2 The Relational Model

Prerequisites for This Section

- Readings:
  - Required: Connolly and Begg, sections 3.1, 3.2, and 3.3.
- Assessments:
  - Multiple-Choice Quiz 2

Section Objectives

In this section you will learn:

① Terminology of relational models.

② How tables are used to represent data.

③ Properties of database relations.

④ How to identify CKs, PK, and FKs.

⑤ Meaning of entity integrity and referential integrity.

Agenda

1. **Data Model**
2. **Tabular Structure: Table Schema**
3. **Properties of Relations**
4. **Keys**
5. **Integrity Constraints**

Data Model

Integrated collection of concepts for describing data, relationships between data, and constraints on the data in an organization.

① A data model defines how the data is organized and manipulated in a database

② A data model comprises:

I. a structural part;

II. a manipulative part;

III. possibly a set of integrity rules.

Relational Model

① The relational model was first proposed by E. F. Codd in 1970.

② The relational model's objectives were specified as follows:

I. To allow a high degree of data independence. Application programs must not be affected by modifications to the internal data representation, particularly by changes to file organizations, record orderings, or access paths.

II. To provide substantial grounds for dealing with data semantics, consistency, and redundancy problem.

III. To enable the expansion of set-oriented data manipulation language.

Relational Model

③ The relational model is currently the most popular model for database systems for two reasons:

I. The simplest and most uniform data structure: data are stored in table

II. The most formal of all data models. It's based on a mathematical concept of a relation in set theory: a table is a relation

④ Relational Model comprises:

I. Tabular structure: table schema

II. Relational operations

III. Integrity constraints

A Sample Instance of Relational Model

Branch

branchNo	street	city	postCode
B005	22 Deer Rd	London	SW1 4EH
B007	16 Argyll St	Aberdeen	AB2 3SU
B003	163 Main St	Glasgow	G11 9QX
B004	32 Manse Rd	Bristol	BS99 1NZ
B002	56 Clover Dr	London	NW10 6EU

Staff

staffNo	fName	lName	position	sex	DOB	salary	branchNo
SL21	John	White	Manager	M	1-Oct-45	30000	B005
SG37	Ann	Beech	Assistant	F	10-Nov-60	12000	B003
SG14	David	Ford	Supervisor	M	24-Mar-58	18000	B003
SA9	Mary	Howe	Assistant	F	19-Feb-70	9000	B007
SG5	Susan	Brand	Manager	F	3-Jun-40	24000	B003
SL41	Julie	Lee	Assistant	F	13-Jun-65	9000	B005

Agenda

1. **Data Model**
2. **Tabular Structure: Table Schema**
3. **Properties of Relations**
4. **Keys**
5. **Integrity Constraints**

Tabular Structure

- In the relational model, all data is logically structured within tables (relations).
- A table is
  - ① used to represent an entity
  - ② used to represent the relationship of entities
  - ③ the result of operating tables

① Represent an Entity

- Entity **Student**

student

sid	lname	fname	class	telephone
1	Jones	Allan	2	555-1234
2	Smith	John	3	555-4321
3	Brown	Harry	2	555-1122
5	White	Edward	3	555-3344

- Entity **Course**

course

cno	cname	croom	time
101	French Ⅰ	2-104	MW2
102	French Ⅱ	2-113	MW3
105	Algebra	3-105	MW2
108	Calculus	2-113	MW4

② Represent a Relationship

- Relationship **Enrollment** between **Student** and **Course**

students

sid	lname	fname	class	telephone
1	Jones	Allan	2	555-1234
2	Smith	John	3	555-4321
3	Brown	Harry	2	555-1122
5	White	Edward	3	555-3344

enrollment

sid	cno	major
1	101	No
1	108	Yes
2	105	No
3	101	Yes
3	108	No
5	102	No
5	105	No

course

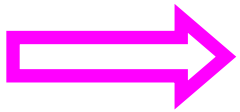
cno	cname	croom	time
101	French Ⅰ	2-104	MW2
102	French Ⅱ	2-113	MW3
105	Algebra	3-105	MW2
108	Calculus	2-113	MW4

③ Represent the Query Result

- List the **sid**s of all students who enrolled in course "101"

enrollment

sid	cno	major
1	101	No
1	108	Yes
2	105	No
3	101	Yes
3	108	No
5	102	No
5	105	No



sid
1
3

Relational Model Terminology

① Relation is a table with columns and rows.

