CS 540:
Introduction to Artificial Intelligence

Final Exam: 1:00-3:30 pm, August 8, 2003
Room 265 Materials Sciences Building

CLOSED BOOK
(two-sided sheet of handwritten 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 you need to make in order to solve the problem. You may use the backs of these sheets for scratch work. Budget your time wisely.

Before you begin, write your name on every page of the exam and read through all the questions (as some have multiple parts and are more involved than others). Make sure your exam contains ***seven (7)*** problems on ***eleven (11)*** pages.

Name _____

Student ID _____

Problem	Score	Max Score
1	_____	20
2	_____	15
3	_____	25
4	_____	15
5	_____	30
6	_____	10
7	_____	10
TOTAL	_____	125

Name: _____

Problem 1: General Search Questions (20 points)

a) Answer each of the following questions *true* or *false*:

i. Tabu search with a horizon of 1 behaves the same as a greedy hill-climbing search.

ii. Simulated annealing with a temperature $T = 0$ also behaves identically to a greedy hill-climbing search.

iii. Breadth-first search is always a complete search method, even if all of the actions have different costs.

iv. When hill-climbing and greedy best first search use the exact same admissible heuristic function, they will expand the same set of search nodes.

v. If two admissible heuristic functions evaluate the same search node n as $h_1(n) = 6$ and $h_2(n) = 8$, we say h_1 dominates h_2, because it is less likely to overestimate the actual cost.

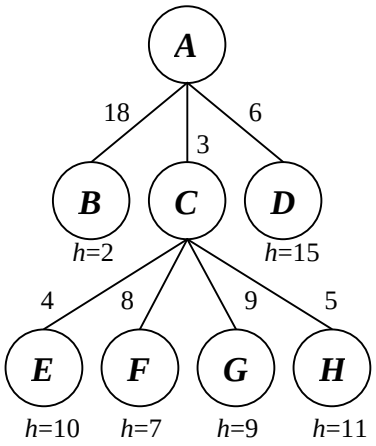
b) Provide a short answer to the following questions:

i. What are the three basic components of any genetic algorithm?

ii. Compare and contrast genetic algorithms to beam search.

Name: _____

c) Consider the following *partial search tree*, where edges are labeled with actual costs of the associated action, and each node is labeled with its heuristic evaluation. Which node will be expanded next by each of the following search methods?



i. Greedy best-first search:

ii. Uniform cost search:

iii. A* search:

d) Your good friend is the groundskeeper at the mansion of mean old Mr. Mathis. In the backyard, there is a huge fountain with a complex network of pipes controlled by over 100 valves. One weekend, Mr. Mathis announces he's going on vacation, and when he returns he wants the fountain to spray as high as it can… but the plans for the pipe network have been lost! Plus, since she only has the weekend, your friend can't possibly try *all* of the valve combinations to find the optimal setting. Which local optimization search method might she want to use *in real life* (since the fountain can't be simulated on a computer) to maximize the height of the fountain? You may assume that a valve is either *on* or *off*, and the water height is easily measured. If you need to make any other assumptions, state them clearly.

Name: _____

Problem 2: Logic and Planning Questions (15 points)

e) Provide a short answer to each of the following questions:

i. Are the literals $P(F(y), y, x)$ and $P(x, F(A), F(v))$ unifiable? If so, show the unifying substitution θ. If not, explain why.

ii. Convert this FOL sentence to conjunctive normal form: $\forall x \exists y\ Owns(x,y) \Rightarrow \neg Like(x,y)$

iii. What is the frame problem in situation calculus? How do we deal with it?

f) Use forward chaining to show that $Bigger(Smaug, Bilbo)$ is entailed by the following KB. Indicate which sentences were used with GMP, and show each substitution θ.

