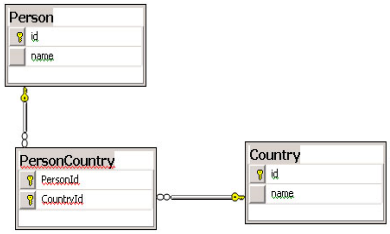
IT2201 / IT2601 / IT2564 / IT2621 / IT2521 / IT2323

Database Management Systems



Unit 2
Relational Data Model

1

Unit Objectives

- At the end of this unit, you should be able to
  - Understand the terminology used in Relational Model.
  - Define the rules which preserve the integrity of a relational database.
  - Describe the operations in relational languages that are used in updating or retrieving data.

2

Relational Model Terminology

- Relational Data Model
  - Based on the mathematical concept of a relation
  - Organizes and represents data in a form of table
- A relation is a table with rows and columns.
- A tuple is a row of a relation.
- An attribute is a named column of a relation.
- The degree of a relation is the number of attributes in the relation.
- The cardinality of a relation is the number of tuples in the relation.
- A domain is a set of allowable values for each attribute.

3

Example of a Relation EMPLOYEE

DOMAINS: String of 9 chars | String of 30 chars | M or F | String of 8 digits | date

EMPLOYEE

EMP_ID	NAME	GENDER	PHONE	BIRTH DATE
F7634567D	John Smith	M	82795544	22-May-76
S7344555A	Franklin Wong	M	92267777	1-May-73
S7688777H	Mhd Ali	M	NULL	19-Jun-76
S7654321Z	Jennier Tan	F	92795544	16-May-76
S7588444J	Joyce Ng	F	62267777	20-Oct-75
F7445345G	James Borg	M	NULL	25-Jun-74

Tuples

Attributes

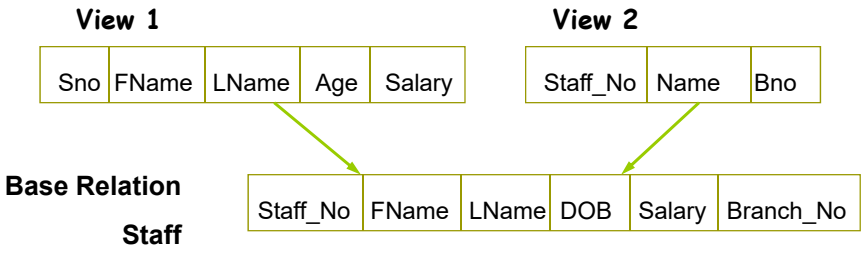
Degree = 5

Cardinality = 6

4

Relational Model Terminology

View vs Base Relation



5

Relational Model Terminology

Base Relation
A named relation corresponding to an entity in conceptual schema, whose tuples are physically stored in database.

View
The dynamic result of one or more relational operations operating on base relations to produce another relation.

A virtual relation that does not necessarily actually exist in the database but is produced upon request, at time of request.

Views are dynamic, meaning that changes made to base relations that affect view attributes are immediately reflected in the view.

Contents of a view are defined as a query on one or more base relations.

6

Alternative Terminology

Table 3.1 Alternative terminology for relational model terms.

Formal terms	Alternative 1	Alternative 2
Relation	Table	File
Tuple	Row	Record
Attribute	Column	Field

7

Properties of Relations

- Relation name is distinct from all other relation names in relational schema.
- Each cell of relation contains exactly one atomic (single) value.
- Each attribute has a distinct name.
- Values of an attribute are all from the same domain.
- Each tuple is distinct; there are no duplicate tuples.
- Order of attributes has no significance.
- Order of tuples has no significance, theoretically.

EMPLOYEE

EMP_ID	NAME	BIRTH DATE
F7634567D	John Smith	22-May-76
S7344555A	Franklin Wong	1-May-73
S7688777H	Mhd Ali	19-Jun-76
S7654321Z	Jennier Tan	16-May-76
S7588444J	Joyce Ng	20-Oct-75
F7445345G	James Borg	25-Jun-74

8

Relational Keys

- Types of relational keys
  - Candidate key
  - Primary key
  - Alternate key
  - Foreign key
- See examples using the following Student table:

