MANIPAL INSTITUTE OF TECHNOLOGY, MANIPAL 576104
(Constituent College of Manipal University)

FIRST SEMESTER M.TECH.(SOFTWARE ENGG.) DEGREE MAKE UP EXAMINATION JAN – 2016
SUBJECT: ADVANCED DATABASE MANAGEMENT SYSTEM – ICT 527
(REVISED CREDIT SYSTEM)

TIME: 3 HOURS   07/01/2016   MAX. MARKS: 50

Instructions to candidates
- Answer any **FIVE FULL** questions.
- Missing data, if any, may be suitably assumed.

1A. Consider a schema R = {A, B, C, D, E}, F=(A->B, BC->E). Check for BCNF and find whether the decomposition is dependency preserving

1B. Explain the different parallel database architecture with necessary diagrams.

1C. What is a weak entity set? Explain with the help of an example.

[5+3+2]

2A. Find all frequent itemsets of the following transaction database using Apriori algorithm with respect to minimum support = 20%.

TID	Item Numbers
1	1,2,5
2	2,4
3	2,3
4	1,2,4
5	1,3
6	2,3
7	1,3
8	1,2,3,5
9	1,2,3
10	1,2

2B. With a neat diagram explain all the steps in Knowledge discovery in databases.

2C. Explain nesting and unnesting in Object based databases with an example.

[5+3+2]

3A. Define and explain data warehouse. What are the major distinguishing features between OLTP and OLAP?

3B. Explain the working of range partitioning sort.

3C. How do you deal when participating site gets failed in distributed databases?

[5+3+2]

4A. Design a Pattern Count tree for the following set of transactions,.

TID	Items Procured			
T1	2	4		
T2	2	3		
T3	1	2	4	
T4	1	3		
T5	2	3		
T6	1	3		
T7	1	2	3	
T8	1	2	3	5
T9	1	2	5	
T10	1	2	4	

4B. Write a FP tree algorithm to find all frequent itemsets
4C. Explain partitioned parallel join with a neat diagram.

[5+3+2]

5A . Find the best splitting attribute when we construct decision tree using information gain for the following training sample :

No.	Outlook	Temperature	Humidity	Class
1	sunny	hot	high	N
2	sunny	hot	high	N
3	overcast	hot	high	P
4	rain	mild	high	P
5	rain	cool	normal	P
6	rain	cool	normal	N
7	overcast	cool	normal	P
8	sunny	mild	high	N
9	sunny	cool	normal	P
10	rain	mild	normal	P
11	sunny	mild	normal	P
12	overcast	mild	high	P
13	overcast	hot	normal	P
14	rain	mild	high	N

5B. Discuss the concept of mean, median, mid range, for the given data distribution. Suppose that the values for a given set of data are grouped into the following intervals. The intervals and corresponding frequencies are as given below.

Age	Frequency
1-10	100
10-20	400
20-30	300
30-40	2500
40-50	800
50-60	150

Compute an approximate median value for the data.

5C. List all the issues that arise during data integration?

[5+3+2]

6A. Suppose that a data warehouse consists of the four dimensions date, spectator, location, and game, and the two measures count and charge, where charge is the fare that a spectator pays when watching a game on a given date. Spectators may be students, adults, or seniors, with each category having its own charge rate.

(a) Draw a star and Snowflake schema diagram for the data warehouse.

(b) Starting with the base cuboid [date, spectator, location, game], what specific OLAP operations should you perform in order to list the total charge paid by student spectators at GM Place in 2015?

6B. Discuss 3-tier data warehouse architecture

6C. Explain the working of Hash partitioning.

[5+3+2]

*****