Databases

Jörg Endrullis

VU University Amsterdam

Overview

1. Functional Dependencies (FDs)
2. Anomalies, FD-based Normal Forms
3. BCNF and 3NF Synthesis
4. Multivalued Dependencies (MVDs) and 4NF
5. Normal Forms and ER Design
6. Denormalization

Introduction

Functional Dependencies (FDs)

- are a **generalization of keys**
- central part of **relational database design theory**

This theory defines when a relation is in **normal form.**

Usually a sign of **bad database design** if a schema contains relations that **violate the normal form.**

If a normal form is violated

- data is stored **redundantly** and
- information about different concepts is **intermixed**

COURSES

CRN	TITLE	INAME	PHONE
22268	Databases I	Grust	7111
42232	Functional Programming	Grust	7111
31822	Graph Theory	Klotz	2418

The phone number for each instructor is stored multiple times!

Introduction

There are different normal forms. The main ones are:

- **Third Normal Form (3NF)**: the standard relational normal form used in practice (and education).
- **Boyce-Codd Normal Form (BCNF)**:
  - a bit more restrictive
  - easier to define
  - better for our intuition of good database design

  Roughly speaking, BNCF requires that **all FDs are keys.**

In rare circumstances, a relation might not have an equivalent BCNF form while preserving all its FDs.

The 3NF normal form always exists (and preserves the FDs).

Introduction

Normalization algorithms can construct good relation schemas from a set of attributes and a set of functional dependencies.

In practice:

- relations are derived from ER models
- normalization is used as an additional check only

When an ER model is **well designed**, the resulting derived relational tables will **automatically be in BCNF**.

Awareness of normal forms can help to detect design errors already in the conceptual design phase.

First Normal Form

The **First Normal Form (1NF)** requires that all **table entries are atomic** (*not* lists, sets, records, relations).

- The relational model all table entries are already atomic.
- All further normal forms assume that tables are in 1NF.

The following are **not violations of 1NF**:

- A table entry contains values with internal structure.
  - e.g. a `CHAR(100)` containing a comma separated list
- List represented by several columns.
  - e.g. columns `value1`, `value2`, `value3`

Nevertheless, these are **bad design**.

Functional Dependencies

COURSES			
CRN	**TITLE**	**INAME**	**PHONE**
22268	Databases I	Grust	7111
42232	Functional Programming	Grust	7111
31822	Graph Theory	Klotz	2418

A **functional dependency (FD)** in this table is

$$\text{INAME} \longrightarrow \text{PHONE}$$

Whenever two rows agree in the instructor name INAME, they **must** also agree in the PHONE column values!

Functional Dependencies

Intuitively, there is a functional dependency

$$\texttt{INAME} \longrightarrow \texttt{PHONE}$$

since the phone number **only depends on the instructor**, not on other course data.

This functional dependency read as

INAME (functionally, uniquely) determines PHONE

A functional dependency is like a **partial key:** uniquely determines some attributes, but not all in general.

A **determinant** is a 'minimal' functional dependency.

INAME is a determinant for PHONE

Functional Dependencies

In general, an **functional dependencies** take the form

$$A_1, \ldots, A_n \to B_1, \ldots, B_m$$

Sequence of attributes is unimportant: formally sets

$$\{A_1, \ldots, A_n\} \to \{B_1, \ldots, B_m\}$$

The **functional dependency (FD)**

$$A_1, \ldots A_n \to B_1, \ldots B_m$$

holds for a relation R in a database state I if and only if

$$t.A_1 = u.A_1 \wedge \cdots \wedge t.A_n = u.A_n$$
$$\Rightarrow \; t.B_1 = u.B_1 \wedge \cdots \wedge t.B_m = u.B_m$$

for all tuples $t, u \in I(R)$:

Functional Dependencies

A functional dependency with m attributes on the right

$$A_1, \ldots A_n \rightarrow B_1, \ldots B_m$$

is **equivalent** to the m FDs:

$$\begin{array}{ccc} A_1, \ldots, A_n & \rightarrow & B_1 \\ \vdots & & \vdots \\ A_1, \ldots, A_n & \rightarrow & B_m \end{array}$$

Thus, in the following it suffices to consider FDs with a single column name on the right-hand side.