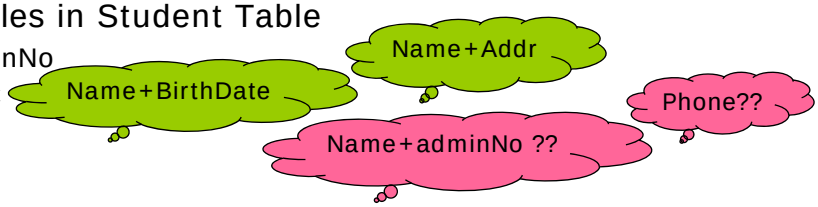

Relational Keys ELO

AdminNo	NAME	GENDER	NRIC	BIRTHDATE
081234A	John Smith	M	F7634567D	22-May-91
081223B	Franklin Wong	M	S9144555A	1-May-91
086633C	Mhd Ali	M	S9188777H	19-Jun-91
089988G	Jennifer Tan	F	S9154321Z	16-May-91
082987C	Joyce Ng	F	S9088444J	20-Oct-90
087623F	James Ng	M	F7445345G	25-Jun-89

9

Candidate Key

- A candidate key is any minimal set of attributes that can uniquely identifies each tuple in a relation.
- A relation may have one or more candidate keys.
- Candidate keys can consist of a single attribute or multiple attributes. Multiple attribute keys are known as composite keys.
- Examples in Student Table
  - AdminNo
  - NRIC



10

Primary Key & Alternate Key

- One candidate key is designated as primary key whose values are used to identify tuples uniquely in a relation.
- Example in Student Table
  - AdminNo
- Candidate keys that are not selected to be primary key are called alternate keys
- Examples in Student Table
  - NRIC
  - Name+birthdate
  - Name+Addr
  - Phone

11

Criteria for choosing a primary key

- Minimal set of attributes.
- Less likely to have its values changed.
- Less likely to lose uniqueness in the future.
- With fewest characters
- Easier to use from the user's point of view.

12

Relational Keys Exercise

Identify candidate keys, primary key and alternate key in the following table:

DEPARTMENT

DNUMBER	DNAME	MGR	MGRSTARTDATE
D1	FINANCE	Jack	1-JUL-1989
D2	HUMAN RESOURCE	Terri	1-May-2000
D3	SALES	Jack	1-Jun-2005
D4	ENGINEERING	Henry	1-Aug-2007
D5	PURCHASING	Mary	1-Sep-2003

Candidate key	Primary key	Alternate key

13

Foreign Key

- A foreign key is an attribute which provides a logical link between tables.
- A foreign key is an attribute, or set of attributes, within one relation(R1) that matches the candidate key of some (possibly the same) relation (R2).

14

Foreign Key Example

Student(R1)

PK FK

AdminNo	stud_name	Sch_cd
111	Steve Sun	SIT
112	Jay Chow	SBM
113	Elaine Chew	Null

School(R2)

PK

sch	Sname
SIT	School of Information Technology
SBM	School of Business Management

The foreign key in Student table is Sch_cd.
Is there any foreign key in School Table?

15

Relational Integrity Rules

- Integrity Rules
  - Can be specified on a relational database schema.
  - The purpose is to ensure accuracy and integrity of the data in the database.
- 2 principal rules
  - **Entity Integrity** - primary key in base relations must not contain any nulls.
  - **Referential Integrity** - If foreign key exists in a relation, either foreign key value must match a candidate key value of some tuple in its home relation or foreign key value must be wholly null.

Referential Integrity

- An example of Violation

PK Student

studNo	name	Sch_cd
121	Steve Sun	SIT
112	Jay Chow	SBM
187	Elaine Chew	Null

?

PK Project FK

ProjNo	ProjDesc	StudNo
1	Magazine Subscription	121
2	Video Rental	879
3	Cinema Booking	Null
4	Library System	187

Violate the Referential Integrity Rule

17

Referential Integrity

- What happens on an attempt to delete the target of a foreign key reference?
  - 3 possibilities:
    - Restrict – Do not allow the action on the referenced row.
      - Eg. In the previous slide, do not allow student 121 to be deleted if there is project done by Student 121.
    - Cascade – Perform the same action to the related rows.
      - Eg. If Student 121 is deleted, delete the related project.
    - Nullify – Set the foreign key of related rows to null.
      - Eg. If student 121 is deleted, set the STUDNO to NULL for project 1.
- What happens on an attempt to update the target of a foreign key reference ?