② Attribute is a named column of a relation.

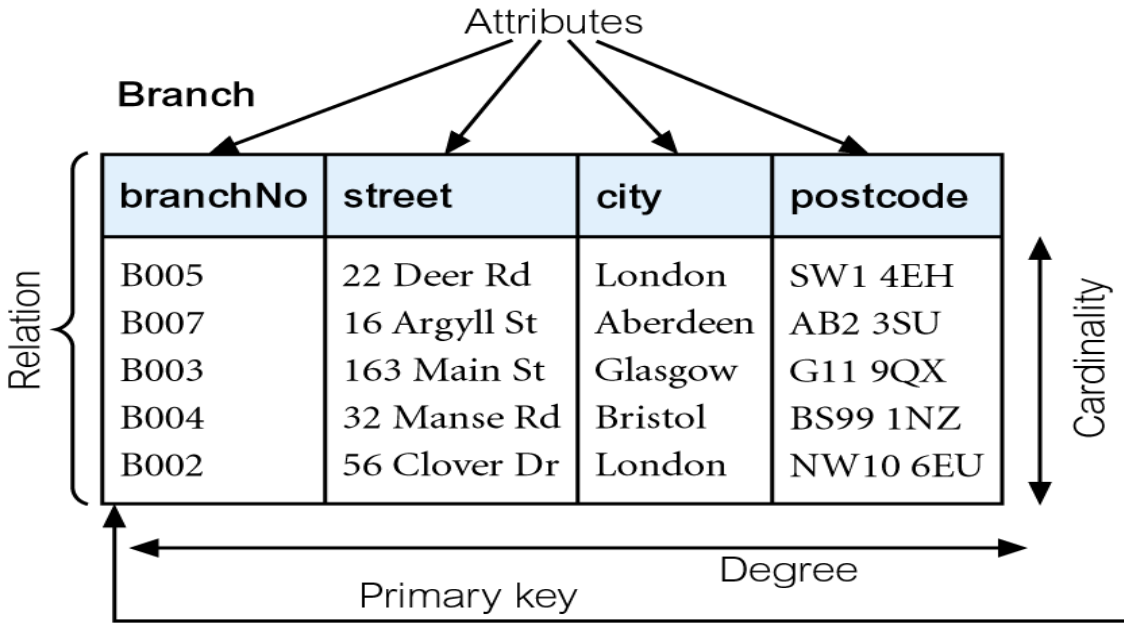
③ Domain is the set of allowable values for one or more attributes.

④ Tuple is a row of a relation.

⑤ Degree is the number of attributes in a relation.

⑥ Cardinality is the number of tuples in a relation.

Instances of Branch and Staff Relations



branchNo	street	city	postcode
B005	22 Deer Rd	London	SW1 4EH
B007	16 Argyll St	Aberdeen	AB2 3SU
B003	163 Main St	Glasgow	G11 9QX
B004	32 Manse Rd	Bristol	BS99 1NZ
B002	56 Clover Dr	London	NW10 6EU

Staff

Relation

staffNo	fName	lName	position	sex	DOB	salary	branchNo
SL21	John	White	Manager	M	1-Oct-45	30000	B005
SG37	Ann	Beech	Assistant	F	10-Nov-60	12000	B003
SG14	David	Ford	Supervisor	M	24-Mar-58	18000	B003
SA9	Mary	Howe	Assistant	F	19-Feb-70	9000	B007
SG5	Susan	Brand	Manager	F	3-Jun-40	24000	B003
SL41	Julie	Lee	Assistant	F	13-Jun-65	9000	B005

Mathematical Relations

I. To understand the true meaning of the term relation, we have to review some concepts from mathematics.

II. Suppose that we have two sets, D_1 and D_2, where D_1 = {2,4} and D_2 ={1,3,5}. The Cartesian product of this two sets, written $D_1 \times D_2$, is the set of all combinations of elements with the first from D_1 and the second from D_2. In our case, we have:
$D_1 \times D_2$= {(2,1), (2,3), (2,5), (4,1), (4,3), (4,5)}

