

# ระบบฐานข้อมูล

# Database System

## ตอนที่ 3 ฐานข้อมูลเชิงสัมพันธ์

## Chapter 3 Relational Database

Written by Thawatchai Suksai

ธวัชชัย สุขสาย

ฐานข้อมูลเชิงสัมพันธ์ เป็นรูปแบบฐานข้อมูลที่นิยมในปัจจุบัน ได้ถูกคิดค้นและพัฒนาโดย E.F.Codd เป็นรูปแบบที่เข้าใจง่าย ไม่ซับซ้อนรวมถึงเป็นรูปแบบที่มีเครื่องมือที่ช่วยในการเรียกใช้ข้อมูลด้วยคำสั่งง่ายๆ เนื่องจากผู้ใช้ไม่จำเป็นต้องทราบเกี่ยวกับการจัดเก็บข้อมูลในระดับกายภาพ เช่น ไม่ต้องทราบว่าข้อมูลถูกจัดเก็บอยู่ที่ใด วิธีการเข้าถึงข้อมูลเป็นแบบใด ตัวอย่าง ฐานข้อมูลเชิงสัมพันธ์ DB2 ORACLE Microsoft SQL Server INGRES FOXPRO PostgreSQL MySQL

บทความนี้จะกล่าวถึง คำศัพท์พื้นฐานที่เกี่ยวข้องกับฐานข้อมูลเชิงสัมพันธ์ ความหมายและข้อดีของฐานข้อมูลเชิงสัมพันธ์ คุณลักษณะการจัดเก็บข้อมูลของรีเลชัน ประเภทของคีย์ และกฎที่เกี่ยวข้องกับคีย์ในฐานข้อมูลเชิงสัมพันธ์

Relational Database is one type of database systems highly favored nowadays. It was invented and developed by E.F. Codd. In this system, users doesn't need prior knowledge the system, such as the location of the stored data or the data accessing procedure. Data manipulations are performed via provided tools that help execute instructions; thus, making the system simple to use. The examples of Relational Database are DB2, ORACLE, Microsoft SQL Server, INGRES, FOXPRO, postgresSQL, and MySQL.

This article will address basic relational database terms, their meanings and advantages, their characteristics, key types, and database key-related rules.

### คำศัพท์พื้นฐานที่เกี่ยวข้องกับฐานข้อมูลเชิงสัมพันธ์

### Basic Relational Database Terms

ศัพท์ทั่วไป (Common Terms)	ศัพท์เทคนิค (Technical Terms)
ตาราง (Table)	รีเลชัน (Relation)
แถว (Row) หรือ เรคคอร์ด (record)	ทูเพิล (Tuple)
คอลัมน์ (Column) หรือ ฟیلด์(Field)	แอททริบิวต์ (Attribute)
จำนวนแถว (Number of Rows)	คาร์ดินาลิตี้ (Cardinality)
ค่าเอกลักษณ์ (Unique Identifier)	คีย์หลัก (Primary key)
ขอบข่ายค่าของค่าข้อมูล (Data range)	โดเมน (Domain)

รูปที่ 1 แสดงคำศัพท์ต่างๆที่เรียกใช้ในฐานข้อมูลเชิงสัมพันธ์

Figure 1 Relational Database Terms



จากรูป 1 แสดงให้เห็นถึงคำศัพท์ต่างๆ ที่เรียกใช้ในฐานะข้อมูลเชิงสัมพันธ์ ซึ่งสรุปได้ดังนี้

**รีเลชัน** คำศัพท์ทั่วไปเรียกว่าตาราง (Table) เป็นค่านามที่แทนข้อมูลของเรื่องใดเรื่องหนึ่ง ที่แสดงถึงรูปแบบตาราง 2 มิติ เช่น รีเลชันฝึกอบรม รีเลชันรหัสฝึกอบรม รีเลชันข้อมูลพนักงาน รีเลชันหน่วยงาน

**ทูเปิล** คือค่าของข้อมูลในแต่ละแถว (Row) หรือ ที่เรียกว่าเรคคอร์ด (Record)

**การดีนาลิตี** คือ จำนวนแถวของข้อมูลในแต่ละรีเลชัน

**แอททริบิวต์** คือ รายละเอียดของข้อมูลในแต่ละคอลัมน์ ซึ่งเป็นสิ่งที่แสดงให้เห็นว่ารีเลชันนั้นๆ ประกอบด้วยรายละเอียดข้อมูลอะไรบ้าง

