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

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)

NUI Galway
OÉ Gaillimh

Semester 1 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 [ ]	No [ ]
MCQ Answersheet	Yes [ ]	No [ ]
Handout	None	
Statistical/ Log Tables	None	
Cambridge Tables	None	
Graph Paper	None	
Log Graph Paper	None	
Other Materials	None	
Graphic material in colour	Yes [ ] No [ ]	

PTO

CT3532 Database Systems 2

Question 1 (25 marks)

a) Explain the term *minimal cover set*. Given a set of functional dependencies F, outline an algorithm to generate a minimal set of functional dependencies. Illustrate the operation of your algorithm on the answer with the following set ***F***.

$$\boldsymbol{F} = \{B \rightarrow CE,\ D \rightarrow E,\ BC \rightarrow D,\ C \rightarrow E\} \qquad (8)$$

b) Given the following relation, R, and functional dependencies, F:

$$R = \{A, B, C, D, E, F, G, H, I, J\}$$

$$F = \{\{A,B,C\} \rightarrow \{D, E, F, G\},\ \{A,C\} \rightarrow \{J\},\ \{B\} \rightarrow \{H\},\ \{H\} \rightarrow \{I\},\ \{J\} \rightarrow \{C\}\}$$

i) Outline, in your words, an approach to identifying a suitable key; illustrate that approach on relation R.
ii) Draw a dependency graph for the set F.
iii) Decompose the relation to BCNF.

(11)

c) Explain, with the use of a suitable example, when de-normalisation may be appropriate. With respect to your example, outline the associated advantages and disadvantages.

(6)

Question 2 (25 marks)

a) Explain the structure of a B+ tree and outline an algorithm for the insertion of items into a B+ tree. Illustrate your algorithm by showing how a B+tree would develop given the following values to insert. You may assume each node in the tree can hold two search values.

34, 25, 16, 30, 15, 19, 7, 18

(10)

b) Linear hashing and dynamic hashing are two approaches to hash values to a dynamically changing file. Briefly outline either approach and illustrate the approach using the following record key values. You may assume each block can hold two records.

34, 25, 16, 15, 19, 9, 30, 24 (10)

c) Suggest an indexing strategy to jointly index a number of attributes *a1, a2, .. an,* where expect queries involving selection criteria on a number of these attributes. (5)

Question 3 (25 marks)

a) Explain what is meant by the *lost update problem* and show with an example schedule of transactions how it may arise. (5)

b) Explain what is meant by *two phase locking* and show how the schedule in a) would proceed under a two phase locking protocol. (6)

c) Prove that two phase locking guarantees conflict-serializability. (8)

d) Outline a general approach to database recovery. Your answer should explain the following concepts: system log, commit point, checkpoint. (6)

Question 4 (25 marks)

a) Outline an efficient approach to implementing sorting of data to satisfy, for example, the ORDER BY command. You may assume that the quantity of data to sort is too large to fit in main memory. Illustrate your approach with a small example. Comment on the efficiency of the approach. (8)

b) Parallel architectures have been used to increase the efficiency of many standard operations used in database management systems. Describe how range partitioning operates and discuss any advantages or limitations that exist with this approach. (9)

c) Explain, with an example, a means to sort the tuples of a relation where the relations has been partitioned across multiple partitions in a parallel database. (8)

Semester 1 Examination, 2019-2020

Exam Code(s)	3BCT1 1EM1
Exam(s)	Third Year Computer Science and IT Erasmus
Module Code(s)	CT3532
Module(s)	Computer Systems and Organization
Paper No.	1
Repeat Paper	No
External Examiner(s)	Dr. Jacob Howe
Internal Examiner(s)	Prof. Michael Madden *Mr. Donal Kelly
Instructions:	Answer any THREE questions. All questions carry equal marks.
Duration:	2 hours
No. of Pages	3
Discipline(s)	Computer Science
Course Co-ordinator(s)	Mr. Donal Kelly
Requirements:	
Release in Exam Venue	Yes
MCQ	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

Q.1.

a) Describe a B+tree leaf node and what it contains. How is it different from a B-tree leaf node?

