Part 1: Multiple Choice (102 points - 3 points per question)

1. Which is an entity in a conference? (A) location (B) session (C) schedule (D) none of the above

2. Which is true? (A) Database system is a collection of related data. (B) Metadata is the state of data. (C) Database instance changes every time the database is updated. (D) none of the above

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

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

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

6. Which is in DBMS-independent design process? (A) transaction Implementation (B) application program design (C) requirement analysis (D) none of the above

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

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

9. In ER diagram what does a double ellipse represent? (A) simple attribute (B) associative attribute (C) derived attribute (D) none of the above

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

11. 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.

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

13. Which of the following is true? (A) The EER is a type of conceptual data models. (B) C. J. Date first proposed the relation model. (C) An entity can exist in the database merely by being a member of a subclass. (D) None of the above

14. Which of the following is true? (A) A relation can have only one candidate key. (B) A primary key is a super key. (C) A candidate key can uniquely identify a row. (D) none of the above

15. Which is persistent data? (A) HTML documents (B) SQL statements (C) work queues (D) none of above

16. Which constraint may the delete operation violate? (A) Referential constraint (B) Entity constraint (C) Key integrity (D) None of the above

17. Which of the following finds those groups meeting stated conditions? (A) group by (B) using by (C) having (D) None of the above

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

19. To eliminate duplicate rows in a query, which can be used? (A) unique (B) index (C) distinct (D) none of the above

20. Which integrity constraints can trigger a sequence of operations? (A) restrict (B) set default (C) set null (D) none of the above

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

22. In the like operator of SQL, which can represent any string? (A) * (B) : (C) _ (D) %

23. In MySQL which command can show the schema of a table? (A) show (B) display (C) present (D) describe
24. In the following EER diagram, which is true?

(A) An account type is determined by Account_no.
(B) An account can be a checking and savings account.
(C) There is no account that does not belong to checking, savings, or loan accounts.
(D) None of the above

25. When you log into your Facebook account, which SQL command will be used?

(A) insert (B) update (C) select (D) delete

26. Which is a benefit of denormalization?

(A) performance improvement (B) higher security (C) less storage (D) none of the above

27. Which is not a reason for nulls?

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

28. In MySQL which is used to execute a SQL script? (A) use (B) \e (C) \. (D) none of the above

29. A rule that database users should obey is called a: (A) regulation. (B) principle. (C) privilege. (D) constraint.

30. A join that is based upon equality between values in two common columns with the same name and where one duplicate column has been removed is called a(n):

(A) equi-join. (B) natural join. (C) multivariate join. (D) inner join.

31. Which is a subset of the database that is presented to one or more users?

(A) correlated subquery (B) derived table (C) view table (D) None of the above.

32. What results would the following SQL statement produce?

```sql
select owner, table_name from dba_tables where table_name = 'customer';
```

(A) A listing of the owner of the customer table (B) A listing of all customers in the customer table (C) A listing of the owner of the customer table as well as customers (D) An error message

33. Which provides a standard software API for using DBMS? (A) CASE (B) ODBC (C) OLAP (D) none of the above

34. The UNION clause is used to:

(A) join two tables together to form one table. (B) find all rows that do not match in two tables. (C) combine the output from multiple queries into a single result table. (D) None of the above.

Part 2: Questions and Answers (120 points)

1. (15 points) Briefly explain these terminologies. If they are acronyms, also write what they stand for.

(a) data model (b) functional dependency (c) weak entity (d) ontology (e) XML

(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) Functional dependency specifies that the value of one attribute (the determinant) determines the value of another attribute in the same table.

(c) A weak entity is an entity whose existence depends on other entity.

(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. (12 points) Briefly explain the four constraints on specialization and generalization and give an example for each case.

- The **disjointness** constraint specifies that the subclasses of the specialization must be disjoint. A graduate student cannot be a undergraduate at the same time.

- The **overlapping** specialization specifies that the subclasses of the specialization can be overlapping. A manager can be a engineer at the same time.
• The **total** specialization specifies that every entity in the superclass must be a member of some subclass. Human beings can be classified into males and females.

• The **partial** specialization specifies that an entity in the superclass is allowed not to belong to any of the subclasses. A student need not to be an undergraduate or a graduate student.

3. (8 points) Explain the key constraint, domain constraint, entity integrity constraint, and referential constraint.

- The key constraint means there is no duplicate key in any relation.
- The domain constraint indicates every value in a tuple must be from the domain of its attribute.
- The entity integrity indicates the values of primary key attributes in a relation cannot be null.
- The referential integrity constraints indicate any attribute of a foreign key in a table can contain only either values from the corresponding parent table’s primary key or the null value.

4. (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.

