MIDTERM #2: INFORMATION SYSTEMS (INDE499B)

Dr. Jennifer Turns

Autumn 2000

Name: _____

General Instructions:

a. Total Time: You will have a total of 50 minutes for this midterm.

b. Point Values: The questions are weighted differently. The point values are listed next to each question. Please take this into account when pacing yourself.

c. Use of Resources: This is a close book, closed note test – with one exception. You may use the two page "logical design" handout (i.e., the one describing the steps required to transform an ER diagram into a relational database schema).

d. A Hint: Be certain to read the questions carefully and respond to all portions of the question.

e. GOOD LUCK

QUESTION 1: THE DATA DESIGN PROCESS (25 points).

Imagine that you have been assigned to a team that will be developing an inventory tracking system. As part of the project startup, your manager has asked each team leader to bring a basic work plan to the next meeting. At that meeting, these work plans will be analyzed to determine the overall project timeframe, costs, personnel requirements and software requirements.

For now, as the team leader for the data design team, you have been asked to bring a work plan that identifies the phases of data design and includes the following information for each phase:

a). a description of the data design phase,
b). the inputs of the phase,
c). the outputs of the phase,
d). a key issue addressed in the phase
e). a challenge that you can anticipate would occur in the phase.

Please prepare the response you will bring to the meeting.

Solution:

Description	Issue	Input	Output	Challenge
a. Conceptual Design Create model that captures major entities, relationships among entities, and attributes of entities required for a particular system.	- Capturing all data - Capturing relationships - Data integrity	- Functional specs - General understanding of problem	- ER diagram	
b. Logical Design Transform the major entity/attribute /relationship requirements into high level specification for database	- Providing location for all data - Data integrity	- ER diagram	- Relational database schema	-
b2 Improving Logical Design Improve the high-level database specification.	- Minimizing redundancy - Minimizing ambiguity	- Relational database schema	- Relational database schema	-
c. Physical Design Transform the high-level specifications for database into detailed specifications for how to construct actual database in a specific relational database software.	- Performance - Data integrity	- Relational database schema - Meaning of data	- Technical specifications for construction of the database	-

Grading:

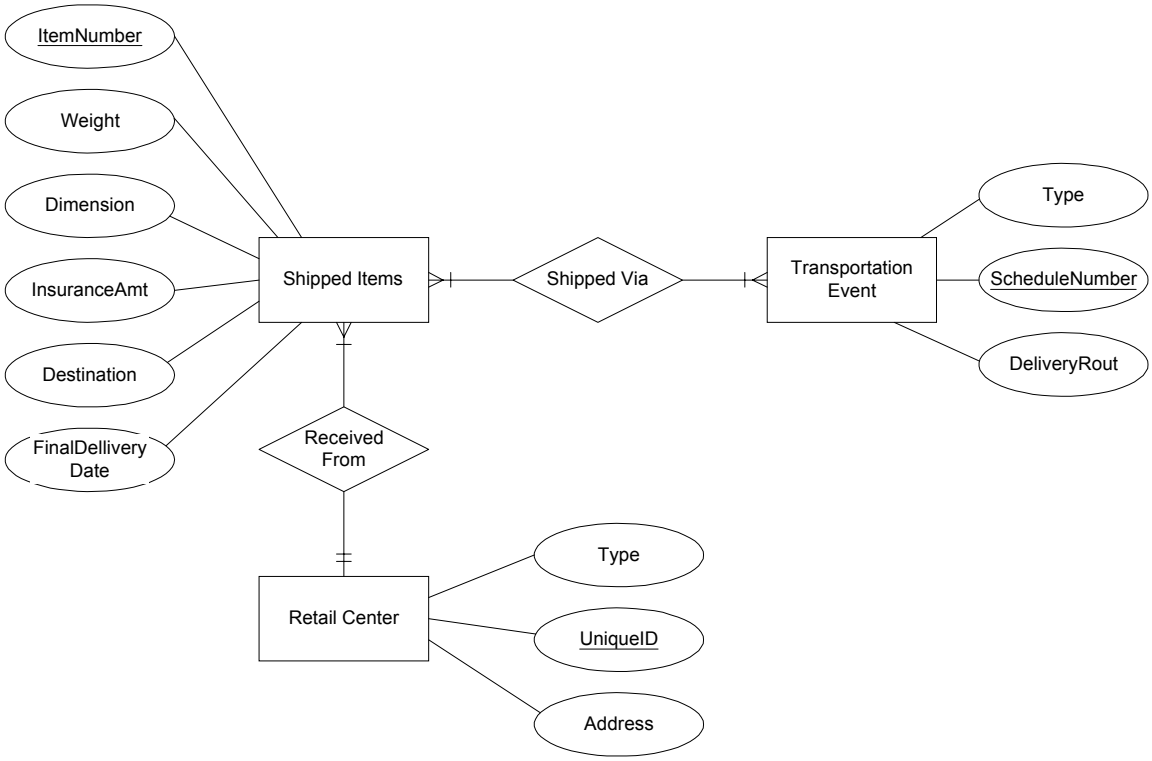
Conceptual Design
Description/issue – 3 pts, input/output – 3 pts, challenge – 1 pt
Logical design:
Description/issue – 3 pts, input/output – 3 pts, challenge – 1 pt
Mentioned improving logical design (can be embedded in the logical design section):
Description/issue - 2 pts, input/outputs - 2 pts
Physical Design:
Description/issue – 3 pts, input/output – 3 pts, challenge – 1 pt

QUESTION 2: CREATING AN ENTITY-RELATIONSHP DIAGRAM (25 points).

