THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA

This exam paper must not be removed from the venue

Venue   _____

Seat Number   _____

Student Number |__|__|__|__|__|__|__|__|__|

Family Name   _____

First Name   _____

School of Information Technology and Electrical Engineering

EXAMINATION

Semester Two Final Examinations, 2017

INFS7903 Relational Database Systems

This paper is for St Lucia Campus students.

Examination Duration:	120 minutes
Reading Time:	10 minutes

Exam Conditions:

This is a Central Examination

This is a Closed Book Examination - specified materials permitted

During reading time - write only on the rough paper provided

This examination paper will be released to the Library

Materials Permitted In The Exam Venue:

(No electronic aids are permitted e.g. laptops, phones)

Calculators - Casio FX82 series or UQ approved (labelled)

Materials To Be Supplied To Students:

Instructions To Students:

Additional exam materials (eg. answer booklets, rough paper) will be provided upon request.

Please answer all questions on the examination paper.

For Multiple Choice Questions, please circle a single answer.

Total Marks: 100 (to be scaled down to 60)

For Examiner Use Only

Question	Mark

Total   _____

Question 1 [4 marks] Which of the following is a false statement about B+ trees?

　A.　B+-trees are balanced

　B.　non-leaf nodes include direct pointers to data records

　C.　insertion of a key can lead to node splitting

　D.　deletion of a key can lead to node coalescing

Question 2 [4 marks] Which of the following factors determines the size of a bitmap index on an attribute *"X"* in relation *"R"*?

　A.　The number of distinct values in *"X"*

　B.　The number of tuples in *"R"*

　C.　The data type for attribute *"X"*

　D.　Answers A & B above

　E.　Answers A & B & C

Question 3 [4 marks] Which of the following is a false statement about cost-based query optimization?

　A.　It selects a query plan in a shorter time than heuristic-based optimization

　B.　It requires estimating the execution cost of a query plan

　C.　It selects the query plan with the minimum execution cost

　D.　All of the above

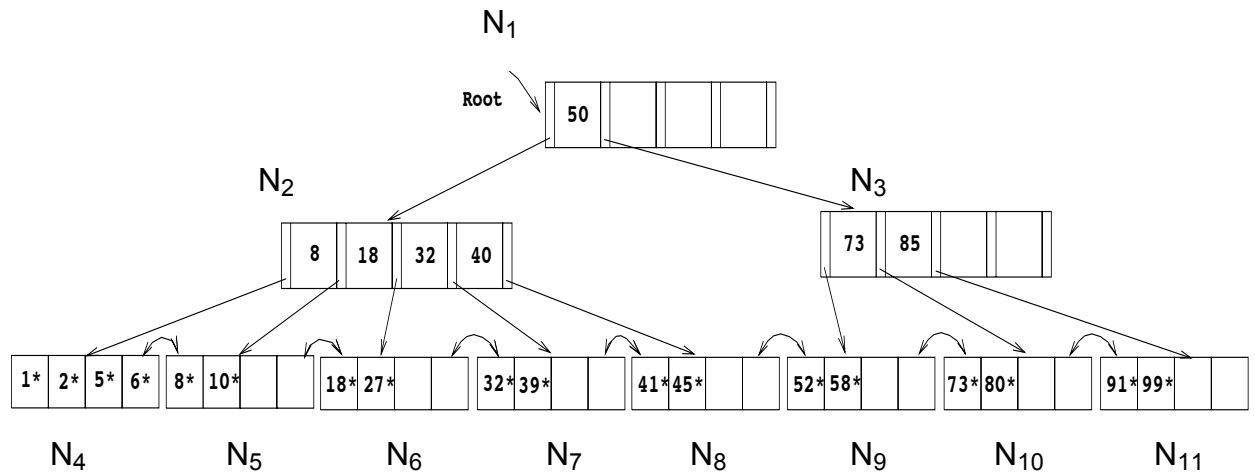**Questions 4-5** Consider the following relation:

Make	Model	Color	Price
Honda	Accord	Blue	Medium
Honda	Civic	Red	Low
Toyota	Corolla	Black	Low
Toyota	Camry	Red	Medium

Question 4 [4 marks] Assume the relation shown above and a bitmap index is created on attribute '*Price*'. The total number of bits required for that index is:

A. 2

B. 4

C. 8

D. 12

Question 5 [4 marks] Again, assume the relation shown above, what is the bitmap corresponding to the value *'Honda'*.

