B02208A/B/C Database Design Final Exam

Part 1: Multiple Choice (93 points - 3 points per question)

(B) 1. Which is an entity in a concert?
   (A) Date (B) Singer (C) Ticket price (D) None of the above

(A) 2. Which is false?
   (A) Database system is a collection of related data.
   (B) Meta-data is the data about data.
   (C) Database instance changes every time the database is updated.
   (D) Data update is a database modification operation.

(C) 3. Which is not a function of a DBMS?
   (A) database definition (B) database construction (C) data abstraction (D) database manipulation

(B) 4. Which is a job of a DBA?
   (A) Defining the database constraints (B) Authorizing access to the database
   (C) Defining the database transactions (D) None of the above

(C) 5. Which is true?
   (A) Database schema refers to the content of a database at a moment in time.
   (B) Schema is also called extension.
   (C) Potential for enforcing standards is an implication of using database approach.
   (D) None of the above

(C) 6. In the three-schema architecture, which schema can have multiple view?
   (A) Internal schema (B) Conceptual schema (C) External schema (D) None of the above

(B) 7. Which is in DBMS-independent design process?
   (A) Transaction Implementation (B) Requirement analysis
   (C) Application program design (D) None of the above

(C) 8. Which stores the presentation layer of the application?
   (A) application server (B) database server (C) Web interface (D) None of the above

(C) 9. Which is usually to represent a relationship in an ER diagram?
   (A) adjective (B) noun (C) verb (D) proposition

(X) 10. In ER diagram what does a double ellipse represent?
   (A) simple attribute (B) associative attribute (C) derived attribute (D) composite attribute

(B) 11. The _____ of a relationship type is the number of participating entity types.
   (A) cardinality (B) degree (C) identification (D) participation

(D) 12. Which is false?
   (A) A relationship type is the schema description of a relationship.
   (B) The relationship set is the current state of a relationship type.
   (C) A relationship can have more than one attribute.
   (D) A relationship set identifies certain relationship constraints.

(B) 13. Which type of relationships between classrooms and courses?
   (A) one-to-many (B) many-to-many (C) many-to-one (D) one-to-one

(D) 14. Which of the following is true?
   (A) The EER is a type of implementation data models.
   (B) An entity can exist in the database merely by being a member of a subclass.
   (C) It is necessary that every entity in a superclass be a member of some subclass.
   (D) None of the above

(B) 15. Which is the process of grouping a set of subclasses to form a superclass?
   (A) Specialization (B) Generalization (C) Aggregation (D) Identification

(A) 16. Which of the following is true?
   (A) A relation can have multiple keys. (B) A superkey is a key.
   (C) The key that is generated by the DBMS is called surrogate key. (D) None of the above

(C) 17. The relational model is a type of _____ data models.
   (A) conceptual (B) physical (C) implementation (D) hierarchical

(D) 18. Which of the following is true?
   (A) C. J. Date first proposed the relation model.
   (B) A relation looks like a table of attributes.
   (C) In a relation each row has a row header called an attribute.
   (D) None of the above

(D) 19. Which constraint may the delete operation violate?
   (A) Domain constraint (B) Entity constraint (C) Key integrity (D) None of the above
20. In the following EER diagram, which is true?

(A) A person can be a camper and a runner.
(B) A person must be a camper, a biker, and a runner.
(C) A person must be a camper, a biker, or a runner.
(D) None of the above

21. Which type of solutions to integrity violation does set null belong to?
(A) Cancel the operation that causes the violation.
(B) Perform the operation but inform the user of the violation.
(C) Trigger additional updates so the violation is corrected.
(D) None of the above

22. Which is an approach to map a n-ary relationship type in the ER diagram to the relational schema?
(A) Merged relation (B) A relationship relation and n foreign keys
(C) A relation set of simple component attributes (D) None of the above

23. To indicate an attribute as a key, which SQL command can be used?
(A) unique (B) index (C) distinct (D) None of the above

24. Which integrity constraints can trigger a sequence of operations?
(A) restrict (B) set default (C) set null (D) cascade

25. Which is a select condition in the following SQL commands?
```
select name from employee, department
where department_name = 'Research' and employee.department_no = department.department_no;
```
(A) department_name = 'Research' (B) employee.department_no = department.department_no
(C) select name from employee, department (D) None of the above

26. In the like operator of SQL, which can represent any string?
(A) * (B) % (C) % (D) None of the above

27. Which is not a aggregation function in SQL? (A) round() (B) sum() (C) avg() (D) max()

28. When you log into your Yahoo account, which SQL command will be used?
(A) insert (B) update (C) select (D) delete

29. Which is a measure of quality for relation schema design?
(A) Semantics of the attributes (B) Reducing the redundant information in tuples
(C) Reducing the NULL values in tuples (D) All of the above

30. Which is not a reason for nulls?
(A) Attribute is not applicable or invalid. (B) Attribute value is unknown (may exist).
(C) The value is beyond the domain range. (D) Value is known to exist, but unavailable.

