

Written by Dr. Thapanapong Rukkanchanunt

09 Data Modeling

OUTLINE

- Conceptual Data Modeling
 - Structure
 - Operation
 - Constraint
- Old Data Model
 - Hierarchal Model
 - Network Model
- Relational Model

Data Modeling

- Data Model คือคำอธิบายของข้อมูลหรือสารสนเทศ โดยส่วนมากแล้วคำอธิบายจะแบ่งออกเป็นสามส่วนใหญ่ๆ
 - โครงสร้างของข้อมูล (Structure) คือโครงสร้างทางกายภาพหรือโครงสร้างข้อมูลที่ใช้ในการเก็บข้อมูลในคอมพิวเตอร์
 - ตัวดำเนินการของข้อมูล (Operation) ประกอบไปด้วยคำสั่งที่เรียกดูข้อมูล (Query) และคำสั่งแก้ไขฐานข้อมูล (Modification)
 - ข้อจำกัดของข้อมูล (Constraint) คือกฎที่ข้อมูลจะประพฤติตาม โดยส่วนมากแล้วข้อจำกัดจะถูกกำหนดโดยความต้องการของแอปที่ใช้ในโลกแห่งความเป็นจริง

Old Types of Data Models (Structure only)

- Hierarchal Model คือ Data Model ที่เรียงข้อมูลในรูปแบบ Tree
 - ข้อมูลถูกจัดเก็บเป็น Record ซึ่งสัมพันธ์กันด้วย Link
 - Link ภายในตารางเดียวกันหรือระหว่างตารางได้



- Network Model ขยายต่อจาก Hierarchal Model โดยอนุญาตให้มีความสัมพันธ์แบบ Many-to-Many นั่นคือโหนดในโครงสร้างสามารถมี Parent ได้มากกว่าหนึ่ง
 - ใช้สองคอนเซ็ปต์หลักคือ Record และ Set
 - Record ประกอบไปด้วย Field
 - Set กำหนดความสัมพันธ์แบบ One-to-Many ระหว่าง Record
 - หนึ่ง Record สามารถอยู่ได้ในหลาย Set



Relational Database

- Relational Database เป็นไอดีที่เริ่มมาจาก E. F. Codd ในปี 1970 ในเปเปอร์ที่ชื่อ “A relational model of data for large shared data banks”
- โดยส่วนมากแล้วข้อมูลใน Relational Database จะอยู่ในรูปแบบของตาราง แต่จริง ๆ แล้วมีการกำหนดเงื่อนไขต่าง ๆ นอกเหนือจากตารางด้วย
- ตัวดำเนินการที่กระทำบนฐานข้อมูลมาจากตัวดำเนินการทางคณิตศาสตร์ มีชื่อเรียกว่า Relational Algebra หรือ Relational Calculus
 - ใช้ภาษา Structured Query Language สำหรับการกระทำบนฐานข้อมูล

Learn from Example

Bank Account เก็บข้อมูลบัญชีธนาคาร

Number	Owner	Balance	Type
101	J. Smith	1000.00	Checking
102	W. Wei	2000.00	Checking
103	J. Smith	5000.00	Savings
104	M. Jones	1000.00	Checking
105	H. Martin	10000.00	Checking

Table Structure

ชื่อตาราง

ชื่อคอลัมน์ (Attribute)

Bank Account เก็บข้อมูลบัญชีธนาคาร

Number	Owner	Balance	Type
101	J. Smith	1000.00	Checking
102	W. Wei	2000.00	Checking
103	J. Smith	5000.00	Savings
104	M. Jones	1000.00	Checking
105	H. Martin	10000.00	Checking

Table Schema

Schema คือโครงสร้างของตาราง ประกอบไปด้วยชื่อตาราง Attribute และประเภทของข้อมูล

Bank Account

Number	Owner	Balance	Type
101	J. Smith	1000.00	Checking
102	W. Wei	2000.00	Checking
103	J. Smith	5000.00	Savings
104	M. Jones	1000.00	Checking
105	H. Martin	10000.00	Checking

