
Mock Exam: CPE241 Database Systems (Module 1)

Lectures 1-5: Advanced Theory & Modeling Analysis

ULTRA HARD MODE - Senior Layout Format

Part 1: Multiple Choice (English) - 20 Marks

1. In the context of Concurrency Control, a "Wait-for Graph" is primarily used to detect:
 - a) Cascading rollbacks
 - b) Deadlock
 - c) Phantom reads
 - d) Dirty writes
2. If you change the underlying data structures (e.g., using B-trees instead of Hashing) without affecting the Conceptual Schema, you are demonstrating:
 - a) Physical Data Independence
 - b) Logical Data Independence
 - c) Internal Consistency
 - d) View Authorization
3. An "Alternate Key" is defined as:
 - a) A Foreign Key that points to a non-primary key
 - b) A Candidate Key not chosen as the Primary Key
 - c) A key that is used only for sorting
 - d) A composite key with redundant attributes
4. What is the critical difference between a Weak Entity and a Strong Entity?
 - a) Weak entities cannot have relationships
 - b) Weak entities lack sufficient attributes to form a primary key of their own
 - c) Weak entities must have double-circle attributes
 - d) Weak entities are only used in 1:1 relationships
5. The SQL **DROP TABLE** statement belongs to which sub-language category?
 - a) Data Manipulation Language (DML)
 - b) Data Definition Language (DDL)
 - c) Data Control Language (DCL)
 - d) Transaction Control Language (TCL)
6. How is a "Multi-valued Attribute" typically handled when mapping to a Relational Schema?

- a) By adding multiple columns to the existing table
 - b) By creating a new table and linking it via a Foreign Key
 - c) By merging it into the Primary Key
 - d) By ignoring the attribute
7. In ER modeling, a relationship with a degree of three represents:
- a) A recursive relationship
 - b) A triple-binary relationship
 - c) A ternary relationship
 - d) A high-level generalization
8. Which notation represents "Total Participation" in a standard Chen ER Diagram?
- a) Single line
 - b) Dotted line
 - c) Double line
 - d) Bold arrow
9. The DBMS component that stores metadata and constraints is known as the:
- a) Query Optimizer
 - b) System Catalog / Data Dictionary
 - c) Transaction Log
 - d) Storage Engine
10. To store a Many-to-Many (M:N) relationship between two entities, how many tables are required at minimum?
- a) 1
 - b) 2
 - c) 3
 - d) 4
11. Which component of a DBMS is primarily responsible for ensuring Atomicity and Durability?
- a) Query Processor
 - b) Buffer Manager
 - c) Transaction Manager
 - d) Network Manager
12. The Referential Integrity rule states that:
- a) A Primary Key must be unique
 - b) A Foreign Key must match a Primary Key in the referenced table, or be NULL
 - c) Every table must have a Primary Key

- d) Data types must be consistent across the database
13. In the 3-Schema Architecture, physical file organization and indexing are defined in the:
- External Schema
 - Conceptual Schema
 - Internal Schema
 - User Schema
14. A "Composite Attribute" is defined as an attribute that:
- Can have multiple values for a single entity
 - Can be further subdivided into smaller component attributes
 - Is derived from other stored attributes
 - Is a combination of two primary keys
15. The process of converting requests between different levels of the database architecture is called:
- Optimization
 - Mapping
 - Normalization
 - Encapsulation
16. In an EER Diagram, a circle with an "o" symbol represents which constraint?
- Disjoint constraint
 - Overlapping constraint
 - Total participation
 - Partial participation
17. How does a "Derived Attribute" differ from a "Stored Attribute"?
- It is not physically stored in the database
 - It is always a primary key
 - It cannot be updated
 - It is only used in weak entities
18. "Data Isolation" in traditional file systems refers to the problem where:
- Data is too secure
 - Data is spread across different files and formats, making it difficult to link
 - Data is duplicated excessively
 - Data is lost during transmission
19. In Crow's Foot notation, a small circle on the relationship line represents:
- Mandatory (1)

- b) Optional (0)
 - c) Many
 - d) Primary Key
20. When a crash occurs, which log is used by the DBMS to perform a recovery?
- a) Error Log
 - b) Audit Log
 - c) Transaction Log (Undo/Redo Log)
 - d) Metadata Log

