Name:____

1	2	3	4	5	6	7	total
20	*15*	*15*	*10*	*10*	*20*	*10*	*100*

UMBC CMSC 671 Final Exam

13 December 2012

Please write all of your answers on this exam. The exam is closed book and has seven problems that add up to 100 points. You have the two hours to work on this exam. Good luck.

1. True/False (20 points)

T F To apply resolution, you must convert your formula to a CNF. False

T F There are eight models for the sentence(a ∨ b ∨ c) among all of the interpretations over the three Boolean variables? False

T F A refutation proof proves that a sentence is entailed by a knowledge base by adding its negation to the KB and deriving a contradiction. True

T F Abduction is a sound form of reasoning but not necessarily complete. False

T F In first order logic, variables can only range over objects and not over relations, predicates or functions. True

T F GRAPHPlan works by transforming a planning problem into a heuristic search problem. False

T F SATPlan works by transforming a planning problem into a constraint satisfaction problem. True

T F The situation calculus is an approach to representing the effects of actions in logic. True

T F The STRIPS planning framework represents actions as constraints. False

T F Random variables A and B are independent if and only if p(A^B) = p(A)*p(B). True

T F If random variables A and B are independent, p(A v B) = p(A) + p(B). False

T F A well-formed Bayesian Belief Network, cannot contain cycles. True

T F Bayes' rule can be used to relate the probability of a cause given it symptoms to the probability of its symptoms give a cause. True

T F In a Bayesian belief network, two nodes are independent if and only if there is no undeirected path path between them. False

T F Overfitting occurs when a machine learning model describes random error or noise instead of the underlying relationships. True

T F SVMs require that each feature have a finite number of values. False

T F While SVMs find a linear combination of features to separate positive and negative examples of a class, the use of the "kernel trick" can introduce non-linear aspects. True

T F Every decision tree can be represented as a set of rules. True

T F The ID3 decision tree learning algorithm finds an optimal decision tree. False

T F Pruning nodes from a decision tree may have no effect on the resulting classifier in practice. True

2. Propositional logic (15 points)

A political pundit says *"A person who is radical (R) is electable (E) if she is conservative (C) and not electable if she is not conservative."*

2.1 Which of the following propositional sentences correctly encode this statement.

yes	no	sentence
	X	(R ^ E) => C
X		R => (E <=> C)
	X	R => ((C => E) v ~E)

2.2 Express each as a set of clauses in conjunctive normal form (e.g., a set of sentences where each is a disjunction of positive or negative atoms).

sentence	clauses
(R ^ E) => C	**~R v ~E v C**
R => (E <=> C)	**~R v ~E v C** **~R v E v ~C**
R => ((C => E) v ~E)	**True** (since it's equivalent to Equivalent to ~R v ~C v E v ~E which reduces to ~R v ~C v True)

2.3 Which of the three can be expressed as a set of horn clauses?

sentence	expressible as horn clauses?
(R ^ E) => C	**yes**
R => (E <=> C)	**yes**
R => ((C => E) v ~E)	**yes**

3. Resolution in FOL (15 points)

Given the following two FOL sentences in conjunctive normal form

C1	~republican(mother(x)) v republican(x)
C2	~republican(y) v likes(y, Romney) v ~livesIn(y, Utah)

2.1 Transform each into a sentence in implicative normal form, i.e., a sentence that is an implication whose left argument is a conjunction of positive literals and whose right argument is a disjunction of positive literals (e.g., p ^ q ^ r => s v t).

C1	republican(mother(x)) => republican(x)
C2	republican(y) ^ livesIn(y, Utah) => likes(y, Romney)

2.2 Give an English paraphrase for each, assuming natural interpretations of the logical predicates republican, livesIn and likes and the function mother.

C1	A person is a republican if their mother is a republican.
C2	A person who is a republican and lives in Utah likes Romney.

2.3 How many ways can C1 and C2 be resolved to produce a new clause C3?

enter a number between 0 and 4 1

For each, give the clause that results, showing it both in CNF and INF.

$C3_1$	~republican(mother(y)) v likes(y, Romney) v ~livesIn(Y, Utah)
	republican(x) ^ livesIn(x, Utah) => likes(x, Romney)
$C3_2$	
$C3_3$	
$C3_4$	

4. Planning (10 points)

4.1 Briefly describe the differences and similarities between problem solving and planning.

Both problem solver and planner are concerned with getting from a start state to a goal using a set of defined operations or actions, typically in a deterministic, discrete, observable environment.