(7)

b) Adding the sorted list `[1,2,3,4,5,6,7,8]` to a Binary Search Tree will give you poor search performance. Will you have the same problem with a B-tree? Show what this list will look like if inserted into a B-tree with order 3. How many comparisons are needed to search for number 3?

(12)

c) What structure does the MySQL InnoDb storage engine use to store its indexes? If given a table `Student` with columns `student_id`, `first_name`, `last_name`, `department_id`, and `email`, suggest an index that could work well for the query `SELECT first_name, last_name FROM student WHERE department_id = 4 AND last_name IN ("O'Reilly", "Reilly", "O' Reilly", "O Reilly");` What syntax would you use to make this index in MySQL?

(6)

Q.2.

a) In hashing, what is a collision? Describe an example. Why can a simple chaining collision resolution policy (CRP) result in poor search performance?

(6)

b) Describe how entries are inserted using a linear probing method. With this CRP, why do buckets need to be marked as *'tombstones'* after items are deleted from them?

(7)

c) Extendible Hashing and Linear Hashing are two kinds of dynamic hashing techniques. What is 'dynamic' about these approaches? From a high-level view, how do extendible and linear dynamic hashing differ? Pick one and describe how it inserts values. You may use pseudocode and/or diagrams.

(12)

PTO

Q.3.

a) Why is concurrency control needed? Describe an example where things can go wrong without concurrency control.

(8)

b) What is the two-phase locking protocol? How does it guarantee serializability?

(8)

c) MySQL uses Multiversion Concurrency Control techniques (MVCC) with its default storage engine. Describe how MVCC works. What does MySQL add to tables to allow it to use MVCC?

(9)

Q.4.

a) How is Boyce-Codd Normal Form (BCNF) different from 3NF? If you have a BCNF violation given table1 with columns `A,B,C` and functional dependencies `AB` $\rightarrow$ `C` and `C` $\rightarrow$ `A`, how should you decompose the table to remove it?

(8)

b) Given the following relation R, and the functional dependencies F, decompose R such that the resulting relations satisfy 3NF.

`R = {A,B,C,D,W,X,Y,Z}`

`F = {AB` $\rightarrow$ `C,C` $\rightarrow$ `D,B` $\rightarrow$ `WX,X` $\rightarrow$ `YZ}`

(8)

c) Given the following set of functional dependencies, F, generate the minimal cover set.
`F = {X` $\rightarrow$ `WD,AB` $\rightarrow$ `Z,A` $\rightarrow$ `BC,W` $\rightarrow$ `D,Z` $\rightarrow$ `C}`

(9)

NUI Galway
OÉ Gaillimh

Autumn Examinations 2018-2019

Course Instance Code(s)	3BCT1, 1EM1
Exam(s)	3Rd B.Sc Computer Science and Information Technology Erasmus
Module Code(s)	CT3532
Module(s)	Database Systems 2
Paper No.	1
External Examiner(s)	Dr Jacob Howe
Internal Examiner(s)	Professor Michael Madden *Dr. Colm O'Riordan

Instructions: Candidates should answer any **THREE** questions.
All questions will be marked equally.

Duration	2 hours
No. of Pages	3
Discipline(s)	Information Technology
Course Co-ordinator(s)	Dr D Chambers

Requirements:

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

PTO

Q.1.

(a) Explain the term conflict-serializability. Describe in detail an approach to guarantee conflict serializability among a set of concurrently running transactions. (10)

(b) For the approach chosen in part a), explain, with examples, any of the undesired side-effects of the approach. (7)

(c) With respect to recovery, explain the notion of a commit point of a transaction and its role in recovery for a system operating under an immediate update protocol. (8)

Q.2.

(a) Describe the structure of a B+tree. Provide pseudo code to outline how insertion into a B+tree occurs.
Illustrate the operation of your algorithm by showing how the following values would be inserted into a B+tree. You may assume a tree of order 3.

21, 12, 16, 18, 22, 23, 24 (12)

(b) Explain briefly what is meant by the term *dynamic hashing*. Outline the advantages of this approach. (6)

(c) Explain with a suitable example an approaches to building a multi-attribute index (7)

Q.3.

(a) Explain, with suitable examples, the process of normalisation in relational database design (10)

(b) Outline an algorithm to generate a minimal cover set. Illustrate, with a suitable example, how your algorithm operates. (9)

(c) Explain the role *Armstrong's axioms* play in the role of database design.

