6.034 Spring 2009

Midterm Solutions

March 18, 2009

Please read these notes before starting the exam.

- This exam has 20 pages. Make sure you have them all.
- This exam is open-note.
- Calculators are allowed, but probably not needed.
- Computers are *not* allowed.

PROBLEM	VALUE	SCORE
1	33	
2	15	
3	30	
4	20	
TOTAL	98	

Problem 1 (33 points)

(a) (2 points) Figure 1 illustrates the boundaries for two decision-tree classifiers trained on the same training set. One of the boundaries corresponds to an unpruned decision tree and the other corresponds to a pruned decision tree.

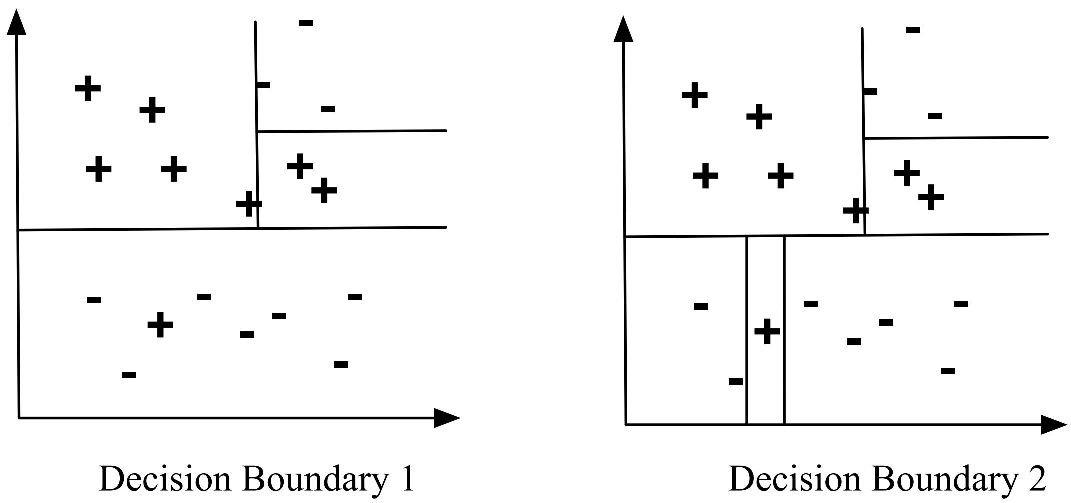

FIGURE 1

Circle the diagram corresponding to the **pruned** decision tree.

The figure on the left is pruned.

(b) (6 points) In class, we defined add-k Laplace correction, which is commonly used for smoothing Bayesian classifiers. Consider training classifiers with two values of k: 1 and 50.

- Would the two resulting classifiers have the same performance on the training set? If "no," which would perform better? Explain briefly.

  Answer: $k = 1$ will perform better on the training set than $k = 50$. When k is small, the classifier will fit the training data closely; as k increases, this fit diminishes.

- Would the two classifiers have the same performance on the test set? If "no," which would perform better? Explain briefly.

  This answer depends on what test set we are using. If the test set is very similar to the training set, then the $k = 1$ classifier will perform better. But if the test set is rather different from the training set, the $k = 50$ classifier will likely perform better.

(c) (6 points) You want to develop a classifier that identifies the disease "pseuditis" based on the results of several tests. Missing a true case of pseuditis has grave consequences for a patient, so it is of crucial importance to minimize false negatives. In contrast, false positive identification of pseuditis is a minor concern because additional tests can easily clarify the confusion. Explain how you can modify the Bayesian classifier presented in class to minimize the number of false negatives.

Answer:

(d) (15 points) Consider an algorithm for decision tree construction that randomly selects an attribute for splitting (rather than using the minimal entropy criterion presented in class).

- Is this algorithm guaranteed to generate a tree that does not test the same attribute twice? Explain briefly.

Answer:

- Is this algorithm guaranteed to correctly separate the training data? Explain briefly.

Answer:

- Is this algorithm guaranteed to provide the same classification accuracy on the *training* set as the entropy-based algorithm presented in class? Explain briefly.

Answer:

- Is this algorithm guaranteed to provide the same classification accuracy on the *training* set as the entropy-based algorithm presented in class, once both algorithms are pruned? Explain briefly.

Answer:

- Is this algorithm guaranteed to provide the same classification accuracy on the *test* set as the entropy-based algorithm presented in class, once both algorithms are pruned? Explain briefly.

Answer:

(e) (4 points) You are given a set A of k points which is not linearly-separable. If the point b is removed from this set, than the set of the remaining $k-1$ points becomes linearly-separable.