III. Any subset of this Cartesian product is a relation.

- R = {(2,1), (4,1)}
- R = {(x, y) | x $\in D_1$, y $\in D_2$, and y = 1 }

Examples of Attribute Domains

Attribute	Domain Name	Meaning	Domain Definition
branchNo	BranchNumbers	The set of all possible branch numbers	character: size 4, range B001–B999
street	StreetNames	The set of all street names in Britain	character: size 25
city	CityNames	The set of all city names in Britain	character: size 15
postcode	Postcodes	The set of all postcodes in Britain	character: size 8
sex	Sex	The sex of a person	character: size 1, value M or F
DOB	DatesOfBirth	Possible values of staff birth dates	date, range from 1-Jan-20, format dd-mmm-yy
salary	Salaries	Possible values of staff salaries	monetary: 7 digits, range 6000.00–40000.00

Alternative Terminology for Relational Model

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

Relational Model Terminology

⑦ **Relation schema**

- **A named relation defined by a set of attribute and domain name pairs.**

⑧ **Relation instance**

- **The tuples of a relation at a specific moment**

⑨ **Relational (database) schema**

- **A set of relation schemas, each with a distinct name.**

⑩ **Relational Database**

- **A collection of normalized relations with distinct relation names.**

A Sample Instance of Relational Model

Branch

branchNo	street	city	postCode
B005	22 Deer Rd	London	SW1 4EH
B007	16 Argyll St	Aberdeen	AB2 3SU
B003	163 Main St	Glasgow	G11 9QX
B004	32 Manse Rd	Bristol	BS99 1NZ
B002	56 Clover Dr	London	NW10 6EU

Staff

staffNo	fName	lName	position	sex	DOB	salary	branchNo
SL21	John	White	Manager	M	1-Oct-45	30000	B005
SG37	Ann	Beech	Assistant	F	10-Nov-60	12000	B003
SG14	David	Ford	Supervisor	M	24-Mar-58	18000	B003
SA9	Mary	Howe	Assistant	F	19-Feb-70	9000	B007
SG5	Susan	Brand	Manager	F	3-Jun-40	24000	B003
SL41	Julie	Lee	Assistant	F	13-Jun-65	9000	B005

Agenda

1. **Data Model**
2. **Tabular Structure: Table Schema**
3. **Properties of Relations**
4. **Keys**
5. **Integrity Constraints**

Relation = Table ?

- **EMPLOYEES** is a table. Is it a relation?

EMPLOYEES

eid	ename	position	dependents
e001	Smith, John	Agent	Michael J., Susan R.
e002	Andrews, David	Superintendent	David M. Jr.
e003	Jones, Franklin	Agent	Andrew K., Mark W., Louisa M.

Properties of Relations

① **Relation name is distinct from all other relation names in relational (database) schema.**

② **Each attribute has a distinct name.**

③ **Order of attributes has no significance.**

④ **Order of tuples has no significance, theoretically.**

Properties of Relations

⑤ **Each cell of relation contains exactly one atomic (single) value (or non multi-valued attributes).**

⑥ **Values of an attribute are all from the same domain (or non internal structure attributes).**

⑦ **Each tuple is distinct; there are no duplicate tuples.**

First Normal Form Rule(= ⑤+⑥): In defining tables, the relational model insists that columns that have multi-valued attributes or have any internal structure are not permitted.

Property ⑤: Atomic (Single) Value

Multi-valued attributes

EMPLOYEES

eid	ename	position	dependents
e001	Smith, John	Agent	Michael J., Susan R.
e002	Andrews, David	Superintendent	David M. Jr.
e003	Jones, Franklin	Agent	Andrew K., Mark W., Louisa M.

