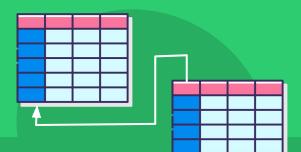


5

The Relational Data Model and Relational Database Constraints

CSF2600700 - BASIS DATA





Outline

1. Relational Model Concept

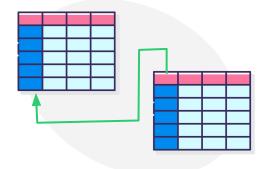
- 2. Characteristics of Relations
- 3. Relational Model Notation
- 4. Relational Model Constraints
- 5. Update Operations, Transactions, and Dealing with Constraint Violations



The Relational Data Model

Relational model

- → Introduced by Ted Codd of IBM Research in 1970
- → The model uses the concept of a mathematical relation
- → First commercial implementations available in early 1980s by IBM and Oracle
- → Has been implemented in a large number of commercial system
- → Popular Relational DBMS: Oracle, DB2, MySQL, PostgreSQL
- → Preceded by hierarchical and network models



Relational Model Concepts

Represents database as a collection of relations

Each relation resembles a table of values

Row

- → Represents a collection of related data values
- → Represents a fact that typically corresponds to a real-world entity or relationship

Table name and column names

→ Interpret the meaning of the values in each row

Formal Terminology:

 $\mathsf{Row} \to \textbf{Tuple}$

Column header → attribute

Table → **Relation**

Relational Model Concepts (cont'd.)

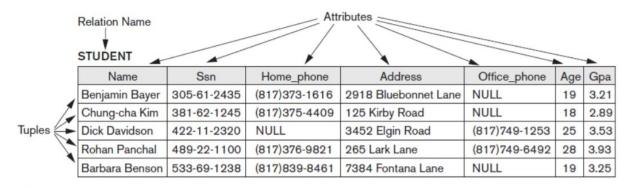


Figure 3.1
The attributes and tuples of a relation STUDENT.

Domains, Attributes, Tuples, and Relations

Domain D

Set of atomic values

Example:

- → GPA: real number between 0 and 4
- → Local_phone_numbers. The set of seven-digit phone numbers valid within a particular area code

Atomic values

→ Each value indivisible

Specifying a domain

→ •Data type specified for each domain

Relation Schema R

- Denoted by $\mathbf{R}(A_1, A_2, ..., A_n)$ Made up of a relation name R and a list of attributes $A_1, A_2, ..., A_n$

Example: **STUDENT**(*Name*, *SSN*, *Home_phone*, *Address*, *Office_phone*, *Age*, *Gpa*)

Attribute A

Name of a role played by some domain D in the relation schema R \rightarrow

Degree (or arity) of a relation

Number of attributes n of its relation schema

Example: **STUDENT**: a relation of degree 7

Relation (or relation state)

Set of n-tuples $\mathbf{r} = \{t_1, t_2, ..., t_m\}$

Each n-tuple t

- Ordered list of n values $t = \langle v_1, v_2, ..., v_n \rangle$ Each value v_i , $1 \le i \le n$, is an element of dom(A_i) or is a special NULL value \rightarrow

Example: t = <'Benyamin Bayer', '305-61-2435', ..., 3.21>

Relation (or relation state) **r(R)**

- \rightarrow Mathematical relation of degree n on the domains dom(A₁), dom(A₂), ..., dom(A_n)
- → Subset of the Cartesian product of the domains that define R:

$$r(R) \subseteq (dom(A_1) \times dom(A_2) \times ... \times dom(A_n))$$

The Cartesian product specifies all possible **combinations** of values from the underlying domains

Example

Given relation schema

$R(A_1, A_2)$

Cartesian product of the domain \rightarrow dom (A₁) X dom (A₂):

A state of R:

is a subset of

Cardinality

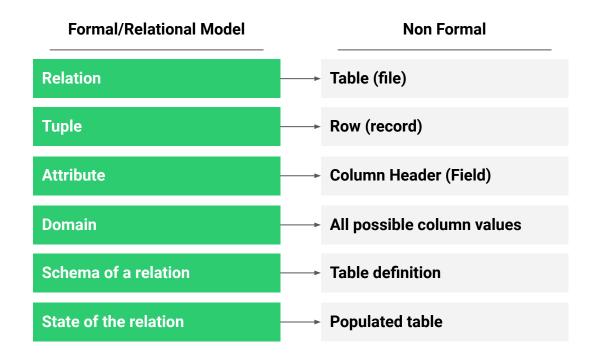
→ Total **number of values** in domain

Current relation state

- → Relation state at a given time
- → Reflects only the valid tuples that represent a particular state of the real world