Part 2: Analytical Short Answer (ภาษาไทย) - 20 Marks

1. จงอธิบายปราการณ์ "Lossless Join" และความสำคัญในการรักษาความถูกต้องของข้อมูล
2. "Deadlock" เกิดขึ้นได้อย่างไรในระบบฐานข้อมูล และ DBMS มีวิธีจัดการเบื้องต้นอย่างไร?
3. "Natural Join" ต่างจาก "Inner Join" ทั่วไปอย่างไรในเชิงโครงสร้าง (Lect. 4)?
4. ในการออกแบบ ERD, การเลือกรหัสว่าง "Attribute" และ "Entity" มีเกณฑ์การตัดสินใจอย่างไร?
5. จงระบุความแตกต่างระหว่าง "Generalization" และ "Specialization" ใน EER Modeling
6. เพาะเหตุใดจึงควรใช้ BCNF แทนที่ 3NF ในบางกรณี (อธิบายหลักการเบื้องต้น)?
7. จงอธิบาย "Data Abstraction" 3 ระดับ และความสัมพันธ์กับ Data Independence
8. "Weak Entity" สามารถมี Weak Entity ของตัวเองซ้อนลงไปอีกขั้นได้หรือไม่? จงยกตัวอย่างประกอบ
9. จงบอกความหมายของ "Surrogate Key" และสถานการณ์ที่ควรหลีกเลี่ยงการใช้งาน
10. ใน EER, "Disjoint Constraint (d)" ช่วยลดปัญหา Inconsistency ได้อย่างไร?

Part 3: Practical Design Challenge (ภาษาไทย) - 20 Marks

ข้อที่ 1. High-Level SQL & Schema Implementation (10 Marks): จงเขียนชุดคำสั่ง SQL เพื่อสร้างโครงสร้างพื้นฐานสำหรับระบบ "Order Management" โดยสังเกตจากสถานะข้อมูลในตารางต่อไปนี้:

Table: CUSTOMERS

cid (PK)	name	email (Unique)
C01	Somsak	s@kmutt.ac.th
C02	Somying	y@kmutt.ac.th

Table: ORDERS

oid (PK)	cid (FK)	total_amt
OR-99	C01	1,500.00
OR-88	C02	2,200.00

Table: LINE_ITEMS (Composite Primary Key)

oid (FK)	item_id	prod_name	qty (CHECK > 0)
OR-99	1	Laptop	1
OR-99	2	Mouse	2
OR-88	1	Keyboard	1

โจทย์: จงเขียน SQL DDL และ DML เพื่อดำเนินการดังนี้:

- i) สร้างตารางทั้ง 3 โดยมีเงื่อนไข: **total_amt** ห้ามน้อยกว่า 0, **email** ห้ามซ้ำ, และถ้าลบ Order ข้อมูลใน Line Items ต้องหายไปด้วย (ON DELETE CASCADE)
- ii) เพิ่มคอลัมน์ **status** (Default = 'Pending') ในตาราง ORDERS โดยใช้คำสั่ง ALTER
- iii) เขียนคำสั่ง INSERT ข้อมูลลงในตาราง LINE_ITEMS สำหรับแถวที่มี **oid** 'OR-99' และ **item_id** 1

พื้นที่เขียนคำตอบ:

```
-- 1. Create Tables
-- 2. Alter Table
-- 3. Insert Data
```

ข้อที่ 2. Advanced EER Modeling (10 Marks): จง画 Conceptual Diagram สำหรับระบบ "Fleet Management":

- Vehicle มี ChassisID, Brand
- แบ่งเป็น Car (Doors) และ Truck (Capacity) แบบ Disjoint & Total (Mandatory)
- ยานพาหนะต้องมี Owner (รหัส, ชื่อ) 1 คน รถ 1 คัน มีเจ้าของเดียวหลายคน (M:N) โดยต้องเก็บวันที่เริ่มครอบครอง (StartDate)
- Maintenance เป็น Weak Entity ขึ้นกับ Vehicle (เลขที่ซ่อม, ค่าใช้จ่าย)

