Part 1: Multiple Choice (78 points - 3 points per question)

(D) 1. Which is true?
   (A) Metadata is a collection of data. (B) XML is a type of structured data.
   (C) The biggest problem with data redundancy is data inconsecutiveness. (D) None of the above

(B) 2. Which of the following is not an advantage of the database approach?
   (A) Planned data redundancy (B) Reduced installation complexity
   (C) Improved data accessibility (D) Improved data sharing

(B) 3. Which is the discovery of new information in terms of patterns or rules from vast amounts of data?
   (A) Data acquisition (B) Data mining (C) Data recongnition (D) None of the above

(C) 4. Which is centralized storehouse of metadata?
   (A) Data inventory (B) Data warehouse (C) Repository (D) None of the above

(D) 5. Which is a DBMS? (A) Apache (B) Microsoft Excel (C) PHP (D) None of the above

(A) 6. Which is an entity in a movie theater? (A) Staff (B) Movie schedule (C) Ticket price (D) None of the above

(A) 7. Which is true?
   (A) Planning, analysis, design, and implementation are four steps to top-down planning.
   (B) Planning matrixes are used to describe interrelationships between planning processes.
   (C) Prototyping is time-consuming but comprehensive.
   (D) None of the above

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

(A) 9. Which type of schema is described in E-R models?
   (A) Conceptual Schema (B) External schema (C) Internal Schema (D) None of the above

(B) 10. Transforming the data specifications into basic, or atomic elements following well-established rules is called:
   (A) design. (B) normalization. (C) typing and cross-matching. (D) implementation.

(C) 11. Which is not a characteristics of identifiers?
   (A) It will not change in value. (B) It will not be null.
   (C) It is better to use intelligent identifiers. (D) Simple keys are better than composite keys.

(B) 12. Which type of relationships between a customer and the order?
   (A) one-to-one (B) one-to-many (C) many-to-one (D) many-to-many

(B) 13. An attribute that is deduced from the other attribute is called a(n) _____ attribute.
   (A) associative (B) derived (C) reqired (D) stored

(A) 14. Which are specific objects or things in the mini-world that are represented in the database.
   (A) Entities (B) Attributes (C) Relationships (D) Descriptions

(C) 15. A relationship that links strong entities to weak entities is called:
   (A) a key relationship (B) a corresponding relationship (C) an identifying relationship (D) an associative entity

(C) 16. One entity related to another of the same entity type is called a(n):
   (A) binary relationship (B) ternary relationship (C) unary relationship (D) none of the above

(C) 17. In the following figure, which attribute is multi-valued? (A) Years_Employed (B) Address (C) Skill (D) Employee_ID

(B) 18. Which of the following is true?
   (A) The EER is a type of internal data models. (B) UML is used for conceptual data modeling.
   (C) A business rule specification should be written strictly in computer code. (D) None of the above

(B) 19. Which is the process of defining a more general entity type from a set of more specialized entity types
   (A) Specialization (B) Generalization (C) Aggregation (D) Identification

(D) 20. Which is false?
   (A) The supertype/subtype relationships are also called IS-A relationships.
   (B) An entity cannot exist in the database merely by being a member of a subtype.
   (C) A member of the supertype can be optionally included as a member of any number of its subtypes.
   (D) It is necessary that every entity in a supertype be a member of some subtypes.
21. An attribute of the supertype that determines the target subtype(s) is called the:
   (A) subtype determinant  (B) subtype discriminator  (C) subtype indicator  (D) subtype identification

22. The property by which subtype entities possess all attributes of a supertype is called attribute:
   (A) identification  (B) generalization  (C) aggregation  (D) inheritance

23. In the SQL language, which statement can be used to implement business rules in databases?
   (A) Create rule  (B) Create operation  (C) Create assertion  (D) None of the above

24. In the following EER diagram, which is true?
   (A) A patient must be either an outpatient or a resident patient.
   (B) A resident patient can be assigned a bed or may not be assigned to any bed.
   (C) A patient must be taken care by at least one responsible physician.
   (D) None of the above

25. Which of the following is true?
   (A) E. F. Codd first proposed the relation model.  (B) Data schema is changed when database is inserted.
   (C) ERP is centralized storehouse of enterprise metadata.  (D) None of the above

Part 2: Questions and Answers (87 points)

1. (28 points) Briefly explain these terminologies. If they are acronyms, also write what they stand for.
   (a) XML  (b) DBA  (c) DBMS  (d) SDLC  (e) CASE  (f) Data warehouse  (g) ISA
      (a) eXtensible Markup Language (XML) is a language for defining markup languages.
      (b) A database administrator (DBA) is the person who takes overall responsibility for data, metadata, and policies about data use.
      (c) A database management system (DBMS) is software system used to create, maintain, and provide controlled access to databases.
      (d) The System Development Life Cycle (SDLC) is a systematic approach to the software development.
      (e) Computer-Aided Software Engineering (CASE) is the scientific application of a set of tools and methods to a software which results in high-quality, defect-free, and maintainable software products.
      (f) Data warehouse is an integrated decision support system derived from various operational databases or a repository of stored data.
      (g) Information System Architecture (ISA) is a conceptual blueprint or plan that expresses the desired future structure for the information systems in an organization.

2. (a) (4 points) What is data model?
      (b) (3 points) Why are data models needed?
      (c) (3 points) What is a relational database?
      (a) A data model is 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) Data models are needed to capture the nature of and relationships among data. Data models are fundamental for effectiveness and efficiency of a database.
      (c) A relational database is a database that represents data as a collection of tables in which all data relationships are represented by common values in related tables.

3. (a) (4 points) Explain the differences among a entity instance and an entity type.
      (b) (3 points) What is the degree of relationships?
      (c) (3 points) What is the cardinality of relationships?
      (a) A entity instance is a single occurrence of an entity type. An entity type is a collection of entities that share common properties or characteristics.
(b) The degree is the number of entity types that participate in a relationship.
(c) The Cardinality specifies the number of instances of one entity that can be associated with each instance of another entity.

4. (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 an undergraduate at the same time.
- The **overlapping** specialization specifies that the subclasses of the specialization can be overlapping.
  A manager can be an 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.

5. (a) (4 points) Illustrate the three-tier client-server architecture.
(b) (3 points) Give an example of software that fits in each tier respectively.
(c) (2 points) List two advantages of the middle tier in the three-tier architecture.

(a) 

(b) Client: firefox, Web server: Apache, Database server: MySQL
(c) Separate the client interface from the database server
    • Enhance Security

6. (8 points) Please draw the ER diagram for a pizza store which can store the information of customers, place, and order and each order can contain different types of pizza. Each of entity in the ER diagram should have at least 3 attributes.

7. (10 points) Draw the EER diagram for players in a game. Players are divided into four groups: student, professional, amateur, and senior. The common attributes for them are: player number, name, age, type of players, and country. The
specific attribute for a student is the school a student is studying. The specific attribute for professional players is the year of profession. Each amateur player can join different clubs and each club can be joined by different amateur players.