In planning, however, we open up the representation of states, goals, and plans, which allows for a wider variety of algorithms that decompose the search space, search forwards or backwards, and use automated generation of heuristic functions.

Thus the search can proceeds through a space of possible plans rather than possible states. This allows subgoals to be planned independently, reducing the complexity of the planning problem

4.2 Describe the differences between a simple linear planner like STRIPS and non-linear partial-order planning. Mentions advantages of each approach.

A linear planner (e.g., STRIPS) solves a problem with a list of goals by finding a sub-plan to achieve each goal goal sequentially and concatenates the sub-plans to produce a single linear sequence of actions. Advantages are that the approach is simple and can be very efficient if the sub-plans for achieving the goals do not interact with one another. Disadvantages include the possibility that the plan found will be suboptimal if the sub-plans interfere with one another and that, while a linear planner is sound (it only finds valid plans), it is incomplete (there are some planning problems that have solutions that will not be found).

A non-linear, partial order planner treats a goal as a set of conditions and can explore the space of plans by interleaving work on each goal in the set. The planner can thus work forward from the initial state and also backward from the desired final state. Moreover, the plan can be represented as a partial ordering of actions. Advantages are that partial order planning can be complete, it allows for algorithms that can produce an optimal (i.e., shorted length) plan, and that the plan representations can be more flexible since they need not be reduced to a total ordering. Disadvantages are the larger size of the search space and the additional complexity of the algorithms. For simple problem domains, a straightforward linear planner may be preferable.

5. Using Bayes' rule (10 points)

Use Bayes' rule to compute the probabilities that, given a patient with a bad cough, (a) the patient has lung cancer, (b) the patient has pneumonia. Show your work!

L = Lung cancer
N = Pneumonia
C = Bad cough

p(C)	0.5
p(L)	0.0001
p(N)	0.01
p(C | L)	0.9
p(C | N)	0.9

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}.$$

Bayes' Rule

p(L | C)	**0.00018**
P(N | C)	**0.018**

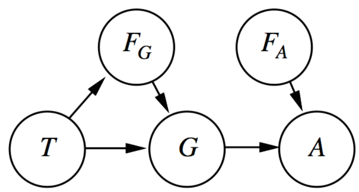
(a) p(L|C) = p(C|L) * p(L) / p(C) = 0.9 * 0.0001 / 0.5 = 0.00018

(b) p(N|C) = p(C|N) * p(N) / p(C) = 0.9 * 0.01 / 0.5 = 0.018

6. Bayesian Belief Networks (20 points)

UMBC's nuclear power plant has a gauge that measures the temperature of the core and an alarm that senses when the gauge reading exceeds a safe threshold. Consider the Boolean variables A (alarm sounds), Fa (alarm is faulty), and Fg (gauge is faulty) and the multivalued nodes G (gauge reading: normal or high) and T (actual core temperature: normal or high).

6.1 Draw a Bayesian network for this domain, given that the gauge is more likely to fail when the core temperature gets too high.



6.2 Is your network a polytree, i.e., a network in which there is at most one undirected path between any two nodes. Why or why not?

The network is not a polytree since the sub-network {T,Fg,G} there are multiple paths between, for example, T and G.

6.3 Assume that the probability that the gauge gives the correct temperature is *x* when it is working (Fg=0) and *y* when it is not (Fg=1). Give the conditional probability table associated with G.

	T=normal		T=high	
	Fg=1	Fg=0	Fg=1	Fg=0
G=normal	**y**	**x**	**1-y**	**1-x**
G=high	**1-y**	**1-x**	**y**	**x**

6.4 Assume the alarm works correctly unless it is faulty, in which case it never sounds. Give the conditional probability table associated with A.

	G=normal		G=high	
	Fa=1	Fa=0	Fa=1	Fa=0
A=1	**0**	**0**	**0**	**1**
A=0	**1**	**1**	**1**	**0**

7. Decision trees (10 points)

Consider the following data set with three binary input attributes (A_1, A_2, and A_3) and one binary output, y.

instance	A_1	A_2	A_3	y
1	1	0	0	0
2	1	0	1	0
3	0	1	0	0
4	1	1	1	1
5	1	1	0	1

Construct a decision tree from this data using the ID3 algorithm that selects the variable at each level that maximizes the information gained.

7.1 What is the variable at the root of the tree.

A_2

7.2 Show the entire decision tree.