or D. _____A___________

(A) 1.0 **(B) log_2(3) = 1.59** **(C) 2.0** **(D) log_2(5) = 2.32**

2.b. (3 pts) Which variable would Information Gain choose as the root? Write one of x_1, x_2, or x_3.

____x_3___________________

2.c. (4 pts) Based on your answer to 2.b. above, draw the decision tree that results. Your answer here will be considered correct if it is correct based upon your answer to 2.b., even if your answer to 2.b. was incorrect.

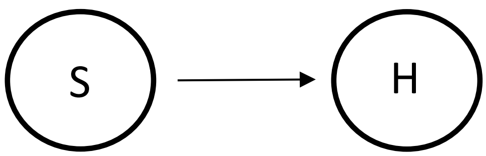


3. (12 pts total, 3 pts each) CROSS-VALIDATION. You have 25 examples, labeled 1 to 25. Each numeric value from 1 to 25 represents an example (= an instance, i.e., a feature vector) in your sample Please construct a k-fold cross validation using k = 5. The first one has been done out for you as an example.

6,7,8,9,10,11 ,12,13,14,15, 16,17,18,19, 20,21,22,23, 24,25	1,2,3,4,5,11, 12,13,14,15, 16,17,18,19, 20,21,22,23, 24,25	1,2,3,4,5,6,7, 8,9,10,16,17, 18,19,20,21, 22,23,24,25	1,2,3,4,5,6,7, 8,9,10,11,12, 13,14,15,21, 22,23,24,25	1,2,3,4,5,6,7, 8,9,10,11,12, 13,14,15, 16,17,18,19, 20
1,2,3,4,5	6,7,8,9,10	11,12,13,14, 15	16,17,18,19, 20	21,22,23,24, 25

4. (14 pts total, 2 pts each) Bayesian Networks. Consider the following Bayesian network, where S = an extreme solar flare occurred and H = your hard drive is damaged.

P(S) = 0.01



P(H | S) = 0.9

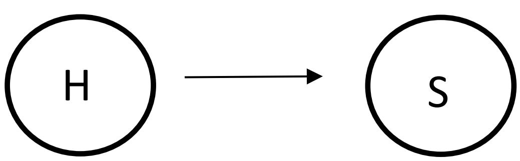
P(H | NOT(S)) = 0.01

4.a. (8 pts total, 2 pts each) Fill in the full joint distribution based upon the conditional probabilities above:

S	H	P(S, H)
T	T	0.009
T	F	0.001
F	T	0.0099
F	F	0.9801

4.b. (6 pts total, 2 pts each) Fill in the probabilities for the following Bayesian network. It is OK to give your answer as a common fraction, without computing the exact numeric value.

P(H) = 0.0189



P(S | H) = 0.009/0.0189

P(S | NOT(H)) = 0.001/0.9811

****** TURN PAGE OVER AND CONTINUE ON THE OTHER SIDE ******

5. (8 pts total, 2 pts each) P(Heads=t) = 1.0. One is a f one coin at random, and flip the fair coin was chosen.

Question 5 is canceled on the grounds that it was confusing to non-native English speakers who do not know what it means to flip a coin. Some thought that it meant simply to turn the coin over, in place. Everyone gets it right, regardless of your answer.

<u>Show your work. Correct answer + no work = no credit. Correct answer + correct work = full credit.</u>

5.a. (2 pts) What is $P(H1=t \wedge H2=t)$?

$P(H1=t \wedge H2=t) = P(H1=t \wedge H2=t \wedge F=t) + P(H1=t \wedge H2=t \wedge F=f)$

$= (0.5 * 0.5 * 0.5) + (1.0 * 1.0 * 0.5)$

$= 0.125 + 0.5 = 0.625$

5.b. (2 pts) What is $P(H1=t \wedge H2=t \mid F=t)$?

$P(H1=t \wedge H2=t \mid F=t) = P(H1=t \wedge H2=t \wedge F=t) / P(F=t)$

$= (0.5 * 0.5 * 0.5) / (0.5)$

$= 0.125 / 0.5 = 0.25$

5.c. (2 pts) What is $P(H1=t \wedge H2=t \mid F=f)$?

$P(H1=t \wedge H2=t \mid F=f) = P(H1=t \wedge H2=t \wedge F=f) / P(F=f)$

$= (1.0 * 1.0 * 0.5) / (0.5)$

$= 0.5 / 0.5 = 1.0$

5.d. (2 pts) What is $P(F=t \mid H1=t \wedge H2=t)$?

