CSC 742
Database Management Systems

Topic #5:
Relational Model

Motivation

A relation is a mathematical abstraction for a table

- The theory of relations provides an elegant basis for databases
- Relational databases can be efficiently implemented on current architectures

Relation Model

- A relation is a table of values

Relation Name

Attribute

Student

Tuple

Name	Student ID	Class	Major
Smith	17	1	CS
Brown	8	2	CS

- Each tuple represents a collection of related data.
- All values in a column are of the same data type.

Is this a relation?

Student

Name	Student ID	Emails	Major
Smith	17	{smith@csc.ncsu.edu, smith@unity.ncsu.edu, smith@eos.ncsu.edu}	CS
Brown	8	{brown@unity.ncsu.edu}	CS

Domain D

- A domain of an attribute specifies all possible values the attribute can have.
- A domain is a set of **atomic** values.
  - Each value in the domain is indivisible.
  - Often expressed as a data type.
- Describe the domains of the following attributes
  - US_Phone_number:
  - Age:
  - Social_Security_Number:
  - Name:

Relation Schema

- A relation schema R(A1 ... An) describes the intension of a relation
  - R: relation name
  - A1, …, An: a list of attributes
  - n: the degree of a relation of this schema.
  - attributes Ai name the roles or columns
  - dom(Ai) = Di: the domain of Ai.
  - attributes are distinct but may share their domain

An Example Schema

- Student(Name, SSN, HomePhone, Address, OfficePhone, Age, GPA)
  - Relation name: ?
  - Attributes: ?
  - Domains of the attributes: ?
  - Degree:

Relation

- A relation r(R) is an instance of the corresponding relation schema R.
  - Set of tuples of the form <v1 ... vn>
  - Each tuple is an ordered list
  - vi belongs to dom(Ai) or is null
  - thus vi is atomic: definition of 1NF
  - Also called *relation extension*

A Mathematical Definition of Relation

- A relation r(R) is a mathematical relation of degree n of the domains dom(A1), dom(A2), …, dom(An), which is a subset of the Cartesian product of the domains that define R:

  $r(R) \subseteq (dom(A1) \times dom(A2) \times \ldots \times dom(An))$

- Mind exercise:
  - What is the number of all possible relations of R?

Achtung!

The ordering of

- attributes is important mathematically, but not in practice
- tuples is not important mathematically, but is in practice

An Alternative Definition of Relation

- A relation schema R={A1, A2, …, An} is a set of attributes.
- A relation r(R) is a finite set of *mappings*

  r={t1, t2, …, tm},

  - where each ti is a mapping from R to D,
  - D is the union of dom(Ai)'s,
  - t[Ai] must be in dom(Ai).
- Each tuple can be considered as a set of attribute-value pairs.

Notation

- Schemas: Q, R, S
- Instances: q, r, s
- Tuples: t, u, v
- t[Ai] = vi, where t = <v1 ... vn>
- t[Ai, Aj, ... Ak] = <vi, vj, ... vk> (subtuple)

Relational Databases

- A relational database schema is
  - A set of relation schemas S={R1, R2, …, Rm} and
  - A set of integrity constraints IC.
- A relational database is a set of relations DB = {r1, r2, …, rm} such that
  - each ri is a relation of Ri,
  - and ri satisfies the constraints in IC.

Constraints

- Constraints that can specified on relational schemas.
  - Domain constraints
  - Key constraints
  - Constraint about NULL
  - Reference integrity constraints

Domain Constraints

- vi belongs to dom(Ai)
- Domains map to standard data types

Key Constraints

- Superkey: subset of {Ai ... An} that uniquely identifies R
- Key: minimal superkey
- Because r(R) must be sets, {Ai ... An} is always a superkey
- Candidate key: key
- Primary key: any single key so designated

An Example

Car

LicenseNumber	EngineSerialNumber	Make	Model	Year
Texas ABC-739	A691324	Ford	Mustang	1996
Florida TVP-234	B43123123	Toyota	Camery	2000
NC 341-1324	2HG32341235	Honda	Civic	1998

- Super Keys:
- Keys:
- Candidate Keys:
- Primary Key:

Entity Integrity Constraints

- No primary key attribute may be null
- Motivation: a violation would be analogous to an uninitialized object

Referential Integrity Constraints

- Specified on two relations
  - Note that the previous constraints are on individual relations.
- Intuition
  - If a tuple in one relation refers to a tuple in another relation, the second tuple should exist.
  - Specified through *foreign keys*.

Foreign Keys

- A set of attributes, FK in relation R1, is foreign key iff
  - Attributes in FK occur as (primary) key in R2
  - FK reference or refer to the relation R2.

An Example

Driver

Name	SSN	BirthDate	Car
Brown	123-45-6789	01/01/50	Texas ABC-739
Smith	222-33-4444	02/22/60	Florida TVP-234
John	444-55-6666	05/01/56	NC 341-1324

Car

LicenseNumber	EngineSerialNumber	Make	Model	Year
Texas ABC-739	A691324	Ford	Mustang	1996
Florida TVP-234	B43123123	Toyota	Camery	2000
NC 341-1324	2HG32341235	Honda	Civic	1998

Referential Integrity Constraints

- If t[FK] in r1 is all non-null, then u with u[FK]=t[FK] exists in r2
- Motivation: a violation is analogous to a dangling pointer

Semantic Integrity Constraints

- Application-specific constraints
  - typically involve restrictions on the values of an attribute with respect to some other attributes
  - some varieties of them may be specified as assertions in SQL2

General Strategy for Operations

- When legal, execute
- When illegal
  - reject
  - restore consistency (referential integrity)
    - change referenced tuple (relation)
    - change referencing tuple

Insert

Insert <v1 ... vn> into R (means current instance)

- execute: just add it
- reject, e.g., when the key is already used
- Correct the problem
  - Ask user to create referenced tuple or change the referencing tuple.
  - Creating referenced tuple may be cascading.

An Example

Student

Name	<u>StudentID</u>	Class	Major
Smith	17	1	CS
Brown	8	2	CS

- Insert
  - <John, “N/A”, 2, “CS”>
  - <John, NULL, 2, “CS”>
  - <John, 17, 2, “CS”>
  - <John, 18, 2, “CS”>

Delete

Delete <v1 ... vn> from R

- execute: remove
- reject, e.g., if other tuples refer to it
- Correct the problem
  - cascade by deleting referring tuples
  - change values to null or default

Modify

- As above

An Example

Grade

StudentID	Course	…	Grade
17	CSC 742	…	A
8	CSC 742	…	B+

Student

Name	StudentID	Class	Major
Smith	17	1	CS
Brown	8	2	CS

- Consider the following operations:
  - Delete the 1st tuple from Student.
  - Insert <18, “CSC742”, …, “A-”> into Grade
  - Change StudentID of the 1st tuple from 17 to 15.
- Options: reject, correct with cascade, correct by changing the values (set to NULL, or a valid value).