Practice problems for the Final
Reasoning about Uncertainty (50pts)

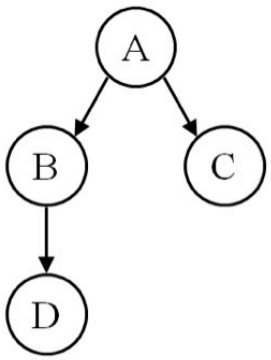
Bayesian Networks (30pts) Given this network calculate the following probabilities. Give both the formula and calculations with values. These questions are designed so that they can be answered with a minimum of computation. If you find yourself doing a copious amount of computation for each part, step back and consider whether there is simpler way to deduce the answer.

$P(A)$

+a	0.1
¬a	0.9

$P(B|A)$

+a	+b	0.5
	¬b	0.5
¬a	+b	0.8
	¬b	0.2



$P(C|A)$

+a	+c	0.4
	¬c	0.6
¬a	+c	0.7
	¬c	0.3

$P(D|B)$

+b	+d	0.9
	¬d	0.1
¬b	+d	0.2
	¬d	0.8

1. [4pts] P(a,¬b,c,¬d)

2. [4pts] P(b)

Bayesian Networks continued

3. [4pts] P(alb)

4. [4pts] P(dla)

5. [4pts] P(dla,c)

(10pts)Given the full joint probability table for the dentist problem, draw the Bayesian network and its corresponding conditional probability tables. Calculate the **P**(*Cavity* | **toothache**) using the conditional probability tables.

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	.108	.012	.072	.008
¬ *cavity*	.016	.064	.144	.576

Probability (20pts)
Circle **True** for each of the following statements if it is always true.

1. [2pts] [True/False] $P(A, B) = P(A)P(B)$

2. [2pts] [True/False] $P(A|B) = P(A)P(B)$

3. [2pts] [True/False] $P(A, B) = P(A)P(B) - P(A|B)$

4.[2pts] [True/False] $P(A,B,C)=P(A|B,C)P(B|C)P(C)$

5. [2pts] [True/False] If $P(A|B,C)= P(B|A,C)$ then $P(A|C)=P(B|C)$

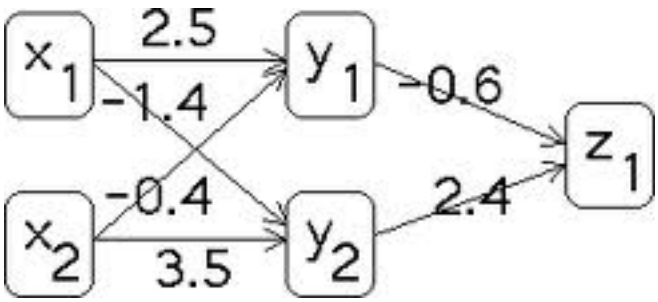
(10pts)Your doctor performs a series of medical tests. You test positive for a serious, but very rare disease. The test is 99% accurate, meaning the probability that it is positive when you do have the disease is 0.99, and the probability that it is negative if you don't is also 0.99. The disease strikes 1 in 10,000.

What is the probability that you have the disease? Give both the formula and the value.

Machine Learning (50pts)

Neural Networks (10pts)

In this network the input layer consists of units X1 and X2. The hidden layer is Y1 and Y2, and the output layer is Z1. Units Y1, Y2 and Z1 all have threshold functions where t>0 for the unit to activate.



1. [1pt] When input X1 = 0 and input X2 = 0, what does the network output?

2. [1pt] When input X1 = 0 and input X2 = 1, what does the network output?

3. [1pt] When input X1 = 1 and input X2 = 0, what does the network output?

4. [1pt] When input X1 = 1 and input X2 = 1, what does the network output?

5. [3pts] Can a weight be changed for the network to compute the Boolean function OR? If yes, which weight and what value should it be changed to?

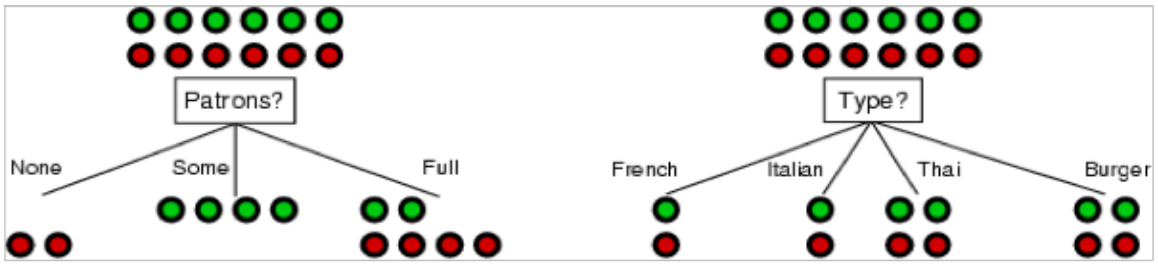
6. [3pts] Can a single weight be changed for the network to compute the Boolean function XOR? If yes, which weight and what value should it be changed to?