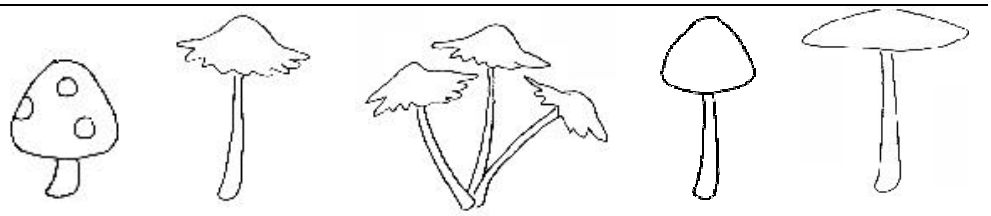
1. $Hobbit(Bilbo)$
2. $Wizard(Gandalf)$
3. $Dragon(Smaug)$
4. $Dragon(x) \wedge Wizard(y) \Rightarrow Bigger(x,y)$
5. $Wizard(x) \wedge Hobbit(y) \Rightarrow Bigger(x,y)$
6. $Bigger(x,y) \wedge Bigger(y,z) \Rightarrow Bigger(x,z)$

Name: _____

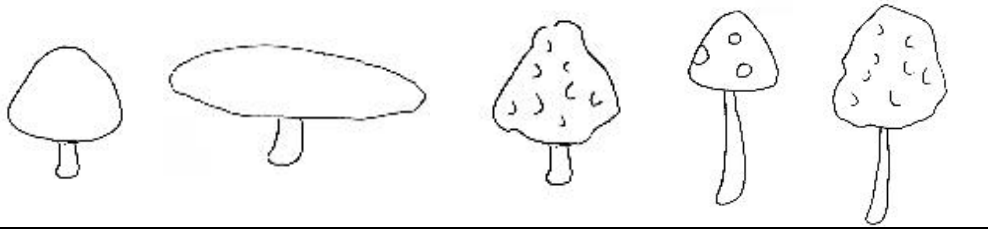
Problem 3: General Learning Algorithms (25 points)

Consider the task of learning to identify mushrooms that are SAFE or POISONOUS to eat based on a set of physical features. Four Boolean and discrete valued features that you could use are: STEM = {short, long}, BELL = {rounded, flat}, TEXTURE = {plain, spots, bumpy, ruffles}, and NUMBER = {single, multiple}. Consider using these features on the following training data:

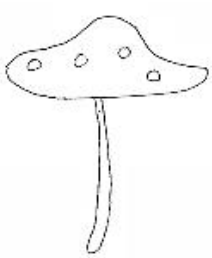
SAFE:



POISONOUS:



a) How would the naïve Bayes classifier label the following example:

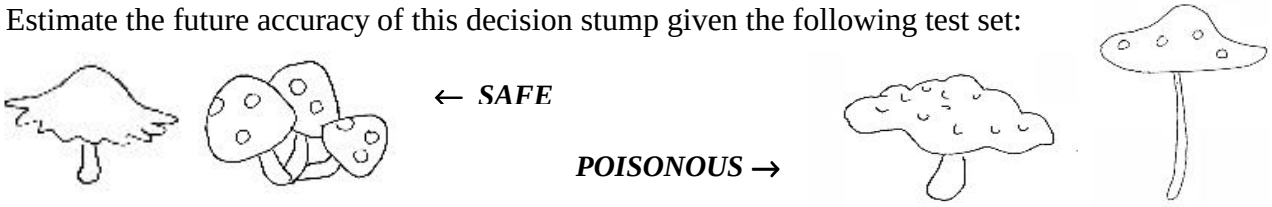


b) How would 3-nearest neighbors, using hamming distance and unweighted voting, classify the same example from part (a)?

Name: _____

c) Use information gain to choose between TEXTURE and NUMBER to use in a *decision stump* (a decision tree with only one internal node at the root). Draw a diagram of the learned stump, and break ties at classification nodes by labeling them as *POISONOUS* (just to be on the safe side). Show all your work for partial credit.

d) Estimate the future accuracy of this decision stump given the following test set:



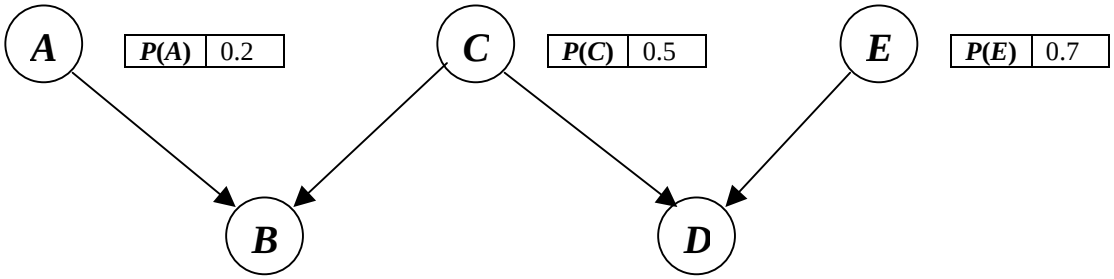
e) Draw a perceptron that might be used to learn this problem (just the structure is fine, you do not need to label with weights).

f) What is the dimensionality of a perceptron's hypothesis space for this problem?