Attribute names: Indicate different roles, or interpretations, for the domain

Equivalent Terminology



Outline

- 1. Relational Model Concept
- 2. Characteristics of Relations
- 3. Relational Model Notation
- 4. Relational Model Constraints
- 5. Update Operations, Transactions, and Dealing with Constraint Violations



Characteristics of Relations

Ordering of tuples in a relation

- → Relation defined as a set of tuples
- → Elements have **no order** among them

Figure 3.2
The relation STUDENT from Figure 3.1 with a different order of tuples.

STUDENT

Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	(817)749-1253	25	3.53
Barbara Benson	533-69-1238	(817)839-8461	7384 Fontana Lane	NULL	19	3.25
Rohan Panchal	489-22-1100	(817)376-9821	265 Lark Lane	(817)749-6492	28	3.93
Chung-cha Kim	381-62-1245	(817)375-4409	125 Kirby Road	NULL	18	2.89
Benjamin Bayer	305-61-2435	(817)373-1616	2918 Bluebonnet Lane	NULL	19	3.21

Ordering of values within a tuple

- → Order of attributes and values is **not that important**
- → As long as correspondence between attributes and values maintained

Alternative definition of a relation

- → Tuple considered as a set of (<attribute>, <value>) pairs
- → Each pair gives the value of the mapping from an attribute A, to a value v, from dom(A,)

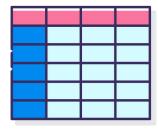


Figure 3.3

Two identical tuples when the order of attributes and values is not part of relation definition.

```
t = < (Name, Dick Davidson),(Ssn, 422-11-2320),(Home_phone, NULL),(Address, 3452 Elgin Road), (Office_phone, (817)749-1253),(Age, 25),(Gpa, 3.53)>
```

$$t = <$$
 (Address, 3452 Elgin Road),(Name, Dick Davidson),(Ssn, 422-11-2320),(Age, 25), (Office_phone, (817)749-1253),(Gpa, 3.53),(Home_phone, NULL)>

Use the first definition of relation

- → Attributes and the values within tuples are ordered
- → Simpler notation

Values in tuples

Each value in a tuple is atomic

Flat relational model

- → Composite and multivalued attributes not allowed
- → First normal form assumption

Multivalued attributes

Example: Favourite color = {red, green}

→ Must be represented by separate relations

Composite attributes

Example: Address can be divided into Street_address, City, State, Zip.

→ Represented only by simple component attributes in basic relational model

NULL Values

Represent the values of attributes that may be unknown or may not apply to a tuple

Meanings for NULL values:

Value unknown

Value exists but is not available

Attribute does not apply to this tuple (also known as value undefined)

Exercise

From the following tables, which one is a **relation** in a relational database?

R₂

R1 A B <u>C</u> D

a2 {b1, b2} c1 d5

a2 b7 c9 d5

a2 b23 c22 d1

.....

A B C D

a2 b2 c6 d1
a2 b7 c9 d5
a2 b7 c9 d5
.....

E# Ename AGE ADDRESS

E2 Diamond 45 1888 Buford Hyw.
E1 Smith 30 3302 Peachtree Rd., Atlanta, GA
E3 Evan null Baker Ct. Atlanta

Outline

- 1. Relational Model Concept
- 2. Characteristics of Relations
- 3. Relational Model Notation
- 4. Relational Model Constraints
- 5. Update Operations, Transactions, and Dealing with Constraint Violations



Relational Model Notation

Examples

Relation schema R of degree n

 \rightarrow Denoted by R(A₁, A₂, ..., A_n)

Uppercase letters Q, R, S

→ Denote relation names

Lowercase letters q, r, s

→ Denote relation states

Letters t, u, v

→ Denote tuples

STUDENT(Name, SSN, Home_phone, Address, Office_phone, Age, Gpa)

STUDENT(Name, SSN, Home_phone, Address, Office_phone, Age, Gpa)

$$r = \{<0,a>, <1,a>, <1,c>\}$$
 $r = \{t_1, t_2, t_3\}$

t = <'Benyamin Bayer', '305-61-2435', ..., 3.21>

Relational Model Notation (Contd.)