Table Rows

Bank Account

Number	Owner	Balance	Type
101	J. Smith	1000.00	Checking
102	W. Wei	2000.00	Checking
103	J. Smith	5000.00	Savings
104	M. Jones	1000.00	Checking
105	H. Martin	10000.00	Checking

แต่ละแถวของตารางเรียกว่า Tuple หรือ Record

Intension vs Extension

Intension

Bank Account

Number	Owner	Balance	Type
101	J. Smith	1000.00	Checking
102	W. Wei	2000.00	Checking
103	J. Smith	5000.00	Savings
104	M. Jones	1000.00	Checking
105	H. Martin	10000.00	Checking

Extension

Size

Degree หรือ Arity คือจำนวนคอลัมน์

Bank Account

Number	Owner	Balance	Type
101	J. Smith	1000.00	Checking
102	W. Wei	2000.00	Checking
103	J. Smith	5000.00	Savings
104	M. Jones	1000.00	Checking
105	H. Martin	10000.00	Checking

Cardinality คือจำนวนแถวของข้อมูล

Database (More Tables)

Account	Number	Owner	Balance	Type
	101	J. Smith	1000.00	Checking
	102	W. Wei	2000.00	Checking
	103	J. Smith	5000.00	Saving
	104	M. Jones	1000.00	Checking
Deposit	AcctNo	Transaction-id	Date	Amount
	102	1	10/22/00	500.00
	102	2	10/29/00	200.00
	104	3	10/29/00	1000.00
Check	AcctNo	Check-number	Date	Amount
	101	924	10/23/00	125.00
	101	925	10/24/00	23.98

แต่ละตารางจะมีคอลัมน์ที่เป็น Key
โดยค่าในคอลัมน์นั้นจะต้องไม่ซ้ำกัน

Table Keys

Account	Number	Owner	Balance	Type
	101	J. Smith	1000.00	Checking
	102	W. Wei	2000.00	Checking
	103	J. Smith	5000.00	Saving
	104	M. Jones	1000.00	Checking
Deposit	AcctNo	Transaction-id	Date	Amount
	102	1	10/22/00	500.00
	102	2	10/29/00	200.00
	104	3	10/29/00	1000.00
Check	AcctNo	Check-number	Date	Amount
	101	924	10/23/00	125.00
	101	925	10/24/00	23.98

Connection between Tables

เป็นไปได้หรือไม่ ถ้าไม่ เราควรจะทำ
ป้องกันอย่างไรได้บ้าง

Account	Number	Owner	Balance	Type
	101	J. Smith	1000.00	Checking
	102	W. Wei	2000.00	Checking
	103	J. Smith	5000.00	Saving
	104	M. Jones	1000.00	Checking
Deposit	AcctNo	Transaction-id	Date	Amount
	102	1	10/22/00	500.00
	102	2	10/29/00	200.00
	104	3	10/29/00	1000.00
Check	AcctNo	Check-number	Date	Amount
	101	924	10/23/00	125.00
	101	925	10/24/00	23.98
	105	926	10/25/00	555.55

เราสามารถกำหนด Referential Integrity Constraint ว่า
Check.AcctNo เป็น Foreign Key ของ Account.Number

Foreign Key

Account	Number	Owner	Balance	Type
	101	J. Smith	1000.00	Checking
	102	W. Wei	2000.00	Checking
	103	J. Smith	5000.00	Saving
	104	M. Jones	1000.00	Checking
Deposit	AcctNo	Transaction-id	Date	Amount
	102	1	10/22/00	500.00
	102	2	10/29/00	200.00
	104	3	10/29/00	1000.00
Check	AcctNo	Check-number	Date	Amount
	101	924	10/23/00	125.00
	101	925	10/24/00	23.98
	105	926	10/25/00	555.55

นอกจาก Check.AcctNo เป็น Foreign Key ของ Account.Number แล้ว Deposit.AcctNo ก็เป็น Foreign Key ของ Account.Number ด้วย

