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

YOUR NAME: _____

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

1. (5 pts) Definition of conditional probability. Write down the definition of P(H | D) in terms of P(H), P(D), P(H ∧ D), and P(H ∨ D).

P(H | D) = P(H ∧ D) / P(D)

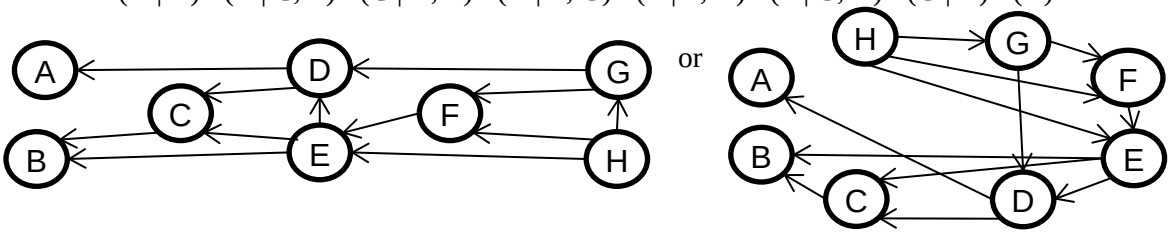
2. (5 pts) Bayes' Rule. Write down the result of applying Bayes' Rule to P(H | D).

P(H | D) = P(D | H) P(H) / P(D)

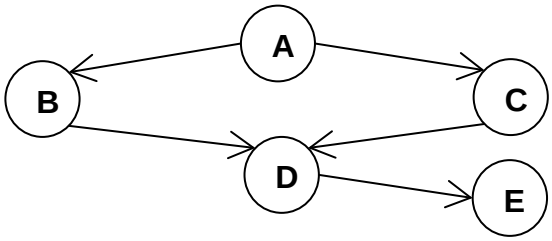
Also OK:
P(H | D) = P(D | H) P(H) / [P(D | H) P(H) + P(D | ¬H) P(¬H)]

3. (15 pts) Bayesian Networks. Draw the Bayesian Network that corresponds to this factored conditional probability expression. Draw left-to-right, i.e., put A and B on the left, G and H on the right.

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



4. (30 pts total) Bayesian Networks. Shown below is a Bayesian network and its probability tables.



P(A)
0.2

A	P(B)
t	0.90
f	0.05

A	P(C)
t	0.50
f	0.35

B	C	P(D)
t	t	0.95
t	f	0.94
f	t	0.29
f	f	0.10

D	P(E)
t	0.60
f	0.25

4.a. (15 pts) Write the factored conditional probability expression that corresponds to this network:

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

4.b. (15 pts) Write down an expression that will evaluate to P(a=T ∧ b=F ∧ c=T ∧ d = F ∧ e = F). 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. **SHOW YOUR WORK.**

P(a=T ∧ b=F ∧ c=T ∧ d = F ∧ e = F)

= P(e=F | d=F) * P(d=F | b = F, c=T) * P(c=T | a=T) * P(b=F | a=T) * P(a=T)

= 0.75 * 0.71 * 0.5 * 0.1 * 0. 2

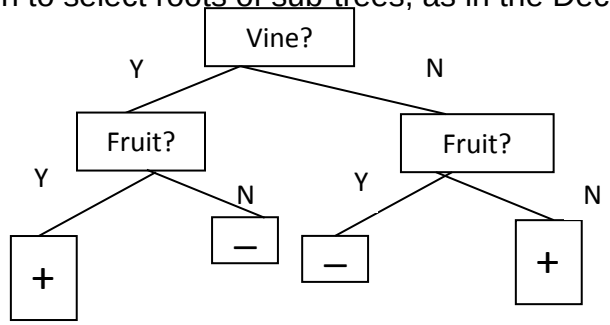
****** TURN QUIZ OVER. QUIZ CONTINUES ON THE REVERSE. ******

5. (15 pts total) Decision Tree Learning.
You are an agricultural robot given the following set of plant examples. Each is assigned a class label of + or — depending on whether or not it is a member of the target class:

Example	Vine?	Fruit?	Leaf?	Class
Watermelon	Yes	Yes	Curly	+
Ivy	Yes	No	Curly	—
Bougainvillea	Yes	No	Flat	—
Kudzu	Yes	No	Flat	—
Maple	No	No	Curly	+
Oak	No	No	Flat	+
Sycamore	No	No	Flat	+
Apple	No	Yes	Curly	—

5.a. (5 pts) Draw the decision tree that would be constructed by recursively applying information gain to select roots of sub-trees, as in the Decision-Tree-Learning algorithm.

Half credit for the correct root; half credit for wrong root but correct classification; full credit for the correct tree.



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

5.b. (5 pts) What class is Grape? (Vine=Yes, Fruit=Yes, Leaf=Curly) ____+____

5c. (5 pt) What class is Orange? (Vine=No, Fruit=Yes, Leaf=Curly) ____—____

6. (30 pts total, 2 pts each) Machine Learning concepts.
For each of the following items on the left, write in the letter corresponding to the best answer or the correct definition on the right.

A .	Learning	A	Improves performance of future tasks after observing the world
J	Information Gain	B	Fixed set, list, or vector of features/attributes paired with a value
M	Decision Boundary	C	Agent learns patterns in the input with no explicit feedback
L	Cross-validation	D	Agent observes input-output pairs & learns to map input to output
N	Linear Classifier	E	Example input-output pairs, from which to discover a hypothesis
B	Factored Representation (Feature Vector)	F	Examples distinct from training set, used to estimate accuracy
D	Supervised Learning	G	Supervised learning with a discrete set of possible output values
F	Test Set	H	Supervised learning with numeric output values
O	Naïve Bayes Classifier	I	Internal nodes test a value of an attribute, leaf nodes=class labels
G	Classification	J	Expected reduction in entropy from testing an attribute value
I	Decision Tree	K	Choose an over-complex model based on irrelevant data patterns
H	Regression	L	Randomly split the data into a training set and a test set
E	Training Set	M	Surface in a high-dimensional space that separates the classes
C	Unsupervised Learning	N	Tests **w**·**f**>0, where **w** is a weight vector and **f** is a feature vector
K	Overfitting	O	Tests $P(C) \prod_i P(X_i \mid C)$, where C is a class label and X_i are features