

YOUR SS# \_\_\_\_\_ (No names please)

## EXAM I

### CS 6400 DATABASE SYSTEM CONCEPTS AND DESIGN

(Navathe, Time : 90 minutes)

Note: Answer the questions as completely as possible. If you need to make any assumptions, **please state them** and proceed. Examples should clearly illustrate the concept under question. (TOTAL POINTS 80 - allow roughly a minute per point to pace yourself).

Q1 and parts of Q4 are answered on the exam and the rest in blue-book.

**Q1. Multiple Choice:** (15 points - 1.5 per question).

1. Which of the following is the best general definition of a database?
  - (a) known facts that can be recorded
  - (b) a collection of related data with some meaning
  - (c) a collection of stored data with some encoding
  - (d) a collection of programs that allow users to create and maintain a database
  
2. Which schema level in the 3- schema (ANSI architecture) hides the details of physical storage structures and concentrates on describing entities, relationships and constraints of the whole database?
  - (a) conceptual level
  - (b) internal level
  - (c) external level
  - (d) global level
  
3. Logical data independence can be best defined as
  - (a) the capacity to change the internal schema without having to change the conceptual or external schema
  - (b) the capacity to change the conceptual schema without having to change external schema or application programs
  - (c) the capacity to change the DBMS software
  - (d) all of the above

4. Which of the following is (are) a capability (capabilities) that should be provided by a database management system?

- (a) maintain consistency and integrity of data
- (b) restrict unauthorized access
- (c) maintain application ( business) rules
- (d) all of the above
- (e) a and b
- (f) a and c

5. Which is not a responsibility of the database administrator?

- (a) authorizing access to the database
- (b) deciding on the storage structures and access strategies
- (c) modifying the DBMS software to support additional features
- (d) monitoring the performance of the database system

6. Which of the following is not a software component of a database management system

- (a) host language compiler
- (b) run-time database processor
- (c) query language interpreter / processor
- (d) all of the above

7. A relation instance is

- (a) a set of attributes in a relation schema
- (b) a set of n-tuples with values drawn from the domains of the attributes in a relation schema
- (c) the name of a relation schema
- (d) the relation intension

8. Information such as database users, description of database transactions, frequencies transactions is stored in

- (a) the database management system
- (b) the data dictionary or catalog
- (c) an entity-relationship diagram
- (d) an application program

9. Given the two relations  $R(A,B,C)$  and  $S(A,B,C)$ , which of the following is always true

- (a)  $R \cap S$  yields the same result as  $S \cap R$
- (b)  $R - S$  yields the same result as  $S - R$
- (c) Projection of  $S$  on  $A$  yields the same result as projection of  $R$  on  $A$
- (d) none of the above

10. A company may not use a DBMS when

- (a) no DBMS can deal with application and data complexity
- (b) no DBMS has a proper interface
- (c) all data in the company is stored in an application-specific software system
- (d) all of the above

**Q2.** (13 pts.)

**A.** (4 pts.) Based on your understanding of data modeling, database design and implementation, pick any two phases of database design for large practical applications. For each, define the inputs, the processing and the outputs of that phase of modeling or design.

**B.** (3 pts.) Show an example ER schema with a binary relationship having a (0,1) min-max constraint on both sides of the relationship. *Explain in words what this diagram implies.*

**C.** (6 pts.) Ternary relationships are used for a valid 3-way relationship between 3 entity types. Show a ternary relationship (other than the car-sale example used in class or the SUPPLY in the book) that is invalid, argue why it is invalid and then show a correct representation of it.

**Q3.** (11 pts.) Draw an EER diagram for the following narrative. Show max and min cardinalities with any notation. For EER, particular attention to show "generalization hierarchies" with appropriate constraints: (a) overlap (o) and disjoint (d) coverage as well as (b) total/partial coverage. Show only essential attributes with special attention to keys. (Supply attributes that you need if they are not given). Identify the "type" attribute for specialization wherever possible.

**NARRATIVE:**

At Scottish Rite Hospital we are creating a database of patients who have been treated here since 1993. Patients are classified into in-patients and out-patients. Out-patients also bear a visit\_date. Every patient gets a Patient-id as soon as admitted. A patient has SSN, name, Address and Phone-no. One important category of in-patients is surgical patients who carry an additional attribute of surgery-type. All patients are children under 18 and have a parent or guardian whose SSN is also recorded. Patients undergo many treatments, each of which is supervised by some physician. Only surgery patients are assigned to operating\_rooms. A physician may order a lab-test on patient if needed. Lab-tests primarily fall into Pathology and Radiology procedures and are identified by type of test and date for each patient. Radiology is further divided into CAT-scans, PET, and MRI. There is also a Primary-physician assigned to each patient.

**Q4. Relational Model. (20 pts.)**

a. Consider the following tables for an order-processing application:

CUSTOMER (Cust#, Cname, City)

ORDER(O#, Odate, C#, Order\_Amt)

ORDER\_ITEM ( O#, Item#, Qty)

ORDER\_PAYMENT ( O#, Paydate, Pay\_Amt, Check#)

OVERDUE\_NOTICE (Order#, Notice\_date, Penalty\_amount)

ITEM (Item#, Item-descr, Unit\_Price)

C# is same as customer number.

Ord\_Amt refers to total amount of the order. Odate is the date order was placed.

Pay\_Amt refers to payment for the order made on a specific date. (Note that payments may come in multiple installments).

Paydate is the date payment was received.

A. Draw a diagram with arrows showing the referential integrity constraints in the above database. - **DRAW ARROWS DIRECTLY ON THE SHEET.** (5 pts.)

B. Assume that we insert < 25110, I10, 30 > in ORDER\_ITEM. What checks must the system perform to guarantee integrity of the above database? Can the quantity be left null in the above? Why or why not? (4 pts.)

C. Consider the following two tables:

TABLE T1

P	Q	R
10	a	5
15	b	8
25	a	6

TABLE T2

A	B	C
10	b	6
25	c	3
10	b	5

Show the results of the following operations: (11 pts.)

1. T1 Equijoin T2 where  $T1.P = T2.A$  (Show either P or A in the result, not both)
2. T1 Equijoin T2 where  $T1.Q < T2.B$  (The inequality is based on character comparison; Show both Q or B in the result)
3. T1 Leftouterjoin T2 where  $T1.P = T2.A$ .
4. T1 Rightouterjoin T2 where  $T1.Q = T2.B$ .
5. T1 Union T2
6. T1 Equijoin T2 where  $T1.P = T2.A$  **and**  $T1.R = T2.C$ .

**Q 5.** (21 pts.) Perform the following queries in algebra against the database in Question 4 (Show step by step operation)

- a. List Cust#, Order# and Order\_amount for all orders placed by customers in Atlanta. (4 pts.)
- b. Draw a query tree for the above query and state it in SQL. (5pts.)
- c. List all items (Item#, Item-descr, qty) ordered by Customer #C50. [ Note: the same item may appear on multiple orders. Multiple listings of the same item in the output are OK.] (5pts.)
- d. List Order#, Total\_Pay\_amt for those orders for which the total of all payments received is less than the order amount. (Hint: use the SUM function). (7pts.)