(6)

Q.4.

(a) With respect to parallel database, outline different approaches to partitioning a large relation across N disks and discuss their suitability for handling both point and range queries. (9)

(b) The join operator is a very common operator in relational database queries. Outline an efficient algorithm for sorting large quantities of data where the size of the data is too large to store in memory. (8)

(c) Outline an approach to performing a parallel join where the data is distributed across several separate machines. (8)

NUI Galway
OÉ Gaillimh

Semester 1 Examination 2018/2019

Exam Code(s)	3BCT1. 1EM1
Exam(s)	3rd B.Sc. (Computer Science and Information Technology) Erasmus
Module Code(s)	CT3532
Module(s)	Database Systems 2
Paper No.	1
Repeat Paper	
External Examiner(s)	Dr. Jacob Howe
Internal Examiner(s)	Prof. Michael Madden Dr. Colm O'Riordan*

Instructions: Candidates should answer any **Three** questions

All questions carry equal marks.

Duration: 2 hours

No. of Pages 3

Discipline(s) Information Technology

Requirements:
MCQ
Handout
Statistical/ Log Tables
Cambridge Tables
Graph Paper
Log Graph Paper
Other Materials

Release to Library: No

Q. 1.

a) Explain the term *minimal cover set*. Given a set of functional dependencies F, outline an algorithm to generate a minimal set of functional dependencies. Illustrate your answer with the following set ***F***.

$$\boldsymbol{F} = \{A \rightarrow BDE, B \rightarrow C, A \rightarrow C, AD \rightarrow G\} \quad (8)$$

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

$$R = \{A, B, C, D, E, F, G, H, I\}$$
$$F = \{ \{A,B,C\} \rightarrow \{D, E, F\} , \{A, B\} \rightarrow \{G\}, \{A,C\} \rightarrow \{J\}, \{C\} \rightarrow \{H\}, \{H\} \rightarrow \{I\}, \{J\} \rightarrow \{C\}\}$$

(10)

c) Explain, with the use of a suitable example, when denormalisation may be appropriate.

(7)

Q. 2.

a) Explain the structure of a B+ tree and outline an algorithm for insertion of items into a B+ tree. Illustrate your algorithm by showing how a B+tree would develop given the following values to insert. You may assume each node in the tree can hold two search values.

32, 23, 15, 29, 12, 18, 9, 21, 22

(10)

b) Outline briefly, with examples, an algorithm for deletion from a B+tree.

(5)

c) Linear hashing and dynamic hashing are two approaches to hash values to a dynamically changing file. Briefly outline either approach and illustrate the approach using the following record key values. You may assume each block can hold two records.

32, 23, 15, 29, 12, 18, 9, 21

(10)

Q.3.

a) Explain the concept of two phase locking and show with an example how it ensures correct concurrent access in databases. (10)

b) Outline the increased difficulties in guaranteeing correct concurrent access in a distributed database. (5)

c) Outline an approach to database recovery using a *system log*. Your answer should explain the following concepts
   i) system log
   ii) commit point
   iii) checkpoint
   iv) an algorithm for recovery (10)

Q.4.

a) Outline an efficient approach to implementing the join operator. (8)

b) Parallel architectures have been used to increase the efficiency of many standard operations used in database management systems. Outline approaches to distributing a relation and discuss the suitability for different types of queries (point and range). (9)

c) Explain, with an example, how to perform a join between two relations when the relations are partitioned across multiple partitions in a parallel database. (8)

NUI Galway
OÉ Gaillimh

Autumn Examinations 2018

Course Instance Code(s)	3BCT1, 1EM1
Exam(s)	3rd B.Sc. (Information Technology) Erasmus
Module Code(s)	CT3532
Module(s)	Database Systems 2
Paper No.	1
External Examiner(s)	Dr. Jacob Howe
Internal Examiner(s)	Professor Michael Madden *Dr. Colm O'Riordan

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

Duration	2 hours
No. of Pages	2
Discipline(s)	Information Technology
Course Co-ordinator(s)	Dr. Desmond Chambers

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

Q.1.

