UNIVERSITI SAINS MALAYSIA

April 2011 Semester II Examination
Academic Session 2010/2011

April/May 2011

EEE 551 – INTELLIGENT SYSTEMS

Time : 3 hours

INSTRUCTION TO CANDIDATE:

Please ensure that this examination paper contains **ELEVEN** printed pages and
SIX questions before answering.

Answer **FIVE** questions.

Distribution of marks for each question is given accordingly.

All questions must be answered in English.

1. (a) By using a suitable diagram, explain the production system and basic structure of a rule-based expert system.
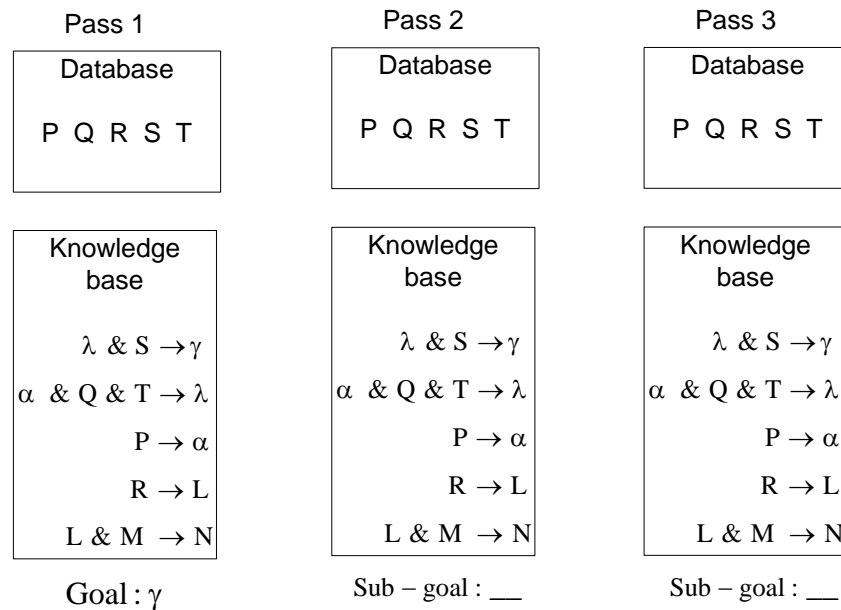
    (20 marks)

    (b) (i) Describe the forward and backward chaining inference process. Give example for each inference process.

    (20 marks)

    (ii) Consider the facts and rules given in the Figure 1 below. Describe the backward chaining process by completing the facts in each database (if any) and adding the arrows to show the matching and firing process until the Goal γ is proved.

    Then, by modifying the rules and facts given in Figure 1, design an expert system for diagnostic purpose.
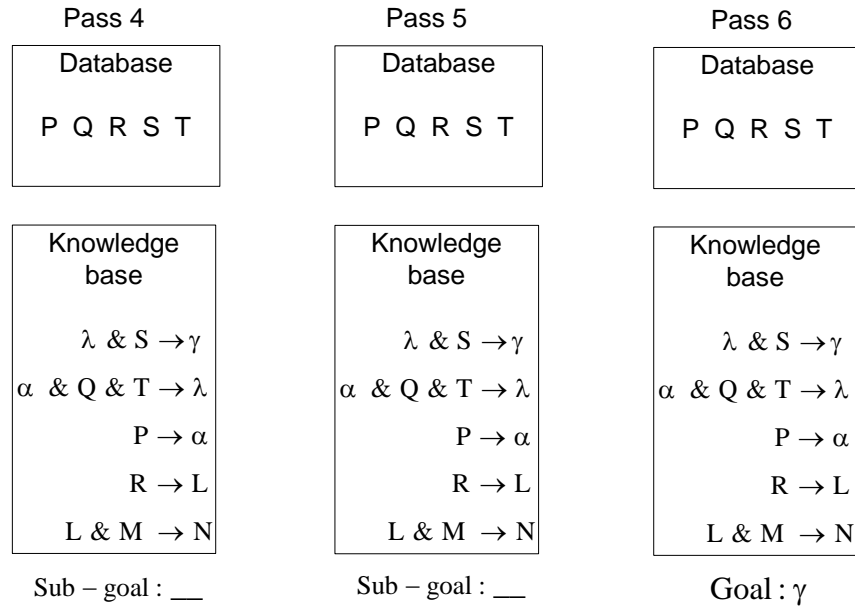
    (40 marks)