EMPLOYEES

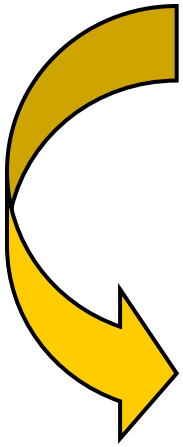
eid	ename	position	dependent1	dependent2	dependent3	…
e001	Smith, John	Agent	Michael J.	Susan R.		
e002	Andrews, David	Superintendent	David M. Jr.			
e003	Jones, Franklin	Agent	Andrew K.	Mark W.	Louisa M.	

Property ⑤: Atomic (Single) Value

Waste space

EMPLOYEES

eid	ename	position	dependent1	dependent2	dependent3	…
e001	Smith, John	Agent	Michael J.	Susan R.		
e002	Andrews, David	Superintendent	David M. Jr.			
e003	Jones, Franklin	Agent	Andrew K.	Mark W.	Louisa M.	



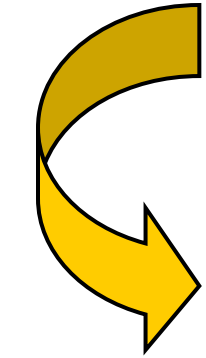
EMPLOYEES

eid	ename	position
e001	Smith, John	Agent
e002	Andrews, David	Superintendent
e003	Jones, Franklin	Agent

DEPENDENTS

eid	dependent
e001	Michael J.
e001	Susan R.
e002	David M. Jr.
e003	Andrew K.
e003	Mark W.
e003	Louisa M.

Property ⑥: Non Internal Structure Attributes



Internal Structure Attributes

EMPLOYEES

eid	ename	position
e001	Smith, John	Agent
e002	Andrews, David	Superintendent
e003	Jones, Franklin	Agent

EMPLOYEES

eid	lname	fname	position
e001	Smith	John	Agent
e002	Andrews	David	Superintendent
e003	Jones	Franklin	Agent

First Normal Form Rules(= ⑤+⑥)

EMPLOYEES

eid	ename		position	dependents
	lname	fname		
e001	Smith	John	Agent	Michael J.
				Susan R.
e002	Andrews	David	Superintendent	David M. Jr.
e003	Jones	Franklin	Agent	Andrew K.
				Mark W.
				Louisa M.

Table

EMPLOYEES

eid	lname	fname	position
e001	Smith	John	Agent
e002	Andrews	David	Superintendent
e003	Jones	Franklin	Agent

DEPENDENTS

eid	dependent
e001	Michael J.
e001	Susan R.
e002	David M. Jr.
e003	Andrew K.
e003	Mark W.
e003	Louisa M.

Relation

Agenda

1. **Data Model**
2. **Tabular Structure: Table Schema**
3. **Properties of Relations**
4. **Keys**
5. **Integrity Constraints**

Relational Keys

Property ⑦ : Each tuple is distinct

- How do we determine if two tuples (or rows) are the same?

  ---- keys

  -- superkey

  -- candidate key

  -- primary key

  -- alternate keys

  -- foreign key

Superkey

- Superkey
  - An attribute, or set of attributes, that uniquely identifies a tuple within a relation.

students

sid	lname	fname	class	telephone
1	Jones	Allan	2	555-1234
2	Smith	John	3	555-4321
3	Brown	Harry	2	555-1122
5	White	Edward	3	555-3344

sid

lname

fname

(lname,fname)

telephone

(lname,fname,telephone)

(sid,fname)

Any set of attributes that includes a superkey is a superkey

Candidate Key

- Candidate Key (CK)
  - Definition: a superkey such that no proper subset is a superkey within the relation.
  - Two properties:
    - Uniqueness -- in each tuple of R, values of CK uniquely identify that tuple.
    - Irreducibility -- no proper subset of CK has the uniqueness property.

students

sid	lname	fname	class	telephone
1	Jones	Allan	2	555-1234
2	Smith	John	3	555-4321
3	Brown	Harry	2	555-1122
5	White	Edward	3	555-3344

Sid is a candidate key

Primary Key and Alternate Key

- Primary Key (PK)
  - A selected candidate key
- Alternate Keys (AK)
  - Candidate keys that are not selected to be primary key.

