DATABASE DESIGN I - 1DL300

Autumn 2012

An Introductory Course on
Database Systems

http://www.it.uu.se/edu/course/homepage/dbastekn/ht12/

Uppsala Database Laboratory
Department of Information Technology, Uppsala University,
Uppsala, Sweden
Introduction to Database Design Using Entity-Relationship Modeling

Elmasri/Navathe ch 7, 9
Padron-McCarthy/Risch ch 2-3

Sobhan Badiozamany
Silvia Stefanova

Department of Information Technology
Uppsala University, Uppsala, Sweden
Outline

1. An example
2. Database design process
3. E-R Model
4. To Summarize
Design the database *University*
DB University, Requirements:

1. Store the names and the personal numbers of the teachers.
2. Store the departments’ names, addresses, the subdivisions and the number of the employed persons.
3. Each teacher can be employed in several departments.
4. Each subdivision has own courses.
5. A teacher can teach several course but only one course per department.
DB Teachers

• Are there any problems with the proposed design?
• Is it a “good” design?

• Think on
  • Clear concepts (clear semantics) ?
  • Needed memory for storing ?
  • Update (insert, update, delete ) ?
  • Search ?
## DB Teachers

### Department

<table>
<thead>
<tr>
<th>department</th>
<th>address</th>
<th>pnumber</th>
<th>teacher</th>
<th>course</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informational Technology</td>
<td>Polack</td>
<td>1234</td>
<td>Sara</td>
<td>Database I</td>
</tr>
<tr>
<td>Engineering</td>
<td>Ång</td>
<td>4567</td>
<td>Erik</td>
<td>Electrical circuits</td>
</tr>
<tr>
<td>Informational Technology</td>
<td>Polack</td>
<td>4567</td>
<td>Erik</td>
<td>Control systems</td>
</tr>
</tbody>
</table>

### Subdivision

<table>
<thead>
<tr>
<th>subdivision</th>
<th>department</th>
<th>num_emp</th>
<th>course</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer Systems</td>
<td>Informational Technology</td>
<td>400</td>
<td>Database I</td>
</tr>
<tr>
<td>Systems and Control</td>
<td>Informational Technology</td>
<td>400</td>
<td>Control systems</td>
</tr>
<tr>
<td>Electrical Engineering</td>
<td>Engineering</td>
<td>1000</td>
<td>Electrical Circuits</td>
</tr>
</tbody>
</table>
To remember!

What we *don’t want*: 
- Anomalies on update, insertion, deletion
- Redundancy (do things several times)
- Errors in our data (violation of integrity)
- Loss of data

What we *do want*: 
- Easy work with data: Do things one time (i.e. update, deletion)!
- Should be logical, easy to understand.
- Assurance: Data is correct, doesn’t get lost

**Conclusion:** Systematic Design Process is needed
Outline

1. An example
2. Database design process
3. E-R Model
4. To Summarize
Database Design Process

Miniworld

Requirements Collection & Analysis

Conceptual Design

Logical Design

Physical Design

Requirements

Conceptual schema

Logical schema

Internal schema

Conceptual model
E-R model

Implementation model
Model of DBMS
(Relational model)
ER-modeling

• High-level conceptual specification of the content in the database.

• History

• Why ER-models?
  o High-level description - easier to understand for non-technicians
  o More formal than natural language - avoid misconceptions and multiple interpretations
  o Implementation independent (of DBMS) - less technical details
  o Documentation
  o Model transformation to an implementation data model
Outline

1. An example
2. Database design process
3. E-R Model
4. To Summarize
Entity type and entity

- **Entity**: Thing in the real world
- **Attributes**: properties describing the entity
- **Entity Type**: collection (set) of entities that have the same attributes
- **Entity types** express the intention, i.e. the meaning of the concept whereas the **set of entities** represents the extension of that type.

Diagram:

- **person**
  - **name**
  - **age**
  - **address**
Attributes

• Every attribute has a *domain* (or value set).
  o A domain specifies the set of allowed values each individual attribute can be assigned.

• There is (at least) six different types of attributes:
  o simple
    sex: M or F
  o composite
    name: (Ior, Karlsson)
  o single-valued
    name: “Ior Karlsson”
  o multivalued
    friends: {Nasse, Puh,...}
  o stored
    birthdate: 980917
  o derived
    age : 0
  o null
Identification of Entities

- **Key attribute(s):** attribute(s) that has unique values for every instance of an entity type
- An entity type can have more than one key.
Relationship type and relationship

- **Relationship:** relation (association) between entities
- **Relationship type:** represents a relationship between a number of entity types.
- A relationship type $R$ is a set of relationships (i.e. relational instances) or tuples.
- A relationship type, $R$, can mathematically be defined as: $R \subseteq E_1 \times E_2 \times \ldots \times E_n$, $E_j$ is a entity type.
- E-R representation: diamond-shaped box
Relationship type and relationship

E1(student): \{Jan, Kalle, Anna\} \hspace{1cm} E2(course): \{Database, Algorithms\}

E1 \times E2: \{(Jan, Database), (Jan, Algorithms),
                     (Kalle, Database), (Kalle, Algorithms)
                     (Anna, Database), (Anna, Algorithms) \}

R \subseteq E1 \times E2 : \{(Jan, Database), (Jan, Algorithms),
                           (Anne, Database) \}
Cardinality ratio of relationship types

- **Cardinality ratio constraint** specifies the number of relational instances that an entity can take part in.

