University of Wisconsin-Madison
Computer Sciences Department

CS 760 — Machine Learning

Spring 1989

Midterm Exam (Open Book)

100 points, 90 minutes

April 27, 1989

Write your answers on these pages and show your work. If you feel that a question is not fully specified, state any assumptions you need to make in order to solve the problem. You may use the backs of these sheets for scratch work. Notice that all questions do not have the same point-value. Divide your time appropriately.

Before starting, write your name on this and all other pages of this exam. Also, make sure your exam contains 4 problems on 6 pages.

Problem	Score	Max Score
1	_______	35
2	_______	20
3	_______	25
4	_______	20
Total	_______	100

1. Similarity-Based Learning (35 points)

Assume you are given the following three nominal features with the possible values shown.

Architecture	$\in$	{Gothic, Romanesque}
Size	$\in$	{Small, Large}
Steeples	$\in$	{Zero, One, Two}

part a (15 points)

Assume Mitchell's version space algorithm is incrementally given the following set of classified examples. Also assume the system is biased to learn purely conjunctive concepts. Report what *S* and *G* are immediately after each example is processed. (You may use the abbreviations which are used to describe the examples.)

Arch = G Sz = S St = 2 +

Arch = R Sz = S St = 2 -

Arch = G Sz = L St = 2 +

Arch = G Sz = L St = 0 -

part b (15 points)

Compute an ID3 tree using Quinlan's max-gain formula for the same set of examples. Show all your work.

part c (5 points)

Compare the final *S* and *G* borders of VS to the decision tree produced by ID3. What similarities are there between the leaf nodes of ID3 and the *S* and *G* borders of VS?

2. Learning without a Teacher (20 points)

part a (15 points)

Assume the Bacon.3 system is given this table of first level data. The independent variables are *x* and *y*.

x	y	z
2.00	7.07	14.14
4.00	2.50	10.00
6.00	2.16	8.16
8.00	0.93	7.07
10.00	3.68	6.32

Describe a formula that Bacon.3 could learn from this data and *roughly* describe the steps it would take to arrive at this result. Show any intermediate data tables produced and describe what motivates their construction.

part b (5 points)

Briefly describe what UNIMEM might do if given the data in the above table.

3. Explanation-Based Learning (25 points)

Consider the following "domain" theory. Terms beginning with ?'s are implicitly universally quantified variables.

A(?x, ?x, ?y) *and* B(red, ?z)	→	C(?w, ?x, ?y, ?z)
D(?z, ?z) *and* D(?y, ?x) *and* E(?x, ?y)	→	A(?x, ?y, ?z)
F(?y, ?x)	→	B(?x, ?y)
G(?x, ?x)	→	D(?x, ?y)

Assume the following problem-specific facts are asserted.

E(eyes, eyes)	F(fire, red)	G(2, 2)	G(eyes, eyes)
E(eyes, ears)	F(tree, green)	G(2, 1)	G(eyes, ears)
E(eyes, nose)	F(snow, white)	G(3, 2)	G(eyes, nose)

part a (5 points)

Explain, with a proof tree, that *C(white, eyes, 2, fire)* is true. Draw to the right of your proof tree the corresponding *explanation structure* (before pruning at operational nodes). Clearly indicate the necessary unifications.

part b (15 points)

Under each of the following two assumptions, what new rule would the EGGS algorithm produce?

Only the predicates E, F, and G are operational.

In addition to E, F, and G, the predicate D is operational.

part c (5 points)

Briefly describe one weakness of the EGGS algorithm.

4. Artificial Neural Networks (20 points)

Consider using an artificial neural network (without hidden units) to learn the following concept:

The (single) output unit is "active" if the first and third (of 3) input units have opposite binary values.

Assume the single output unit is a linear threshold unit, whose threshold is initially zero (assume the node is "active" if it equals or exceeds the threshold). Also assume all weights are initially zero and that η=0.25.

part a (15 points)

Draw the network after each of the following training examples.

Input			**Output**
0	1	0	0
1	0	0	1
1	1	1	0
0	0	1	1

part b (5 points)

Is this concept learnable by the delta rule? Why or why not?