Foreign Keys

Account	Number	Owner	Balance	Type
	101	J. Smith	1000.00	Checking
	102	W. Wei	2000.00	Checking
	103	J. Smith	5000.00	Saving
	104	M. Jones	1000.00	Checking
Deposit	AcctNo	Transaction-id	Date	Amount
	102	1	10/22/00	500.00
	102	2	10/29/00	200.00
	104	3	10/29/00	1000.00
Check	AcctNo	Check-number	Date	Amount
	101	924	10/23/00	125.00
	101	925	10/24/00	23.98
	105	926	10/25/00	555.55

Underlining Keys

โดยปกติแล้ว Key มักจะถูกขีดเส้นใต้

Account	<u>Number</u>	Owner	Balance	Type
	101	J. Smith	1000.00	Checking
	102	W. Wei	2000.00	Checking
	103	J. Smith	5000.00	Saving
	104	M. Jones	1000.00	Checking
Deposit	<u>AcctNo</u>	Transaction-id	Date	Amount
	102	1	10/22/00	500.00
	102	2	10/29/00	200.00
	104	3	10/29/00	1000.00
Check	<u>AcctNo</u>	<u>Check-number</u>	Date	Amount
	101	924	10/23/00	125.00
	101	925	10/24/00	23.98

Back to Schema

- อย่างที่เราทราบแล้วว่า Schema คือโครงสร้างของตาราง ประกอบไปด้วยชื่อตาราง Attribute และประเภทของข้อมูล โดยลำดับการประกาศ Schema มีดังนี้
 1. กำหนดชื่อให้ตาราง
 2. แต่ละคอลัมน์ของตาราง กำหนด Domain ของคอลัมน์นั้น
 3. เลือกคอลัมน์เป็น Key (เลือกได้มากกว่าหนึ่งคอลัมน์)
 4. กำหนด Foreign Key ตามความเหมาะสม

Column Domain

- Schema จะต้องระบุว่าข้อมูลของแต่ละคอลัมน์มีค่าได้แก่แบบบ้าง เช่น
 - Number ต้องเป็นเลขสามหลัก
 - Owner ต้องเป็นอักขระ 30 ตัวอักษร
 - Type ต้องเป็น Checking หรือ Saving เท่านั้น

Account	<u>Number</u>	Owner	Balance	Type
	101	J. Smith	1000.00	Checking
	102	W. Wei	2000.00	Checking
	103	J. Smith	5000.00	Saving

- ช่วงข้อมูลที่เป็นไปได้หรือค่าที่เป็นไปได้จะเรียกว่า Domain

Example Database Schema

- ใน Relational Model เราจะเขียนเป็น relation(attribute:domain)
 - Account(Number:int, Owner:string, Balance:float, Type:string)
 - Deposit(AcctNo:int, Transaction-id:int, Date:date, Amount:float)
 - Check(AcctNo:int, Check-number:int, Date:date, Amount:float)

Account	<u>Number</u>	Owner	Balance	Type
Deposit	<u>AcctNo</u>	Transaction-id	Date	Amount
Check	<u>AcctNo</u>	<u>Check-number</u>	Date	Amount

- Schema ข้างต้นไม่รวม Constraint / Domain

Operation

- ตัวดำเนินการบนฐานข้อมูลจะถูกกำหนดโดยภาษาที่ใช้ติดต่อกับฐานข้อมูล
- สำหรับ Relational Model เรามักจะใช้ Structured Query Language (SQL)
 - DDL (Data Definition Language) CREATE, ALTER, DROP
 - DML (Data Manipulation Language) SELECT, INSERT, UPDATE, DELETE
- ในบทเรียนนี้เราจะทำความรู้จักกับ Single-Relation SQL Query และ Multi-Relation SQL Query

Single-Relation SQL Query

- รูปแบบของภาษาที่ใช้ Query ความสัมพันธ์แบบเดี่ยวมีลักษณะดังนี้

SELECT <เลือก Attribute ที่ต้องการ>

FROM <เลือกตาราง>

WHERE <เงื่อนไขสำหรับ Tuple ในตาราง>

Example Query

```
SELECT Number, Owner  
FROM Account  
WHERE Type="saving"
```

Number	Owner
103	J. Smith

Account	<u>Number</u>	Owner	Balance	Type
	101	J. Smith	1000.00	Checking
	102	W. Wei	2000.00	Checking
	103	J. Smith	5000.00	Saving
	104	M. Jones	1000.00	Checking

