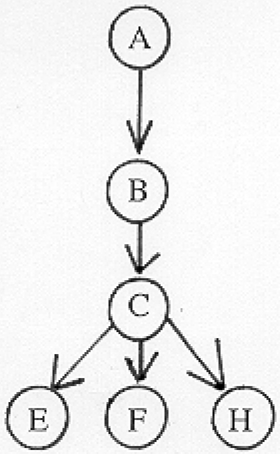
C1. ARTIFICIAL INTELLIGENCE (25 POINTS)

a) Bayesian Nets (20 points)



B, C, E, F, and H are Boolean variables.
A is a variable that can take on the values 0, 1, or 2.

A	P(A)
0	.2
1	.3
2	.5

A	P(B)
0	.6
1	.5
2	.3

B	P(C)
t	.3
f	.8

C	P(E)	P(F)	P(H)
t	.1	.2	.3
f	.4	.5	.6

What is P(B|E,F,H)?

Show all your work and justify each step.

1b) Search (5 points)
Suppose that each state in a problem has m successor states. How does local beam search differ from hill climbing?

C2. ARTIFICIAL INTELLIGENCE (25 POINTS)

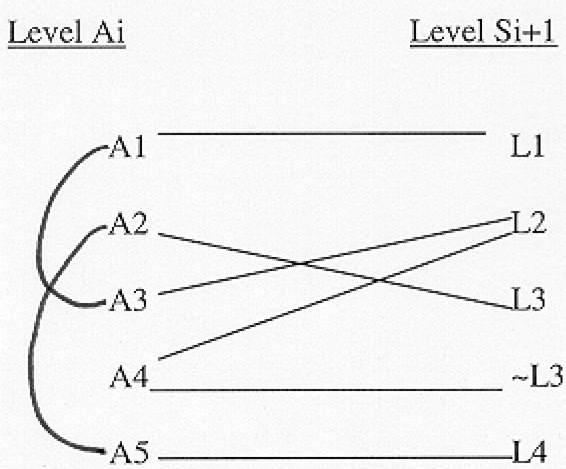
a) (5 points) Does the arc consistency algorithm used for constraint satisfaction problems reveal every possible inconsistency? If yes, prove it. If no, give an example showing how it fails to identify an inconsistency.

b) (5 points) In what two ways are heuristics used in constraint satisfaction problems? Give a sample heuristic for each of these.

c) PLANNING GRAPHS (6 POINTS)

The following displays level Ai of a planning graph, with solid lines between actions showing pairs of actions that are mutex. Solid lines between actions at level Ai and literals at level Si+1 give the effects of actions.

Give each pair of literals at level Si+1 that are mutex and explain why the pair is mutex.



d) Search (9 points)

Explain clearly how a genetic algorithm works, giving particular attention to 1) how a state is selected to participate in reproduction and 2) how new elements are constructed from the states selected for reproduction. (Assume that a state is represented as a string of 32 bits.)

C3 Artificial Intelligence (25 points)

Planning. (16 points) Consider the domain of planning for the new iRobot Kachooba Robotic Kitchen Cleaner.

3a. (6 pts) Write a set of STRIPS-style operators for this domain. When you describe the operators, take into account the following considerations:

- Cleaning the STOVE or the REFRIGERATOR will get the FLOOR DIRTY.
- To clean the OVEN, it is necessary to APPLY OVEN CLEANER and then to REMOVE THE CLEANER.
- Before the FLOOR can be WASHED, it must be SWEPT.
- Before the FLOOR can be SWEPT, the GARBAGE must be taken out.
- Cleaning the REFRIGERATOR generates GARBAGE and messes up the COUNTERS
- WASHING the COUNTERS or the FLOOR gets the SINK DIRTY.

3b. (3 pts) Write a description using your operators of a likely initial state of a kitchen in need of cleaning.

3c. (3 pts) Write a description of a desirable goal state.

3d. (4 pts) Describe a possible threat and threat resolution

Decision Trees. (9 points) For the following questions, your answers should fit in a few sentences, possibly one. Scenario: Imagine you grow a very large, complex decision tree from a training set which contains many features (attributes).

3e. (3 points) You find that the training set accuracy for your tree is very high, but the test set accuracy (as measured on held out data) is very low. Explain why the accuracy on the training data could be so much higher than on the test data.

3f. (3 points) You decide to use pruning to create a series of successively smaller subtrees from your full tree (e.g. χ-squared pruning, truncating the tree at smaller depths, using a pruning set). As you prune more and more of the tree, the test accuracy goes up at first, but then starts to drop again after extensive pruning. Explain the rise and then fall of test accuracy in terms of specific terminology and concepts discussed in the decision tree literature.

3g. (3 points) Why is it important to use accuracy on held out data rather than accuracy on training data in order to decide when to stop pruning a decision tree?

C4 Artificial Intelligence (25 points)

Game Playing. (18 points) Consider the game of 2×2 tic tac toe where each player has the additional option of *passing* (i.e., marking no square). Assume X goes first.

4a.(6 points) Draw the full game tree down to depth 2. You need not show nodes that are rotations or reflections of siblings already shown. (Your tree should have five leaves.)

4b. (4 points) Suppose the evaluation function is the number of Xs minus the number of Os. Mark the values of all leaves and internal nodes.

4c. (4 points) Circle any node that would not be evaluated by alpha–beta during a left-to-right exploration of your tree.

4d. (4 points) Suppose we wanted to solve the game to find the optimal move (i.e., no depth limit). Explain why alpha–beta with an appropriate move ordering would be much better than minimax.

Decision Trees. (7 points)
Consider decision trees with continuous input attributes $A_1, \ldots, A_n$ and a Boolean output attribute Y. In such trees, the test at each internal node is an inequality of the form $A_i > c$, where c, the split point, may be any real number (to be chosen by the learning algorithm). The value at each leaf is true or false. In a *test-once* tree, each attribute may be tested at most once on any path in the tree; in a *test-many* tree, each attribute may be tested more than once on a path.

4e. (4 points) Can every non-noisy training set be classified correctly by a test-**many** tree? Why or why not?

4f. (3 points) Can every non-noisy training set be classified correctly by a test-**once** tree? Why or why not?