**คีย์หลัก** คือ แอททริบิวต์ที่มีค่าของข้อมูลเป็นเอกลักษณ์หรือเฉพาะเจาะจง จะเป็นค่าที่ไม่ซ้ำกันในแต่ละ ทูเปิล

**โดเมน** คือ ขอบเขตค่าของข้อมูลที่ควรจะเป็น ในแต่ละแอททริบิวต์

เพื่อให้เกิดความเข้าใจศัพท์เทคนิคดังกล่าว จะใช้ตัวอย่างในรูป 2 ซึ่งเป็นฐานข้อมูลการพัฒนาพนักงาน ประกอบด้วย 4 รีเลชัน(2.1-2.4)

Figure 1, conclusion can be made concerning above terms as followings.

**Relation** is often referred to a common term as *Table*. It is a set of data in form of 2 dimensional tables such as Training Relation, Training Code Relation, Staff Relation, and Department Relation.

**Tuple** is the value of data within each row. Sometimes it is called *Record*.

**Cardinality** is the number of row in each relation.

**Attribute** is detail of data in each column. It shows what that relation is consisted of.

**Primary key** is an attribute which is always unique for each individual tuple.

**Domain** is the set of value which attribute can take form.

For better understanding, Figure 2 Staff improvement Program Database (consist of 4 relations i.e.2.1-2.4) are use to explain the mentioned.

## 2.1 รีเลชันฝึกอบรม (Training Relation)

รหัสฝึกอบรม (Training Code)	หลักสูตร (Class)	วันที่เริ่มต้น (Started Date)	วันที่สิ้นสุด (Ended Date)	จำนวนวัน (Number of Days)	หน่วยงานรับผิดชอบ (Organizer)	หมายเหตุ (Note)
11	การสอบเทียบพีเอช มิเตอร์ (pH meter Calibration)	8/12/2548	9/12/2548	2	สำนักพัฒนา ศักยภาพ นักวิทยาศาสตร์ฯ (Bureau of Laboratory Personnel Development)	
12	ข้อกำหนด ISO 9001:2000 (ISO 9001:2000 Requirements)	5/9/2548	6/9/2548	2	สำนักบริหารและ รับรอง ห้องปฏิบัติการ (Bureau of Laboratory Accreditation)	

13	การเขียนโปรแกรมเชิงวัตถุด้วยภาษาจาวา (Object Oriented Programming with Java)	4/9/2548	8/9/2548	5	คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์ (School of Applied Statistic, National Institute of Development Administration)
14	ข้อกำหนด ISO/IEC 17011 (ISO/IEC 17011 Requirements)	5/5/2548	5/5/2548	1	สำนักบริหารและรับรองห้องปฏิบัติการ (Bureau of Laboratory Accreditation)
15	การวางแผนและการบริหารเชิงกลยุทธ์ (Strategic Planning and Management)	12/1/2548	12/1/2548	1	สำนักงานปลัดกระทรวงวิทยาศาสตร์และเทคโนโลยี (Office of the Permanent Secretary)
16	ORACLE 9i PL/SQL Programming	11/7/2548	15/7/2548	5	ศูนย์เทคโนโลยีสารสนเทศ (Information Technology Center)



## 2.2 รีเลชันรหัสฝึกอบรม (Training Identification Relation)

รหัสฝึกอบรม (Training ID)	รหัสพนักงาน (Staff ID)
11	21
12	19
13	22
14	26

## 2.3 รีเลชันข้อมูลพนักงาน (Staff Information Relation)

รหัสพนักงาน (Staff ID)	ชื่อ-สกุล (Name-Surname)	ตำแหน่ง (Position)	รหัสหน่วยงาน (Department ID)
19	แสนดี มากมี (Sandee Markmee)	เจ้าหน้าที่บริหารงาน ธุรการ (Clerical Service Administration Staff)	M1
21	ชวน ใจชื้อ (Chuan Chaisue)	โปรแกรมเมอร์ (Programmer)	M2
22	ทิพย์ ดีพร้อม (Thip Deeprome)	นักวิเคราะห์ระบบ (System Analyst)	M4
26	สนั่น ดวงดี (Sanun Duangdee)	นักวิทยาศาสตร์ (Scientist)	M5