Name: _____

Problem 4: Bayesian Networks (15 points)

Consider the following Bayesian network where all variables are Boolean:



$P(A)$	0.2

$P(C)$	0.5

$P(E)$	0.7

	A,C	$\neg A,C$	$A,\neg C$	$\neg A,\neg C$
$P(B)$	0.9	0.6	0.4	0.5

	C,E	$\neg C,E$	$C,\neg E$	$\neg C,\neg E$
$P(D)$	0.3	0.1	0.8	0.7

a) How large would the full-joint probability table (FJPT) have to be for this problem?

b) What is the joint probability that all five variables are simultaneously true?

c) Compute the probability $P(\neg C \mid A, \neg B, D, \neg E)$

Name: _____

Problem 5: Machine Learning Questions (30 points)

Provide a short answer to each of the following questions:

a) What's the difference between *eager* and *lazy* learning algortihms? Give an example of each.

b) Give an example of a *regression* learning task. Which of the machine learning algorithms we've discussed can learn this kind of concept well?

c) Give one advantage and one disadvantage of Bayesian belief networks over naïve Bayes.

d) Would the hypotheses learned by an inductive logic programming (ILP) system be considered *symbolic* or *connectionist* AI? Breifly explain.

e) What is the *curse of dimensionality*? Which learning methods are most affected by it? Describe one approach to dealing with this problem.

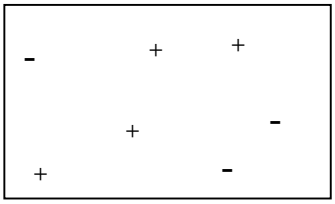
f) What relationship does the *minimum description length* principle have with the ID3 learning algorithm? What about backpropagation learning for neural networks?

g) Ensemble learning algorithms often benefit from the combined knowledge of its constituent agents, which learn from *slightly* different subsets of examples. Imagine that we want to create and ensemble learner where each constituent agent learns from subsets that are *very* different. How might we do this? How might the ensemble then classify new examples? Clearly state any assumptions that you need to.

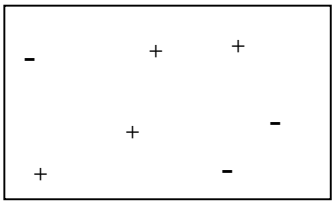
Name: _____

Problem 6: Feature Space (10 points)

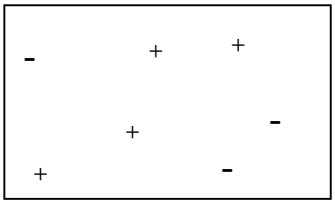
The following Venn diagrams represent training examples for a Boolean-valued concept function plotted in *feature space*. Show how each of the following machine learning algorithms *might* partition the space based on these examples. Briefly explain to the right of each diagram why that algorithm would partition the data that particular way. (No need for calculations… just give a qualitative answer.)



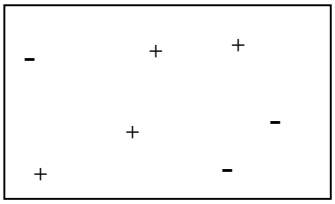
Naïve Bayes



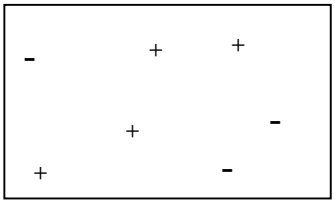
1-Nearest Neighbor



ID3



Neural Network with 1 hidden layer of 2 units



Neural Network with unlimited hidden layers/units

Name: _____

Problem 7: Miscellaneous Definitions (10 points)

Briefly define the following concepts and explain the significance of each to A.I.
(Write your answer *below* the phrase.)

α-β Search

Occam's Razor

Tuning Sets

Promotion / Demotion

The Markov Assumption

Have a good rest of the summer!!