18

Enterprise Constraints

- Definition
  - Additional rules that are specified by the users or database administrators that the data must satisfy.
- Examples
  - Upper limit for exam marks is 100
  - Business rules such as no more than 25 students to be assigned to a tutorial group

19

Quiz (1)

1. Constraints that applied to data items to make sure that the data is valid and consistent is called _____.
   a) monitoring facility
   b) transaction management
   c) integrity rules
   d) security rules

2. Identify the Primary keys and Foreign keys in the following relational schema:

Student

StudName	StudNo	SAddress	School

StudentModule

ModuleNo	StudNo	Grade

20

Quiz (1)

3. Which row in the following relation violates entity integrity rule?

Emp_no(PK)	Emp_name	Department	position
1002	Jane	HR	Null
Null	Mary	IS	Supervisor
1004	Null	FN	Accountant
1005	Philip	Null	Instructor

4. Which row violates referential integrity rule?

Course		
CourseNo	**course_title**	**StudNo**
1	Java Programming	120
2	Database Management	900
3	Web Design	null
4	Operating Systems	435

→

Student	
StudNo	**Name**
120	John
125	Mary
400	Peter
435	Mark

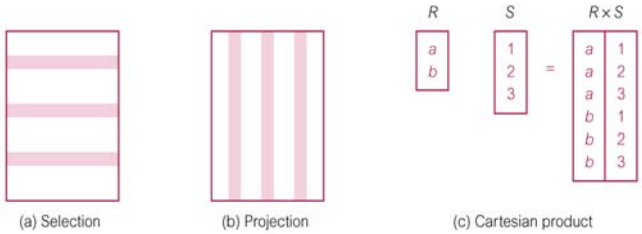
21

Relational Algebra

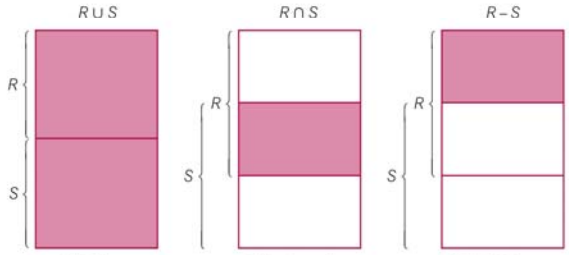
- Relational algebra operations work on one or more relations to define another relation without changing the original relations.
- Both operands and results are relations, so output from one operation can become input to another operation.
- Uses of Relational Algebra
  - To define a scope of retrieval.
  - To define a scope for update.
  - To define a view.
  - To define access rights.
  - To define integrity constraints.

22

Relational Algebra Operations



(a) Selection (b) Projection (c) Cartesian product



(d) Union (e) Intersection (f) Set difference

Relational Algebra Operations

T

A	B
a	1
b	2

U

B	C
1	x
1	y
3	z

$T \bowtie U$

A	B	C
a	1	x
a	1	y

$T \triangleright_B U$

A	B
a	1

$T \bowtie_C U$

A	B	C
a	1	x
a	1	y
b	2	

(g) Natural join (h) Semijoin (i) Left Outer join

Relational Algebra

- The SELECT operation
  - Use to select a subset of the tuples in a relation that satisfy a selection condition.
  - Example
    - To select the subset of EMPLOYEE tuples who work in department 4.

EMP#	NAME	DEPT
1	TOM	1
2	DICK	4
3	HARRY	2
4	JOHN	4
5	PAUL	2

EMP#	NAME	DEPT
2	DICK	4
4	JOHN	4

$\sigma_{DEPT = 4}(EMPLOYEE)$

25

Relational Algebra

- The PROJECT operation
  - Use to select certain attributes from the relation and discards the other attributes.
  - Example

    To list each employee's last name and salary.

EMP#	NAME	LNAME	DEPT	SALARY
1	TOM	TAN	1	$1,000
2	DICK	LEE	4	$1,500
3	HARRY	LIM	2	$2,000
4	JOHN	YEO	4	$3,500
5	PAUL	ONG	2	$1,890

LNAME	SALARY
TAN	$1,000
LEE	$1,500
LIM	$2,000
YEO	$3,500
ONG	$1,890

$\pi_{LNAME, SALARY}(EMPLOYEE)$

26

Relational Algebra

- The UNION Operation
  - The result is a relation that includes all tuples that are either in relation R or in relation S or in both. Duplicate tuples are eliminated.