Pass 1	Pass 2	Pass 3
Database	Database	Database
P Q R S T	P Q R S T	P Q R S T
Knowledge base	Knowledge base	Knowledge base
λ & S $\to \gamma$	λ & S $\to \gamma$	λ & S $\to \gamma$
α & Q & T $\to \lambda$	α & Q & T $\to \lambda$	α & Q & T $\to \lambda$
P $\to \alpha$	P $\to \alpha$	P $\to \alpha$
R $\to$ L	R $\to$ L	R $\to$ L
L & M $\to$ N	L & M $\to$ N	L & M $\to$ N
Goal : γ	Sub – goal : ___	Sub – goal : ___

Pass 4	Pass 5	Pass 6
Database	Database	Database
P Q R S T	P Q R S T	P Q R S T

Knowledge base	Knowledge base	Knowledge base
λ & S $\rightarrow \gamma$	λ & S $\rightarrow \gamma$	λ & S $\rightarrow \gamma$
α & Q & T $\rightarrow \lambda$	α & Q & T $\rightarrow \lambda$	α & Q & T $\rightarrow \lambda$
P $\rightarrow \alpha$	P $\rightarrow \alpha$	P $\rightarrow \alpha$
R $\rightarrow$ L	R $\rightarrow$ L	R $\rightarrow$ L
L & M $\rightarrow$ N	L & M $\rightarrow$ N	L & M $\rightarrow$ N
Sub – goal : __	Sub – goal : __	Goal : γ

Figure 1 : Backward Chaining

(c)　List the advantages and disadvantages of rule-based expert systems.

(20 marks)

2.　(a)　Refer to Figure 2 below. If the occurrence of event A depends on only two mutually exclusive events, i.e. *B* and NOT *B*, the Bayesian rule can be simply written as

$$P(A|B) = \frac{p(B|A) \text{ x } P(A)}{P(B|A) \text{ x } P(A) + P(B|\rightarrow A) \text{ x } P(\rightarrow A)}$$
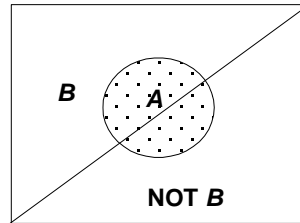
Figure 2 : The Joint Probability

From this basic probability theory, explain the Bayesian reasoning to manage uncertainty in expert system.   Suppose all rules in the knowledge base are represented as follows:

> IF E is true
> THEN H is true [with probability P]

Where,   H denotes a hypothesis
              E denotes evidence to support this hypothesis.

(30 marks)

(b)   What is the difference between a crisp set and a fuzzy set. Represent crisp and fuzzy sets of low, medium and high temperature using mathematical notations and computer representation.   You can create your own variables and membership functions to explain your answer.

(20 marks)

(c)   Two types of biometric information from speech signal and face modalities are fused in order to maintain the biometric system performance (%) under adverse conditions.

Speech signal quality is measured based on signal to noise ratio (SNR) while face quality depends on quality density (Q) imposed during data collection.

Based on the following trapezoidal/triangular membership functions and rules in Table 1, calculate the biometric system performance (%) if the current speech signal quality and face quality density are 35 SNR and 20 Q, respectively.

Please use max method for OR, min method for AND, maximum for aggregation and centre of gravity (COG) for defuzzification.

