

OLLSCOIL NA hÉIREANN
The National University Of Ireland

NATIONAL UNIVERSITY OF IRELAND, GALWAY

SUMMER EXAMINATIONS 2000

SECOND YEAR EXAMINATION IN COMPUTING STUDIES
[CS201 Paper II]

CS206

METHODOLOGY and INFORMATION SYSTEMS

Prof. D. Bell
Dr. G. Lyons
Ms. J. Griffith

Candidates are required to answer **FIVE** questions.
Two questions must be answered from each section
All questions carry equal marks.
Time allowed: **Three hours**

SECTION A

- Q. 1**
- i) Describe the main steps that should be followed when developing software. Distinguish between the waterfall life cycle model and the spiral life cycle model for developing software.
 - ii) Outline the main types of errors that can lead to a reduction in software quality. Comment on the usefulness of the following as an aid to testing:
 - meaningful variable names.
 - comments in source code.
 - documentation.
 - iii) Describe what is meant by a well-defined recursive procedure. Develop a recursive function `sum` which finds the sum of the first n integers.
(e.g. `sum(4)` returns 10 as $1 + 2 + 3 + 4 = 10$).
Trace through your code with `sum(4)`.

- Q. 2**
- i) Distinguish between stack and queue data structures, outlining algorithms for insertion and deletion in both data structures.
 - ii) Outline algorithms to search for an item in the following data structures where elements in both structures are sorted in non-decreasing order:
 - array
 - linked list
 - iii) Describe a Binary Search Tree data structure. Outline a search algorithm to search for an item in a Binary Search Tree. Compare the algorithm you develop with the two algorithms you developed in part ii).
- Q. 3**
- i) Discuss why compression techniques are required. Distinguish between lossy and lossless compression techniques detailing where each can be used.
 - ii) Outline metrics that can be used to judge the quality of compression algorithms. With the aid of an example describe run-length encoding, outlining its strengths and weaknesses.
 - iii) Huffman encoding is an example of a prefix encoding algorithm. Outline the steps involved in Huffman's algorithm for compression. Given the following characters and the associated frequency of their occurrence in a file, use Huffman's algorithm to produce codes for the characters:

S	A	E	R	T	O
4	15	25	5	8	16

- Q. 4**
- i) Insertion Sort and Selection Sort are two comparison-based techniques to sort a list of elements. Outline the steps involved in *either* insertion sort or selection sort. Include pseudocode.
 - ii) Divide and Conquer techniques are another approach to sorting.
 - a) Describe what is meant by a divide and conquer technique.
 - b) Describe any algorithm which uses a divide and conquer approach.
 - iii) Compare Count Sort with the algorithms developed in part i) and part ii) by discussing count sort under the following headings:
 - approach taken.
 - assumptions made.
 - efficiency in comparison to algorithms developed in part i) and ii).

SECTION B

- Q. 5** i) Outline the advantages and disadvantages of the sequential file organisation. Give details of approaches that can be adopted when inserting new records into a sorted sequential file.
- ii) What is meant by a hashed file organisation? With the aid of examples describe the following hash functions:
- folding method.
 - division method.
- iii) Explain why a collision resolution technique is required when using hashed file organisation. By inserting the data (Key value) given in the example below (with associated hash address) describe the operation of the following collision resolution techniques:
- linear probing.
 - chaining.

Key value:	A	B	C	D	E	X	Y	Z
Hash address:	3	7	1	10	3	10	4	0

(using an array of size 11, with locations from 0 to 10).

Outline the advantages and disadvantages of each technique.

- Q. 6** i) Distinguish between the traditional file processing approach and the database approach.
- ii) Why is a database approach not adequate for all information management needs? What is meant by the term *information overload*?
- iii) With the aid of a diagram describe the main components of an Information Retrieval system. How is the accuracy of an Information Retrieval system measured?
- Q. 7** i) Discuss criterion that can be used to classify Database Management Systems.
- ii) Describe the integrity constraints that are considered part of the relational data model and give examples of cases where the DBMS must check to ensure that such integrity constraints are not violated by update operations.

(Question 7 continued on next page)

- Q. 7 iii) Given the following relational schema and interpretation with keys underlined:

TOURS(Code, Type, Location, DepartDate, Duration, Cost)
LOCATION(LocationName, SummerTemp, WinterTemp, TimeZone)
CUSTOMER(Name, TourCode, Address, No_of_People)

A company offers a number of TOURS, where each tour has an associated, unique code (Code). Each tour is of a certain type (e.g. skiing, beach, hiking, etc.), and has a fixed departure date (DepartDate), duration (number of days of tour) and associated cost. Each tour takes place in a certain LOCATION which has an average summer and winter temperature and is located in a certain time zone. The details held on customers in the CUSTOMER relation are: name and address of customer, the code of the tour they have booked (TourCode) and the number of people booked by the named customer for that tour.

It is assumed that no customer will book the same tour twice.

Develop SQL queries to satisfy the following information needs:

- List the location and departure date of all skiing tours not longer than 3 days.
- List the customer name and number of people booked by that customer for a beach tour where the summer temperature is greater than 22°C.
- List the total number of people who are booked for hiking tours.
- List the name and address of all customers who have booked the same tour as "James Flynn".

- Q. 8 i) Outline the desirable properties of a transaction.
- ii) Describe what is meant by the *two-phase locking protocol*. Discuss the problem that occurs with the following transactions operating under the two-phase locking protocol. Detail any approach which can be used to deal with the situation.

T1	T2
write_lock(Z);	
write_item(Z);	
	write_lock(X)
	write_item(X);
write_lock(X);	
	write_lock(Z);

- iii) Discuss why recovery techniques are required. Describe the role the system log plays in an error recovery technique, discussing the main entries in the system log. Explain how each of the following recovery techniques operate: deferred update and immediate update.