BINDURA UNIVERSITY OF SCIENCE EDUCATION

FACULTY OF SCIENCE AND ENGINEERING

DEPARTMENT OF COMPUTER SCIENCE

BACHELOR OF SCIENCE HONORS DEGREE IN COMPUTER SCIENCE/INFORMATION TECHNOLOGY

CS413: ADVANCED DATABASE DESIGN AND IMPLEMENTATION

DURATION: 2HOURS 30 MINUTES TOTAL MARKS: 100

MAY 2024

INSTRUCTIONS TO CANDIDATES

The paper containts five questions. Answer ALL questions.

Question 1

a) Explain the following terms in data warehousing.
   i. Snowflake Schema,
   ii. Fact Constellation,
   iii. Data Marts. [3x2]

b) Examine deferred modification and immediate modification technique for recovery explaining how does recovery takes place in case of a failure in these techniques. [10]

c) Given R (A, B, C, D, E) with the set of FDs, F {AB → CD, A → E, C → D}. Is the decomposition of R into R1 (A,B,C), R2 (B,C,D) and R3(C,D,E) lossless? Prove. [10]

Question 2

a) Discuss the key characteristics of a data warehouse and how it differs in content, structure and function from an on-line transaction processing (OLTP) database. You should support your discussion with suitable diagrams and examples. [12]

b) Describe the five types of knowledge produced from data mining giving an example of each. [10]

Question 3

a) Consider the following three linked tables that contain information about employees and the projects they work on:

employees (emplD, name, salary)
project (projNbr, title, budget)
workload (emplD*, projNbr*, duration)

Consider the following query:

```
SELECT P.title, E.name
FROM employees E, project P, workload W
WHERE E.emplD = W.emplD
AND P.projNbr = W.projNbr
AND E.salary > 15000
AND W.duration < 20;
```

i. Draw an initial relational algebra tree for the above query. [4]

ii. Apply a series of transformations to the tree obtained in part (i) to make the query more efficient. Discuss each step and state the heuristic used. [12]

b) Security is of paramount concern when using multi-user systems. Briefly explain the following security issues that arise in a multi-user system:

i. authentication of users

ii. user privileges

iii. confidentiality of data [3x2]

Question 4

a) Suppose you are given a relation R = (A,B,C,D,E) with the following functional dependencies: {CE → D, D → B, C → A}.

i. Identify the best normal form that R satisfies. [2]

ii. If the relation is not in BCNF, decompose it until it becomes BCNF. At each step, identify a new relation, decompose and re-compute the keys and the normal forms they satisfy. [4]

b) Prove the Armstrong's union rule. [4]

Question 5

The academic world is an interesting example of international cooperation and exchange. This problem is concerned with modelling of a database that contains information on researchers, academic institutions, and collaborations among researchers. A researcher can either be employed as a professor or a lab assistant. There are three kinds of professors: Assistant, associate, and full professors. The following should be stored:

- For each researcher, his/her name, year of birth, and current position (if any).
- For each institution, its name, country, and inauguration year.
- For each institution, the names of its schools (e.g. School of Law, School of Business, School of Computer Science,. . .). A school belongs to exactly one institution.
- An employment history, including information on all employments (start and end date, position, and what school).
- Information about co-authorships, i.e., which researchers have co-authored a research paper. The titles of common research papers should also be stored.
- For each researcher, information on his/her highest degree (BSc, MSc or PhD), including who was the main supervisor, and at what school.
- For each professor, information on what research projects (title, start date, and end date) he/she is involved in, and the total amount of grant money for which he/she was the main applicant.

Design and draw an ER diagram for the data sets described above. [20]

**********END OF PAPER**********