(50 marks)

Table 1 : Trapezoidal/triangular membership functions and rules

Speech Signal (in SNR) Range [0, 50]

Low	{5, 1	20, 0}	
Medium	{10, 0	25, 1	40, 0}
High	{30, 0	40, 1}	

Face Quality Density (in Q) Range [0, 50]

In adequate	{10, 1	35, 0}
Adequate	{15, 0	40, 1}

System Performance (in %) Range [0, 100]

Bad	{25, 1	40, 0}	
Average	{25, 0	50, 1	75, 0}
Excellent	{60, 0	75, 1}	

<u>Rules</u>

IF	speech signal quality is low
OR	face quality density is inadequate
THEN	performance is bad

IF	speech signal quality is medium
AND	face quality density is adequate
THEN	performance is average

IF	speech signal quality is high
OR	face quality density is adequate
THEN	performance is excellent

----

3.   (a)   (i)   Genetic algorithm and genetic programming are two evolutionary computing methods. Describe the main operations that are similar for both methods. Discuss a suitable application area for each method.

(20 marks)

(ii)   Describe the chromosome structure of genetic programming. Discuss the special features of the chromosome structure of genetic programming in comparison with that of a genetic algorithm

(20 marks)

(iii)   One of the selection functions in genetic programming is greedy over-selection, which is used in very large populations. Explain this selection function.

(10 marks)

(b) Figure 3 shows two tree structures of genetic programming.
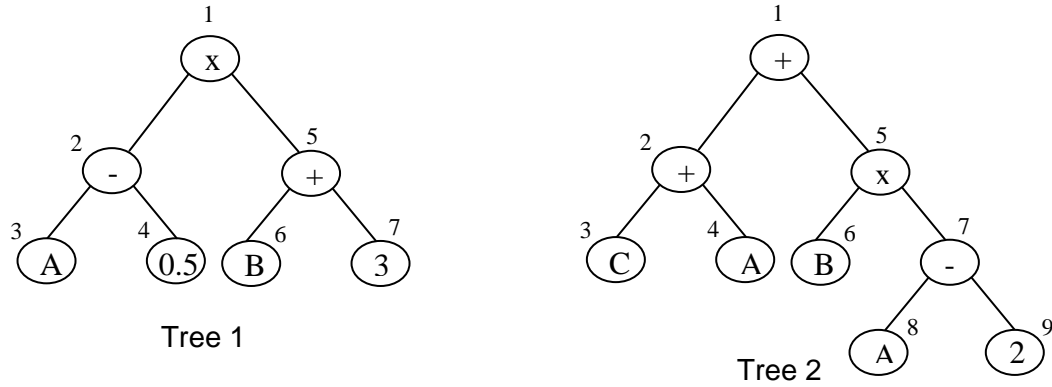

Tree 1

Tree 2

Figure 3

(i) Determine the terminal and function sets of Tree 1 and Tree 2.

(10 marks)

(ii) Determine the formulae of Tree 1 and Tree 2.

(10 marks)

(iii) Assume that crossover between Tree 1 (subtree at point 2) and Tree 2 (subtree at point 5) occurs and produces offspring 1 and offspring 2. Draw the new tree structures of offspring 1 and offspring 2, and determine the formulae of offspring 1 and offspring 2.

(30 marks)

4. (a) (i) Explain two main differences between a genetic algorithm and the traditional search and optimization methods.

(10 marks)

(ii) By using a suitable flow chart, discuss the main steps of a genetic algorithm.

(10 marks)

(iii) Describe four possible termination criteria of a genetic algorithm.

(10 marks)

(iv) Biological terms of allele, locus, genotype, and phenotype are commonly used in a genetic algorithm. Explain the meaning of these terms in a genetic algorithm.

(10 marks)

(v) By using suitable examples, discuss the operations of crossover and mutation of a genetic algorithm.