5. (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. 1NF is the relation that has no composite attributes, multivalued attributes, and nested relations.
      ii. In 2NF, 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. Third normal form (3NF) is a normal form in which there is no transitive functional dependency in the relation.

6. (8 points) A player relation has 4 attributes: player_id, name, birthdate, and email. No two players have the same player_id and email.

   (a) (6 points) List keys, 3 superkeys, and primary key for the player relation.
   (b) (2 points) Explain the reason of choosing the primary key.

   (a)
   • keys: player_id, email (simple keys)
   • superkeys: (any three of the following) player_id, email, (player_id, name), (player_id, email), (player_id, birthdate), (name, email), (birthdate, email), (player_id, name, email), (player_id, name, birthdate), (player_id, birthdate, email), (name, birthdate, email), (player_id, name, birthdate, email)

   (b) The player_id is choosen because it can uniquely identify each tuple in the student relation and the email might have the null value or be changed.

7. (27 points) Consider the following library database:

<table>
<thead>
<tr>
<th>borrower table</th>
<th>book_loan table</th>
<th>book table</th>
</tr>
</thead>
<tbody>
<tr>
<td>card_no</td>
<td>name</td>
<td>birthdate</td>
</tr>
<tr>
<td>B97022</td>
<td>Taylor Swift</td>
<td>03-5123456</td>
</tr>
<tr>
<td>B97145</td>
<td>Lily Allen</td>
<td>0912-123456</td>
</tr>
<tr>
<td>B97262</td>
<td>Lady Gaga</td>
<td>0928-342512</td>
</tr>
</tbody>
</table>

where primary keys are underlined. card_no and book_id in the book_loan table are foreign keys referencing to the borrower and book 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 ('B97145', 'Wu Bai', '03-5168168') into the borrower table.
ii. (2 points) Remove the row ('B97145', 'Lily Allen', '0912-123456') from the borrower table.
iii. (2 points) Change the book_id of 'Java' in the book table from '123688' to '123123'.
iv. (2 points) Change the card_no in the borrower table from 'B97262' to 'B97623'.
v. (2 points) Change the title value in the book table from 'Java' to 'Java Programming'.
(b) Use SQL to answer the following questions.
   i. (3 points) Create the borrower table.
   ii. (2 points) Add a attribute publisher into the book table.
   iii. (2 points) Insert ('B07168', 'Justin Bieber', '03-5186417') into the borrower table.
   iv. (2 points) Change 'Database' to 'Database Design' in the book table.
   v. (2 points) Remove all books borrowed by 'Lady Gaga' from the book table.
   vi. (3 points) List all book title that are borrowed on June 6, 2010.
   vii. (3 points) Count the number of books borrowed by each borrower.

(a) i. It violates the key constraint because the card_no 'B97145' already existed.
   ii. It violates the referential integrity because the foreign key, card_no 'B97145' in the book_loan table will have no primary key in the borrower table to reference to.
   iii. It violates the referential integrity because the foreign key, book_id '123688' in the book_loan table will have no primary key in the book table to reference to.
   iv. It violates the referential integrity because the foreign key, card_no 'B97145' in the book_loan table will have no primary key in the borrower table to reference to.
   v. It violates no constraint.

(b) i. create table borrower ( 
   card_no char(6) primary key not null,  
   name varchar(30),  
   birthdate date)  
   ii. alter table book add publisher varchar(30);  
   iii. insert into borrower values ('B07168', 'Justin Bieber', '03-5186417')  
   iv. update book set title = 'Database Design' where title = 'Database'  
   v. delete from book where book_id in (select book_id from borrower, book_loan where borrower.card_no = book_loan.card_no and name = 'Lady Gaga')  
   vii. select name, count(*) from borrower, book_loan where borrower.card_no = book_loan.card_no group by book_id

8. (14 points) Consider the world cup. There are teams, player, and matches. Each entity should have a least 3 attributes.
   (a) (7 points) Add the necessary relationships to draw a ER/EER diagram to represent the relationships among teams, players, and matches.
   (b) (7 points) Transform the above ER/EER diagrams to 3NF relations and show the 3NF relation schema.

(a) 

(b) 

player: | player_no | country | name | birthday | position |
--- | --- | --- | --- | --- | --- |

team: | country | coach | group |
--- | --- | --- | --- |

match: | country | date | match_country | time | location |
9. (14 points) Consider the following schema for the movie information. Normalize it to 3NF relations.

```
movie(movie_id, title, year, star_id, star_name, gender, birthdate, director_id, director_name)
```

where primary keys are underlined.

(a) (7 points) Draw the ER diagram.

(b) (7 points) Show the relation schema.

(a) [Diagram of ER model]

(b) star: |
| star_id | star_name | gender | birthday |

movie: |
| star_id | director_id | movie_id | title | year |

director: |
| director_id | director_name |