UPS prides itself on having up-to-date information on the processing and current location of each shipped item. To do this, UPS relies on a company-wide information system. Shipped items are the heart of the UPS product tracking information system. Shipped items can be characterized by item number (unique), weight, dimensions, insurance amount, destination, and final delivery date. Shipped items are received into the UPS system at a single retail center. Retail centers are characterized by their type, uniqueID, and address. Shipped items make their way to their destination via one or more standard UPS transportation events (i.e., flights, truck deliveries). These transportation events are characterized by a unique scheduleNumber, a type (e.g, flight, truck), and a deliveryRoute.

Please create an Entity Relationship diagram that captures this information about the UPS system. Be certain to indicate identifiers and cardinality constraints.

Solutions:

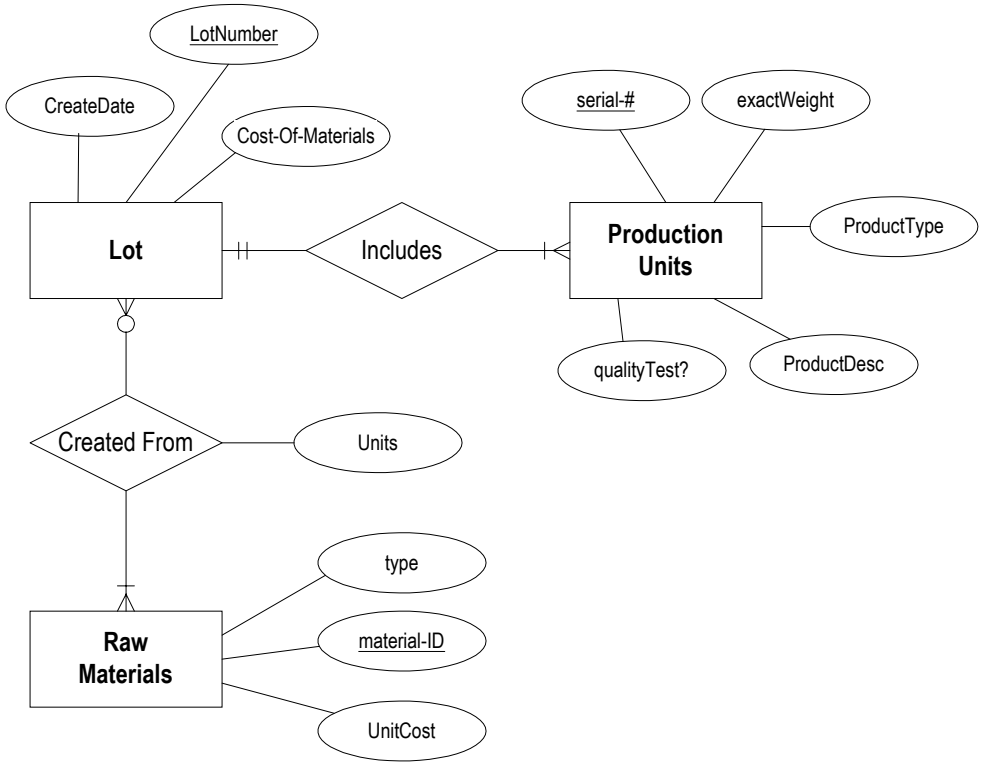


Grading:

Entities correctly identified: 5
Attributes correctly identified: 5
Primary keys correctly identified: 5
Relationships and cardinality correctly identified: 10