31. What is a virtual table in SQL? (A) view (B) vision (C) screen (D) None of the above

Part 2: Questions and Answers (111 points)

1. (15 points) Briefly explain these terminologies. If they are acronyms, also write what they stand for.
(a) entity integrity (b) functional dependency (c) referential integrity (d) ontology (e) XML

(a) The entity integrity indicates the values of primary key attributes in a relation cannot be null.
(b) Functional dependency specifies that the value of an attribute in a table determine the value of other attribute in the same table.
(c) The referential integrity constraints indicate any attribute of a foreign key in a table can contain only either values from the corresponding main table’s primary key or the null value.
(d) Ontology means using conceptual modeling and other tools to develop “a specification of a conceptualization”.
(e) EXtensible Markup Language (XML) is a language used to specify the data content.

2. (a) (4 points) What is data model?
(b) (2 points) Why do we need mappings between schema levels?
(c) (6 points) Explain the differences among an entity, an entity type, and an entity set.

(a) A set of concepts to describe the structure of a database, the operations for manipulating these structures, and certain constraints that the database should obey.

(b) We need mappings to achieve data independence.

(c) i. An entity is a specific object or thing in the mini-world that are represented in the database.
   ii. An entity type is a collection of all entities that share common properties or characteristics.
   iii. An entity set is the collection of entities in the database.

3. (a) (4 points) What does SQL stand for? Explain it.

(b) (6 points) Based on the functions how can SQL be classified into three categories?

(a) Structured Query Language (SQL) is a standard language used to retrieve, update and delete data from relational database management systems (DBMS).

(b) Data Definition Language (DDL) is used to define databases.
   Data Manipulation Language (DML) is used to manipulate databases.
   Data Control Language (DCL) is used to control databases.

4. (a) (3 points) Explain normalization for a relational database.

(b) (9 points) Explain the 1NF, 2NF, and 3NF.

(a) The process of structuring relations by decomposing their attributes into smaller relations.

(b) i. The relation has no composite attributes, multivalued attributes, and nested relations.
   ii. Every non-prime attribute is fully functionally dependent on the primary key in the relation.
      Other possible answers:
      There are no two keys in the relation.
      There is no partial functional dependency in the relation.
   iii. There is no transitive functional dependency In the relation .

5. (8 points) A customer relation has 4 attributes: customer_id, name, email, and address. No two customers have the same customer_id and email.

(a) (6 points) List keys, 3 superkeys, and primary key for the customer relation.

(b) (2 points) Explain the reason of choosing the primary key.

(a) • keys: customer_id, email
   • superkeys: (any three of the following) customer_id, email, (customer_id, name), (customer_id, email),
     (customer_id, address), (name, email), (email, address), (customer_id, name, email),
     (customer_id, name, address), (customer_id, email, address), (name, email, address),
     (customer_id, name, email, address)
  • customer_id

(b) The customer_id is choosen because it can uniquely identify each tuple in the doctor relation and the email might have the null value.

6. (20 points) Consider the following course database:

<table>
<thead>
<tr>
<th>student table</th>
<th>class table</th>
</tr>
</thead>
<tbody>
<tr>
<td>student_no</td>
<td>employee_no</td>
</tr>
<tr>
<td>B94022</td>
<td>E92012</td>
</tr>
<tr>
<td>B94145</td>
<td>E93008</td>
</tr>
<tr>
<td>B94262</td>
<td>E94012</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>instructor table</th>
<th>course table</th>
</tr>
</thead>
<tbody>
<tr>
<td>employee_no</td>
<td>name</td>
</tr>
<tr>
<td>E92012</td>
<td>David Tao</td>
</tr>
<tr>
<td>E93008</td>
<td>Julius Tsai</td>
</tr>
<tr>
<td>E94012</td>
<td>Ge Cao</td>
</tr>
</tbody>
</table>

where primary keys are underlined. employee_no, student_no, and course_no in the class table are foreign keys referencing to the instructor, student, and course table respectively.

(a) If the following operations are taken, check if domain constraints, key constraints, entity integrity, or referential integrity is violated. If there is any violation, explain it.
   i. (2 points) Insert (‘E94012’, ‘Wu Bai’, ‘Professor’, MUSIC) into the instructor table.
iii. (2 points) Change the student_no value in the class table from 'B94262' to 'B94623'.
iv. (2 points) Change the department value in the instructor table from 'IM' to 'CIS'.

(b) Use SQL to answer the following questions.
i. (3 points) Create the student table.
ii. (3 points) Add a constraint of $0 \leq \text{credit} \leq 6$ in the course table.
iii. (2 points) Insert ('CS410', 'Data Mining', 3) into the course table.
v. (2 points) Remove all classes taught by 'Ge Cao' from the class table.

(a) i. It violates the key constraint because the employee_no 'E94012' already existed.
ii. It violates the referential integrity because the foreign key, student_no 'B94145' in the class table will have no primary key to reference to in the student table.
iii. It violates the referential integrity because the foreign key, student_no 'B94623' in the class table will have no primary key to reference to in the student table.
iv. It violates no constraint.