## 2.4 รีเลชันหน่วยงาน (Department Relation)

รหัสหน่วยงาน (Department ID)	ชื่อหน่วยงาน (Department Name)
M1	ฝ่ายบริหาร (Clerical Service Administration)
M2	ฝึกอบรม (Training)
M3	ฝ่ายเทคโนโลยีสารสนเทศ (Information Technology)
M4	ฝ่ายประมวลข้อมูล (Data Processing)
M5	ฝ่ายห้องปฏิบัติการ (Laboratory)

รูปที่ 2 ฐานข้อมูลการพัฒนาพนักงาน

Figure 2 Staff Improvement Program Database

จากรูป 2 จะเห็นว่าในรีเลชันฝึกอบรม (2.1) ประกอบด้วย 7 แอททริบิวต์ ซึ่งรีเลชันนี้จะมีคาร์ดินาลิตี้ (Cardinality) เท่ากับ 6 ทูเปิล (Tuple) สำหรับคีย์หลัก (Primary key) เป็นค่าของแอททริบิวต์ที่มีค่าของข้อมูลเป็นเอกลักษณ์ ไม่มีความซ้ำซ้อนของข้อมูลในแต่ละทูเปิล นั่นคือ รหัสฝึกอบรม ส่วนโดเมน (Domain) เป็นขอบเขตหรือค่าที่ควรจะเป็นของข้อมูลในแอททริบิวต์หนึ่ง เช่น แอททริบิวต์ (วันที่เริ่มต้น) ควรจะเป็นค่าในวันที่ในปฏิทิน

### ฐานข้อมูลเชิงสัมพันธ์

เป็นการจัดเก็บรวบรวมข้อมูลเป็นแถว และ คอลัมน์ในลักษณะตารางสองมิติ ที่ประกอบด้วยแอททริบิวต์ ที่แสดงคุณสมบัติของรีเลชันหนึ่งๆ โดยที่รีเลชันต่างๆ ได้ผ่านกระบวนการทำรีเลชันให้เป็นบรรทัดฐาน (Normalization) ระหว่างการออกแบบเพื่อลดความซ้ำซ้อน และความขัดแย้งของข้อมูล

ฐานข้อมูลเชิงสัมพันธ์เป็นฐานข้อมูลในระดับภายนอก (External Level) และระดับแนวคิด (Conceptual Level) รีเลชันต่างๆ ในฐานข้อมูลจะประกอบด้วยแอททริบิวต์ต่างๆ ที่ออกแบบเพื่อลดความซ้ำซ้อนของการเก็บข้อมูล และสามารถเรียกใช้ข้อมูลอย่างมีประสิทธิภาพ

### ข้อดีของฐานข้อมูลเชิงสัมพันธ์

1. ฐานข้อมูลเชิงสัมพันธ์เป็นกลุ่มของรีเลชันหรือตารางที่ข้อมูลถูกจัดเก็บเป็นแถวและคอลัมน์ ซึ่งทำให้ผู้ใช้เห็นภาพของข้อมูลได้ง่าย
2. ผู้ใช้ไม่ต้องรู้ว่าข้อมูลถูกจัดเก็บอย่างไร รวมถึงวิธีการเรียกใช้ข้อมูล

Figure 2, Training Relation (2.1) has 7 attributes and cardinality of 6 tuples. The primary key: the attribute whose value within the same tuple is always unique in this relation is *Training ID*. The domain: a set of value the attribute can be for Started Date should have the 'Date' data type.

### Relational Database

Relational Database organizes data in the form of 2 dimensional tables consisting of attributes which characterize any particular relations. Relations go through the process of data normalization during the design period as to reduce data redundancy and to obtain a more effective data manipulation.

### Advantage

1. Relational Databases are collections of data or tables that are organized into rows and columns in order to help users visualize the concept easier.
2. There are no needs for users to have priori knowledge on data storing or data accessing method.