  For binary relationship types:
  - one-to-one (1:1)
  - one-to-many (1:N)
  - many-to-many (M:N)
Structural constraints for relationship types

- **Participation constraint**
  - specifies whether *the entity existence is dependent of another entity* via a relationship type
  - Partial participation: the entity can exist without this relationship
  - Total participation: the entity requires this relationship in order to exist.
Roles of relational types

- A role name specifies what role an entity type plays in a specific relationship
- Especially for recursive relationships
Attributes for relationship types

- A relationship type can have attributes
- If the relation is a 1:1 or 1:N relation, the attribute can be stored at one of the participating entities.
- When the relation is of the type M:N one must store the attributes with the instance of the relation.

![Diagram of relationship types]

- Working hours
- Employee
- Project
- N
- 1
- Works_on
Attributes for relationship types – Transform!

- **Working hours**
  - Employee (N) \(\rightarrow\) Works_on \(\rightarrow\) Project (1)

- **Supervisor**
  - Master student (1) \(\rightarrow\) Works_on \(\rightarrow\) Master thesis (1)

- **Percentage**
  - Student (N) \(\rightarrow\) Study \(\rightarrow\) Course (M)
Weak entity type

- **Weak entity types**: those that are meaningless without an owner entity type.
- Do not have an own key but **partial key**
- The relationship to the owner is called the *identifying relationship*.
ER-notation (Elmasri/Navathe Figure 7.14)

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
<th>Figure 7.14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Line</td>
<td>Entity</td>
<td>Summary of the notation for ER diagrams.</td>
</tr>
<tr>
<td>Diamond</td>
<td>Weak Entity</td>
<td></td>
</tr>
<tr>
<td>Diamond</td>
<td>Relationship</td>
<td></td>
</tr>
<tr>
<td>Small Diamond</td>
<td>Identifying Relationship</td>
<td></td>
</tr>
<tr>
<td>Ellipse</td>
<td>Attribute</td>
<td></td>
</tr>
<tr>
<td>Underlined Ellipse</td>
<td>Key Attribute</td>
<td></td>
</tr>
<tr>
<td>Dotted Ellipse</td>
<td>Multivalued Attribute</td>
<td></td>
</tr>
<tr>
<td>Ellipse</td>
<td>Composite Attribute</td>
<td></td>
</tr>
<tr>
<td>Dotted Line</td>
<td>Derived Attribute</td>
<td></td>
</tr>
</tbody>
</table>

- Total Participation of $E_2$ in $R$
- Cardinality Ratio $1:N$ for $E_1:E_2$ in $R$
- Structural Constraint $(\text{min, max})$ on Participation of $E$ in $R$
Summary

• Database design process
• E-R model
  o Entity and entity type
  o Relationship and relationship type
  o Cardinality ratio and structural constraints for relationship types
  o Attributes for relationship types
  o Weak entity type
Example ER-modeling

- An enterprise consists of a number of departments. Each department has a name, a number, a manager, and a number of employees. The starting date for every department manager should also be registered. A department can have several office rooms.
- Every department finances a number of projects. Each project has a name, a number and an office room.
- For each employee, the following information is kept: name, social security number, address, salary and sex. An employee works for only one department but can work with several projects that can be related to different departments. Information about the number of hours (per week) that an employee work with a project should be stored. Information about the employees manager should also be stored.
Example ER-modeling

Entity types:
- ET1 with attributes …
- ET2 with attributes …
- ET3 with attributes …
Example ER-modeling

• Relationship types:
  o An employee MANAGES a department ratio?
  o Every department FINANCES at least one project ratio?
  o Every employee WORKS-FOR a department ratio?
  o Every employee WORKS –WITH one (or several) project(s) ratio?
  o An employee IS-MANAGER over several employees ratio?
Example ER-modeling

EMPLOYEE

- fstname
- name
- sex
- salary
- ssn
- famname
- address

IS-MANAGER

1

WORKS-WITH

- hours

N

WORKS-FOR

1

MANAGES

1

DEPARTMENT

1

FINANCES

N

PROJECT

N

name
number
room

startdate

1

M

1

N

N

N