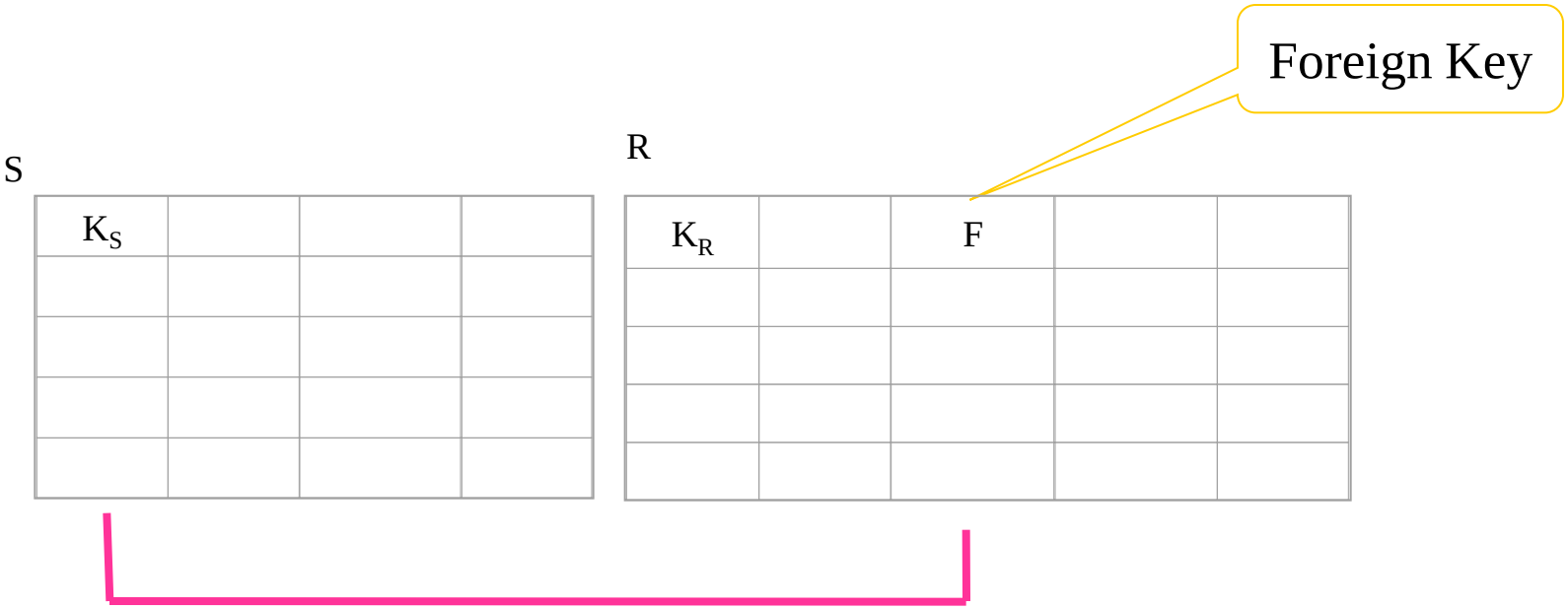
Example -- Primary Key and Alternate Key

CK? PK? AK?