3. ภาษาที่ใช้ในการเรียกดูข้อมูล เป็นลักษณะคล้ายภาษาอังกฤษ และไม่จำเป็นต้องเขียนเป็นลำดับขั้นตอน (Procedural) เช่น ภาษา SQL

4. การเรียกใช้หรือเชื่อมโยงข้อมูลสามารถทำได้ง่าย โดยใช้โอเปอเรเตอร์ทางคณิตศาสตร์ เช่น SELECT, INTERSECTION, PROSECTION เป็นต้น

3. The syntax of the languages used in accessing the data is similar to English language and need not be written following procedure. Such language is SQL.

4. Data accessing or linking is easily performed via mathematical operators such as SELECT, INTERSECTION, PROSECTION.

### คุณลักษณะการจัดเก็บข้อมูลของรีเลชัน

1. ข้อมูลในแต่ละทูเปิลจะไม่มีข้อมูลซ้ำกัน ทั้งนี้เพราะการจัดการฐานข้อมูลที่ดีไม่ควรจะมีข้อมูลซ้ำซ้อนกันปรากฏอยู่

2. การเรียงลำดับของแอททริบิวต์จะเรียงลำดับก่อนหลังอย่างก็ได้ ไม่มีการระบุว่า แอททริบิวต์ (หลักสูตร) คือแอททริบิวต์แรก หรือ แอททริบิวต์ (รหัสฝึกอบรม) คือแอททริบิวต์สุดท้าย ทั้งนี้เพราะการอ้างอิงถึงแอททริบิวต์ใดจะใช้ชื่อของแอททริบิวต์นั้นๆ ในการอ้างอิง ไม่ใช่ลำดับที่แอททริบิวต์นั้นปรากฏอยู่

3. การเรียงลำดับของข้อมูลในแต่ละแถวไม่เป็นสาระสำคัญ การจัดเก็บข้อมูลในรีเลชัน ถูกจัดเก็บตามลำดับบนสื่อบันทึกข้อมูล แต่การเรียกใช้ข้อมูลในรีเลชันสามารถเรียกใช้ตามที่ต้องการได้

4. ค่าของข้อมูลในแต่ละแอททริบิวต์ของทูเปิลหนึ่งๆ จะบรรจุข้อมูลได้เพียงค่าเดียว ไม่ใช่กลุ่มของข้อมูลที่แสดงค่าที่มากกว่าหนึ่งค่า

5. ค่าของข้อมูลในแต่ละแอททริบิวต์จะบรรจุค่าของข้อมูลประเภทเดียวกัน เช่น แอททริบิวต์ (วันที่เริ่มต้น) จะมีค่าของข้อมูลเป็นวันเดือนปี ในปฏิทิน ไม่ใช่จำนวนวัน หรือแอททริบิวต์ (หน่วยงานรับผิดชอบ)

### Data Storing Characteristics

1. Data within the same tuple is never duplicated since data duplication is forbidden in good database management.

2. The order of attribute can be either ascending or descending. It was not specified whether *Class* should be the first attribute or *Training ID* should be the last. Any particular attribute is referred to by its name, not by its order.

3. The order of data within the row is of no importance. Data is stored base on its order in the storage; however, data manipulation is performed at user's will.

4. There can be only one value, not a group of data that hold more than one value, assigned to a particular attribute in each individual tuple.

5. Each attribute will contain the value of the same type. For instance, *Started Date* has the 'date' data type, not 'integer' data type or 'organizer'

### Type of keys

Detailed attributes structure relation when one of the attribute in that relation acts as a key. The key is used either to show or refer to the value in one of the tuples in that relation or to provide linkage with data from other relations. Keys can be divided into 2 categories as following.

#### 1. Primary key

Primary key is attribute whose value is non-duplicable or unique. Due to its uniqueness, primary key is used to identify the tuple to which it belongs. For an example, *Staff ID* of the value '19' is used to identify a Staff called 'Sande Markmee'. The value of *staff ID* is never duplicated.

#### 2. Foreign key

Foreign key is attribute or group of attributes that acts as primary key in another relation. From the *Staff Improvement Program Database* in figure 2, the *Department ID* in Staff Relation serves as primary key in Department Relation; hence, *Department ID* is Foreign Key.