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

R = {A, B, C, D, E, F, G}
F= { {A,B, C} -> {D, E} , {B} -> {F}, {E} -> {G}, {D}->{A} }

(10)

(b) With respect to functional dependencies explain the following terms: *Armstrongs's axioms, closure, cover set.* (6)

(c) Outline an algorithm to generate a minimal cover set. Illustrate, with a suitable example, how your algorithm operates. (9)

Q.2.

(a) Linear hashing is one approach to hashing values to a dynamically changing file. Briefly outline this approach and illustrate the approach using the following record key values.
You may assume that each block can contain two records and that the initial file contains two blocks.

29, 28, 49, 43, 26, 11 (8)

(b) Outline an algorithm for deleting an item from a linearly hashed file. (8)

(c) Describe the structure of a B+tree. Explain, with an example, how insertion into a B+tree operates. (9)

Q.3.

(a) Explain the types of problems that arise without correct concurrency control mechanisms. (7)

(b) Explain the term conflict-serializability. Describe in detail an approach to guarantee conflict serializability among a set of concurrently running transactions. (10)

(c) With respect to recovery, explain the notion of a commit point of a transaction and its role in recovery for a system operating under an immediate update protocol. (8)

Q.4.

(a) Outline different approaches to partitioning a large relation across N disks and discuss their suitability for handling both point and range queries. (9)

(b) The join operator is a very common operator in relational database queries. Outline an efficient algorithm for sorting large quantities of data where the size of the data is too large to store in memory. (8)

(c) Outline an approach to performing a parallel join where the data is distributed across several separate machines. (8)

NUI Galway
OÉ Gaillimh

Semester 1 Examination 2017/2018

Exam Code(s)	3BCT1, 1EM1
Exam(s)	3rd B.Sc. (Information Technology) Erasmus
Module Code(s)	CT3532
Module(s)	Database Systems 2
Paper No.	1
Repeat Paper	
External Examiner(s)	Dr. Jacob Howe
Internal Examiner(s)	Dr. Michael Schukat Dr. Colm O'Riordan*
Instructions:	Candidates should answer any **Three** questions All questions carry equal marks.
Duration	2 hours
No. of Pages	3

Requirements:
MCQ
Handout
Statistical/ Log Tables
Cambridge Tables
Graph Paper
Log Graph Paper
Other Materials

Release to Library: No

Q.1

(a) Explain what is meant by the term *minimal cover set.* Illustrate an approach to generate a minimal cover set from the following set of functional dependencies:

F = {x -> y, x->z, xy->w, y->z}. (8)

(b) Define the term non-additive join. Show, with a suitable example, how you would check if a decomposition of a relation *R* into *R1* and *R2* has the non-additive join property. (8)

(c) Given the following relation *R* and set of functional dependencies *F*, decompose *R* into relations that satisfy Boyce-Codd normal form.

R = (A,B,C,D,E,F,G,H,I,J)
F = { {A,B,C}→{D,E,F}, {A,B}→{G}, {A,C} → {J},
{C} →{H}, {H}→{I},{J} →{C} } (9)

Q.2

(a) Linear hashing is one approach to hashing values to a dynamically changing file. Briefly outline this approach and illustrate the approach using the following record key values. You may assume that each block can contain two records and that the initial file contains two blocks.

29, 28, 49, 43, 26, 11 (8)

(b) Outline an algorithm for deleting an item from a linearly hashed file. (8)

(c) Describe the structure of a B+tree. Explain, with an example, how insertion into a B+tree operates. (9)

Q.3

(a) Explain, with an example, the *temporary update problem*. (4)

(b) Explain the term conflict-serializability. Outline an approach to guarantee conflict serializability among a set of concurrently running transactions. Illustrate the approach on the example schedule you used in part (a). (7)

(c) Prove that the approach outlined in b) guarantees conflict serializability. (7)

(d) With respect to recovery, explain the notion of a *commit point* of a transaction and its role in recovery for a system operating under an immediate update protocol. (7)

Q.4.