$P(F=t \mid H1=t \wedge H2=t) = P(H1=t \wedge H2=t \wedge F=t) / P(H1=t \wedge H2=t)$

$= P(H1=t \wedge H2=t \wedge F=t) / [\ P(H1=t \wedge H2=t\ \hat{}\ F=t) + P(H1=t \wedge H2=t\ \hat{}\ F=f)\ ]$

$= (0.5 * 0.5 * 0.5) / [\ (0.5 * 0.5 * 0.5) + (1.0 * 1.0 * 0.5)\ ]$

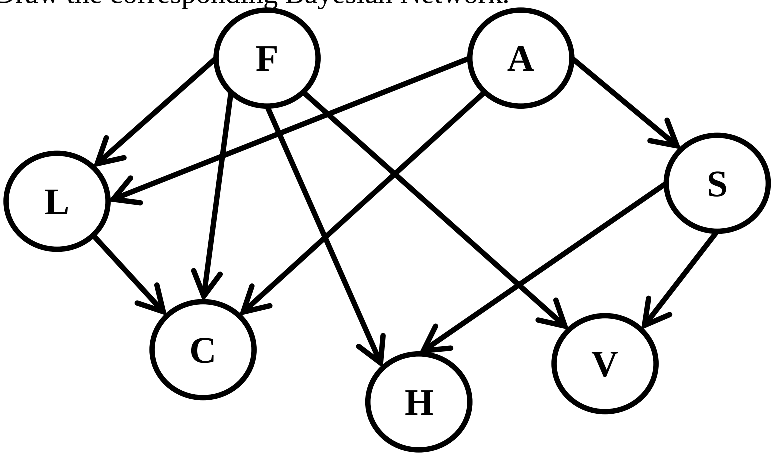
$= 0.125 / 0.625 = 0.2$

6. (8 pts total, 4 pts each) Bayesian Networks. Consider these diseases and symptoms:

S = (S)inus congestion   L = (L)ung infection   F = (F)lu   A = A)llergy
H = (H)eadache   V = Fe(v)er   C = (C)oughing.

You are told that (F)lu and (A)llergy cause (L)ung infection, and (A)llergy also causes Sinus congestion. (F)lu causes (H)eadache, Fe(v)er and (C)oughing. (S)inus congestion causes (H)eadache and Fe(v)er; (A)llergy causes (C)oughing. (L)ung infection causes (C)oughing.

6.a. (4 pts) Draw the corresponding Bayesian Network.



6.b. (4 pts) Write down the corresponding factored form of the joint probability distribution.

P(F,) = P(C | L, F, A) P(H | F, S) P(V | F, S) P(L | F, A) P(S | A) P(A) P(S)

7. (4 pts total, 2 pts each) Naive Bayes Classifier. We have a dataset of patients who got Flu before. Our goal is to have a Naive Bayes Classifier predict if a new patient has flu. The dataset has only values of (N)o and (Y)es.

	C = (C)ough	**H = (H)eadache**	**V = Fe(v)er**	**F = (F)lu**
#1	**N**	**N**	**N**	**N**
#2	**Y**	**N**	**N**	**Y**
#3	**N**	**Y**	**N**	**Y**
#4	**Y**	**Y**	**N**	**Y**
#5	**Y**	**N**	**Y**	**N**

You see a new patient having symptoms C=y, H=n, V=n. Assume that symptoms are conditionally independent given disease status. Use a Naive Bayes classifier to predict whether or not the patient as the flu. Write your answer as the product of simple fractions. You do not need to do the arithmetic to produce a numerical answer. You may use α as a normalizing constant.

7.a. (2 pts) P(F=y | C=y, H=n, V=n) = α P(F=y) P(C=y | F=y) P(H=n | F=y) P(V=n | F=y)

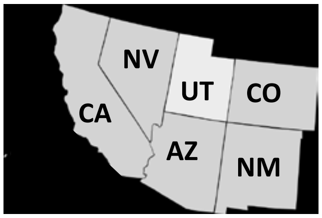
= α 3/5 * 2/3 * 1/3 * 3/3

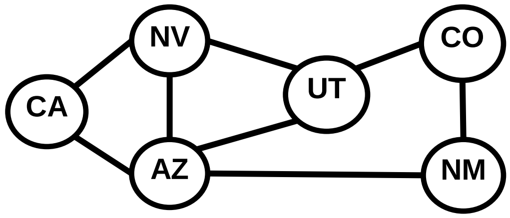
7.b. (2 pts) P(F=n | C=y, H=n, V=n) = α P(F=n) P(C=y | F=n) P(H=n | F=n) P(V=n | F=n)

= α 2/5 * 1/2 * 2/2 * 1/2

****** TURN PAGE OVER AND CONTINUE ON THE OTHER SIDE ******

8. (10 points total, 2 pts each) Constraint Satisfaction Problems. See Chapter 6.





You are a map-coloring robot assigned to color this Southwest USA map. Adjacent regions must be colored a different color (R=Red, B=Blue, G=Green). The constraint graph is shown.

8a. (2pts total, -1 each wrong answer, but not negative) FORWARD CHECKING. Cross out all values that would be eliminated by Forward Checking, after variable AZ has just been assigned value R as shown:

CA	NV	AZ	UT	CO	NM
~~R~~ G B	~~R~~ G B	R	~~R~~ G B	R G B	~~R~~ G B

8b. (2pts total, -1 each wrong answer, but not negative) ARC CONSISTENCY.
CA and AZ have been assigned values, but no constraint propagation has been done. Cross out all values that would be eliminated by Arc Consistency (AC-3 in your book).

