UPPSALA UNIVERSITET

DATABASE DESIGN I - 1DL300

Spring 2012

An Introductory Course on Database Systems

http://www.it.uu.se/edu/course/homepage/dbastekn/vt12/

Erik Zeitler
Uppsala Database Laboratory
Department of Information Technology, Uppsala University,
Uppsala, Sweden

UPPSALA UNIVERSITET

Normalization

Elmasri/Navathe ch 14
Padron-McCarthy/Risch ch 11

Silvia Stefanova

Department of Information Technology
Uppsala University, Uppsala, Sweden

UPPSALA UNIVERSITET

Outline

1. Normalization - Why and what ?
2. Guidelines for better database design
3. Normal forms
   - 1 NF
   - Functional dependency (FD), Full functional dependency (FFD)
   - 2 NF
   - 3 NF
   - BCNF
4. To Summarize

UPPSALA UNIVERSITET

Outline

1. Normalization - Why and what ?
2. Guidelines for better database design
3. Normal forms
   - 1 NF
   - Functional dependency (FD), Full functional dependency (FFD)
   - 2 NF
   - 3 NF
   - BCNF
4. Summarization

UPPSALA UNIVERSITET

“*Good*” database design (“*good*” relations)???

How to measure ?

The Database *Employee*

1. Names and personal numbers of the employees are known.
2. Each employee can be employed in several departments.
3. Each department has an address and several subdivisions. The number of employees in a department is known.
4. Every subdivision has own projects.
5. An employee can work on several projects but only on one project per department.

Employee

department	**pnumber**	**daddress**	**name**	**projects**
Informational Technology	1234	Polack	Sara	Project_IT_CS
Engineering	4567	Ång	Erik	Project_E_EE
Informational Technology	4567	Polack	Erik	Project _IT_SysC

Subdivision

subdivision	**department**	**num_emp**	**projects**
Computer Systems	Informational Technology	400	Project_IT_CS
Systems and Control	Informational Technology	400	Project _IT_SysC
Electrical Engineering	Engineering	1000	Project_ E_EE

The Database *Employee*

- Are there 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 ?

UPPSALA UNIVERSITET

Outline

1. Normalization - Why and what ?
2. Guidelines for better database design
3. Normal forms
   - 1 NF
   - Functional dependency (FD), Full functional dependency (FFD)
   - 2 NF
   - 3 NF
   - BCNF
4. To Summarize

UPPSALA UNIVERSITET

Outline

1. Normalization - Why and what ?
2. Guidelines for better database design
3. Normal forms
   - 1 NF
   - Functional dependency (FD), Full functional dependency (FFD)
   - 2 NF
   - 3 NF
   - BCNF
4. To Summarize

First Normal Form, 1 NF

1NF: Relations should not have multivalued attributes or nested relations.

Ssn	Ename	Pnumber	Hours
1234	Smith	1	12
		2	7
4534	Wong	3	40
		2	26

Normalization:

Form new relations for each multivalued attribute or nested relation.

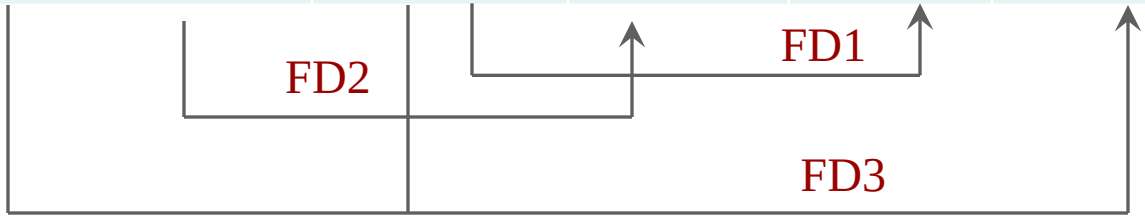
UPPSALA UNIVERSITET

Outline

1. Normalization - Why and what ?
2. Guidelines for better database design
3. Normal forms
   - 1 NF
   - Functional dependency (FD), Full functional dependency (FFD)
   - 2 NF
   - 3 NF
   - BCNF
4. Exercises

FD and FFD

X	**Y**	**A**	**P**	**XY**
X1	Y1	A1	P1	X1Y1
X1	Y2	A1	P2	X1Y2
X2	Y3	A2	P3	X2Y3



FDs : FD1, FD2

FFDs : FD3

The Database *Employee*

Employee

department	pnumber	daddress	name	projects
Informational Technology	1234	Polack	Sara	Project_IT_CS
Engineering	4567	Ång	Erik	Project_E_EE
Informational Technology	4567	Polack	Erik	Project _IT_SysC

Subdivision

subdivision	department	num_emp	projects
Computer Systems	Informational Technology	400	Project_IT_CS
Systems and Control	Informational Technology	400	Project _IT_SysC
Electrical Engineering	Engineering	1000	Project_ E_EE

? FDs :

? FFDs :