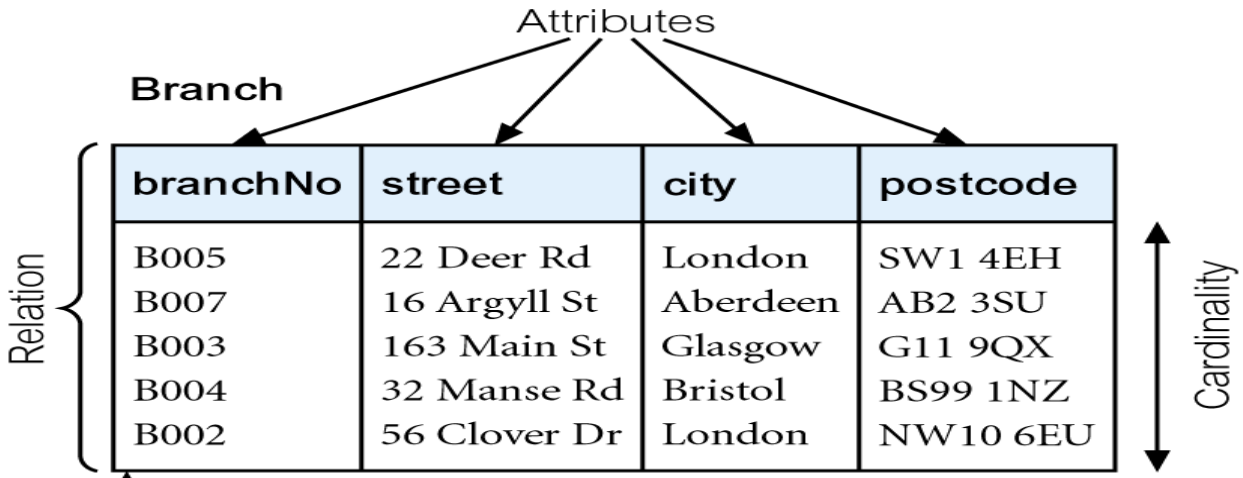
branchNo	street	city	postcode
B005	22 Deer Rd	London	SW1 4EH
B007	16 Argyll St	Aberdeen	AB2 3SU
B003	163 Main St	Glasgow	G11 9QX
B004	32 Manse Rd	Bristol	BS99 1NZ
B002	56 Clover Dr	London	NW10 6EU

Foreign Key

- Foreign Key (FK)
  - Attribute, or set of attributes, within one relation that matches candidate key of some (possibly same) relation.



Example – Foreign Key



Staff

Relation

staffNo	fName	lName	position	sex	DOB	salary	branchNo
SL21	John	White	Manager	M	1-Oct-45	30000	B005
SG37	Ann	Beech	Assistant	F	10-Nov-60	12000	B003
SG14	David	Ford	Supervisor	M	24-Mar-58	18000	B003
SA9	Mary	Howe	Assistant	F	19-Feb-70	9000	B007
SG5	Susan	Brand	Manager	F	3-Jun-40	24000	B003
SL41	Julie	Lee	Assistant	F	13-Jun-65	9000	B005

Agenda

1. **Data Model**
2. **Tabular Structure: Table Schema**
3. **Properties of Relations**
4. **Keys**
5. **Integrity Constraints**

Integrity Constraints

- Integrity constraints are used to ensure accuracy and consistency of data in a relational database
- Integrity constraints include:
  - ① Domain constraints
    - Every attribute has an associated domain
  - ② Entity integrity
  - ③ Referential integrity
  - ④ General (or Enterprise) constraints

Null

- Represents the value for an attribute that is currently unknown or not applicable for tuple.
- Deals with incomplete or exceptional data.
- Represents the absence of a value and is not the same as zero or spaces, which are values.

Entity Integrity

- Base relation
  - A named relation corresponding to an entity in the conceptual schema, whose tuples are physically stored in the database
- Entity integrity: in a base relation, no attribute of a primary key can be null.
  - Why is the rule restricted to base relations?
  - Why does the rule apply only to primary key and not more generally to candidate keys?

Example -- Primary Key and Candidate Key

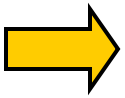
CK? PK?

branchNo	street	city	postcode
B005	22 Deer Rd	London	SW1 4EH
B007	16 Argyll St	Aberdeen	AB2 3SU
B003	163 Main St	Glasgow	G11 9QX
B004	32 Manse Rd	Bristol	BS99 1NZ
B002	56 Clover Dr	London	NW10 6EU

Examples of Entity Integrity

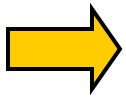
enrollment

sid	cno	major
1	101	No
1	108	Yes
2	105	No
3	101	Yes
3	108	No
5	102	No
5	105	No



enrollment

sid	cno	major	first
1	101	No	Y
1	108	Yes	Y
2	105	No	Y
3	101	Yes	Y
3	108	No	Y
5	102	No	Y
5	105	No	Y



enrollment

sid	cno	major	time
1	101	No	1
1	108	Yes	1
2	105	No	1
3	101	Yes	1
3	108	No	1
5	102	No	1
5	105	No	1

× sid?