(วาด EER Diagram ตรงนี้ - เน้นความถูกต้องของ Hierarchy และ Weak Entity)

เฉลยละเอียด (ULTRA HARD MODE Senior Solution)

Part 1: MCQ Explanations

No.	Ans	Explanation
1	b	Wait-for Graph เป็นกราฟที่แสดงการรอทัวพยากรณ์ระหว่าง Transaction ถ้าเป็นวงกลม (Cycle) แสดงว่าเกิด Deadlock
2	a	Physical Data Independence คือการเปลี่ยนโครงสร้างระดับ Internal (เช่น Index, File structure) โดยไม่กระทบ Conceptual
3	b	Alternate Key หรือ Candidate Key ที่มีคุณสมบัติระบุว่าหนึ่งเดียวแล้วแต่กูกเลือกเป็น Primary Key
4	b	Weak Entity ขาด identity ของตัวเอง จำเป็นต้องอาศัย Owner Entity และ Partial Key เพื่อสร้างคีย์หลักชุดสมบูรณ์
5	b	DDL (Data Definition Language) จัดการโครงสร้าง Schema; DROP คือการถอนโครงสร้างทั้งหมดออกจากระบบ
6	b	1NF กำหนดให้ข้อมูลในคอลัมน์ต้องเป็น Atomic การมี Multi-valued ต้องแยกตารางเพื่อลด Redundancy
7	c	Degree = จำนวน Entity Types ที่เข้าร่วมใน Relationship; 3 = Ternary
8	c	เส้นคู่ (Double Line) แทนการมีส่วนร่วมแบบเบ็ดเสร็จ (Total Participation) หรือ Mandatory
9	b	System Catalog/Data Dictionary เก็บ Metadata ทั้งหมดของฐานข้อมูล
10	c	ตาราง A + ตาราง B + ตาราง Junction (Link Table) สำหรับ Many-to-Many = 3 ตาราง
11	c	Transaction Manager ดูแลเรื่องความถูกต้องของธุรกรรมตามกฎ ACID (Atomicity, Isolation, Durability)
12	b	Referential Integrity บังคับความสัมพันธ์ระหว่างคีย์ FK ต้องอ้างไปเจอ PK ที่มีอยู่จริงเพื่อความปลอดภัย
13	c	Internal Schema จัดการเรื่องการเก็บข้อมูลทางภาษาพาร์เซอร์ เช่น การระบุที่อยู่ของ Record บน Disk
14	b	Composite attribute คือคุณลักษณะที่ยังสามารถแบ่งย่อยลงได้อีก เช่น ชื่อ (ประกอบด้วยชื่อจริงและนามสกุล)
15	b	Mapping คือการระบุความซึ่อมโยงระหว่างชั้นต่างๆ ของ 3-Schema Architecture
16	b	Overlap (o) หมายถึงข้อมูลหนึ่งทัวสามารถเป็นสมาชิกของหลาย Subclass ได้ในเวลาเดียวกัน
17	a	Derived Attribute คือค่าที่ได้จากการคำ算法 (Runtime calculation) จึงไม่เก็บลงดิสก์จริง
18	b	Data Isolation ในระบบเก่าคือข้อมูลที่จะจัดกระจายอยู่ในหลายไฟล์และหลายฟอร์แมต ทำให้ข้อมูลยาก
19	b	วงกลม (Circle) ใน Crow's Foot หมายถึงค่า 0 หรือ Optional
20	c	Transaction Log บันทึกทุกการเปลี่ยนแปลงเพื่อใช้ในการ Undo/Redo เมื่อระบบล้มหรือต้องการ Rollback

Part 2: Analytical Short Answer Key

1. **Lossless Join:** คือเงื่อนไขที่เมื่อเรา Decompose ตารางหนึ่งออกเป็นสองตารางแล้ว เมื่อนำมา Join กันกลับด้วยเงื่อนไขที่ถูกต้อง ผลลัพธ์ต้องได้เหมือนเดิมเป็น หากเกิดข้อมูลเกิน (Spurious Tuples) จะเป็น Lossy ซึ่งทำลายความถูกต้องของข้อมูล