UPPSALA UNIVERSITET

Outline

1. Normalization - Why and what ?
2. Guidelines for better database design
3. Normal forms
   - 1 NF
   - Functional dependency (FD), Full functional dependency (FFD)
   - 2 NF
   - 3 NF
   - BCNF
4. Exercises

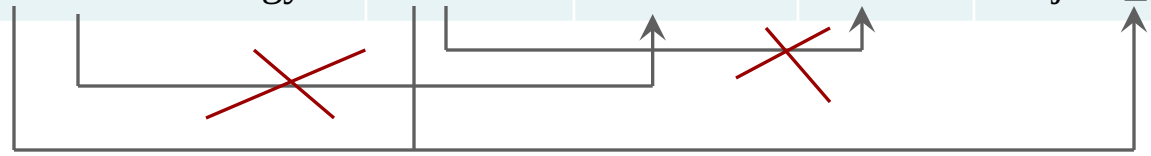
Second Normal Form - 2NF

2NF: A relation is in 2NF if:

- It is in 1NF
- Every non-key attribute in the relation is FFD of each candidate key .

Employee

department	**pnumber**	**daddress**	**name**	**projects**
Informational Technology	1234	Polack	Sara	Project_IT1
Engineering	4567	Ång	Erik	Project_EE0
Informational Technology	4567	Polack	Erik	Project _IT2



Normalization:

Decompose the relation, set up a new relation for each partial key with its dependent attribute(s).

UPPSALA UNIVERSITET

Outline

1. Normalization - Why and what ?
2. Guidelines for better database design
3. Normal forms
   - 1 NF
   - Functional dependency (FD), Full functional dependency (FFD)
   - 2 NF
   - 3 NF
   - BCNF
4. To Summarize

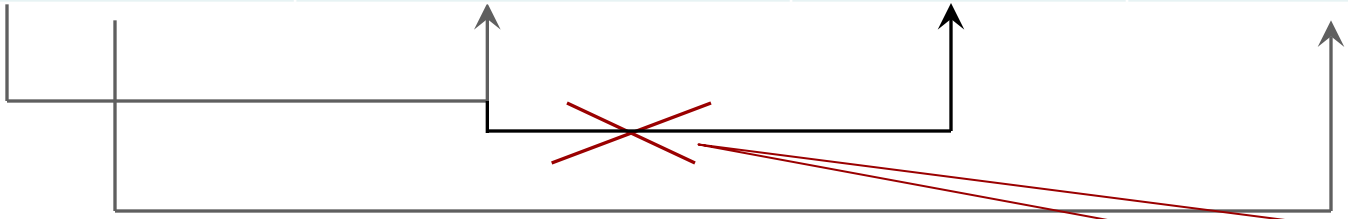
Third Normal Form - 3NF

3NF: A relation is in 2NF if:

- It is in 2NF
- No non-key attribute in a relation is allowed to be FFD on other non-key attribute.

Subdivision

subdivision	**department**	**num_emp**	**Projects**
Computer Systems	Informational Technology	400	Project_IT_CS
Systems and Control	Informational Technology	400	Project _IT_SysC
Electrical Engineering	Engineering	1000	Project_ E_EE



transitive FD on the primary key

Normalization:

Decompose the relation, set up a new relation including the non-key attribute(s) that is FD on other non-key attribute(s).

UPPSALA UNIVERSITET

Outline

1. Normalization - Why and what ?
2. Guidelines for better database design
3. Normal forms
   - 1 NF
   - Functional dependency (FD), Full functional dependency (FFD)
   - 2 NF
   - 3 NF
   - BCNF
4. To Summarize

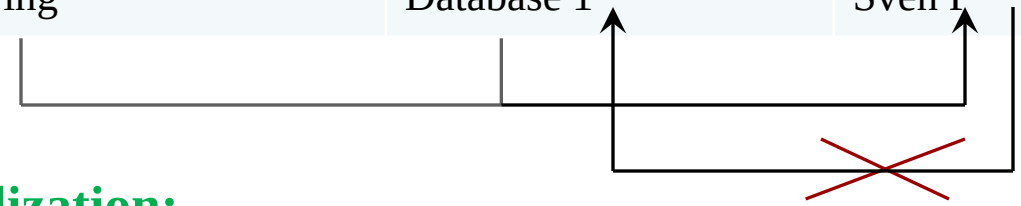
Boyce-Codd Normal Form - BCNF

BCNF: A relation is in BCNF if:

- It is in 1NF
- Every determinant is a candidate key.

Teach

department	**course**	**teacher**
Informational Technology	Database 1	Sara S
Informational Technology	Database 2	Sara S
Engineering	Signals and Systems	Peter E
Engineering	Database 1	Sven P



Normalization:

Decompose the relation so that after joining the new relations *spurious tuples* will not be generated (*lossless join decomposition*)

UPPSALA UNIVERSITET

Summary

- Normalization
- Redundancy
- Functional dependency (FD)
- Full functional dependency (FFD)
- 1 NF
- 2 NF
- 3 NF
- BCNF
- Spurious tuples