For each question on Quiz #4, “Zero” below gives the fraction of students who scored zero, “Partial” gives the fraction who got partial credit, and “Perfect” gives the fraction who scored 100%.

Problem 1
Zero: 0.6% (~1 students), Partial: 18% (~30 students), Perfect: 82% (~139 students)

Problem 2
Zero: 1.8% (~3 students), Partial: 52% (~88 students), Perfect: 46% (~79 students)

Problem 3
Zero: 0% (~0 students), Partial: 11% (~19 students), Perfect: 89% (~151 students)

CS-171, Intro to A.I. — Quiz#4 — Fall Quarter, 2014 — 20 minutes

YOUR NAME: _____

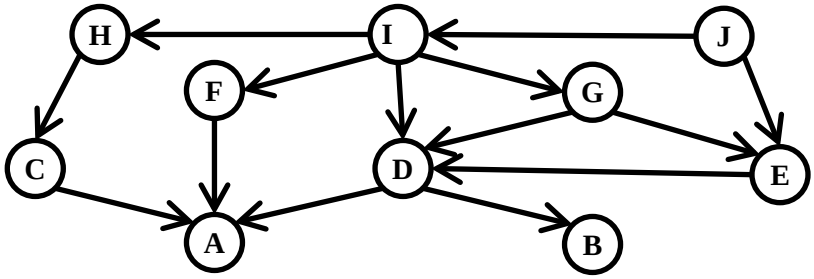
YOUR ID: _____ ID TO RIGHT:_____ ROW:_____ SEAT NO.: _____

1. (45 pts total, 15 pts each, -1 each error, but not negative) Bayesian Networks

See Section 14.1-2.

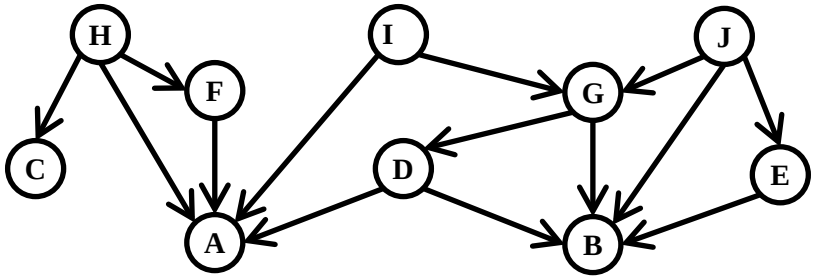
1a. (15 pts) Write down the factored conditional probability expression that corresponds to the graphical Bayesian Network shown.

P(A | C,D,F) P(B | D) P(C | H) P(D | E,G,I) P(E | G,J) P(F | I) P(G | I) P(H | I) P(I | J) P(J)

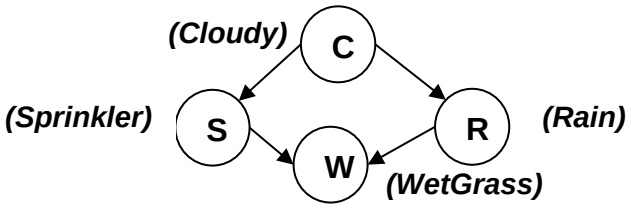


1b. (15 pts) Draw the Bayesian Network that corresponds to this conditional probability:

P(A | D,F,H,I) P(B | D, E,G, J) P(C | H) P(D | G) P(E| J) P(F | H) P(G | I, J) P(H) P(I) P(J)



1.c. (15 pts) Below is the Bayesian network for the WetGrass problem [Fig. 14.12(a) in R&N].



P(C)
.5

C	P(S)
t	.1
f	.5

C	P(R)
t	.8
f	.2

S	R	P(W)
t	*t*	.99
t	*f*	.90
f	*t*	.90
f	*f*	.00

Write down an expression that will evaluate to P(C=t ∧ R=t ∧ S=f ∧ W = t).
The probability tables show the probability that variable is True, e.g., P(M) means P(M=t). Express your answer as a series of numbers (numerical probabilities) separated by multiplication symbols. You do not need to carry out the multiplication to produce a single number (probability). **SHOW YOUR WORK.**

P(C=t ∧ R=t ∧ S=f ∧ W = t)
= P(W=t | R=t ∧ S=f) * P(R=t | C = t) * P(S=f | C=t) * P(C=t)
= .90 * .8 * .9 * .5

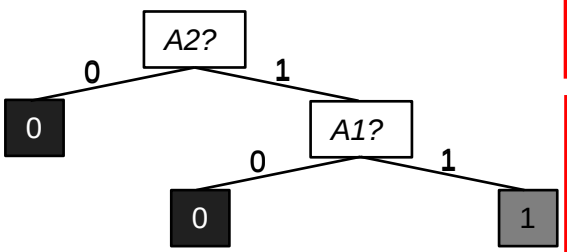
****** TURN PAGE OVER. QUIZ CONTINUES ON THE REVERSE ******

2. (30 pts total) Decision Tree Learning. Consider the following data set consisting of three binary input attributes (*A1*, *A2*, and *A3*) and one binary output (*y*):

Example	*A1*	*A2*	*A3*	Output *y*
x1	1	0	0	0
x2	1	0	1	0
x3	0	1	0	0
x4	1	1	1	1
x5	1	1	0	1

See Section 18.3 and Exercise 18.6, page 764.

2.a. (15 pts) Use the Decision-Tree-Learning algorithm to draw a decision t



If root is A1:
0 branch = 0, 1 branch = 0011
If root is A2:
0 branch = 00, 1 branch = 011
If root is A3:
0 branch = 001. 1 branch = 01

(Assume root is A2, as above.)
If attribute tested is A1:
0 branch = 0, 1 branch = 11
If attribute tested is A3:
0 branch = 01, 1 branch = 1

2.b. (5 pts) How would your tree classify *A1*=0, *A2*=0, *A3=0*? (Write 0 or 1.) 0

2.c. (5 pts) How would your tree classify *A1*=0, *A2*=0, *A3=1*? (Write 0 or 1.) 0

2.d. (5 pts) How would your tree classify *A1*=0, *A2*=1, *A3=1*? (Write 0 or 1.) 0

Full credit if your answers are right for the tree you drew, even if the tree itself is wrong.

3. (25 pts total, 5 pts each) Machine Learning. Label the following statements T (true) or F (false).

3a. T A decision tree can learn and represent any Boolean function.

See Section 18.3.2.

3b. F The information gain from an attribute A is how much classifier accuracy improves when attribute A is added to the example feature vectors in the training set.

See Section 18.3.4.

3c. T Overfitting is a general phenomenon that occurs with most or all typ

See Section 18.3.5.

3d. F Cross-validation is a way to improve the accuracy of a learned hypothesis by reducing over-fitting using Ockham's razor.

See Section 18.4.1.

3e. T An agent is learning if it improves its performance on future tasks after making observations about the world.

See the first sentence of Chapter 18.