CA	NV	AZ	UT	CO	NM
B	~~R~~ G ~~B~~	R	~~R~~ ~~G~~ B	R G ~~B~~	~~R~~ G B

8c. (2pts total, -1 each wrong answer, but not negative) MINIMUM-REMAINING-VALUES HEURISTIC. Consider the assignment below. NV is assigned and constraint propagation has been done. List all unassigned variables that might be selected by the Minimum-Remaining-Values (MRV) Heuristic: **CA, AZ, UT** .

CA	NV	AZ	UT	CO	NM
R B	G	R B	R B	R G B	R G B

8d. (2pts total, -1 each wrong answer, but not negative) DEGREE HEURISTIC. Consider the assignment below. (It is the same assignment as in problem 8c above.) NV is assigned and constraint propagation has been done. List all unassigned variables that might be selected by the Degree Heuristic: **AZ** .

CA	NV	AZ	UT	CO	NM
R B	G	R B	R B	R G B	R G B

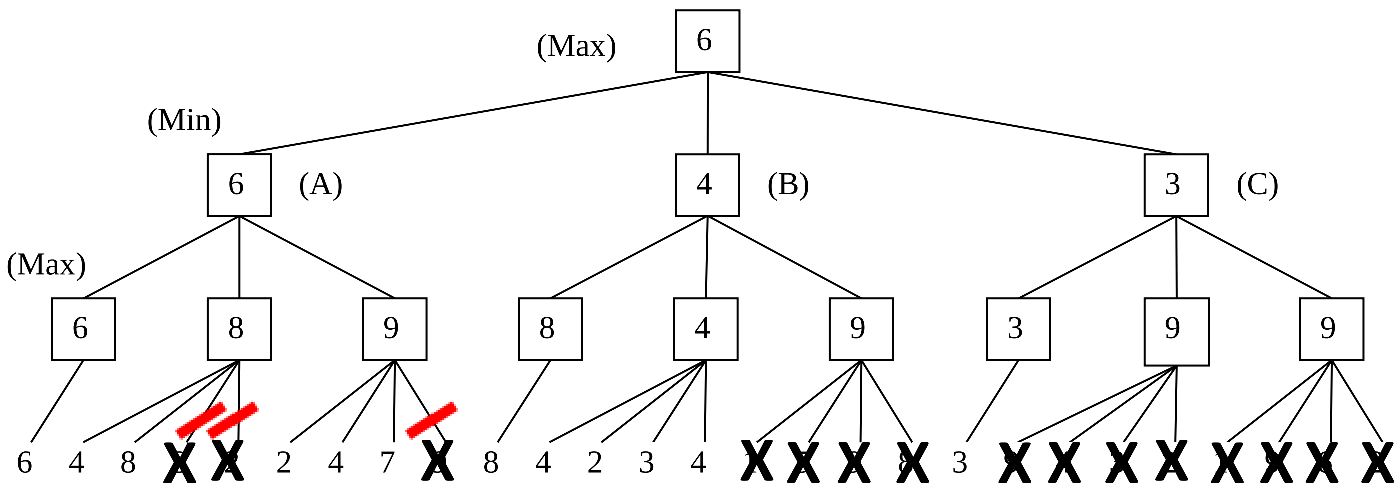
8e. (2pts total) MIN-CONFLICTS HEURISTIC. Consider the complete but inconsistent assignment below. AZ has just been selected to be assigned a new value during local search for a complete and consistent assignment. What new value would be chosen below for AZ by the Min-Conflicts Heuristic? **R** .

CA	NV	AZ	UT	CO	NM
B	G	**?**	G	G	B

9. (10 pts total, -1 for each error, but not negative) Alpha-Beta Pruning. In the game tree below it is **Max**'s turn to move. Inside each leaf node is the estimated score of that resulting position as returned by the heuristic static evaluator.

(1) Perform Mini-Max search and label each branch node with its value.

(2) Cross out each leaf node that would be pruned by alpha-beta pruning.



10. (4 pts total, 1 pt each) Task Environment. Your book defines a task environment as a set of four things, with acronym PEAS. Fill in the blanks with the names of the PEAS compo[nents] See Section 2.3.1

Performance (measure)   Environment   Actuators   Sensors

11. (10 pts total, 1/2 pt each) Search Properties. Fill in the values of the four evaluation criteria for each search strategy shown. Assume a tree search where b is the finite branching factor; d is the depth to the shallowest goal node; m is the maximum depth of the search tree; C* is the cost of the optimal solution; step costs are identical and equal to some positive ε; and in Bidirectional search both directi[ons] See Figure 3.21 first search.

Note that these conditions satisfy all of the footnotes of Fig. 3.21 in your book.

Criterion	Complete?	Time complexity	Space complexity	Optimal?
Breadth-First	Yes	O(b^d)	O(b^d)	Yes
Uniform-Cost	Yes	O(b^(1+floor(C*/ε))) O(b^(d+1)) also OK	O(b^(1+floor(C*/ε))) O(b^(d+1)) also OK	Yes
Depth-First	No	O(b^m)	O(bm)	No
Iterative Deepening	Yes	O(b^d)	O(bd)	Yes
Bidirectional (if applicable)	Yes	O(b^(d/2))	O(b^(d/2))	Yes

****** THIS IS THE END OF THE FINAL EXAM ******