Grad. Comp. Exam: Artificial Intelligence (Spring 2014) Student ID: ____________

Answer all questions on the exam. You may use the back for additional space. Total: 100 points. Good Luck.

1. (25 pts) Search
   (a) What is the difference between Hill Climbing and Greedy Best-First Search?
   (b) Design an example to illustrate that Hill Climbing might not yield the optimal solution.
   (c) Design an example to illustrate that Greedy Best-First Search might not yield the optimal solution.

2. (25 pts) Learning

   (a) Consider only boolean attributes, explain why a decision tree can represent all possible boolean functions.

   (b) Consider n boolean attributes, derive the number of possible boolean functions a decision tree can represent. Explain your answer.

   (c) Consider the decision tree learning (DTL; aka ID3) algorithm is given as input all four possible examples for the $A \wedge B$ boolean function in this table:

example	A	B	$A \wedge B$
x_1	t	t	t
x_2	t	f	f
x_3	f	t	f
x_4	f	f	f

   i. Consider the following decision trees representing $A \wedge B$ (at each tree node, the left branch represents true and the right branch represents false):

```
      A              A              B              B
     / \            / \            / \            / \
    B   B          B   f          A   A          A   f
   / \ / \        / \            / \ / \        / \
  t  f f  f      t   f          t  f f  f      t   f

   Tree 1         Tree 2         Tree 3         Tree 4
```

   Which of the given decision tree(s), if any, can be the output of the DTL algorithm? Explain your answer.

   ii. For learning a correct decision tree for the $A \wedge B$ function, which example(s) from the table above, if any, may be removed from the input? Explain your anser.

3. (25 pts) Constraint Satisfaction Problems

   (a) Define the constraint satisfaction problem. Specify and give examples for all its components.

   (b) Describe a general search algorithm that can be applied to solving CSPs.

   (c) CSPs are frequently used for modelling the problem of scheduling classes. How would you model as a CSP the problem of scheduling 3 classes (C1, C2, C3) in 2 classrooms such that each class has to meet two times a day and each class-room is available for 2 slots a day. Two classes (C1, C2) are taught by the same professor, and C1 is a prerequisite of C3.

4. (25 pts) Alice (A) and Bob (B) witness on each other's identity. Let A, B be the Boolean random variables representing these hidden states, and let A_B, B_A be the evidence random variables describing the actual witness by Alice for the identity of Bob, the one by Bob for the identity of Alice. Independence statements for this problem are:

- A and B are independent.
- A_B and B_A are independent given A and B.

Assume that $P(A) = 0.8$, $P(B) = 0.5$, while $P(A_B|A, B)$ and $P(B_A|A, B)$ for the various combinations of values for A and B are given in the following tables:

A	B	$P(A_B\|A, B)$
T	T	0.9
T	F	0.1
F	T	0.5
F	F	0.5

A	B	$P(B_A\|A, B)$
T	T	0.9
T	F	0.5
F	T	0.1
F	F	0.5

(a) What is the difference between absolute independence and conditional independence of random variables?

(b) How do you write with probabilities the two independence statements in the problem description above?

(c) What is the chain rule, and where is it useful?

(d) What is a full joint probability distribution and why is it important?

(e) Compute the full joint probability for the problem above.

(f) Compute $P(B|A, A_B, \neg B_A)$