



Spring Examinations 2012 / 2013

Exam Code(s)	3BCT1 3IF1
Exam(s)	3 rd B.Sc. Computer Science and Information Technology 3 rd B.Sc. (Information Technology)
Module Code(s)	CT332
Module(s)	Database Systems II
Paper No.	1
External Examiner(s)	Prof. Michael O'Boyle
Internal Examiner(s)	Prof. Gerard Lyons Dr. Michael Madden Dr. Colm O'Riordan*
<u>Instructions:</u>	Answer any 3 questions
Duration	3 hours
No. of Pages	4 including this one
Department(s)	Information Technology
Requirements	None

PTO

Q. 1.

- a) Given the following relation, R, and functional dependencies, F, decompose the relation such that the resulting relations satisfy BCNF.

- i. $R = \{A, B, C, D, E, F, G\}$
- ii. $F = \{ \{A, B, C\} \rightarrow \{D, E\}, \{B\} \rightarrow \{F\}, \{E\} \rightarrow \{G\}, \{D\} \rightarrow \{A\} \}$

(8)

- b) Explain, with reference to types of de-normalisation, when the process of de-normalisation should be used. (5)

- c) Explain, with the use of a suitable example, what is meant by a total, overlapping specialisation. Suggest a suitable mapping to relational schema for the model example you've provided. (10)

- d) Show how the following ER model constructs can be mapped to a suitable object-oriented database model. (10)

- i. Entities
- ii. Relationships
- iii. Attributes
- iv. Specialisations

Q. 2.

- a) Linear hashing and dynamic hashing are two approaches to hashing values to a dynamically changing file. Briefly outline either approach and illustrate the approach using the following record key values. You may assume a blocking factor of two.

120, 121, 408, 405, 407, 411, 102, 106

(13)

- b) Explain the structure of a B+ tree. Discuss the advantage of a B+ tree index over a hash index. Your answer should refer to the following types of queries: range queries, point queries. (8)

- c) Outline a suitable approach to implementing the join operator. Illustrate the algorithm with a simple example. Suggest an algorithm for a join operator given a parallel architecture where the data can be partitioned across multiple disks.

(12)

PTO

Q. 3.

- a) Explain, with examples, the importance of concurrency control in databases. (6)
- b) For the following schedule of transactions, show how the schedule would proceed under either a two phase locking protocol *or* a time-stamping protocol.

T1	T2	T3
read_item(X)		
read_item(Y)		
	read_item(Z)	
	read_item(Y)	
		read_item(X)
		read_item(Z)
write_item(X)		
write_item(Y)		
	write_item(Z)	
		write_item(X)

(12)

- c) What is meant by conflict-serializability? Choosing either time-stamping or two phase locking, show that the protocol guarantees conflict serializability.

(7)

- d) With respect to recovery, explain how the notion of a commit point of a transaction can be extended to function in a distributed database.

(8)

PTO

Q.4.

- a) Given the following fragment of a database schema:

EMPLOYEE (SSN, FName, LName, Salary, Address, Dno)

DEPARTMENT (Dno, Dname, Location, Description)

PROJECT (Pno, Pname, Budget, P_Location)

WORKS_ON (SSN, Pno, Hours)

Develop an SQL query to satisfy the following information need:

List all employees with last name "Smith" who have worked on a project located in "Galway" or who worked for a department based in "Galway".

Outline how heuristic optimisation can be used to improve the efficiency of a query. Illustrate the approach on the SQL query you developed. (15)

- b) The operators in the relational algebra can be implemented in logic databases. With reference to your query developed in part (a), show how the operators could be implemented in Datalog. (9)
- c) Assuming the relations EMPLOYEE and DEPARTMENT specified in part a) are stored at sites *s1* and *s2* respectively, and a query is run that involves returning the last name of all employees working for a department named "Testing", show how the semi-join operator may be used to evaluate the query. Specify which attributes and tuples are shipped at each stage. (9)