Databases

O.O.Bogomolets National Medical University

October 26, 2012

Data and databases

Definition

Data *refers to qualitative or quantitative attributes of a variable or set of variables, stored in some way.*

Definition

Database management system (DBMS) *is a software system specifically designed to hold databases*

Database Management System (DBMS) provides efficient, reliable, convenient, and safe multi-user storage of and access to massive amounts of persistent data.

DBMS

Key concepts:

- Data model
- Schema versus data
- Data definition language (DDL)
- Data manipulation language (DML)

Types

Definition

Data type *is a classification that determines the* **possible values** *for that type; the* **operations** *that can be done on values of that type; the meaning of the data; and the way values of that type can be stored.*

Relational Databases

The Relational Model:

- Very simple model
- Used by all major commercial database systems
- Efficient implementations

- **Instance** – actual contents at given point in time
- **Database** – set of named relations (tables)
- Each **relation** has a set of named *attributes* (columns)
- Each **tuple** (row) has a *value* for each attribute
- Each **attribute** has a *type* (domain)

Definition

Key – *attribute(or set of attributes) whose value is unique (in each tuple)*

Indexes

- Primary mechanism to get improved performance on a database
- Persistent data structure, stored in database
- difference between full table scans and immediate location of tuples
- Many DBMSs build indexes automatically on primary key and/or unique attributes

Relational Design Theory

- Usually many designs possible
- Some are better than others!

Design by decomposition:

- Start with mega relations containing everything
- Decompose into smaller, better relations with same info
- System decomposes based on properties

Boyce-Codd Normal Form

Definition

Normal *forms of relational database theory provide criteria for determining a table's degree of vulnerability to logical inconsistencies and anomalies.*

Boyce-Codd Normal Form

First normal form	Table faithfully represents a relation and has no repeating groups
Second normal form	No non-prime attribute in the table is functionally dependent on a proper subset of any candidate key
Third normal form	Every non-prime attribute is non-transitively dependent on every candidate key in the table. The attributes that do not contribute to the description of the primary key are removed from the table. In other words, no transitivity dependency is allowed.
BoyceCodd normal form	Every non-trivial functional dependency in the table is a dependency on a superkey

Dixi