(a) Sorting is a common operation in database systems; outline an efficient algorithm for sorting large quantities of data where the size of the data is too large to store in memory. (8)

(b) Explain how such a sorting approach can be extended to work on a parallel architecture where N machines are available to perform the processing. (8)

(c) Outline different approaches to partitioning a large relation across N disks and discuss their suitability for handling batch processing and point and range queries. (9)

Autumn 2017

	3BCT1, SWB, EM
Exam Code(s)	Erasmus
Exam(s)	3rd B.Sc. Computer Science and Information Technology Science without Borders
Module Code(s)	CT3532
Module(s)	Database Systems 2
Paper No.	
External Examiner(s)	Dr. John Power
Internal Examiner(s)	Dr. Michael Schukat Dr. Colm O'Riordan*

Instructions: Answer any 3 questions

Duration	2 hours
No. of Pages	3 including this one
Department(s)	Information Technology
Requirements	None

PTO

Q. 1.

a) Explain the term *minimal cover set.* Given a set of functional dependencies F, outline an algorithm to generate a minimal set of functional dependencies. Illustrate your answer with the following set F.

$$F = \{A \rightarrow BC, B \rightarrow D, A \rightarrow DQ, B \rightarrow Q\} \quad (11)$$

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

$$R = \{A, B, C, D, E, F, G, H, I, J\}$$
$$F = \{ \{A,B, C\} \rightarrow \{D, E, F\} , \{A, B\} \rightarrow \{G\}, \{A,C\} \rightarrow \{J\}, \{C\} \rightarrow \{H\}, \{H\} \rightarrow \{I\}, \{J\} \rightarrow \{C\} \}$$

(12)

c) Define the term *non-additive join*. Show with a suitable example how you would check if a decomposition of a relation R into R1 and R2 has the *non-additive join property.* (10)

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.

125, 124, 409, 412, 411, 415, 106, 110

(13)

b) Compare the structure of a B tree with the structure of a B+tree. (7)

c) Illustrate the growth of a B+tree when the values in part (a) are inserted. You may assume an order of 3 for all nodes. (13)

PTO

Q. 3.

a) Explain, with an example, the *lost update problem*. (6)

b) Explain the term *conflict-serializability.* Outline an approach to guarantee conflict serializability among a set of concurrently running transactions. Illustrate the approach on the example schedule you used in part a). (9)

c) Prove the approach outlined in b) guarantees conflict serializabilty. (10)

d) With respect to recovery, explain the notion of a checkpoint of a transaction and its role in recovery for a system operating under an immediate update protocol. (8)

Q.4.

a) NOSQL database model have become popular as an alternative model to the relational database model. Discuss the main motivations behind these models and discuss how they differ from relational models. Your answer should include a discussion of the data models that are adopted and differences in terms of integrity enforcement. (11)

b) The join operator is a commonly performed operator in SQL. Outline an efficient approach to implementing this operator. (11)

c) Parallel architectures have been used to increase the efficiency of many standard operations used in database management systems. Outline an algorithm for efficiently performing a join between two relations R and S. (11)

NUI Galway
OÉ Gaillimh

Semester 1 2017

Exam Code(s)	3BCT1, SWB, EM Erasmus
Exam(s)	3rd B.Sc. Computer Science and Information Technology Science without Borders
Module Code(s)	CT3532
Module(s)	Database Systems 2
Paper No.	
External Examiner(s)	Dr. John Power
Internal Examiner(s)	Dr. Jim Duggan Dr. Colm O'Riordan

Instructions: Answer any 3 questions

Duration	2 hours
No. of Pages	3 including this one
Department(s)	Information Technology
Requirements	None

PTO

Q. 1.

a) Explain the term *minimal cover set.* Given a set of functional dependencies F, outline an algorithm to generate a minimal set of functional dependencies. Illustrate your answer with the following set F.

$$F = \{X \rightarrow YW, Y \rightarrow Z, X \rightarrow ZQ, Y \rightarrow Q\} \quad (11)$$

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

$$R = \{A, B, C, D, E, F, G, H, I\}$$
$$F = \{ \{A,B, C\} \rightarrow \{D, E, F\}, \{A, B\} \rightarrow \{G\}, \{A,C\} \rightarrow \{J\}, \{C\} \rightarrow \{H\}, \{H\} \rightarrow \{I\}, \{J\} \rightarrow \{C\} \}$$

(12)

c) Define the term *non-additive join*. Show with a suitable example how you would check if a decomposition of a relation R into R1 and R2 has the *non-additive join property*. (10)

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.

