Ollscoil na hÉireann, Gaillimh GX 1522
National University of Ireland, Galway

Semester II Examinations, 2003/2004

Exam Code(s)	4IF1, 4BS2, 1MF2
Exam(s)	B.Sc. in Information Technology B.Sc. (Hons.) M.Sc. Software Design & Development (Stream II)
Module Code(s)	CT433
Module(s)	Advanced Studies in Information Technology
Paper No.	1
Repeat Paper	N Special Paper N
External Examiner(s)	Prof. P. Nixon
Internal Examiner(s)	Prof. G. Lyons, Mr S. Hill, Dr M. Schukat, Mr C. O'Riordan, Dr M. Madden

Instructions:

Answer questions from TWO sections only:
- Attempt Section A or Section B
- Attempt Section C or Section D

Use a separate answer book for each section.
In Section A, answer Q1 and one other question.
In Sections B, C or D, answer any two questions.
All questions carry equal marks.

Duration	3 hrs
No. of Answer books	2

Requirements:

Handout	
MCQ	
Statistical Tables	
Graph Paper	
Log Graph Paper	
Other Material	
No. of Pages	6
Department(s)	Information Technology

Section A: Evolutionary Computation

NOTE: In this Section, answer Question 1 and one other question

Q1

i) Given a clearly defined problem to be solved and a symbol string representation for candidate solutions, outline briefly how a simple GA works. (5)

ii) List the essential characteristics of Evolution Strategies. (5)

iii) Briefly outline a mutation operator most widely associated with a "meta-EP". (5)

iv) Outline two ways in which a classifier system can learn. (5)

v) Briefly describe the three steps which Genetic Programs follow to breed computer programs in order to solve problems. (5)

Q2

Choice of representation in GAs influences the types of crossover and mutation operators used, discuss. (25)

Q3

Outline the similarities and differences that evolutionary programming and evolution strategies have towards self-adaptation. (25)

Section B: Compiler Theory and Compiler Design

Q4

a) Distinguish between the terms *compiling, cross-compiling* and *bootstrapping.* (6)

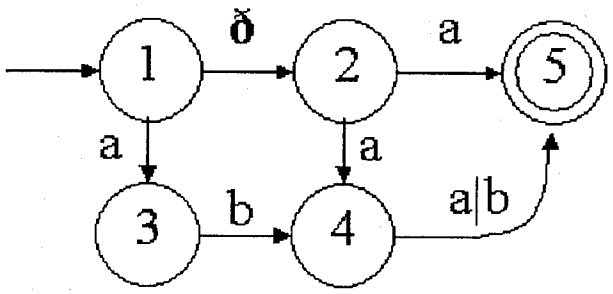
b) Explain in some detail the conceptual structure of a compiler and the different steps that are performed to translate a source code into instructions for a target microprocessor. (10)

c) Distinguish between *regular expressions* and *context free grammars* and show in which context both are typically used. (9)

Q5

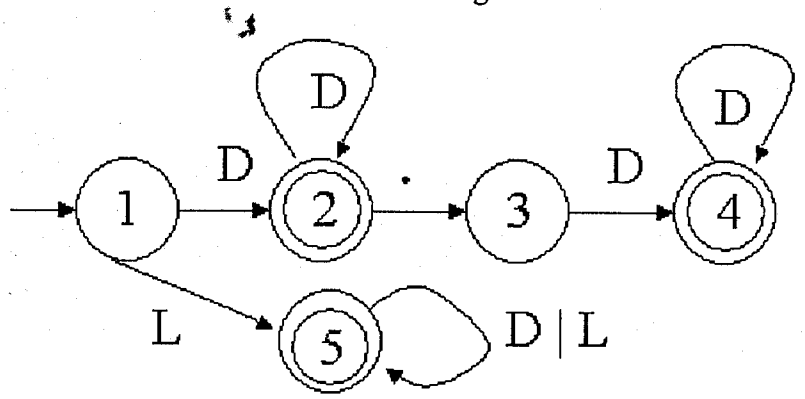
a) Describe in some detail how *finite automata (FA)* can be used to implement a scanner. Use an example to illustrate your answer. (8)

b) Distinguish between *deterministic finite automata (DFA)* and *non-deterministic finite automata (NDFA).* Show how the NDFA below can be converted into a DFA.