A. 1 1 0 0

B. 1 1

C. 1 0 1 0

D. 0 0

N_1

Root

| 50 | | | | |

N_2

| 8 | 18 | 32 | 40 |

N_3

| 73 | 85 | | |

| 1* | 2* | 5* | 6* | | 8* | 10* | | | 18* | 27* | | 32* | 39* | | 41* | 45* | | 52* | 58* | | 73* | 80* | | 91* | 99* | |

N_4 N_5 N_6 N_7 N_8 N_9 N_{10} N_{11}

Questions 6-8: Consider a B+ tree index as shown in figure, where index nodes are labeled: N_1, N_2, …, N_{11}. Also, assume the following rule applies for redistributing keys after a leaf node split: **Three keys** stay in the old leaf node and the remaining keys move to a new leaf node.

Question 6 [4 marks] Which nodes in the B+ tree index that must be fetched to answer the query: "Get all records with key greater than 30 and less than 75"

     A. N_1 N_2 N_6 N_7 N_8 N_9 N_{10}

     B. N_1 N_2 N_7 N_8 N_9

     C. N_1 N_3 N_7 N_8 N_9

     D. N_1 N_3 N_7 N_8 N_9 N_{10}

Question 7 [4 marks] What is the number of leaf nodes after inserting an entry with key "3"?

     A. 8

     B. 9

     C. 11

     D. 12

Question 8 [4 marks] What is the number of non-leaf nodes after inserting an entry with key "3"?

     A. 3

     B. 4

     C. 11

     D. 12

Questions 9-11 Given two relations R1 and R2, where R1 contains N1 tuples, R2 contains N2 tuples, and N2 > N1 > 0, answer the following questions.

Question 9 [4 marks] The minimum and maximum number of tuples produced from R1 ∪ R2 is:

    A. minimum 0, and maximum N1+N2

    B. minimum N1, and maximum N2

    C. minimum N1, and maximum N1+N2

    D. minimum N2, and maximum N1+N2

Question 10 [4 marks] The minimum and maximum number of tuples produced from R1 X R2 is:

    A. minimum 0, and maximum N1*N2

    B. minimum N1, and maximum N1+N2

    C. minimum N1*N2, and maximum N1*N2

    D. minimum N2, and maximum N1+N2

Question 11 [4 marks] Assume relation R1 contains an attribute named x, the minimum and maximum number of tuples produced from $\sigma_{x=5}$ (R1) is:

    A. minimum 0, and maximum N1

    B. minimum N1, and maximum N1

    C. minimum 1, and maximum N1

    D. minimum N1, and maximum N2

Questions 12-13: Suppose we have two unary (one attribute only) relations, R and S as shown below. Use R for the outer loop and S for the inner loop.

R
7
2
9
8
3
9
1
3
6

S
8
4
2
1
3
2
7
3

Question 12 [4 marks] Assume a natural join between R and S using **Nested Loop** join (one tuple at a time). The first five results of that join in the **order** that they would be produced by the nested loop is:

    A. 7, 2, 8, 3, 1

    B. 7, 2, 2, 8, 3

    C. 7, 2, 8, 3, 3

    D. None of the above

Question 13 [4 marks] Again, assume a natural join between R and S using **Nested Loop** join (one tuple at a time). The number of iterations needed to finish this join operation is:

    A. 1

    B. 8

    C. 9

    D. 5

Questions 14-15 Consider the relation `Student(Id, Major, Status)`, which has:

- A B+ tree index on `Major` and no other indexes.

- 10,000 tuples of data spread over 100 different blocks.

- The domain of `Status` has 9 values and that of `Major` has 10 different values.

Question 14 [4 marks] What is the estimated number of results returned by the expression $\sigma_{Major='IT'}$ (Student):

    A. 10

    B. 100

    C. 1,000

    D. 10,000

    E. None of the above

Question 15 [4 marks] What is the selectivity of the expression $\sigma_{Major='IT'\ OR\ Major="CS"}$ (Student):

    A. 0.0

    B. 0.1

    C. 0.2

    D. 1

    E. None of the above

Question 16 [4 marks] Which of the following is a correct statement about transactions?

    A. Redo is needed for atomicity

    B. Undo is needed for durability

    C. Concurrency Control is realized using Triggers and Assertions

    D. Deadlocks do not occur in serial executions

    E. None of the above.

Question 17 [4 marks] Which of the following transaction schedules does not contain conflicting operations? Recall that r = read and w = write.