Consider training a perceptron algorithm on the original set of k points for n iterations. Would this algorithm correctly classify $k-1$ points in the set A-$\{b\}$? Explain briefly.

Answer:

Problem 2 (15 points)

DECISION TREE

(a) (5 points) Write an expression for the average entropy of the test $x_2 > -0.5$ for the data below. You do not need to find the final numerical value, but do not leave any variables in your expression. (You can leave logs in your answer.)

i	x_1	x_2	y
1	-2	0	1
2	2	2	1
3	-2	-1	-1
4	-2	3	-1

Answer:

(b) (5 points) For the same data, write a test (of the form $x_i > value$) considered by the decision tree algorithm that has average entropy equal to 1.

Answer:

(c) (5 points) The following table shows a data set and the number of times each point is misclassified during a run of the perceptron algorithm. What is the equation of the separating line found by the algorithm, as a function of x_1 and x_2? Assume that the learning rate is 1 and the initial weights are all 0.

i	x_1	x_2	y	times misclassified
1	-3	2	+1	1
2	-1	1	+1	0
3	-1	-1	-1	2
4	2	2	-1	1
5	1	-1	-1	0

Answer:

Problem 3 (30 points)

In this problem, we are going to tackle the nursery rhyme Three Blind Mice as a FOL problem. For those of you unfamiliar, here are the words to the nursery rhyme:

Three blind mice. Three blind mice.
See how they run. See how they run.
They all ran after the farmer's wife,
Who cut off their tails with a carving knife,
Did you ever see such a sight in your life,
As three blind mice?

We are going to convert English sentences that encapsulate the characters and situations in this nursery rhyme into FOL sentences, convert these FOL sentences to clausal form, and then do a resolution-refutation proof concerning the mice in this nursery rhyme.

Please carefully follow the directions, making sure that you do everything asked in each part. Sometimes you will need to express an English sentence as FOL and convert to clausal form. Sometimes you will only need to convert a FOL sentence to clausal form. And sometimes we will give you the English sentence, the FOL sentence, and the clauses.

Also, be sure to copy the clauses from this page to the proof where applicable. We have noted the number of clauses you should get from each sentence.

For each English sentence below, fill in the missing steps in converting the English sentence to FOL logic to clausal form. Two have been done for you. Use the following predicates:

Mouse(x) =	x is a mouse
Blind(x) =	x is blind
FarmersWife(x) =	x is a farmer's wife
Run(x) =	x runs
Chase(x, y) =	x chases y
Cut(x, y) =	x cuts the tail off y
Tail(x) =	x has a tail
Eq(x, y) =	x is equal to y

(a) There are exactly three blind mice.

FOL:

Answer: $\exists x \, \exists y \, \exists z$ Mouse(x) $\wedge$ Blind(x) $\wedge$ Mouse(y) $\wedge$ Blind(y) $\wedge$ Mouse(z) $\wedge$ Blind(z) $\wedge$ $\neg$Eq(x,y) $\wedge$ $\neg$Eq(y,z) $\wedge$ $\neg$Eq(x,z) $\wedge$ ($\forall$w Mouse(w) $\wedge$ Blind(w) $\rightarrow$ Eq(w,x) $\vee$ Eq(w,y) $\vee$ Eq(w,z))

Clauses (10 clauses):

Answer:
Mouse(Huey)
Blind(Huey)
Mouse(Dewey)
Blind(Dewey)
Mouse(Louie)
Blind(Louie)
¬ Eq(Huey,Dewey)
¬ Eq(Dewey,Louie
¬ Eq(Huey,Louie)
¬Mouse(w) ∨ ¬Blind(w) ∨ Eq(w,Huey) ∨ Eq(w,Dewey) ∨ Eq(w,Louie)

(b) (4 points) All blind mice run.

FOL:

Answer:

Clauses (1 clause):

Answer:

(c) (5 points) There is only one farmer's wife.

FOL:

Answer:

Clauses (2 clauses):

Answer:

(d) (3 points) All blind mice chase a farmer's wife.

FOL:

Answer: $\forall x$ Mouse(x) $\land$ Blind(x) $\rightarrow$ $\exists y$ FarmersWife(y) $\land$ Chase(x,y)

Clauses (2 clauses):

Answer:

(e) (5 points) All farmer's wives cut the tail off blind mice that chase them.

FOL:

Answer:

Clauses (1 clause):

Answer:

(f) If there is someone who cuts off your tail, you have no tail.

FOL:

Answer: $\forall x\ \exists y\ Cut(y,x) \rightarrow \neg Tail(x)$

Clauses (1 clause):

Answer: $\neg Cut(f(x),x) \vee \neg Tail(x)$

(g) (5 points) We want to prove "All blind mice do not have tails." Write this sentence as a FOL sentence.