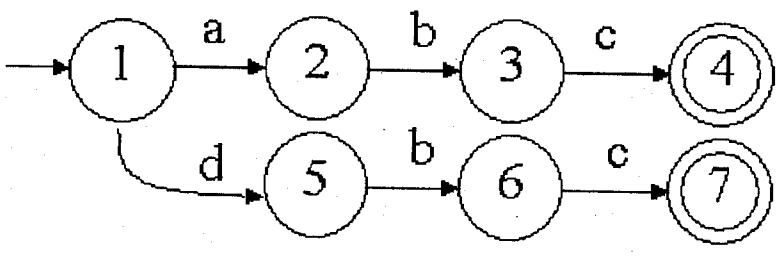
(7)

c) Construct the transition table for the following DFA:



(4)

d) Optimize the following DFA by reducing the number of required states:



(6)

Q6

a) Based on the grammar below (whereby terminal symbols are printed in **bold** letters),

(1) <Program> → **begin** <Stmts> **end $**
(2) <Stmts> → <Stmt> ; <Stmts>
(3) <Stmts> → ð(4)<Stmt> → **SimpleStmt**
(5) <Stmt> → **begin** <Stmts> **end**

perform a top-down parse of the following code fragment:

begin SimpleStmt ; end $ (8)

b) Calculate the *First(), Follow()* and *Predict()* sets of the grammar as described in part a) of this question. (7)

c) The following grammar describes assignment statements for simple parenthesized expressions (terminal symbols are printed in **bold** letters and are embedded in single quotes):

(1) <Header> → <Ident> '=' <Expression>
(2) <Ident> → **'A'** | **'B'** | **'C'** | **'D'** | **'E'** | **'F'** | **'G'** | **'H'** | **'I'** | **'J'**
(3) <Expression> → <Digit> | '(' <Expression> <Operator> <Expression> ')'
(4) <Operator> → '+' | '*'
(5) <Digit > → **'0'** | **'1'** | **'2'** | **'3'** | **'4'** | **'5'** | **'6'** | **'7'** | **'8'** | **'9'**

Describe a data type called AST_NODE, which allows syntactically correct assignment statements to be represented by an Abstract Syntax Tree (AST). Based on your node design draw the AST for the following code fragment:

B = ((3 * 4) + (6 * 7)) (10)

Section C: Multi-Agent Systems

Q7 (a) The concept of an agent is usually defined by listing the properties that agents exhibit. Identify and explain the properties you would associate with the concept of agents. (8)

(b) Discuss the properties of multi-agent systems and the advantages of adopting a multi-agent systems approach to designing and deploying multi-agent systems. (8)

(c) Explain the term *"Speech Act Theory"*. Discuss the role of this theory in the design of agent communication languages. (9)

Q8 (a) Discuss the advantages and limitations of adopting game-theoretic approaches to analysing, and reasoning about, agent interactions. (8)

(b) Define the *prisoner's dilemma* and the *iterated prisoner's dilemma*. Write a short note discussing scenarios where cooperation emerges as a rational choice for agents participating in the dilemma. (9)

(c) Trust and reputation schemes have been employed in multi-agents systems to help prevent exploitation of agents. Discuss the benefits and limitations of such an approach. (8)

Q9 (a) Auctions and auction theory have been used as a means to allow agents reach an agreement . Identify the main type of auctions that can be used, discussing their relative strengths and weaknesses. (10)

(b) Negotiation among a set of agents usually involves a *negotiation protocol* and *a negotiation object.* Explain with the use of a suitable example, the underlined terms. (6)

(c) What are the minimum requirements agents should have to facilitate negotiation? What additions are usually required to ensure that negotiation occurs more efficiently? (9)

Section D: Machine Learning

Q10 (a) Explain the Information Gain measure. Describe the ID3 algorithm, which uses Information Gain for decision tree induction. (7)

(b) "ROC curves allow comparison of classifiers independently of misclassification costs." Explain, with examples, the concept of unequal misclassification costs. Describe how ROC curves may be used to compare classifiers. (8)

(c) You have received the following email message from a colleague. Prepare a detailed response.

I have been working recently on a new classification algorithm that learns logic rules. I think it should be superior to decision tree learning, but I'd like advice on how to test this. What kind of performance measures should I use? What datasets should I use for the comparison and how do I use them to test whether my new algorithm is better? (10)

Q11 (a) Discuss what supervised, unsupervised and semi-supervised Machine Learning tasks all have in common and what the differences between them are. As part of your discussion, provide an example of each task and a technique that might be appropriate for it. (9)

(b) Explain the operation of the Q-learning algorithm for Reinforcement Learning, including an explanation of each of the learning parameters. (9)

(c) Explain what is meant by the "curse of dimensionality". Outline two approaches to dealing with it. (7)

Q12 (a) Describe the operation of the Naive Bayes classification algorithm. (6)

(b) What is meant by conditional independence of two variables? In a Bayesian network, how is conditional independence represented? (5)

(c) Bayesian networks can be used for both data exploration and classification. Explain this, using appropriate examples. (4)

(c) A biomedical scientist is developing a new test for cervical cancer. She has analysed tissue from 200 test samples, 100 with the disease and 100 without. In each case, she has measured 50 different attributes of the tissue. Because of the cost of measurements, she would like to identify a subset of 10 or fewer attributes that are good predictors of the disease, and has approached you for advice. What are your recommendations? (10)