Decision Trees (10pts)

Given these examples from the Restaurant domain, answer the questions below:

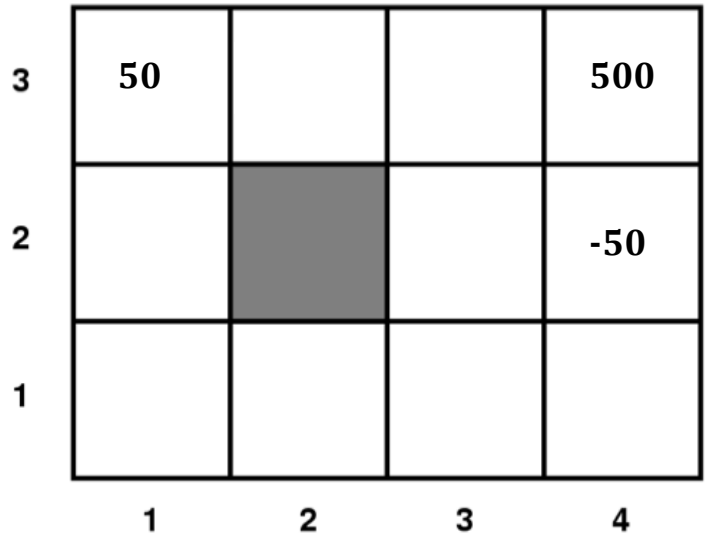
Example	Attributes										Target
	Alt	*Bar*	*Fri*	*Hun*	*Pat*	*Price*	*Rain*	*Res*	*Type*	*Est*	*Wait*
X_1	T	F	F	T	Some	$$$	F	T	French	0–10	T
X_2	T	F	F	T	Full	$	F	F	Thai	30–60	F
X_3	F	T	F	F	Some	$	F	F	Burger	0–10	T
X_4	T	F	T	T	Full	$	F	F	Thai	10–30	T
X_5	T	F	T	F	Full	$$$	F	T	French	>60	F
X_6	F	T	F	T	Some	$$	T	T	Italian	0–10	T
X_7	F	T	F	F	None	$	T	F	Burger	0–10	F
X_8	F	F	F	T	Some	$$	T	T	Thai	0–10	T
X_9	F	T	T	F	Full	$	T	F	Burger	>60	F
X_{10}	T	T	T	T	Full	$$$	F	T	Italian	10–30	F
X_{11}	F	F	F	F	None	$	F	F	Thai	0–10	F
X_{12}	T	T	T	T	Full	$	F	F	Burger	30–60	T

1) (5pts) Which one of these two attributes, **Patrons** or **Type?** would be chosen first by the Decision Tree algorithm? Why?



2) (5pts) Draw a graph similiar to the ones in problem 1 for the attribute **Est.** for the wait time estimate. Would the Decision Tree algorithm choose **Est.** before **Patrons**? before **Type**?

MDP (10pts)
Given a Gridworld domain, where terminal states (1,3), (4,3), and (4,2) have rewards 50, 500, and -50 respectively, the set of possible actions are {N,E,S,W, or X for terminal states}, the agent moves deterministically, all V and Q values for non terminal states have been initialized to 0, answer the questions below.



Circle the letter that corresponds to the best answer for the question

1. [4pts] What are the optimal values, V^* of each state in the above grid if $\gamma = 0.5$, c(a)=0, R(s)=0 for non terminal states?

   (Remember $\mathbf{V_{t+1}(s) = R(s) + Max_{a\varepsilon A}\ \{c(a)+\gamma\Sigma_{s',\varepsilon S}\ Pr(s'|a,s)\ V_t(s')\}}$)

   a. $V_{(1,1)}$=15.75, $V_{(1,2)}$=25, $V_{(2,1)}$=31.25, $V_{(2,3)}$=125, $V_{(3,1)}$=62.5, $V_{(3,2)}$=125, $V_{(3,3)}$=250, $V_{(4,1)}$=25
   b. $V_{(1,1)}$=12.5, $V_{(1,2)}$=25, $V_{(2,1)}$=31.25, $V_{(2,3)}$=125, $V_{(3,1)}$=62.5, $V_{(3,2)}$=125, $V_{(3,3)}$=250, $V_{(4,1)}$=31.25
   c. $V_{(1,1)}$=15.625, $V_{(1,2)}$=25, $V_{(2,1)}$=31.25, $V_{(2,3)}$=125, $V_{(3,1)}$=62.5, $V_{(3,2)}$=125, $V_{(3,3)}$=250, $V_{(4,1)}$=31.25
   d. $V_{(1,1)}$=12.5, $V_{(1,2)}$=25, $V_{(2,1)}$=25, $V_{(2,3)}$=25, $V_{(3,1)}$=50, $V_{(3,2)}$=100, $V_{(3,3)}$=250, $V_{(4,1)}$=25
   e. None of the above

2. [4pts] What are the Q values of state (3,2) in the above grid if $\gamma = 0.5$, c(a)=0, R(s)=-2 for non terminal states?

   (Remember $\mathbf{Q_{t+1}(a,s) = R(s) + c(a)+\gamma\Sigma_{s',\varepsilon S}\ Pr(s'|a,s)\ Q_t(a's')\}}$)

   a. $Q_{((3,2),N)}$=122, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=59
   b. $Q_{((3,2),N)}$=122, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=27.5
   c. $Q_{((3,2),N)}$=125, $Q_{((3,2),E)}$=-25, $Q_{((3,2),S)}$=62.5
   d. $Q_{((3,2),N)}$=120, $Q_{((3,2),E)}$=-27, $Q_{((3,2),S)}$=31.5
   e. None of the above

Short Answer (20pts)

Keep your answers brief, one or two sentences.

1. [6pts] List 2 reasons why machine learning is needed?

2. [8pts] What is Ockham's razor? How is it used in decision tree learning?

3. [6pts] What does the size of a hypothesis class have to do with generalization and overfitting?