Name of a relation schema: STUDENT

→ Indicates the current set of tuples in that relation

Notation: STUDENT(Name, Ssn, ...)

→ Refers only to relation schema

Attribute A can be qualified with the relation name R to which it belongs Using the dot notation $\mbox{R.A}$



Relational Model Notation (Contd.)

n-tuple t in a relation r(R)

- Denoted by $t = \langle v_1, v_2, \dots, v_n \rangle$
- v, is the value corresponding to attribute A.

Component values of tuples

- $t[A_i]$ and $t.A_i$ refer to the value v_i in t for attribute A_i $t[A_u, A_w, ..., A_z]$ and $t.(A_u, A_w, ..., A_z)$ refer to the subtuple of values $v_u, v_w, ..., v_z$ from t corresponding to the attributes specified in the list

Example

```
the tuple t = <'Barbara Benson', '533-69-1238', '(817)839-8461', '7384 Fontana Lane', NULL, 19, 3.25> from the STUDENT relation
t[Name] = <'Barbara Benson'>, and
t[Ssn, Gpa, Age] = <'533-69-1238', 3.25, 19>.
```

Outline

- 1. Relational Model Concept
- 2. Characteristics of Relations
- 3. Relational Model Notation
- 4. Relational Model Constraints
- 5. Update Operations, Transactions, and Dealing with Constraint Violations



Relational Model Constraints

Constraints

- Restrictions on the actual values in a database state
- → Derived from the rules in the miniworld that the database represents

Constraints Categories:

Inherent model-based constraints or implicit constraints

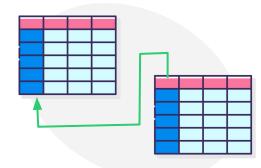
→ Inherent in the data model

Schema-based constraints or explicit constraints

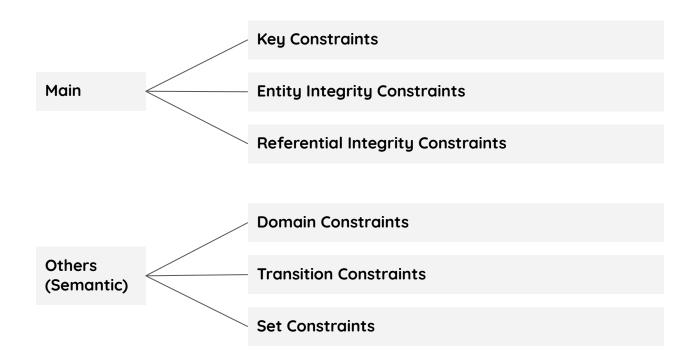
→ Can be directly expressed in schemas of the data model or in DDL

Application-based or **semantic constraints** or **business rules**

- → Cannot be directly expressed in schemas
- → Expressed and enforced by application program



Classification of Relational Integrity Constraints



Domain Constraints

Specify that within each tuple, the value of each **attribute A** must be an atomic value from the **domain dom(A)**

Typically include:

- → Numeric data types for integers and real numbers Characters
- → Booleans
- → Fixed-length strings
- → Variable-length strings
- → Date, time, timestamp
- → Money
- → Other special data types



Key Constraints and Constraints on NULL Values

Key Constraints

No two tuples can have the same combination of values for all their attributes

Superkey (SK)

SK: an attribute or set of attributes that guarantee that no two distinct tuples in any state r of R can have the same value for SK

Key Constraints and Constraints on NULL Values

Key

- → Superkey of R
- Removing any attribute A from K leaves a set of attributes K that is not a superkey of R any more

Key satisfies two properties:

- → Two distinct tuples in any state of relation cannot have identical values for (all) attributes in key
- → Minimal superkey: Cannot remove any attributes and still have uniqueness constraint in above condition hold

Key Constraints and Constraints on NULL Values (Contd.)

Candidate Key

→ Relation schema may have more than one key

Primary key of the relation

- → Designated among candidate keys
- → Underline attribute

Other candidate keys are designated as unique keys / alternate keys

Key Constraints and Constraints on NULL Values (Contd.)

CAR

Figure 3.4
The CAR relation, with two candidate keys: License_number and Engine_serial_number.