2. **Deadlock:** เกิดเมื่อ T1 รอ B (ที่ T2 ถือ) และ T2 รอ A (ที่ T1 ถือ) วนเป็นวงกลม วิธีแก้คือ Deadlock Detection (Wait-for graph) แล้วทำ Rollback ตัวที่มีค่าความเสียหายน้อยที่สุด (Victim selection)
3. **Natural Join vs Inner Join:** Natural Join จะทำการเบรี่ยบเทียบคอลัมน์ที่ซ้ำกันโดยอัตโนมัติ และกำจัดคอลัมน์ที่ซ้ำกันให้เหลือเพียงอันเดียว ต่างจาก Inner Join ที่มักจะคงไว้ทั้งสองฝั่งหากไม่ได้ระบุ SELECT
4. **Attribute vs Entity:** พิจารณาจาก ”รายละเอียด” หากต้องการเก็บข้อมูลอยู่ๆ ของสิ่งนั้น (Attributes) และมันมีความสัมพันธ์กับคนอื่น สิ่งนั้นควรเป็น Entity แต่หากเป็นเพียงค่าเดียว สิ่งนั้นเป็น Attribute
5. **Gen vs Spec: Specialization** (บันลงล่าง) แยกกลุ่มใหญ่ออกเป็นกลุ่มย่อยที่เจาะจง Generalization (ล่างขึ้นบน) รวมกลุ่มย่อยที่มีจุดร่วมกันขึ้นเป็นกลุ่มฐานข้อมูลเดียวกัน
6. **BCNF vs 3NF:** 3NF ยอมให้ Determinant ไม่เป็น Superkey ได้หากเป็นส่วนของคีย์ CK แต่ BCNF บังคับว่า ”ทุก Determinant ต้องเป็น Superkey” แห่งนั้นเพื่อกำจัดความซ้ำซ้อนระดับสูง
7. **Data Abstraction:** External (หน้าจอส榭อร์), Conceptual (โครงสร้างเนื้อหา), Internal (การเก็บจริง) การแยกขั้นชั้วยield แก้ไขส่วนล่างได้โดยไม่กระทบส่วนบน (Data Independence)
8. **Multi-level Weak Entity:** ได้ เช่น Building (Strong) -> Room (Weak) -> Computer (Weak ขึ้นกับ Room) โดย Computer ต้องมี PK ที่รวมมาจากทั้ง Building และ Room
9. **Surrogate Key:** คีย์ที่ระบบสร้างให้เอง (Auto-increment) ข้อดีคือเร็วและแน่นอน แต่ควรเลี่ยงหากข้อมูลมี Natural Key ที่มั่นคงและสื่อความหมายชัดเจนอยู่แล้ว เพื่อลดภาระการ JOIN ตาราง
10. **Disjoint (d):** บังคับว่า Instance หนึ่งย้ายไปอยู่ได้เพียง Subclass เดียวเท่านั้น ป้องกันการทำข้อมูลซ้ำซ้อนและซ้ำๆให้การ Check Inconsistency ทำได้ง่ายขึ้น

Part 3: Practical Solution Details

1. Advanced SQL Solution:

```
-- i) Create Tables with Constraints
CREATE TABLE CUSTOMERS (
    cid CHAR(3) PRIMARY KEY,
    name VARCHAR(50) NOT NULL,
    email VARCHAR(100) UNIQUE
);

CREATE TABLE ORDERS (
    oid CHAR(5) PRIMARY KEY,
    cid CHAR(3) REFERENCES CUSTOMERS(cid),
    total_amt DECIMAL(10,2) CHECK (total_amt >= 0)
);

CREATE TABLE LINE_ITEMS (
    oid CHAR(5) REFERENCES ORDERS(oid) ON DELETE CASCADE,
    item_id INT,
    prod_name VARCHAR(50),
    qty INT CHECK (qty > 0),
    PRIMARY KEY (oid, item_id)
);

-- ii) Alter Table
ALTER TABLE ORDERS ADD status VARCHAR(20) DEFAULT 'Pending';

-- iii) Insert Data
INSERT INTO LINE_ITEMS (oid, item_id, prod_name, qty)
VALUES ('OR-99', 1, 'Laptop', 1);
```

2. EER Diagram Solution (TikZ Logic):