### ประเภทของคีย์

เค้าร่างของรีเลชัน ประกอบด้วยรายละเอียดของแอททริบิวต์ต่างๆ ซึ่งแอททริบิวต์ใดแอททริบิวต์หนึ่งในรีเลชันจะมีคุณสมบัติเป็นคีย์ โดยคีย์นั้นๆ สามารถใช้ในการแสดงหรือบ่งบอกถึงค่าของทูเปิลใด ทูเปิลหนึ่ง ในรีเลชันนั้น หรือใช้ในการเชื่อมโยงข้อมูลอื่นในอีกรีเลชันหนึ่ง ประเภทของคีย์แบ่งเป็น 2 ประเภทหลัก ดังนี้

#### 1. คีย์หลัก (primary key)

เป็นแอททริบิวต์ที่มีคุณสมบัติของข้อมูลที่มีค่าเป็นเอกลักษณ์หรือ มีค่าที่ไม่ซ้ำกัน คุณสมบัติดังกล่าวจะสามารถระบุว่าข้อมูลนั้นเป็นข้อมูลของทูเปิลใด ตัวอย่างเช่น รหัสพนักงาน เลขที่ 19 สามารถระบุว่า เป็นรหัสของพนักงานชื่อ แสนดี มากมี และค่ารหัสพนักงานแต่ละคนจะมีค่าไม่ซ้ำกัน

#### 2. คีย์นอก (foreign key)

เป็นแอททริบิวต์หรือกลุ่มของแอททริบิวต์ที่อยู่ในรีเลชันหนึ่งๆ ที่มีค่าของแอททริบิวต์นั้น ไปปรากฏเป็นคีย์หลักในอีกรีเลชัน ตัวอย่างจากฐานข้อมูลพัฒนาพนักงาน รูปที่ 2 จะพบว่าแอททริบิวต์รหัสหน่วยงานในรีเลชันข้อมูลพนักงานจะไปปรากฏเป็นคีย์หลักอยู่ใน รีเลชันหน่วยงาน ดังนั้น จึงกล่าวได้ว่าแอททริบิวต์รหัสหน่วยงานเป็นคีย์นอก



### กฎที่เกี่ยวข้องกับคีย์ในฐานะข้อมูลเชิงสัมพันธ์

เนื่องจากรีเลชันต่างๆ ในฐานะข้อมูลเชิงสัมพันธ์อ้างอิงความสัมพันธ์ของข้อมูลโดยใช้คีย์ กฎที่เกี่ยวข้องกับคีย์ในฐานะข้อมูลเชิงสัมพันธ์ มีดังนี้

1. กฎความบูรณาภาพของเอนทิตี (The Entity Integrity Rule)  
กฎนี้ระบุว่า แอททริบิวต์ใดที่เป็นคีย์หลัก ข้อมูลในแอททริบิวต์นั้นจะเป็นค่าว่าง (Null) ไม่ได้ ความหมายของการเป็นค่าว่างไม่ได้ในที่นี้หมายความว่าข้อมูลของแอททริบิวต์ที่เป็นคีย์หลักจะไม่ทราบค่าที่แน่นอน หรือ ไม่มีค่าไม่ได้ (ต้องมีค่าของข้อมูล)

2. กฎความบูรณาภาพของการอ้างอิง (The Referential Integrity Rule)

การอ้างอิงข้อมูลระหว่างรีเลชันในฐานะข้อมูลเชิงสัมพันธ์ ใช้คีย์นอกของรีเลชันหนึ่ง ไปตรวจสอบกับค่าของแอททริบิวต์ที่เป็นคีย์หลักของอีกรีเลชันหนึ่ง เพื่อเรียกดูข้อมูลอื่นที่เกี่ยวข้อง

ในกรณีที่มีการลบข้อมูลหรือแก้ไขข้อมูลของแอททริบิวต์ที่เป็นคีย์หลักในรีเลชันหนึ่งซึ่งมีคีย์นอก จากรีเลชันหนึ่งอ้างอิงถึง จะทำการลบหรือแก้ไขข้อมูลหรือไม่นั้น ขึ้นอยู่กับการออกแบบฐานข้อมูลว่า ได้ระบุให้แอททริบิวต์นั้นๆ มีคุณสมบัติอย่างไร ซึ่งมีโอกาสเป็นไปได้ 4 ทางเลือก