(b) i. create table student (
    student_no char(6) primary key not null,
    name varchar(30),
    major varchar(10),
    phone_no char(10));
ii. alter table course add constraint credit_constraint check (0 <= credit and credit <= 6)
iii. insert into course values ('CS410', 'Data Mining', 3)
v. delete from class where employee_no = ( select employee_no from instructor where name = 'Ge Cao')

7. (8 points) Consider a database in a university with the following functional dependencies:

\[
\begin{align*}
\text{StdSSN} & \rightarrow \text{StdCity}, \text{StdClass} \\
\text{OfferNo} & \rightarrow \text{Semester}, \text{Year}, \text{CourseNo}, \text{CourseName} \\
\text{CourseNo} & \rightarrow \text{CourseName} \\
\text{StdSSN}, \text{OfferNo} & \rightarrow \text{Score}
\end{align*}
\]

Decide which type of normalized form it is. If it is not a 3NF, transform it to a 3NF.

(a) (3 points) course (OfferNo, Semester, Year, CourseNo, CourseName)
(b) (5 points) grade_report (stdSSN_no, StdClass, OfferNo, Year, CourseNo, CourseName, score)

(a) 2NF, offer((OfferNo, Semester, Year, CourseNo), course(CourseNo, courseName)
(b) 1NF, student(StdSSN, StdClass), offer((OfferNo, Year, CourseNo), score(stdSSN, CourseNo, score), course(CourseNo, courseName)

8. (12 points) Consider the following bank schema.

\[
\begin{align*}
\text{account} & (\text{account_id}, \text{branch_name}, \text{balance}) \\
\text{borrower} & (\text{customer_id}, \text{loan_number}) \\
\text{branch} & (\text{branch_id}, \text{branch_name}, \text{branch_city}, \text{assets}) \\
\text{customer} & (\text{customer_id}, \text{customer_name}, \text{customer_street}, \text{customer_city}) \\
\text{depositor} & (\text{customer_id}, \text{account_number}) \\
\text{loan} & (\text{loan_number}, \text{branch_id}, \text{amount})
\end{align*}
\]

Write SQL commands for the following query:

(a) (3 points) Retrieve all different branch names.
(b) (3 points) Retrieve all loan number that falls between 1000000 and 2000000.
(c) (3 points) Retrieve all customer names in the 'East Gate' branch.
(d) (3 points) Retrieve the branch name and number of accounts for each branch.

(a) select distinct branch_name from branch
(b) select loan_number from loan where amount >= 1000000 and amount <= 2000000
(c) select customer_name from borrower, loan

where borrower.loan_number = loan.loan_number and branch_name = 'East Gate'
(d) select branch_name, count(account_id) from account group by branch_name

9. (14 points) Consider the following order table.

<table>
<thead>
<tr>
<th>order_no</th>
<th>date</th>
<th>customer_no</th>
<th>customer_name</th>
<th>item</th>
<th>product_no</th>
<th>product_name</th>
<th>unit</th>
<th>unit_price</th>
<th>amount</th>
<th>price</th>
</tr>
</thead>
<tbody>
<tr>
<td>0610248</td>
<td>2006-5-30</td>
<td>VICRP</td>
<td>Victor Corp.</td>
<td>1</td>
<td>10001042</td>
<td>rice</td>
<td>3 Kg/bag</td>
<td>150</td>
<td>10</td>
<td>1500</td>
</tr>
<tr>
<td>0610248</td>
<td>2006-5-30</td>
<td>VICRP</td>
<td>Victor Corp.</td>
<td>2</td>
<td>10001072</td>
<td>coke</td>
<td>24 cans/box</td>
<td>480</td>
<td>5</td>
<td>2400</td>
</tr>
<tr>
<td>0610249</td>
<td>2006-6-8</td>
<td>DONDI</td>
<td>Dondi Corp.</td>
<td>1</td>
<td>10001014</td>
<td>milk</td>
<td>24 bottles</td>
<td>600</td>
<td>9</td>
<td>5400</td>
</tr>
<tr>
<td>0610249</td>
<td>2006-6-8</td>
<td>DONDI</td>
<td>Dondi Corp.</td>
<td>2</td>
<td>10001051</td>
<td>corn chips</td>
<td>24 bags/box</td>
<td>720</td>
<td>5</td>
<td>3600</td>
</tr>
<tr>
<td>0610249</td>
<td>2006-6-15</td>
<td>JENRE</td>
<td>Jenren Corp.</td>
<td>1</td>
<td>10001002</td>
<td>beer</td>
<td>24 cans/box</td>
<td>480</td>
<td>10</td>
<td>4800</td>
</tr>
</tbody>
</table>

where the keys are underlined.

(a) (7 points) Normalize the above table to the 3 NF and draw the relational schema diagram and indicate the primary keys and the referential constraints.

(b) (7 points) Based on the above schema diagram draw the ER diagram.

(a) 

(customer: customer_no, customer_name) 

order: order_no, customer_no, date 

detail: order_no, item, product_no, amount 

(b) 

ER Diagram:

- Customer
- Order
- Detail
- Product