Name:
Student ID:

<u>CSE 473 Autumn 2006: Take-Home Final Exam</u>

Total: 150 points, 5 questions
Open book, open notes
Due: Wednesday Midnight, December 13, 2006
Email your answers to both Raj and Abhay

<u>Instructions</u>:

1. Type in your answers using your favorite word processor.
2. Type your name and student ID at the top.
3. Keep your answers brief but provide enough details and explanations to let us know you understand the concepts involved.
4. If you need to draw something or write equations by hand as part of your answer, be sure to scan or photograph the page and include the image as part of your answer.

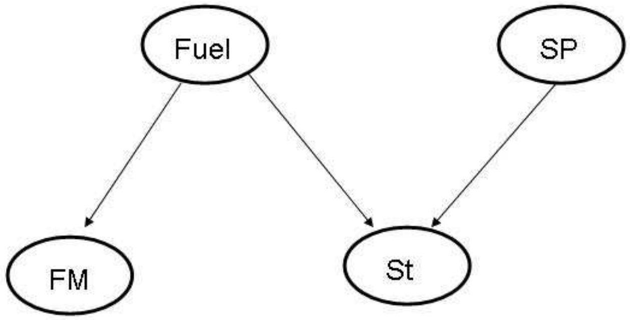
1) (25 points: 5, 10, 10 points) Planning

A monkey wants to eat some bananas hanging from a ceiling. There is a box in the vicinity of the bananas. Initially, the monkey is *Hungry* and is at height *Low* while the bananas are at height *High*. If the monkey climbs onto the box, it will be at height *High*. If the monkey is at the same height as an object, the monkey can eat the object. After eating, the monkey is not *Hungry* anymore. The actions available to the monkey are *ClimbUp* onto a *Climbable* object and *Eat* an *Eatable* object. The box is *Climbable* and the bananas are *Eatable*.

a) Write down the initial state description for the above planning problem.
b) Write down STRIPS-style definitions of the actions.
c) Suppose the monkey's goal is to be not *Hungry*. Show the steps of partial-order planning for achieving this goal using a format similar to the example in pages 391-393 in the text.

2) (20 points: 5, 5, 10 points) Bayesian Networks

The following Bayesian network captures some of the causal dependencies pertaining to whether your car will start in the morning:



The random variables *Start?* (St), *Fuel?* (Fuel), and *Clean-Spark-Plugs*? (SP) can each take on the values Yes or No, while *Fuel-Meter* (FM) can take on the values Full, Half, or Empty.

You know **P**(Fuel) = ⟨0.98,0.02⟩ and **P**(SP) = ⟨0.96,0.04⟩. You also know the following CPTs for **P**(FM | Fuel) and **P**(St | Fuel, SP):

Fuel	P(FM = Full)	P(FM = Half)
Yes	0.50	0.30
No	0.001	0.001

Fuel	SP	P(St = Yes)
Yes	Yes	0.99
Yes	No	0.01
No	Yes	0
No	No	0

a) Write down an expression for the full joint distribution over the above four random variables as a product of conditional probabilities based on the above Bayesian network structure.
b) Compute the joint probability P(Fuel = Yes, SP = No, FM = Half, St = No).
c) Use either inference by enumeration or the variable elimination algorithm to compute **P**(St | FM = Empty). Show all the steps involved.

3) (15 points) Decision Trees

Suppose a problem domain is described by the attributes A, B, and C, where A and B can each assume the values *Yes* or *No*, and C can assume the values *Yes*, *No*, or *Maybe*. Based on the decision tree learning algorithm discussed in class and in the textbook (best attribute at each step chosen according to information gain), construct a decision tree for this problem using the following set of training examples:

Example	**A**	**B**	**C**	**Goal (Class)**
1	*Yes*	*Yes*	*Maybe*	*Yes*
2	*No*	*Yes*	*Maybe*	*No*
3	*Yes*	*No*	*No*	*No*
4	*No*	*No*	*No*	*No*
5	*Yes*	*No*	*Yes*	*Yes*

4) (30 points: 5, 10, 10, 5 points) Neural Networks

In class, we discussed neural networks (perceptrons) that have threshold and sigmoid activation functions. Consider networks whose neurons have *linear activation functions*, i.e., each neuron's output is given by $g(x) = bx+c$, where x is the weighted sum of inputs to the neuron, and b and c are two fixed real numbers.

a) Suppose you have a single neuron with a linear activation function g as above and input $\mathbf{x} = x_0,\ldots,x_n$ and weights $\mathbf{W} = W_0,\ldots,W_n$. Write down the squared error function for this input if the true output is y.

b) Write down the weight update rule for the neuron based on gradient descent on the above error function.

c) Now consider a network of linear neurons with one hidden layer of m units, n input units, and one output unit. For a given set of weights w_{kj} in the input-hidden layer and W_j in the hidden-output layer, write down the equation for the output unit as a function of w_{kj}, W_j, and input $\mathbf{x}$. Show that there is a single-layer linear network with no hidden units that computes the same function.

d) Given your result in (c), what can you conclude about the computational power of N-hidden-layer linear networks for N = 1, 2, 3, …?

5) (60 points: 10 points each) Important Concepts and Techniques in AI

Provide brief summaries (100-200 words each) of why the following concepts/techniques are important in AI:

a) A* search

b) Simulated annealing

c) Alpha-beta pruning

d) Resolution

e) Cross-validation in machine learning

f) Temporal difference learning