× cno?

√ (sid,cno) ?

√ (sid,cno,major) ?

(sid,cno,first)?

(sid,cno,time)?

Referential Integrity

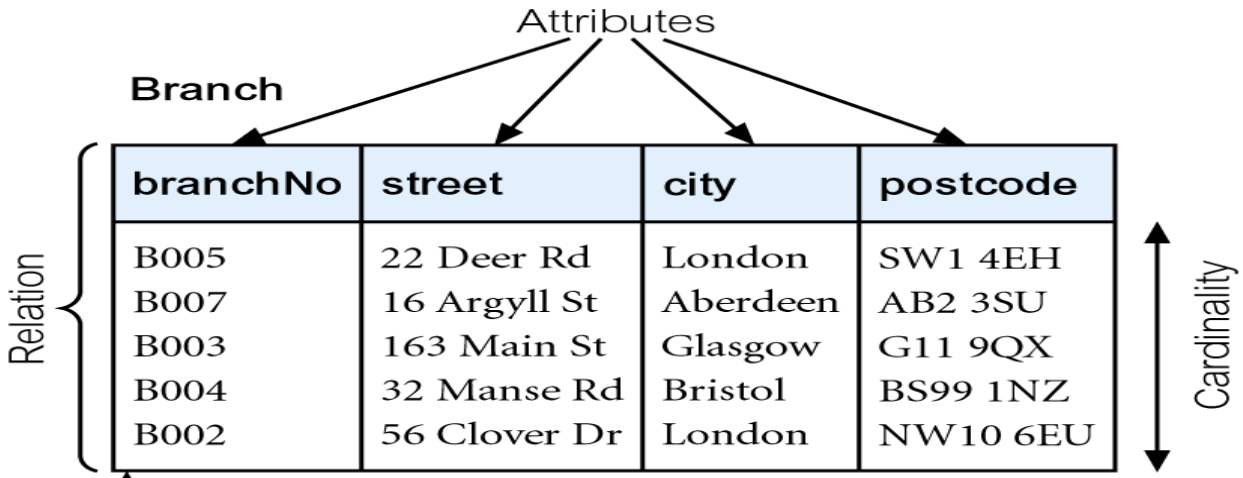
- If a foreign key exists in a relation, either foreign key value must

  ① match a candidate key value of some tuple in its home relation

  or

  ② be wholly null.

Examples of Referential Integrity



Branch

branchNo	street	city	postcode
B005	22 Deer Rd	London	SW1 4EH
B007	16 Argyll St	Aberdeen	AB2 3SU
B003	163 Main St	Glasgow	G11 9QX
B004	32 Manse Rd	Bristol	BS99 1NZ
B002	56 Clover Dr	London	NW10 6EU

Staff

staffNo	fName	lName	position	sex	DOB	salary	branchNo
SL21	John	White	Manager	M	1-Oct-45	30000	B005
SG37	Ann	Beech	Assistant	F	10-Nov-60	12000	B003
SG14	David	Ford	Supervisor	M	24-Mar-58	18000	B003
SA9	Mary	Howe	Assistant	F	19-Feb-70	9000	B007
SG5	Susan	Brand	Manager	F	3-Jun-40	24000	B003
SL41	Julie	Lee	Assistant	F	13-Jun-65	9000	B005

General (Enterprise) Constraints

- Additional rules specified by users or database administrators of a database that define or constrain some aspect of the enterprise.
- Example
  - An upper limit of 20 has been placed upon the number of staff that may work at a branch office

Section Objectives

In this section you will learn:

① Terminology of relational models.

② How tables are used to represent data.

③ Properties of database relations.

④ How to identify CKs, PK, and FKs.

⑤ Meaning of entity integrity and referential integrity.

Questions?

Assignments

- **Exercise 1: Part Ⅱ, and Part Ⅲ**

Prerequisites for Next Section

- **Readings:**
  - **Required:** Connolly and Begg, Section 2.2
  - **Required:** Connolly and Begg, section 4.1
  - **Required:** Connolly and Begg, sections 7.1 and 7.2

- **Assessments**
  - Multiple-Choice Quiz 2