

Lesson 4: Relational Data Integrity

Relational Data Integrity

- Database-specific
- General

Database-Specific

- Student id must be of the form 999999999
- Age must be in between 18 - 25 years
- Students cannot take course more than 18 units
- There are three semesters: semester I, semester II, and semester summer
- Etc.

Apply to just one specific database

General Integrity

It is relevant to the integrity of *every* database

General Integrity

- กฏ General Integrity จะมีส่วนเกี่ยวข้องกับ
- candidate (and primary) keys
 - foreign keys

Candidate Keys

Basic Definition



RECORD IDENTIFIER

Candidate Keys

COURSE

CODE	DESC	UNIT
AC201	Accounting	3
FN201	Finance	3

REGISTER

ID	CODE	GRADE
101	AC201	3
101	FN201	2
102	AC201	3

Candidate Keys

A candidate key for R is a subset of the set of attributes of R , say K , such that:

Uniqueness property:

No two distinct tuples of R have the same value for K

...

Candidate Keys

...

Irreducibility property:

No proper subset of K has the uniqueness property

Composite and Simple Candidate keys

- A candidate key that involves exactly one attribute is said to be *simple*
- A candidate key that involves more than one attribute is said to be *composite*

What are candidate keys for?

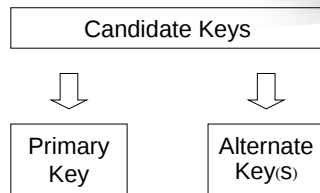
- They provide the basic *tuple-level addressing mechanism* in a relational system.

Primary Keys and Alternate Keys

- A given base relation may have more than one candidate keys.
- One of those candidate keys must be chosen as the *primary key*.
- The remainder, if any, are then called *alternate keys*.



Primary Key and Alternate Keys



Foreign Keys

<u>COURSE</u>		
CODE	DESC	UNIT
AC201	Accounting	3
FN201	Finance	3

<u>REGISTER</u>		
ID	CODE	GRADE
101	AC201	3
101	FN201	2
102	AC201	3

An arrow points from the 'CODE' column in the REGISTER table to the 'CODE' column in the COURSE table, indicating a foreign key relationship.

Foreign Keys

Let R_2 be a base relation. Then a **foreign key** in R_2 is a subset of the set of attributes of R_2 , say FK , such that:

- There exists a base relation R_1 (R_1 and R_2 not necessarily distinct) with a candidate key CK , and

...

Foreign Keys

...

- For all time, each value of FK in the current value of R_2 is identical to the value of CK in some tuple in the current value of R_1 .

Foreign keys

จุดที่น่าสนใจ

- Foreign key ต้องมี candidate key มารองรับ แต่ในทางกลับกันไม่จำเป็น
- Foreign key สามารถเป็นได้ทั้ง Composite และ Simple

Foreign keys

จุดที่น่าสนใจ (ต่อ)

- Foreign key และ candidate key โดยส่วนใหญ่จะมาจาก domain เดียวกัน
- Foreign key ไม่จำเป็นต้องเป็นส่วนหนึ่งของ candidate key ของ relation ที่มันอยู่

Foreign keys

จุดที่น่าสนใจ (ต่อ)

- เราเรียก relation ที่ foreign key อยู่ว่า Referencing Relation (ไปข้างอิง)
- เราเรียก relation ที่ candidate key อยู่ว่า Referenced หรือ Target Relation (ถูกข้างอิง)

Foreign keys

จุดที่น่าสนใจ (ต่อ)

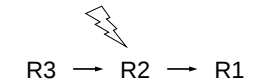
- เราสามารถเขียน Referential Diagram ได้ดังนี้



Foreign keys

จุดที่น่าสนใจ (ต่อ)

- Relation หนึ่งๆ สามารถเป็นได้ทั้ง Referenced และ Referencing Relation



Foreign keys

จุดที่น่าสนใจ (ต่อ)

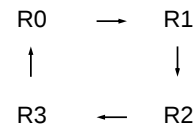
- R1 และ R2 อาจเป็น relation เดียวกันก็ได้

emp#	name	mgr#
101	Tom	
102	Mary	101
103	Tony	102
104	Mike	102

Foreign keys

จุดที่น่าสนใจ (ต่อ)

- บางกรณีอาจเกิด Referential Cycle ได้



Foreign keys

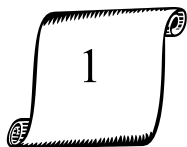
จุดที่น่าสนใจ (ต่อ)

- Foreign-to-candidate-key คือ กาว ซึ่งทำให้ database เกาะติดกัน



Referential Integrity

The database must not contain any unmatched foreign key values.



