NATIONAL UNIVERSITY OF SINGAPORE

SCHOOL OF COMPUTING

EXAMINATION FOR
Semester 2 AY2011/2012

CS3243 — INTRODUCTION TO ARTIFICIAL INTELLIGENCE

Apr / May 2012 Time Allowed: 2 Hours

INSTRUCTIONS TO CANDIDATES

1. This examination paper consists of **SEVEN (7)** questions and comprises **NINE(9)** printed pages, including this page.
2. Answer **ALL** questions.
3. Write your answers within the space provided in this booklet. You may use pen or pencil to write your answers.
4. This is an **OPEN BOOK** exam. The maximum mark is **100**.
5. Non-programmable calculators are allowed, but not electronic dictionaries, laptops, PDAs, or other computing devices.
6. Do not look at the questions until you are told to do so.
7. Please write your **Matriculation number** below.

This portion is for examiner's use only.

Question	Marks	Remarks
Q1		
Q2		
Q3		
Q4		
Q5		
Q6		
Q7		
Total		

1. **[15 marks]** A best-first search is performed on a search tree with the following weighted evaluation function

$$f(n) = (1 - \omega)g(n) + \omega h(n)$$

where $0 \leq \omega \leq 1$ and $h(n)$ is an admissible function.

(a) Comment on the optimality of search in each of the following cases:

- $\omega = 0$
- $\omega = 0.5$
- $\omega = 1$

(b) What range of values of ω would ensure that best-first search using the weighted evaluation function is optimal? Explain your answer.

2. [**10 marks**] Given the following formulation of a constraint satisfaction problem:

Variables: A, B, C

Domains: $D_A = D_B = D_C = \{0, 1, 2, 3, 4\}$

Constraints:

$$\begin{aligned} C_{AB} &: A = B + 1 \\ C_{BC} &: B = 2 \times C \end{aligned}$$

(a) Construct a constraint graph for this problem.

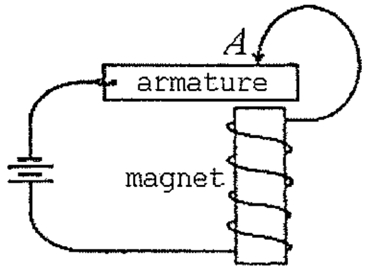
(b) Show a trace of the AC-3 algorithm on this problem. Assume that initially, the arcs in queue are in the order $\{(A, B), (B, A), (B, C), (C, B)\}$. What are the final domain values of D_A, D_B and D_C?

3. **[10 marks]** Given the following passage in the blocks world domain.

> A block x that is on top of another block y is said to be above the block y. If block x is above block y and y is above block z, then x is above z. Now, block B is on top of block A and A is on top of the table.

Prove using resolution that block B is above the table.

4. **[10 marks]** The following diagram shows a buzzer that is rigged such that when the armature makes contact at point A, current flows and causes the magnet to be activated, causing the armature to be drawn away from A thereby breaking the circuit. The loss of current deactivates the magnet causing the armature to return to point A thereby allowing current to flow and the cycle repeats.



The following events can be observed.

- *If contact is made at A, then magnet is activated;*
- *If magnet is activated, then contact at A is broken;*
- *If contact at A is broken, then magnet is deactivated;*
- *if magnet is deactivated, then contact is made at A;*

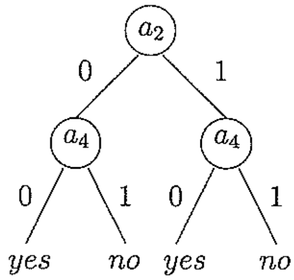
If we represent the above as predicate statements, what is the problem faced during the process of logical inference? Why did it happen? Suggest an alternative inference mechanism to tackle this problem.

5. [**25 marks**] You are given the following training data set.

id	a_1	a_2	a_3	a_4	*class*
1	1	1	1	1	*yes*
2	1	0	0	0	*yes*
3	1	1	0	1	*yes*
4	1	0	1	0	*yes*
5	0	1	1	0	*no*
6	0	1	0	1	*yes*
7	0	0	0	1	*no*

(a) Construct the complete decision tree using the information theoretic measure of attribute selection. You need only explicitly work out the information gain at the root of the decision tree. (Note: $\log_2 \frac{x}{y} = \log_2 x - \log_2 y$ and $\log_2 1 = 0$, $\log_2 2 = 1$, $\log_2 3 = 1.585$, $\log_2 4 = 2$, $\log_2 5 = 2.322$, $\log_2 6 = 2.585$, $\log_2 7 = 2.807$)

(b) Compare the tree constructed in question 5a with the one below.



Comment on the depth of the tree generated using the information theoretic measure. Does the information theoretic measure guarantee the generation of the shallowest decision tree?

(c) Suppose we consider the *id* as another attribute when selecting the split at the root of the decision tree. What is the information gain, *Gain*(*id*)? Would you choose this attribute over all other attributes? Why? Briefly suggest how the information gain criteria can be improved.

6. [**15 marks**] You are given the following training data set.

size	color	shape	class
med	blue	brick	yes
small	red	wedge	no
small	red	sphere	yes
large	red	wedge	no
large	green	pillar	yes
large	red	pillar	no
large	green	sphere	yes

(a) Using the data above, fill in the prior and posterior probabilities in the table below. Leave your answers as fractions.

	yes	*no*
$P(C_i)$		
$P(small\|C_i)$		
$P(blue\|C_i)$		
$P(brick\|C_i)$		

(b) Given an unseen instance $< small, blue, brick >$, deduce its classification using the probabilistic values above. Assume that the attributes are conditionally independent given the class value.

7. **[15 marks]** A three-valued logic system is one in which a value can signify absolute truth (1), absolute falsity (-1) or indeterminism (0). We define a **priority** function over a three-valued logic system consisting of n inputs $in_1, in_2, \ldots, in_n$ such that the output $out = in_k$ if $in_k \neq 0$ and all preceding inputs $in_j = 0,\ j < k$. An example of the **priority** function over 2 inputs is given below.

in_1	in_2	out
1	1	1
1	-1	1
1	0	1
-1	1	-1
-1	-1	-1
-1	0	-1
0	1	1
0	-1	-1
0	0	0

Notice that in_1 has "priority" over in_2. Represent the **priority** function using a **single-layer perceptron** over n input nodes and 1 output node. Justify your choice of the weights and activation function.

— END OF PAPER —