Reg. No.										

प्रज्ञानं ब्रह्म

Manipal
INSPIRED BY LIFE

MANIPAL INSTITUTE OF TECHNOLOGY, MANIPAL 576104

(Constituent College of Manipal University)

KNOWLEDGE IS POWER
MANIPAL INSTITUTE OF TECHNOLOGY

FIRST SEMESTER M.TECH.(SOFTWARE ENGG.) DEGREE END SEMESTER EXAMINATION DEC – 2015
SUBJECT: ADVANCED DATABASE MANAGEMENT SYSTEM – ICT 527
(REVISED CREDIT SYSTEM)

TIME: 3 HOURS **05/12/2015** **MAX. MARKS: 50**

Instructions to candidates

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

1A. Find the minimal cover for the following set of functional dependencies

F={A → B, AB→E , CD → E, EF → GH, AC→EG , DF → EG }

1B. Let R = (*A, B, C*)

F = {A → B, B → C} and let (R1, R2) be the decomposition of R such that R_1 = (A, B), R_2 = (B, C) check whether the relation in the decomposition are in BCNF and also verify if the decomposition is dependency preserving.

1C. Distinguish between intraoperation parallelism and interoperation parallelism.

[5+3+2]

2A. Compute the frequent items using FP-Tree technique, considering a minimum support count of 2, for the following set of transactions :

{T1={I1,I2,I5},T2={I2,I4},T3={I2,I3},T4={I1,I2,I4},T5={I1,I3},T6={I2,I3},T7={I1,I3}, T8={I1,I2,I3,I5},T9={I1,I2,I3}}

2B. Explain the working of two–phase commit protocol.

2C. How do you represent composite attributes in structure types? Explain with an example.

[5+3+2]

3A. With a neat diagram explain the transaction server process architecture.

3B. Discuss all partitioning techniques to partition a relation.

3C. Discuss the advantages and disadvantages of data replication.

[5+3+2]

4A. Use Pincer search algorithm on the following database to find all maximal frequent itemsets with respect to the minimum support 40%

TID	A	B	C	D	E
T1	1	1	1	0	0
T2	1	1	1	1	1
T3	1	0	1	1	0
T4	1	0	1	1	1
T5	1	1	1	1	0

4B. Write the dynamic itemset counting technique algorithm to find all frequent itemsets.

4C. Distinguish between supervised learning and unsupervised learning.

[5+3+2]

5A. The resting pulse rates from all the students of a class are as follows.

68,76,84,80,76,72,60,68,80,68,80,64,64,72,76,72,68,56,88,80,76,68,56,64,60,92,72,84,72

i) Find measures of central tendency mean, mode, midrange and median.

ii) Draw the boxplot for this data.

5B. Suppose that the data for analysis includes the attribute age. The age values for the data tuples are (in increasing order) 13, 15, 16, 16, 19, 20, 20, 21, 22, 22, 25, 25, 25, 25, 30, 33, 33, 35, 35, 35, 35, 36, 40, 45, 46, 52, 70

(i) Use min-max normalization to transform the value35for age on to the range [0.0..1.0].

(ii) Use z-score normalization to transform the value 35 for age, where the standard deviation of age is 12.94 years

5C. Discuss the structure of XML data and give one example for nested XML representation.

[5+3+2]

6A. Suppose that a data warehouse for a Big University consists of the four dimensions (student, course, semester, and instructor) and two measures count and avg grade. At the lowest conceptual level the avg grade measure stores the actual course grade of the student. At higher conceptual levels, avg grade stores the average grade for the given combination.

i) Draw a snowflake schema diagram for the data warehouse.

ii) Starting with the base cuboid [student, course, semester, instructor], what specific OLAP operations should you perform in order to list the average grade of CS courses for each Big University student.

iii) If each dimension has five levels (including all), how many cuboids will this cube contain (including the base and apex cuboids)?

6B. Illustrate the working of Asymmetric Fragment and Replicate Join.

6C. Explain 3-4-5 rules for discretization.

[5+3+2]

*****