License_number	Engine_serial_number	Make	Model	Year
Texas ABC-739	A69352	Ford	Mustang	02
Florida TVP-347	B43696	Oldsmobile	Cutlass	05
New York MPO-22	X83554	Oldsmobile	Delta	01
California 432-TFY	C43742	Mercedes	190-D	99
California RSK-629	Y82935	Toyota	Camry	04
Texas RSK-629	U028365	Jaguar	XJS	04

Example

SSN	FName	LName	BirthDate	Sex	Address
150668001	Izuku	Midoriya	01-05-1997	L	Tanegashima
150668002	Shouto	Todoroki	02-06-1997	L	Kyoto
150668003	Natsu	Hinata	03-05-2000	Р	Bali
150668004	Chiaki	Morisawa	01-05-1997	L	Jakarta
150668005	Shouyou	Hinata	02-07-1997	L	Sendai
150668006	Tobio	Kageyama	01-01-1998	L	Sendai

Super key: SSN, {SSN, Lname}, {FName, BirthDate}, {FName, Sex}, ...

Candidate key: SSN, FName



Alternate key: FName

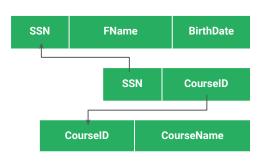
Key Constraints and Constraints on NULL Values (Contd.)

Constraints on NULL Value

- → For an attribute, we can specify whether NULL values are or are not permitted
- → For example: every STUDENT tuple must have a valid, non-NULL value for the Name attribute → then Name of STUDENT is constraint to be NOT NULL



Key Constraints and Constraints on NULL Values (Contd.)



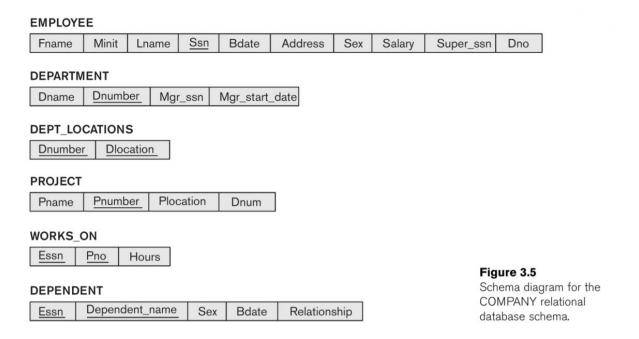
- Set of relation schemas $S = \{R_1, R_2, ..., R_m\}$
- Set of integrity constraints IC

SS	SN	FName			BirthDate	
12	23	Arima Kousei			02-02-1997	
		SSN		CourseID		
			12	3	1	
	C	CourseID		CourseName		
	1		Database			

Relational database state

- Set of relation states **DB** = $\{r_1, r_2, ..., r_m\}$ Each ri is a state of R_i and such that the r_i relation states satisfy integrity constraints specified in IC

Relational Databases and Relational Database Schemas



Relational Databases and Relational Database Schemas (Contd.)

Figure 3.6

One possible database state for the COMPANY relational database schema.

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	В	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	М	30000	333445555	5
Franklin	Т	Wong	333445555	1955-12-08	638 Voss, Houston, TX	М	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	М	38000	333445555	5
Joyce	Α	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	٧	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	М	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	М	55000	NULL	1

DEPARTMENT

Dname	Dnumber	Mgr_ssn	Mgr_start_date
Research	5	333445555	1988-05-22
Administration	4	987654321	1995-01-01
Headquarters	1	888665555	1981-06-19

DEPT_LOCATIONS

Dnumber	Dlocation
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

Relational Databases and Relational Database Schemas (Contd.)



Invalid State

Does not obey all the integrity constraints



vallu State

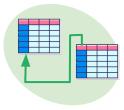
Satisfies all the constraints in the defined set of integrity constraints IC

Relational Databases and Relational Database Schemas (Contd.)