125, 124, 409, 412, 411, 415, 106, 110

(13)

b) Outline, with suitable examples, an approach to delete items from a dynamically hashed file. (7)

c) Explain the structure of a B+ tree and an approach to inserting an item in a B+ tree. Discuss the advantages of a B+ tree index over a hash index. Your answer should refer to the following types of queries: range queries, point queries.

(13)

Q. 3.

a) Explain, with an example, the *incorrect summary problem*. (6)

b) Explain the term *conflict-serializability*. Outline an approach to guarantee conflict serializability among a set of concurrently running transactions. Illustrate the approach on the example schedule you used in part a). (9)

c) Prove the approach outlined in b) guarantees conflict serializabilty. (10)

d) With respect to recovery, explain the notion of a commit point of a transaction and its role in recovery for a system operating under an immediate update protocol. (8)

Q.4.

a) NOSQL database model have become popular as an alternative model to the relational database model. Discuss the main motivations behind these models and discuss how they differ from relational models. Your answer should include a discussion of the data models that are adopted and differences in terms of integrity enforcement. (11)

b) Sorting is commonly performed procedure in SQL. Outline an efficient approach to implementing this operator on a relation. (11)

c) Parallel architectures have been used to increase the efficiency of many standard operations used in database management systems. Outline an algorithm for parallel sorting a file. (11)

Semester 1, 2015

Exam Code(s)	3BCT, SWB, EM
Exam(s)	3rd year Information Technology and Computer Science Science Without Borders Erasmus
Module Code(s)	CT3532
Module(s)	Database Systems 2
Discipline(s)	Information Technology
Internal Examiner(s)	Dr. Michael Madden Dr. Colm O'Riordan*
External Examiner(s)	Dr. John Power
No. of Pages	3 (including cover sheet)
Duration	2 hours
Instructions:	Answer any three questions

PTO

Q.1.

i) Let R(A, B, C, D, E) be a relational schema with five attributes. Let F be the set of functional dependencies as follows: F = {A → B, {E, D} → A, {B, C} → E} List all candidate keys for R. (6)

ii) Explain, with an example, how a relation may not satisfy Boyce-Codd normal form. Show how you decompose the relation such that the resulting relations satisfy Boyce Codd normal form.
(8)

iii) Explain the process of database design by synthesis. Your answer should include an explanation of functional dependencies, closure and cover sets. Illustrate your answer with an example. Outline any disadvantages associated with this approach.
(13)

iv) Describe briefly the properties of the schema resulting from the design by synthesis approach
(6).

Q.2.

i) Explain the term *conflict-serializability* and explain its importance in concurrency control. (6)

ii) Outline the time-stamping approach to guaranteeing conflict-serializability in database systems and explain how it guarantees conflict serializibility. (10)

iii) With respect to recovery in database systems explain the role of both *commit points* and *checkpoints.* (8)

iv) For a system operating under the immediate update protocol, describe the main entries that are stored in the system log. Explain how the recovery process might proceed under the immediate update protocol. (9)

Q.3.

i) Describe the structure of a B+ tree. Outline an algorithm for inserting values into a B+ tree; illustrate the operation of your algorithm with suitable examples. Discuss briefly how deletion from a B+ tree might proceed. (12)

ii) Discuss the advantages of B+ trees over B trees as means to index data. (8)

iii) Explain what is meant by *linear hashing*. Outline a suitable algorithm to illustrate the operation of inserting values into a dynamic file via linear hashing. Illustrate your algorithm with the following values. You may assume each block has capacity for two records:

24, 29, 11, 13, 10, 15, 26 (13)

Q.4.

i) The join operator is a commonly used operator. Outline an efficient algorithm to implement a join between two relations *R* and *S*. Discuss the efficiency of the algorithm. (11)

ii) With respect to object-oriented databases, discuss the main differences between the classical relational model and the object-oriented model. Your answer should include a discussion of the comparative advantages of the two models. (11)

iii) NOSQL databases have been proposed as an alternative to relational models. Discuss the main motivations behind these models and discuss how they differ from relational models. Your answer should include a discussion of the data models that are adopted and differences in terms of integrity enforcement. (11)