Referential Integrity

- กฎ Referential Integrity จะต้องรักษาอยู่ ณ ทุกสถานะของระบบฐานข้อมูล (database states)
- การรักษากฎ ณ ทุกสถานะจะก่อให้เกิดปัญหากับการ delete และ update ข้อมูลกับ referenced table

Referential Integrity

- วิธีการแก้ไขคือ ยอมให้ database states อยู่ในสถานะที่ไม่ถูกต้องขณะหนึ่ง และมีการกระทำเพิ่มเติมเพื่อให้ database กลับมาอยู่ในสถานะที่ถูกต้อง

Foreign Key Rules

การใช้ Foreign key rules จะแบ่งเป็นสองกรณีคือ

- Update
- Delete

Foreign Key Rules

กรณีมีการ Update ที่ referenced table แล้ว
referencing table จะมีการเปลี่ยนแปลงสองกรณีคือ

- Restricted
- Cascades

Foreign Key Rules

กรณีมีการ Delete ที่ referenced table แล้ว
referencing table จะมีการเปลี่ยนแปลงสองกรณีคือ

- Restricted
- Cascades

Foreign Key Rules

กรณีอื่นๆ ที่อาจนำมาใช้

- Nullifies
- Default

Foreign keys and null

- Let R_2 be a base relation. Then a *foreign key* in R_2 is a subset of the set of attributes of R_2 , say, FK , such that:
 - There exists a base relation R_1 (R_1 and R_2 not necessarily distinct) with a candidate key CK , and

...

Foreign keys and null

...

- For all time, each value of FK in the current value of R_2 **either is null or** is identical to the value of CK in some tuple in the current value of R_1 .

จุดที่น่าสนใจ Foreign keys and null

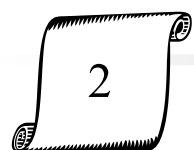
- ผู้ออกแบบระบบจะต้องตัดสินใจว่า foreign จะรับค่า null ได้หรือไม่
- หาก foreign key เป็น composite ค่าจะต้องเป็น wholly null or non null
- หาก foreign key สามารถรับค่า null Foreign key rule อาจเพิ่ม option: nullifies

สรุปกฎ Foreign keys

- Restricted
- Cascade
- Nullifies
- Default

จุดที่น่าสนใจของ Foreign key Rules

- การกำหนด Cascade หรือ Restricted ทราบได้จาก การได้พูดคุยกับลูกค้า
- $R_3 \xrightarrow{\text{Delete: cascade}} R_2 \xrightarrow{\text{Delete: cascade}} R_1$
Delete: Restricted



Entity Integrity

- No component of the primary key of a base relation is allowed to accept nulls.



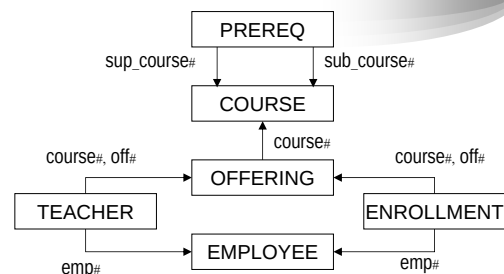
Primary key

- Theory
- User
- Programmer

Exercises

- course (course#, title)
- prereq (sup_course#, sub_course#)
- offering (course#, off#, offdate, location)
- teacher (course#, off#, emp#)
- enrollment (course#, off#, emp#, grade)
- employee (emp#, ename, job)

ต้องการ หา primary key, เขียน referential diagram, กำหนด foreign key rule



- DEPT (DEPT#, ..., MGR_EMP#, ...)
- EMP(EMP#, ..., DEPT#, ...)

- ทุก department ต้องมี manager (MGR_EMP#)

- พนักงานทุกคนต้องมีฝ่ายอยู่

ต้องการ เขียน referential diagram และ กำหนด foreign key rule ที่เหมาะสม

UPDATE: CASCADE
DELETE: CASCADE

UPDATE: CASCADE
DELETE: CASCADE

EMP	
emp# ...	dept#
101	AC
102	AC
103	AC
104	AC
105	FN
106	FN
107	MK

DEPT	
dept# ...	mgr#
AC	101
FN	105
MK	107

- EMP (EMP#, ..., JOB, ...)
- PGMR (EMP#, ..., LANG, ...)

- โปรแกรมเมอร์ทุกคนต้องเป็นพนักงาน
- พนักงานทุกคนไม่จำเป็นต้องเป็นโปรแกรมเมอร์

ต้องการ หา primary key, เขียน referential diagram, กำหนด foreign key rules