DD1368 Database Technology Lecture notes

Zixian Song

January 24, 2016

KTH Computer Science
and Communication

Contents

1 Lecture 1: Intro

2016-01-20
Introduction

1.1 Content of databases

- Design of databases
- E/R model, relation model, semistructured model, XML
- Database programing
- SQL, XPath, XQuery
- Relational algebra, Tuple Calculus, FOL.
- ODBC JDBC or etc ...

1.2 Interesing Stuff About Database

- Websearch
- Data mining
- Scientific and medical databases
- integrationg information.
- **NSA! wikileaks HeathCare.gov**
- Still more. ...

1.3 And More

Database often have unique concurrency-control problems.

- Many activities (transactions) at the database at all times.
- Must not confuse actions, e.g. two withdrewal from the same account must each debit the account.
- ATC :

1.4 What is a Data Model

1. Mathematical represention of data.
   - Examples:relational model = tables;
   - Semistruced model = trees,graphs
2. Operations on data
3. Constraints

1.5 Schema

- Relation schema = relation name and attribute list.

1.6 Our Running Example

Beer(<u>name</u>, manf)
Bars(<u>name</u>, addr, license)
Drinkers(<u>name</u>, addr, phone)
Underline = *key* (tuples cannot have the same value in all key attributes).

- Excellent

1.7 Example: Multiattribute Key

```
CREATE TABLE Sells (
    bar CHAR(20),
    beer VARCHAR(20),
    price REAL,
    PRIMARY KEY (bar, beer)
    );
```

1.8 Semistrucured Data

- Another data model, based on trees graphs.
- **Motivation**: flexible represention of data
- **Motivation**: sharing of *documents* among systems and databases

2 Lecture 2: Relational Algebra and Tuple Calculus

2.1 Ideal Picture

- Tuple Calculus: Declarative(logical) expression of what the user wants
- Relational Algebra: Procedural Expression that can obtain answers
- Opetimized Relational Algebra: Can be efficienty executed over database

2.2 Example

"Names of companies in Sweden that have a supplier from China"

```
{x.name | company(x) AND x.country = 'Sweden' AND
(Ey)(Ex)( supplies(y) AND company(z) AND
z.contry = 'China' AND x.id = y.to AND y.from = z.id)}
```

2.3 What is Relational Algebra

- An algebra whose operands are relations or variables that represent relations.
- Operators are designed to do the most common things that we need to do with relations in a database.
- The result is an algebra that can be used as a query language for relations.

2.4 Core Relational Algebra

- Union, intersection, and difference.
  - Usual set operations, but *both operands must have the same relation schema.*
- **Selection**: picking certain rows.
- **Projection**: picking certain columns.
- **Products** and **joins**: compositions of relations.
- **Renaming** of relations and attributes.

2.5 Selection

$$R1 := \sigma_c(R)$$

- C is a condition (as in if statements) that refers to attributes of R2.
- R1 is all those tuples of R2 that satisfy C.

2.6 Selection:Example

Relation Sells:

bar	beer	price
Joes	Bud	2.50
Joes	Miller	2.75
Sues	Bud	2.50
Sues	Miller	3.00

$JoeMenu := \sigma_{bar = "joe's"}(Sells)$:

bar	beer	price
Joes	Bud	2.50
Joes	Miller	2.75

2.7 Projection

$$R2 := \pi_L(R) :$$

- L is a list of attributes from the schema of R2.
- R1 is constructed by looking at each tuple of R2, extracting the attributes on list L, in the order specified, and creating from those components a tuple for R1.
- Eliminate duplicate tuples, if any.

2.8 Product

$$R3 := R1 \times R2$$

- Pair each tuple t1 of R1 with each tuple t2 of R2.
- Concatenation t1 t2 is a tuple of R3.
- Schema of R3 is the attributes of R1 and then R2, in order.
- But beware attribute A of the same name in R1 and R2: use R1.A and R2.A.

2.9 Theta-Join

$$R3 := R1 \bowtie_{\theta\ c} R2$$

2.10 example

Names if companis who supply another company:
$\{x.name | company(x) \wedge (\exists y)(Supplies(Y) \wedge y.from = x.name)\}$