Keys are Functional Dependencies

A **key** uniquely determines **all** attributes of its relation.

There are never two distinct rows with the same key, so the **functional dependency condition is trivially satisfied**.

COURSES			
CRN	**TITLE**	**INAME**	**PHONE**
22268	Databases I	Grust	7111
42232	Functional Programming	Grust	7111
31822	Graph Theory	Klotz	2418

We have the following functional dependencies:

- CRN $\longrightarrow$ TITLE, INAME, PHONE

or equivalently:

- CRN $\longrightarrow$ TITLE
- CRN $\longrightarrow$ INAME
- CRN $\longrightarrow$ PHONE

Functional Dependencies are Partial Keys

Functional dependencies are **constraints** (like keys).

COURSES			
CRN	**TITLE**	**INAME**	**PHONE**
22268	Databases I	Grust	7111
42232	Functional Programming	Grust	7111
31822	Graph Theory	Klotz	2418

In this example state, the functional dependency

$$\text{TITLE} \longrightarrow \text{CRN}$$

holds. But this is probably **not true in general**!

For the database design, the only interesting functional dependencies are those that **hold for all possible states.**

Functional Dependencies are Partial Keys

Functional dependencies are a **generalisation of keys**.

$A_1, \ldots, A_n$ is a key of relation $R(A_1, \ldots, A_n, B_1, \ldots, B_m)$

$$\iff$$

the functional dependency $A_1, \ldots, A_n \rightarrow B_1, \ldots B_m$ holds.

COURSES

CRN	TITLE	INAME	PHONE
22268	Databases I	Grust	7111
42232	Functional Programming	Grust	7111
31822	Graph Theory	Klotz	2418

Here CRN $\rightarrow$ TITLE, INAME, PHONE.

Functional Dependencies are Partial Keys

Functional dependencies are partial keys.

The functional dependency

$$A_1, \dots A_n \to B_1, \dots B_m$$

holds for a relation R if $\{ A_1, \dots A_n \}$ is a key for the relation obtained by restricting R to the columns $\{A_1, \dots A_n, B_1, \dots B_m\}$.

The restriction of the table `COURSES` to $\{$ `INAME`, `PHONE` $\}$ is:

COURSES	
INAME	**PHONE**
Grust	7111
Klotz	2418

The attribute `INAME` is a key of this table.

The **goal of database normalization is to turn FDs into keys**.

The DBMS is then able to enforce the FDs for the user.

Example: Books and Authors

BOOKS				
AUTHOR	**NO**	**TITLE**	**PUBLISHER**	**ISBN**
Elmasri	1	Fund. of DBS	Addison-W.	0805317554
Navathe	2	Fund. of DBS	Addison-W.	0805317554
Silberschatz	1	DBS Concepts	Mc-Graw H.	0471365084
Korth	2	DBS Concepts	Mc-Graw H.	0471365084
Sudarshan	3	DBS Concepts	Mc-Graw H.	0471365084

- a book may have multiple authors, one author per row
- attribute NO is used to indicate the order of the authors

- ISBN $\longrightarrow$ TITLE, PUBLISHER (ISBN uniquely identifies a book)
- ISBN $\longrightarrow$ AUTHOR ? Does not hold.
- AUTHOR $\longrightarrow$ TITLE ? Does not hold in general.
- TITLE $\longrightarrow$ nothing (There may be books with the same title)
- ISBN, NO $\longrightarrow$ AUTHOR
- ISBN, AUTHOR $\longrightarrow$ NO ? questionable (e.g. Smith & Smith)
- PUBLISHER, TITLE, NO $\longrightarrow$ AUTHOR ? questionable
  Authorship sequence might change in a new edition of a book!