Entity integrity constraint

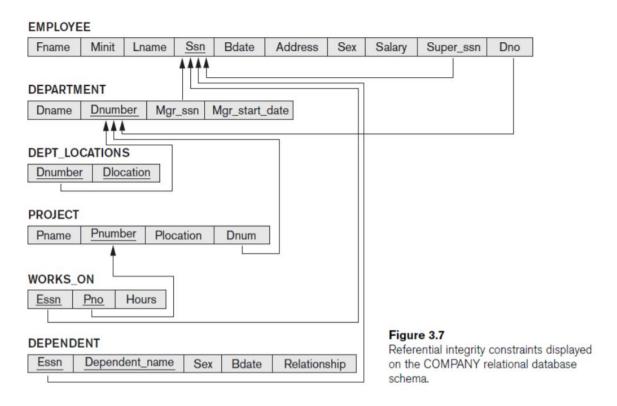
→ No primary key value can be NULL



Referential integrity constraint

- → Specified between two relations
- → Maintains consistency among tuples in two relations

Integrity, Referential Integrity, and Foreign Keys



Integrity, Referential Integrity, and Foreign Keys (Contd.)

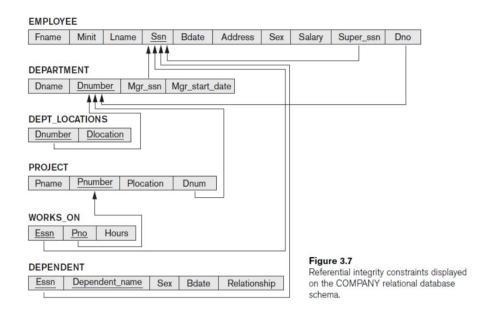
Foreign key rules:

- 1. The attributes in FK have the same domain(s) as the primary key attributes PK
- 2. Value of **FK** in a tuple t_1 of the current state $r_1(R_1)$ either occurs as a value of **PK** for some tuple t_2 in the current state $r_2(R_2)$ or is **NULL**

Integrity, Referential Integrity, and Foreign Keys (Contd.)

Diagrammatically display referential integrity constraints

- → Directed arc from each foreign key to the relation it references
- → All integrity constraints should be specified on relational database schema



Other Types of Constraints

Semantic integrity constraints

- → May have to be specified and enforced on a relational database ∘
- → Use triggers and assertions
- → More common to check for these types of constraints within the application programs

Latihan

Basis data yang memproses order (pemesanan) pada sebuah perusahaan memiliki 6 relasi berikut:

CUSTOMER (Cust#, Cname, City)
ORDER (Order#, Odate, Cust#, Ord_Amt)
ORDER_ITEM (Order#, Item#, Qty)
ITEM (Item#, Unit_price)
SHIPMENT (Order#, Warehouse#, Ship_date)
WAREHOUSE (Warehouse#, City)

Ord_Amt mengacu pada jumlah harga pada satu kali order. Odate menyatakan tanggal pemesanan dilakukan, Ship_date menyatakan tanggal pengiriman barang yang dipesan customer dari gudang.

Asumsikan bahwa suatu order dapat mengambil barang dari beberapa gudang (warehouse). Nyatakan foreign key yang mungkin untuk skema basis data ini.

Outline

- 1. Relational Model Concept
- 2. Characteristics of Relations
- 3. Relational Model Notation
- 4. Relational Model Constraints
- 5. Update Operations, Transactions, and Dealing with Constraint Violations



Update Operations, Transactions, and Dealing with Constraint Violations

Operations of the relational model can be categorized into retrievals and updates Basic operations that change the states of relations in the database:

Insert

Delete

Update (or Modify)

Update Operations, Transactions, and Dealing with Constraint Violations

Figure 3.6

One possible database state for the COMPANY relational database schema.

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	В	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	М	30000	333445555	5
Franklin	Т	Wong	333445555	1955-12-08	638 Voss, Houston, TX	М	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	М	38000	333445555	5
Joyce	Α	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	٧	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	М	55000	NULL	1

DEPARTMENT

Dname	Dnumber	Mgr_ssn	Mgr_start_date
Research	5	333445555	1988-05-22
Administration	4	987654321	1995-01-01
Headquarters	1	888665555	1981-06-19

DEPT_LOCATIONS

Dnumber	Diocation
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

Update Operations, Transactions, and Dealing with Constraint Violations

Figure 3.6
One possible database state for the COMPANY relational database schema.