    A. r1 (A), r2 (A), w1 (C), r1 (B), r2 (B)

    B. r1 (A), r1 (B), w1 (A), r2 (B), r2 (A), w2 (A)

    C. r1 (A), w1 (A), r1 (B), w1 (B), r2 (A), w2 (A), r2(B), w2(B)

    D. r1(A), w1(A), r2(B), w2(B), r2(A), w2(A), r1 (B), w1 (B)

    E. All (A)-(D) contain conflicts

Question 18 [4 marks] The write-ahead logging (WAL) protocol simply means that:

    A. writing of a data item should be done ahead of any logging operation.

    B. the log record for an operation should be written before the actual data is written.

    C. all log records should be written before a new transaction begins execution.

    D. the log never needs to be written to disk.

Question 19 [4 marks] If a database system supports ACID properties for transaction execution, which of the following pairs of values is a possible result for A and B, after executing the below transactions T1 and T2 concurrently, with an initial value of A=100 and B=200?

T1:
read(B)
B=B+50
write(B)
read(A)
A=A-50
write(A)

T2:
read(B)
tmp=B*0.1
B=B-tmp
write(B)
read(A)
A=A+tmp
write(A)

A. A=70 , B=230

B. A=50 , B=180

C. A=120 , B=250

D. A=50 , B=250

E. None of the above

Question 20 [4 marks] If a steal/no-force buffer management policy is in place, which of the following is true about system recovery?

A. Both the Redo and Undo operations are needed

B. Neither the Redo nor the Undo operations is needed

C. Redo is needed but Undo is not needed

D. Redo is not needed but Undo is needed

Question 21 [5 marks]:

You are given the following tables:

```
Student(StudId, Name, Addr, Status)

Transcript(Id, CrsCode, Semester, Grade)
```

Consider a query that outputs the names of all students who took INFS2200 in 2017. An execution plan for that query can be expressed as follows:
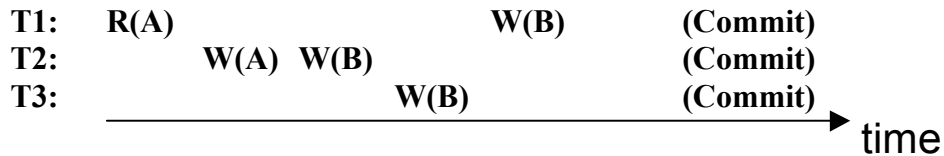
$$\pi_{Name}(\sigma_{Id=StudId\ AND\ CrsCode='INFS2200'\ AND\ Semester='2017'}\ (Student \times Transcript))$$

In the following, fill in the missing subscripts of the different operators (i.e., Π, σ, I) so that to create an optimized plan that is equivalent to the one above.

Π_{NAME} [

   $(\Pi_{_____}$ Student$) \bowtie_{_____}$

              $(\sigma_{_____}$ $(\Pi_{_____}$ Transcript$))$

     ]

Question 22 [5 marks]::

Consider the following sequences of actions, listed in the order they are submitted by transactions T1, T2, and T3.

```
T1:   R(A)                        W(B)        (Commit)
T2:         W(A)  W(B)                        (Commit)
T3:                     W(B)                  (Commit)
      ────────────────────────────────────────►  time
```

Describe how **strict 2PL** with deadlock detection (assume wait-for-graph is used) executes this sequence of actions. Specifically, <u>complete</u> the sequence listed below to show the **lock** and **unlock** requests made by these transactions as well as the blocked and unblocked operations. If a transaction is blocked, assume that all its actions are queued until it is resumed; the DBMS continues with the next action of an unblocked transaction.

T1 acquires shared-lock on A.

T2 blocks waiting for an exclusive-lock on A.

…

Question 23:

For each of the following schedules: 1) construct a precedence graph, 2) determine if the schedule is conflict serializable, and 3) determine the equivalent serial schedule.

(a) **[3 marks]** r1(X); r3(X); w1(X); r2(X); w3(X)

(b) **[3 marks]** r3(X); r2(X); w3(X); r1(X); w1(X)

Question 24:

Consider the join $R \bowtie_{R.a=S.b} S$, given the following information about the relations to be joined. The **cost** metric is the number of block read operations.

Relation R contains 10,000 tuples and has 10 tuples per block.

Relation S contains 2000 tuples and also has 10 tuples per block.

Attribute b of relation S is the **primary** key for S.

Both relations are stored as simple heap files.

(a) [2 marks] What is the cost of joining R and S using a **page-oriented nested loop** join?

(b) [2 marks] How <u>many tuples</u> does the join of R and S produce, at most? and how <u>many blocks</u> are required to store the result of the join back on disk?

END OF EXAMINATION