OLLSCOIL NA GAILLIMHE
UNIVERSITY OF GALWAY

Autumn Examinations 2022-2023

Course Instance Code(s)	3BCT1
Exam(s)	Third B.Sc. Computer Science and IT
Module Code(s)	CT3532
Module(s)	Database Systems 2
Paper No.	1

External Examiner(s) Dr. R. Trestian
Internal Examiner(s) Professor M. Madden
*Dr. C. O'Riordan

Instructions: Answer any 3 questions. All questions are equally weighted.

Duration	2 hours
No. of Pages	3
Discipline(s)	Computer Science
Course Co-ordinator(s)	Dr. C. O'Riordan

Requirements:

Release in Exam Venue	Yes
Handout	None
Statistical/ Log Tables	None
Cambridge Tables	None
Graph Paper	None
Log Graph Paper	None
Other Materials	None

PTO

CT3532 Database Systems 2

Question 1

(a) Explain what is meant by a *minimal cover set*. Outline, with an example how you would generate a minimal cover set from a given set of functional dependencies. Explain the importance of minimal cover sets in database design. (8)

(b) Decomposing to *Boyce-Codd normal form* is not a *dependency preserving* decomposition. Explain, with an example, how you would decompose a relation to BCNF and highlight which functional dependencies are not preserved. (8)

(c) Explain what is meant by the term *denormalisation*. Outline, with the aid of examples, two separate forms of denormalisation. Discuss scenarios where denormalisation can be used. (9)

Question 2

(a) In the context of concurrency control, explain with a suitable example what is meant by the *incorrect summary problem.* (5)

(b) Outline the timestamping approach to concurrency control and show that it guarantees conflict serializability. (10)

(c) With respect to database recovery, what is meant by a *commit point*? In distributed databases, the database items are distributed across a number of sites with some items replicated across a number of site. Outline an approach that could be used to commit transactions in such a system. (10)

Question 3

(a) A B tree is a commonly used data structure used for efficient access to data. Given a B tree built on some attribute a_i, write pseudo-code to return all occurrences where:

i) of a_i equal to k (4)

ii) of a_i in a range between values k1 and k2 (6)

(b) In the context of parallel databases, compare round-robin and range partitioning techniques. Discuss the relative merits of these approaches for handling range queries. (7)

(c) Discuss the motivations for adopting a dynamic hashing approach. Describe, with the aid of an example, any approach to hashing to a dynamic file. (8)

Question 4

(a) Outline an efficient algorithm for sorting the tuples in a relation. Discuss the efficiency of your algorithm. (8)

(b) Explain how the sort algorithm in (a) can be improved given a parallel architecture. Discuss the improvement in efficiency. (8)

(c) Explain, with the use of suitable example, an efficient means to jointly index a number of attributes. Discuss the efficiency of your approach. (9)

END