Order of Evaluation

SELECT Number, Owner (3)

FROM Account (1)

WHERE Type="saving" (2)

Number	Owner
103	J. Smith

Account	<u>Number</u>	Owner	Balance	Type
	103	J. Smith	5000.00	Saving

Account	<u>Number</u>	Owner	Balance	Type
	101	J. Smith	1000.00	Checking
	102	W. Wei	2000.00	Checking
	103	J. Smith	5000.00	Saving
	104	M. Jones	1000.00	Checking

Renaming Attributes

- หากเราต้องการเปลี่ยนชื่อ Attribute ที่แสดงผล เราสามารถใช้ “AS <ชื่อใหม่>” ได้

```
SELECT Number, Balance/1000 AS Balance(K)  
FROM Account  
WHERE Type="checking"
```

Number	Balance(K)
J. Smith	1
W. Wei	2
M. Jones	1

WHERE Syntax

- นอกจาก = แล้ว เรายังสามารถใช้เครื่องหมายดำเนินการอื่น ๆ ได้ดังนี้
- การเปรียบเทียบ =, <>, <, >, <=, >=
- ตัวดำเนินการทางคณิตศาสตร์พื้นฐาน +, -, *, /
- ตัวดำเนินการเกี่ยวกับ String =, LIKE
- ตัวดำเนินการเกี่ยวกับการเปรียบเทียบวันเวลา

LIKE Pattern

- สำหรับตัวดำเนินการ LIKE ของ string มีวิธีการใช้งานเพิ่มเติมดังนี้
- รูปแบบพื้นฐาน `<Attribute> LIKE <Pattern>`
`<Attribute> NOT LIKE <Pattern>`
- ส่วนของ Pattern จะอยู่ในเครื่องหมายคำพูด (Quotation)
 - % แทน string อะไรก็ได้
 - _ แทน character อะไรก็ได้
 - เช่น “J%” คือชื่อที่ขึ้นต้นด้วย J

Ordering the Result

- เรียงลำดับผลลัพธ์(จากน้อยไปมาก)ด้วย ORDER BY
- ถ้าต้องการเรียงจากมากไปน้อยให้เพิ่ม DESC
- ถ้าต้องการ Tie Break ให้เพิ่ม Attribute ใน ORDER BY

SELECT *

FROM Account

WHERE Name LIKE 'J%'

ORDER BY Balance

Account	<u>Number</u>	Owner	Balance	Type
	101	J. Smith	1000.00	Checking
	103	J. Smith	5000.00	Saving

Multi-Relation Query

- Query ที่น่าสนใจมักจะเกิดจากการรวมข้อมูลหลายตาราง
- เราสามารถระบุตารางทั้งหมดใน FROM
- หากหลายตารางมีชื่อ Attribute เหมือนกัน เราสามารถแบ่งแยกได้โดย <ตาราง>.<Attribute>

Correlation Name

- เพื่อให้ Query สั้นลง เราสามารถใช้ Correlation Name แทนแต่ละตารางใน FROM

```
SELECT A.Owner, A.Balance
```

```
FROM Account A, Deposit D
```

```
WHERE D.AcctNo = A.Number and A.Balance > 1000;
```

- วิธีการคำนวณเหมือนกับ Single-Relation Query คือ FROM -> WHERE -> SELECT
- จำนวนคอลัมน์ในขั้นตอน WHERE จะเท่ากับจำนวนคอลัมน์ของทุกตารางรวมกัน

Beyond this Course

- เนื้อหาเพิ่มเติมเช่น Subquery, Outer JOIN, Inner JOIN, Cross Product, EquiJOIN, Natural JOIN, Boolean Operation, Set Operation, Grouping, Aggregation จะได้เรียนในวิชา Database
- Query อื่น ๆ เช่น UPDATE, DROP จะไม่ Cover ในบทเรียนนี้