(20 marks)

(b) A genetic algorithm is used to solve the problem of maximizing a mathematical function, i.e. $f(x) = x^3$, between the integer interval of [0, 63]. A six-bit binary scheme is used to represent the strings of the genetic algorithm for the numbers from 0 to 63. Table 4 shows a randomly generated population of the genetic algorithm.

(i) Determine the x value of each string, as in column 3 of Table 4.

(10 marks)

(ii) Assume that the fitness function is the same as the function to be maximized. Determine the fitness value of each string, as in column 4 of Table 4.

(10 marks)

(iii) Assume that the Roulette Wheel selection method is used for reproduction. Determine the probability of selection (*p_select*) for each string, as in column 5 of Table 4.

Draw a biased Roulette Wheel that can be used for reproduction of the genetic algorithm.

(20 marks)

Table 4

String No	Initial population	x	$f(x)$	p_select
1	1 0 1 1 0 1			
2	0 1 1 0 0 0			
3	1 0 1 0 0 0			
4	0 1 0 1 0 0			
5	1 1 0 1 1 1			
6	0 1 0 0 1 1			

5. Answer all the following questions:

(a) What is the objective of training an artificial neural network? (10 marks)

(b) Explain the term *overfitting* in artificial neural network training process, i.e. what is it, how it occurs and how to prevent it.

(20 marks)

(c) Explain how the problem of ANN *overtraining* and immaturity can be eliminated.

(10 marks)

(d) What is meant by *generalization* in artificial neural network? Why is it important and what are the factors that need proper consideration in order to achieve *generalization*?

(20 marks)

(e)   Explain the problem of *linearly non-separable* in an artificial neural network. Use a diagram to illustrate the phenomenon. Name an artificial neural network model which suffers from the problem and an elementary logical function that make a good example for demonstrating the problem. Finally, explain how the problem can be eliminated.

(40 marks)

6.   Answer the following questions:

(a)   Differentiate a Self-Organizing Feature Map from a Hopfield neural network in terms of their:
        (i)   architecture,
        (ii)   learning scheme, and
        (iii)   speed, memory capacity and efficiency of learning   (10 marks)

(b)   A company requests you to intelligently solve a failure detection problem of a control system given by the following conditions:

Machine Conditions				Machine State
Temperature, T > 100ºC	Pressure, P > 100kPa	Liquid Level, L > 1 m	Power Dissipation, D > 1000kW	
No	No	No	No	Good
No	No	No	Yes	Good
Yes	No	Yes	No	Fail
No	No	Yes	No	Good
No	No	Yes	Yes	Fail
No	Yes	No	No	Good
No	Yes	No	Yes	Fail
Yes	No	No	No	Good
Yes	Yes	Yes	Yes	Fail

You decide to use a Multi-Layer Perceptron (MLP) neural network to solve the problem.

(i)   What would be the machine's state if: T=Yes, P=Yes, L=No, D=Yes?

(5 marks)

(ii)   Draw a schematic diagram of the MLP neural network for the above problem assuming that the ANN has two hidden nodes. Completely label all the input and output nodes, biases and their connection links.

(35 marks)

(iii)   Suppose you decide to use the value 0.5 for all initial weights and 1.0 for all biases. Calculate the MLP output after the first iteration assuming that the hidden and output neurons employ the hyperbolic tangent and logarithmic activation functions, respectively.

(50 marks)

ooo0ooo

EEE 551/4 – INTELLIGENT SYSTEMS

Errata Sheet

Question 6(b)

(iii)   Suppose you decide to use the value 0.5 for all initial weights and 1.0 for all biases. Calculate the MLP output **for the inputs in 6(b)(i)** <u>after the first iteration</u> assuming that **No=0 and Yes=1**, and the hidden and output neurons employ the hyperbolic tangent and logarithmic activation functions, respectively.

Given:

Hyperbolic tangent = $\dfrac{e^x - e^{-x}}{e^x + e^x}$ and   Logarithmic function = $\dfrac{1}{1+e^x}$