NUI Galway
OÉ Gaillimh

Autumn Examinations 2021-2022

Course Instance Code(s)	3BCT
Exam(s)	3rd year Computer Science and Information Technology
Module Code(s)	CT3532
Module(s)	Database Systems 2
Paper No.	1
External Examiner(s)	Dr. Ramona Trestian
Internal Examiner(s)	Professor M. Madden *Dr. C. O'Riordan

Instructions: Answer any 3 questions. All questions carry equal marks

Duration	2 hours
No. of Pages	4
Discipline(s)	Computer Science
Course Co-ordinator(s)	Dr. Colm O'Riordan

Requirements:

Release in Exam Venue	Yes
MCQ Answersheet	No
Handout	None
Statistical/ Log Tables	None
Cambridge Tables	None
Graph Paper	None
Log Graph Paper	None
Other Materials	None
Graphic material in colour	No

PTO

CT3532 Database Systems 2

Question 1 (25 marks)

a) Explain what is meant by term *minimal cover set* and explain the importance of *minimal cover sets* in database design. (10)

b) Explain, with suitable examples, what is meant by a *non-additive join property*. Outline how to test for such a property. (8)

c) When decomposing to BCNF, certain functional dependencies are no longer explicitly represented in the relational schema. Illustrate this with an example and outline approaches to deal with this issue. (7)

Question 2 (25 marks)

a) Outline, with examples, an algorithm for inserting values into a B+tree. (8)

b) Linear hashing is an approach to allow insertion and retrieval of records into a dynamically sized files. Outline an algorithm for retrieval of a record from the file given a key value. (8)

c) Suggest an approach jointly index more than one attribute. Outline the advantages and disadvantages of your approach. (9)

Question 3 (25 marks)

a) Consider the following schedule of three transactions:

T1	T2	T3
Read_item(X)		
Write_item(X)		
	Read_item(Y)	
	Write_item(Y)	
		Read_item(Z)
		Write_item(Z)
Read_item(Y)		
Read_item(Z)		
	Write_item(X)	
		Write_item(Y)

i) Show if this schedule exhibits any problems from a concurrency control perspective. (4)

ii) Show how the schedule would proceed under 2PL (you may assume shared and exclusive locks). (7)

iii) Show how the schedule would proceed under a time-stamping protocol.(7)

b) Explain what is meant by *conflict-serializability* and explain any mechanism to ensure that a schedule is conflict-serializable. (7)

Question 4 (25 marks)

a) Outline an efficient algorithm for performing a join between two relations. Specify the efficiency of your algorithm. (8)

b) Explain how the join algorithm described in (a) can be improved given a parallel architecture. Specify the improvement in efficiency. (8)

c) Explain how the GROUP BY and HAVING commands can be efficiently implemented. Your answer should outline an algorithm, give a simple example and estimate the efficiency of your approach. (9)