Northeastern University
Khoury College of Computer and Information Sciences

Relational Databases

15

Key Topics

- Tables & Relations
- Primary and Foreign Keys
- Referential Integrity
- Relational Operations
- Data Retrieval Approach
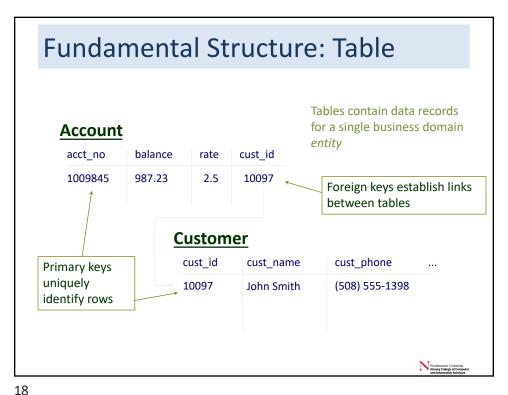
Ars Doceo
Northeastern University

16

Northeastern
Khoury Colleg
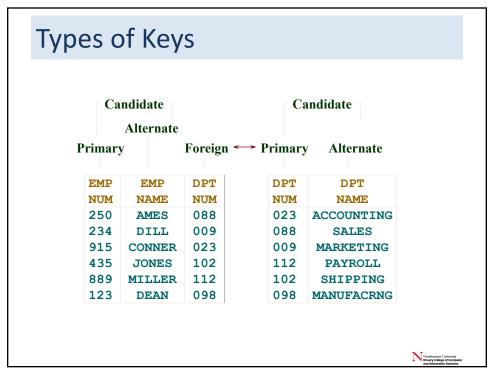**Computer Scie**

Relational Database

A relational database S is a set of relations such that $S = (R_1, R_2, \ldots, R_n)$ and $R_i(A_{i1}, A_{i2}, \ldots, A_{ik})$.

Each relation instance must satisfy the integrity constraints specified in the relational database schema.

17

Fundamental Structure: Table

Tables contain data records for a single business domain *entity*

Account

acct_no	balance	rate	cust_id
1009845	987.23	2.5	10097

Foreign keys establish links between tables

Customer

cust_id	cust_name	cust_phone	...
10097	John Smith	(508) 555-1398	

Primary keys uniquely identify rows

18

Types of Keys

	Candidate			Candidate	
	Alternate				
Primary			Foreign ⟷ Primary		Alternate

EMP NUM	EMP NAME	DPT NUM	DPT NUM	DPT NAME
250	AMES	088	023	ACCOUNTING
234	DILL	009	088	SALES
915	CONNER	023	009	MARKETING
435	JONES	102	112	PAYROLL
889	MILLER	112	102	SHIPPING
123	DEAN	098	098	MANUFACRNG

Northeastern University
Khoury College of Computer and Information Sciences

19

Database Integrity

- There are four types of integrity constraints for relational databases:
  1. *domain integrity*
  2. *entity integrity*
  3. *referential integrity*
  4. *semantic integrity*

Northeastern University
Khoury College of Computer and Information Sciences

20

Northeastern
Khoury Colleg
Computer Sci

Domain Integrity

- Domain integrity is applied on the attribute and requires that every attribute A_i is a value from its domain D_i.
- More formally, for a relation $R(A_1, A_2, \ldots, A_n)$, each instance of the relation satisfies:

$$\{<A_1{:}V_1, A_2{:}V_2, \ldots, A_n{:}V_n> \mid V_1 \in D_1, V_2 \in D_2, \ldots, V_n \in D_n\}$$

21

Entity Integrity

- Entity integrity defines restrictions on the tuples of a relation which ensures the accuracy and consistency of the database.
- It requires that each tuple in a relation is unique.
- The primary key must be a set of attribute(s) that cannot have duplicate values or be *null*.

22

Primary Key Properties

- Primary keys must be:
  - *unique*
  - *not reusable*
  - *unchangeable*
  - *void of semantic meaning*
  - *invisible to the user*

Northeastern University
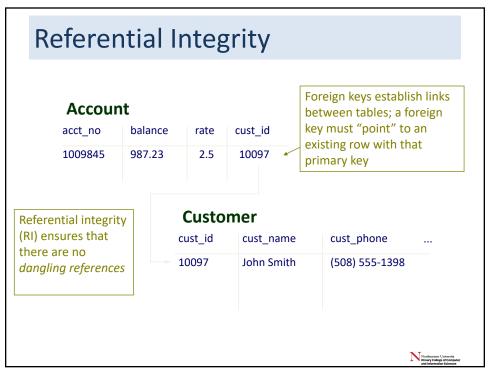Khoury College of Computer
and Information Sciences

23

Defining Primary Keys

- A primary key is a subset of attributes whose values are unique for a relation.
- Often there is no subset of the attributes that satisfies all the above constraints, so a *surrogate* or *artificial* key is created as a new attribute.
- For some relations there is a single attribute that is unique: *natural key*.

Northeastern University
Khoury College of Computer
and Information Sciences

24

Northeastern
Khoury Colleg
Computer Sci

Referential Integrity

Account

acct_no	balance	rate	cust_id
1009845	987.23	2.5	10097

Foreign keys establish links between tables; a foreign key must "point" to an existing row with that primary key

Referential integrity (RI) ensures that there are no *dangling references*

Customer

cust_id	cust_name	cust_phone	...
10097	John Smith	(508) 555-1398	

Northeastern University
Khoury College of Computer and Information Sciences

25

Semantic Integrity

- Semantic integrity ensures that business and other logical rules are enforced.

- Examples of semantic integrity rules:
  - *Number of projects a project manager manages is greater than 0*
  - *A project cannot have a start date after its end date*
  - *A project must have a project manager*
  - *A project manager cannot be a resource on the project that the project manager manages*

Northeastern University
Khoury College of Computer and Information Sciences

26

Northeastern
Khoury Colleg
Computer Sci

Relational Table Rules

The following rules govern tables:

1. All rows are formatted identically
2. Rows are unsorted
3. Columns are in no particular order
4. Each row in a table is a unique instance
5. There are no duplicate column (attribute) names
6. Within the table, values represent data
7. Each column (attribute) value is atomic
8. Primary keys should be non-compound and minimal

Northeastern University
Khoury College of Computer
and Information Sciences

27

Understanding Tables and Keys

1. Consider the table *Sections* below.

   a. What is the likely primary key?

   b. Is the primary key natural or surrogate?

   c. What kind of a key is "*instructor*"?

Table: Sections

crn	title	hours	instructor
1009845	DS4100	4.0	10097
1009861	DS4400	4.0	10097
1009893	CS1200	1.0	10461

Northeastern University
Khoury College of Computer
and Information Sciences

28

Northeastern
Khoury Colleg
Computer Sci

Creating a Relational Schema

Suppose that you need a relational database for an investment company that must track portfolios, investments (such as stocks, bonds, CDs, etc.), and trades in the portfolios, *e.g.*, buying some stock at some price on some day. Sketch a data model for this with some necessary attributes and keys. Map the model to a database schema, *i.e.*, tables, primary keys, foreign keys.

29

Summary, Review, & Questions…

30