WORKS ON

Essn	Pno	Hours
123456789	1	32.5
123456789	2	7.5
666884444	3	40.0
453453453	1	20.0
453453453	2	20.0
333445555	2	10.0
333445555	3	10.0
333445555	10	10.0
333445555	20	10.0
999887777	30	30.0
999887777	10	10.0
987987987	10	35.0
987987987	30	5.0
987654321	30	20.0
987654321	20	15.0
888665555	20	NULL

PROJECT

Pname	Pnumber	Plocation	Dnum
ProductX	1	Bellaire	5
ProductY	2	Sugarland	5
ProductZ	3	Houston	5
Computerization	10	Stafford	4
Reorganization	20	Houston	1
Newbenefits	30	Stafford	4

DEPENDENT

Essn	Dependent_name	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	М	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	М	1942-02-28	Spouse
123456789	Michael	М	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

The Insert Operation

Provides a list of attribute values for a new tuple t that is to be inserted into a relation R **Can violate** any of the four types of constraints if an insertion violates one or more constraints

→ Default option is to reject the insertion



The Delete Operation

Can violate only referential integrity ff tuple being deleted is referenced by foreign keys from other tuples

Restrict

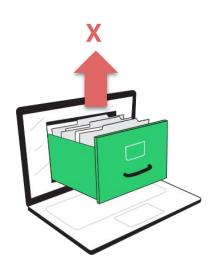
→ Reject the deletion

Cascade

Propagate the deletion by deleting tuples that reference the tuple that is being deleted

Set null or set default

→ Modify the referencing attribute values that cause the violation



The Update Operation

Necessary to specify a condition on attributes of relation

→ Select the tuple (or tuples) to be modified

If attribute not part of a primary key nor of a foreign key

→ Usually causes no problems

Updating a primary/foreign key

→ Similar issues as with Insert/Delete



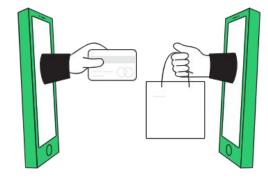
The Transaction Concept

Transaction

- → Executing program
- → Includes some database operations
- → Must leave the database in a valid or consistent state

Online transaction processing (OLTP) systems

→ Execute transactions at rates that reach several hundred per second



Latihan

Apakah ada constraints yang dilanggar pada operasi berikut?

- 1. **Insert** < 'ProductA', 4, 'Bellaire', 2 > into PROJECT.
- 2. Insert < '677678989', null, '40.0' > into WORKS_ON.
- 3. Delete the WORKS_ON tuples with ESSN= '333445555'.
- 4. **Delete** the EMPLOYEE tuple with SSN= '987654321'.
- 5. **Modify** the SUPERSSN attribute of the EMPLOYEE tuple with SSN= '999887777' to '943775543'.

Figure 5.6

One possible database state for the COMPANY relational database schema.

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	В	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	М	30000	333445555	5
Franklin	Т	Wong	333445555	1955-12-08	638 Voss, Houston, TX	М	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	М	38000	333445555	5
Joyce	Α	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	М	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	М	55000	NULL	1

DEPARTMENT

Dname	Dnumber	Mgr_ssn	Mgr_start_date
Research	5	333445555	1988-05-22
Administration	4	987654321	1995-01-01
Headquarters	1	888665555	1981-06-19

DEPT_LOCATIONS

Dnumber	Diocation
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

WORKS ON

Essn	Pno	Hours	
123456789	1	32.5	
123456789	2	7.5	
666884444	3	40.0	
453453453	1	20.0	
453453453	2	20.0	
333445555	2	10.0	
333445555	3	10.0	
333445555	10	10.0	
333445555	20	10.0	
999887777	30	30.0	
999887777	10	10.0	
987987987	10	35.0	
987987987	30	5.0	
987654321	30	20.0	
987654321	20	15.0	
888665555	20	NULL	

PROJECT

Pname	Pnumber	Plocation	Dnum
ProductX	1	Bellaire	5
ProductY	2	Sugarland	5
ProductZ	3	Houston	5
Computerization	10	Stafford	4
Reorganization	20	Houston	1
Newbenefits	30	Stafford	4

Essn	Dependent_name	Sex	Bdate	Relationship
333445555	Alice	F	1986-04-05	Daughter
333445555	Theodore	М	1983-10-25	Son
333445555	Joy	F	1958-05-03	Spouse
987654321	Abner	М	1942-02-28	Spouse
123456789	Michael	М	1988-01-04	Son
123456789	Alice	F	1988-12-30	Daughter
123456789	Elizabeth	F	1967-05-05	Spouse

Summary

Characteristics differentiate relations from ordinary tables or files

Classify database constraints into:

→ Inherent model-based constraints, explicit schemabased constraints, and application-based constraints

Modification operations on the relational model:

→ Insert, Delete, and Update