FOL:

Answer:

Because we are going to use resolution refutation, negate this sentence and use the negated sentence to create the clauses that we will use in our proof.

Clauses (3 clauses from the *negation* of the FOL sentence you wrote above):

Answer:

(h) (8 points) Now fill all the clauses you derived above into the table below and do a resolution refutation proof to derive a contradiction. Write the numbers of the two clauses (from the step column) you are unifying into P1 and P2, and fill in the unifier you are using.

The table is longer than you should need.

Step	P1	P2	Clause	Unifier
1	From (a)		Mouse(Huey)	N/A
2	From (a)		Blind(Huey)	N/A
3	From (a)		Mouse(Dewey)	N/A
4	From (a)		Blind(Dewey)	N/A
5	From (a)		Mouse(Louie)	N/A
6	From (a)		Blind(Louie)	N/A
7	From (a)		¬Eq(Huey,Dewey)	N/A
8	From (a)		¬Eq(Dewey,Louie)	N/A
9	From (a)		¬Eq(Huey,Louie))	N/A
10	From (a)		¬Mouse(w) ∨ ¬Blind(w) ∨ Eq(w,Huey) ∨ Eq(w,Dewey) ∨ Eq(w,Louie)	N/A
11	From (b)			N/A
12	From (c)			N/A
13	From (c)			N/A
14	From (d)			N/A
15	From (d)			N/A
16	From (e)			N/A
17	From (f)		¬Cut(f(x),x) ∨ ¬Tail(x)	N/A
18	From (g)			N/A
19	From (g)			N/A
20	From (g)			N/A
21				
22				

Step	P1	P2	Clause	Unifier
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				

Problem 4 (20 points)

(a) (4 points) Consider the following axioms in predicate calculus, where we obey the usual naming conventions, so that A, B, C, and D are constant symbols, x and y are variables, and P and R are predicate symbols:

$$\forall x.R(x, A) \leftrightarrow \neg\exists y.P(y) \wedge R(x, y) \quad (1)$$

$$\forall x.\neg R(A, x) \quad (2)$$

Is this set of sentences *valid*, *satisfiable*, or *unsatisfiable*, according to model theory for first-order predicate calculus? If satisfiable, give an interpretation in which each sentence holds. If valid, sketch a proof that it would hold in all possible interpretations, and if invalid, that it would hold in no interpretations.

Answer:

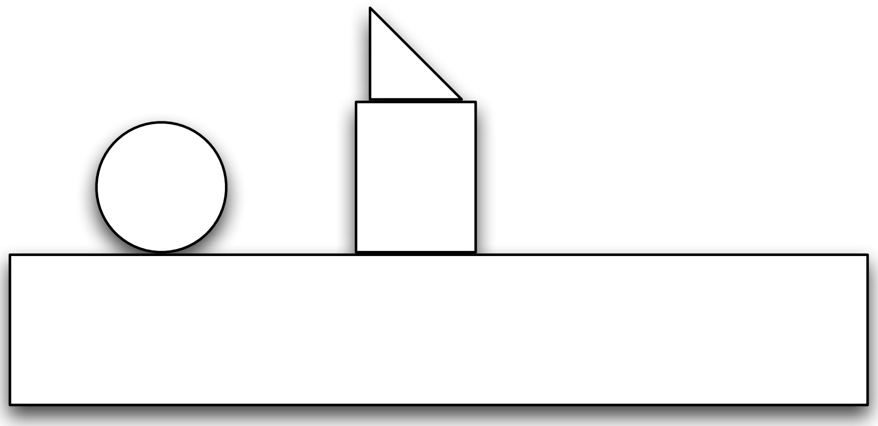

FIGURE 2. A table and three shapes.

(b) (4 points) If we add the following sentences,

$$\neg P(A) \tag{3}$$

$$P(B) \tag{4}$$

$$P(C) \tag{5}$$

$$P(D) \tag{6}$$

then we can read these as describing a simple blocks world scenario represented by Figure 2, which shows a table and three other shapes: a circle, a rectangle, and a triangle.

Give a "natural" interpretation of A, B, C, D, P and R to represent the scene in Figure 2.

Answer:

(c) (4 points) What is the meaning of sentence 1 in this blocks world?

Answer:

(d) (4 points) How could you prove that the table cannot support itself?

Answer:

(e) (4 points) A set of formulas can often describe models very different from the ones intended. Give an interpretation that satisfies sentences 1–6 in terms of the natural numbers and an arithmetic interpretation of R.

Answer: