CS 540: Introduction to Artificial Intelligence

Final Exam: 12:25-2:25pm, December 16, 2002
Room 168 Noland

CLOSED BOOK
(two sheets of notes and a calculator allowed)

Write your answers on these pages and show your work. If you feel that a question is not fully specified, state any assumptions that you need to make in order to solve the problem. You may use the backs of these sheets for scratch work.

Write your name on this and all other pages of this exam. Make sure your exam contains six problems on ten pages.

Name ____

Student ID ____

Problem	Score	Max Score
1	______	**28**
2	______	**12**
3	______	**20**
4	______	**10**
5	______	**12**
6	______	**18**
TOTAL	______	**100**

Name: ______________________________

Problem 1 – Representing and Reasoning with Logic (28 points)

a) Convert each of the following English sentences into *First-Order Predicate Calculus* (FOPC), using reasonably named predicates, functions, and constants. If you feel a sentence is ambiguous, clarify which meaning you're representing in logic. (Write your answers *below* each English sentence.)

All birds can fly except for penguins and ostriches or unless they have a broken wing.

There was a student in CS 540 Fall 1999 who was born in a country in South America.

John sold Mary his CS 540 textbook (and, hence, this book that John formerly owned is now owned by Mary). [You must use *situation calculus* here.]

Name: ______________________________

b) Provide a formal *interpretation* that shows that the following translation from English to FOPC is *incorrect*. Be sure to *explain your answer formally* using the interpretation you provide.

A book of Sue's is missing. $\forall x\ [\ book(x) \wedge owner(x, Sue)\ ] \rightarrow missing(x)$

c) What is the *most-general unifier* (mgu) of these two wff's? ______________________
Show your work.

$P(?x, ?x, f(?y))$ $P(g(?a, ?b), g(1, ?b), ?b)$

d) Why is *And Elimination* a legal inference rule but *Or Elimination* is not?

Name: ______________________________

Problem 2 – Neural Networks (12 points)

a) Consider a *perceptron* that has two real-valued inputs and an output unit with a *step function* as its activation function. All the initial weights and the bias ("threshold") equal 0.1. Assume the teacher has said that the output should be 0 for the input *in1 = 5* and *in2 = -3*.

Show how the perceptron learning rule would alter this neural network upon processing this training example. Let η (the learning rate) be 0.2 and be sure to adjust the output unit's bias during training.

Perceptron BEFORE Training

Perceptron AFTER Training

b) Qualitatively draw a (2D) picture of *weight space* where the backprop algorithm is likely to

i. do well

ii. do poorly

Be sure to explain your answers.

Name: ______________________________

Problem 3 –Miscellaneous Questions (20 points)

a) What do you feel are the two (2) most important design choices you would need to make if you used *CBR* to choose the location of your next vacation? Briefly justify your answers**.**

i. ____

ii. ____

b) In a weird dream, you're the *simulated annealing* algorithm. Currently you're at node *A* in a search space; *g(A)* = 7 and *h(A*) = 5. You next randomly select node *B*; *g(B)* = 9 and *h(B)* = 8. The temperature is a Wisconsin-like 10 degrees.

Do you move to node *B*? __________ Show your work. (*Lower* h values are better.)

c) Show an example of a *cross over* for a GA whose individuals/entities are 6-bits long.

Name: ______________________________

d) On your way out of the hit feature *To Build a Decision Tree*, you are surprised to find out the movie theater is giving away prizes. You watch the people ahead of you choose their prize either from behind Door #1 or Door #2. Of those who chose Door #1, half received \$5, 1% got a new bike worth \$1000, and the rest got a worthless movie poster. Everyone who chose Door #2 got \$10.

Assuming you want to maximize the likely dollar value of your prize, what door should you choose? _______________ Why?

e) Consider the *joint probability distribution* below.

A	**B**	**C**	**P(A, B, C)**
False	False	False	0.05
False	False	True	0.10
False	True	False	0.03
False	True	True	0.25
True	False	False	0.15
True	False	True	0.02
True	True	False	0.07
True	True	True	0.33

i. What is $P(A = true)$? _______________ Show your work below.

ii. What is $P(A \rightarrow B)$? _______________ Explain.

Name: ______________________________

Problem 4 – Important AI Concepts (10 points)

Describe each of the following AI concepts and briefly explain its most significant aspect. (Write your answers in the space *below* the AI concept.)

Soundness

Overfitting

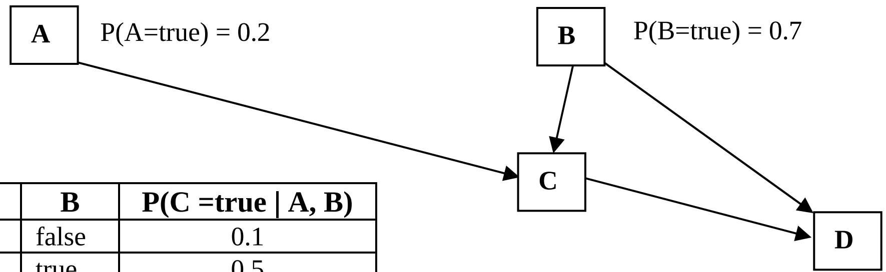
Fitness Functions

Vector-Space Model

Negation by Failure

Problem 5 – Bayesian Networks (12 points)

Consider the following Bayesian Network, where variables **A-D** are all Boolean-valued:



A	B	P(C =true | A, B)
false	false	0.1
false	true	0.5
true	false	0.4
true	true	0.9

B	C	P(D=true | B, C)
false	false	0.8
false	true	0.6
true	false	0.3
true	true	0.1

a) What is the probability that all four of these Boolean variables are false? _______________

b) What is the probability that ***C*** is true, ***D*** is false, and ***B*** is true? ________________

c) What is the probability that ***C*** is true *given* that ***D*** is false and ***B*** is true? ________________

Name: ____________________________

Problem 6– More Probabilistic Reasoning (18 points)

a) Imagine that 99% of the time *RE Disease* (RED) causes red eyes in those with the disease, at any point in time 2% of all people have red eyes, and at any point in time 1% of the population has RED.

You have red eyes. What is the probability you have RED? ______________

b) Assume we have one diagnostic random variable (call it ***D***) and two measurement variables (call them ***M1*** and ***M2***). For simplicity, assume that the ***M***'s variables have three possible values (e.g., *low*, *medium*, and *high*) and that ***D*** is Boolean-valued.

We collect data on 300 episodes and find out the following:

D was ***true*** 100 times and for these cases:

M1=low 50 times, ***M1=med*** 30 times, and ***M1 = high*** 20 times
M2=low 10 times, ***M2=med*** 80 times, and ***M2 = high*** 10 times

D was ***false*** 200 times and for these cases:

M1=low 20 times, ***M1=med*** 80 times, and ***M1 = high*** 100 times
M2=low 180 times, ***M2=med*** 10 times, and ***M2 = high*** 10 times

Making the assumption that ***M1*** and ***M2*** are *conditionally independent* given ***D***,

i. Show how Bayes rule can be used to compute ***P(D | M1, M2)*** *given the data above and under the stated assumptions.* [Do this **algebraically** – i.e., as an equation.**]**

ii. On a new episode we find ***M1=low*** and ***M2=low***. What is the most likely diagnosis? _________________ This time justify your answer numerically.

iii. Draw the Bayesian network that one would construct from the above data (do *not* add any “pseudo” counts to the above statistics; we won’t worry about dealing with probabilities equaling zero). Be sure to explain your solution.

Have a good vacation!