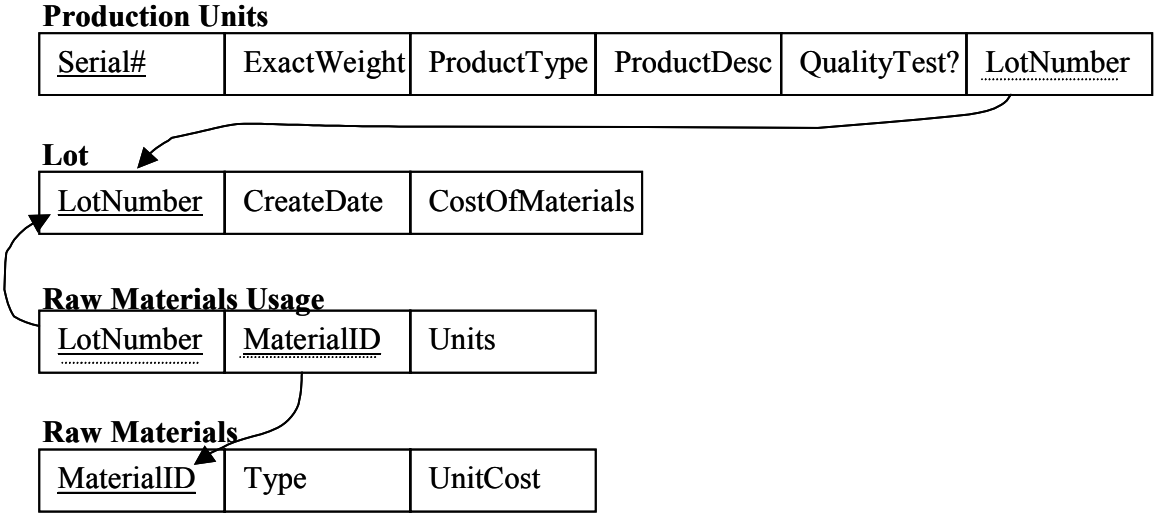
QUESTION 3: CREATING A RELATIONAL DATABASE SCHEMA (37 points).

Production tracking is important in many manufacturing environments (e.g., the pharmaceuticals industry, children's toys, etc.). The following ER diagram captures important information in the tracking of production. Specifically, the ER diagram captures relationships between production lots (or batches), individual production units, and raw materials.



a. Please convert the ER diagram into a relational database schema. Be certain to indicate primary keys and referential integrity constraints.
(25 pts)

Solutions:



Grading:

Strong Entities – Production Units, Lot, Raw Materials: 3 points each
Exists, Has attributes, Has Primary key defined.

One to Many Relationship – Lot Number as Foreign Key on Production Unit Entity: 4 points
Exists or not

Many to Many Relationship - Relation for Raw Material Usage: 7 points
Relation exists, Has primary key correctly identified, has additional attribute

Referential Integrity Constraints Correct - 5 points

Note: The following questions refer back to the above ER diagram for problem 3.

b. Please identify an attribute in the above ER diagram that might represent a composite attribute, and explain why/how it might represent a composite attribute (**3 pts**).

Answer: Many of the attributes **could** actually represent composite attributes:
- Weight might be stored as pounds and ounces.
- Product Description might have several components
- Product type might have several components
- Create date might be stored as both date and time.

c. Please identify an attribute in the ER diagram that could represent a derived attribute and explain why/how it might represent a derived attribute (**3 pts**).

Answer: Cost-of-Materials (associated with the Lot entity) most likely represents a derived attribute. The cost of materials could be computed based on the materials unit cost (from the raw materials entity) and the number of units required for a lot (on the relationship).

d. The ER diagram/relational database schema contains several instances of data redundancy. Please identify one instance where a data redundancy issue exists. (**3 pts**).

Answer: Examples include:
- Both production description and product type are being stored for each and every production unit. The production description could be stored elsewhere.
- Production type and production description are being stored on each and every production unit, when these parameters are most likely the same for the entire lot. They could be stored on the Lot entity.

e. The current ER diagram has the following relationship, "raw materials are used in 0 to many lots." Please explain, in the context of the manufacturing environment, how the meaning changed if the minimal cardinality is changed to "1" (i.e., the relationship becomes "raw materials are used in 1 to many lots.") (**3 pts**)

Answer: The current representation of "raw materials are used in 0 to many lots" implies that raw materials can be in the system without being designated for a specific lot. If the minimum cardinality was changed to 1, this would imply that all raw materials must be designated as related to at least one lot.

QUESTION 4: PROBING (13 points).

a. **Data Integrity**: Ensuring data integrity is an important issue in data design. Data integrity issues are addressed in each and every stage of data design. According to the book, integrity constraints "facilitate maintaining the accuracy of data in the database." According to Merriam Webster's dictionary, synonyms for integrity include incorruptibility, soundness, and completeness.

Please identify two features of data design that embody data integrity concerns. For each, describe how the feature is related to the general issue of data integrity. The examples can come from the same and/or different phases of the data design process. (**8 pts - 4 pts each**)

Examples of answers include:

- Minimum cardinality constraints (identified during conceptual design) indicate the minimum number of relationships that MUST exist.
- During physical design, the definition of each attribute can include constraints on the values of the attribute.
- During logical design, the specification of referential integrity constraints indicate the required relationships among foreign keys and primary keys.

Grading:
Is each example related to integrity? 2 points
Is the relationship to data integrity well explained? 2 points

b. **Normalization**. Normalization is a process within logical design. What is the general goal of normalization? (Hint: Why is it important to identify and remove partial and transitive functional dependencies during normalization?) (**5 points**)

Answer:
The general goal is to remove redundancy in the data!

Grading:
Full credit if redundancy is explicitly or implicitly mentioned.
Partial credit if on the right track.