#### 1. การลบหรือแก้ไขข้อมูลแบบมีข้อจำกัด (Restrict)

การลบหรือแก้ไขข้อมูลจะกระทำได้เมื่อข้อมูลของคีย์หลักในรีเลชันหนึ่ง ไม่มีข้อมูลที่ถูกอ้างอิงโดยคีย์นอกของอีกรีเลชันหนึ่ง เช่น รหัสหน่วยงาน ในรีเลชัน หน่วยงาน จะถูกแก้ไขหรือลบทิ้งต่อเมื่อไม่มีพนักงาน สังกัดอยู่หน่วยงานนั้น

#### 2. การลบหรือแก้ไขข้อมูลแบบต่อเนื่อง (Cascade)

การลบหรือแก้ไขข้อมูลจะกระทำแบบลูกโซ่ คือ หากมีการแก้ไขหรือลบข้อมูลของคีย์หลักในรีเลชันหนึ่งระบบจะทำการลบหรือแก้ไขข้อมูลของคีย์นอกในอีกรีเลชันหนึ่งที่เกี่ยวข้องถึงข้อมูลของคีย์หลักที่ถูกลบด้วย เช่น กรณีที่ยกเลิกหน่วยงาน M5 ในรีเลชัน หน่วยงาน ข้อมูลของพนักงานที่อยู่หน่วยงาน M5 ในรีเลชันข้อมูลพนักงาน จะถูกลบออกไปด้วย

#### 4. การลบหรือแก้ไขข้อมูลแบบใช้ค่าปริยาย (Default)

การลบหรือแก้ไขข้อมูลของคีย์หลักสามารถทำได้ โดยถ้าคีย์นอกที่อ้างอิงถึงคีย์หลักที่ถูกลบหรือแก้ไขก็จะทำการปรับค่าของคีย์นอกนั้น เป็นค่าโดยปริยาย (Default Value) ที่ถูกกำหนดขึ้น เช่น กรณีที่ยกเลิก M5 ในรีเลชัน หน่วยงาน ข้อมูลของพนักงานที่อยู่ M5 ในรีเลชัน ข้อมูลพนักงานจะถูกเปลี่ยนค่าเป็น “00” ซึ่งเป็นค่าโดยปริยายที่หมายความว่าไม่ได้สังกัดหน่วยงานใด

พบกันใหม่ฉบับหน้า.....สวัสดีครับ

### Key-Related Relational Database Rules

Relations in Relational Database are referred to using Keys. The rules concerning the keys are as followings.

#### 1. The Entity Integrity Rule

It is stated that for any attribute acting as Primary Key, its value can not be Null. In this case, Null is data of an unidentifiable or non-measurable value.

#### 2. The Referential Integrity Rule

Referencing of data between relations in relational database uses a foreign key from one relation to compare with the value of the attribute acting as a Primary Key from another relation in order to access other related data.

In case there is need for a deletion or edition of the attribute acting as a Primary key in one relation and Foreign key in another relation; whether the mentioned actions shall be performed or not, depends on the database's design. This kind of attribute can be assigned either one of these following 4 characteristics.

#### 1. Restrict

Deletion or alteration of data can be performed only when the value of the primary key in one relation is not referred to as foreign key by another relation. For an example, *Department ID* can be altered or deleted only when there is no staff left under that department.

#### 2. Cascade

The deletion or alteration of data is a chained action that is when there is deletion or alteration of data in one relation, the system will automatically apply or alter the data of the foreign key which is referred to by that primary key in another relation as well. For instance, if M5 in Department Relation is deleted, the staff working under that department in Staff Relation will be deleted as well.

#### 3. Nullify

The deletion or alteration of data can be performed only when the referred Foreign Key has been first changed to Null. The value of Staff in M5 must be changed to Null before deletion or alteration of the department of ID M5 in Department Relation can be performed.

#### 4. Default

When deletion or alteration of Primary Key is performed, the value of the Foreign Key that refers to that particular Primary Key will be changed to the preset Default Value. In case M5 in Department Relation is deleted, the value of staff working under M5 in Staff Relation shall automatically be assigned the value “00” which was the preset Default Value that means the certain Staff works under no department in particular.

See you in the next edition. Good bye.