R
EMP#
1
2
3
4

S
EMP#
1
3
5

R ∪ S
EMP#
1
2
3
4
5

27

Relational Algebra

- The INTERSECTION Operation
  - The result is a relation that includes all tuples in both relation R and S.

R
EMP#
1
2
3
4

S
EMP#
1
3
5

R ∩ S
EMP#
1
3

28

Relational Algebra

- The DIFFERENCE Operation
  - The result is a relation that includes all tuples that are in relation R but not in relation S.

R
EMP#
1
2
3
4

S
EMP#
1
3
5

R - S
EMP#
2
4

29

Relational Algebra

- The CARTESIAN PRODUCT
  - This operation is used to combine tuples from two relations so that related tuples can be identified.

EMP

EMP#	NAME
1	TOM
2	DICK
3	HARRY

CHILD

CHILD	DAD
A	1
B	1
C	2

EMP X CHILD

EMP#	NAME	CHILD	DAD
1	TOM	A	1
1	TOM	B	1
1	TOM	C	2
2	DICK	A	1
2	DICK	B	1
2	DICK	C	2
3	HARRY	A	1
3	HARRY	B	1
3	HARRY	C	2

30

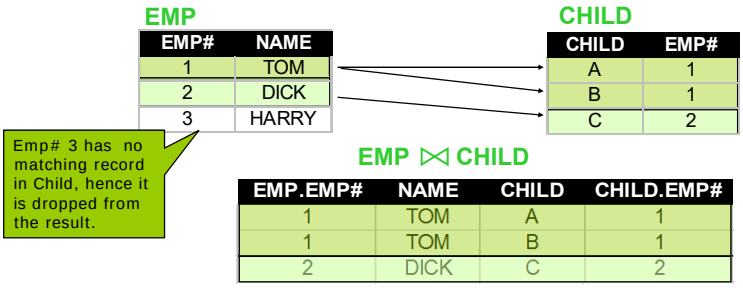
Relational Algebra

- The JOIN Operation
  - This operation is used to combine related tuples from two relations into a single tuple.
  - Types of Joins
    - Inner Join: combines tuples from two relations using comparison operators in a condition.
      - Equijoin
      - Natural join
    - Outer Join: combine tuples from two relations and keep all the tuples in the result even if there is no matching tuple in the two relations.

31

Relational Algebra

- EquiJoin
  - Equijoin combine tuples with same value in the join condition.
  - Example - To retrieve the names of employees with their children.



- Natural join is similar to equijoin except that one of the common columns is eliminated.

32

Relational Algebra

- OUTER JOIN
  - combine tuples from two relations and keep all the tuples in the result even if there is no matching tuple in the two relations.
  - Eg Left outer join

EMP

EMP#	NAME
1	TOM
2	DICK
3	HARRY

CHILD

CHILD	DAD
A	1
B	1
C	2

EMP ⟕ CHILD

EMP#	NAME	CHILD	DAD
1	TOM	A	1
1	TOM	B	1
2	DICK	C	2
3	HARRY		

Even if Emp# 3 has no matching records in Child, it will be preserved in the result set with the remaining columns set to NULL.

33

Quiz (2)

- Identify types of algebra operations used in the following data retrieval :
  - List all the order details placed by the customer number 104.
  - List order number and order date for orders dated in May 2004.
  - List order number, order date, customer name for all the orders.

ORDER_NUM	ORDER_DATE	CUST_NUM
1001	20-May-04	104
1002	21-May-04	101
1003	22-Jun-04	104
1004	22-Jun-04	106

CUST_NUM	FNAME	LNAME
101	Ludwig	Pauli
102	Carole	Sadler
103	Philip	Currie
104	Anthony	Higgins
105	Raymond	Vector
106	George	Watson

34

Summary

- Relational Terminology
  - Relation, Attribute, Domain, Tuple, Degree, Cardinality, View, Base Relation
- Relational Keys
  - Candidate key, Primary key, Foreign key
- Relational Integrity
  - Entity integrity, Referential integrity, Enterprise Constraints
- Relational Algebra Operation
  - Select, Project, Union, Intersection, Difference, Cartesian, Join

35

Reference

1. Database Systems, Connolly, Ch 4 & 5

36