Example: Books and Authors

BOOKS				
AUTHOR	**NO**	**TITLE**	**PUBLISHER**	**ISBN**
Elmasri	1	Fund. of DBS	Addison-W.	0805317554
Navathe	2	Fund. of DBS	Addison-W.	0805317554
Silberschatz	1	DBS Concepts	Mc-Graw H.	0471365084
Korth	2	DBS Concepts	Mc-Graw H.	0471365084
Sudarshan	3	DBS Concepts	Mc-Graw H.	0471365084

- a book may have multiple authors, one author per row
- attribute NO is used to indicate the order of the authors

During database design, **only unquestionable conditions should be used as functional dependencies.**

Database normalization **alters the table structure** depending on the specified functional dependencies.

Later hard to change: needs creation/deletion of tables!

Quiz

A table with homework grades:

HOMEWORK_RESULTS					
STUD_ID	**FIRST**	**LAST**	**EX_NO**	**POINTS**	**MAX_POINTS**
100	Andrew	Smith	1	9	10
101	Dave	Jones	1	8	10
102	Maria	Brown	1	10	10
101	Dave	Jones	2	11	12
102	Maria	Brown	2	10	12

- Which FDs should hold for this table in general?
- Identify FDs that hold in this table but not in general.

Implication of Functional Dependencies

If $A \longrightarrow B$ and $B \longrightarrow C$ hold, then $A \longrightarrow C$ is holds automatically.

COURSES

CRN	TITLE	INAME	PHONE
22268	Databases I	Grust	7111
42232	Functional Programming	Grust	7111
31822	Graph Theory	Klotz	2418

Note that CRN $\longrightarrow$ PHONE is a consequence of

CRN $\longrightarrow$ INAME and INAME $\longrightarrow$ PHONE

FDs of the form $A \longrightarrow A$ always hold.

PHONE $\longrightarrow$ PHONE holds, but is not interesting

Implication of Functional Dependencies

A set of FDs Γ **implies** an FD $\alpha \longrightarrow \beta$

$$\Longleftrightarrow$$

every DB state which satisfies all FDs in Γ, also satisfies $\alpha \longrightarrow \beta$.

Implication of Functional Dependencies

The DB designer is normally not interested in all FDs, but only in a **representative FD set** that implies all other FDs.

Armstrong Axioms

- **Reflexivity:**
  If $\beta \subseteq \alpha$, then $\alpha \rightarrow \beta$.
- **Augmentation:**
  If $\alpha \rightarrow \beta$, then $\alpha \cup \gamma \rightarrow \beta \cup \gamma$.
- **Transitivity:**
  If $\alpha \rightarrow \beta$ and $\beta \rightarrow \gamma$, then $\alpha \rightarrow \gamma$.

Use the Amstrong axioms to show that

$$\begin{aligned} \texttt{ISBN} &\rightarrow \texttt{TITLE}, \texttt{PUBLISHER} \\ \texttt{ISBN}, \texttt{NO} &\rightarrow \texttt{AUTHOR} \\ \texttt{PUBLISHER} &\rightarrow \texttt{PUB\_URL} \end{aligned}$$

implies $\texttt{ISBN} \rightarrow \texttt{PUB\_URL}$.

Implication of Functional Dependencies

Simpler way to **check whether** $a \rightarrow \beta$ **is implied by an FD set**:

- compute the **cover** α^+ of α, and
- then check if $\beta \subseteq \alpha^+$.

Cover

The **cover** $\alpha^+_{\mathcal{F}}$ of

- a set of attributes α
- with respect to an FD set $\mathcal{F}$

is the set of all attributes B that are uniquely determined by α:

$$\alpha^+_{\mathcal{F}} := \{ B \mid \mathcal{F} \text{ implies } \alpha \rightarrow B \}$$

Implication Check

A set of FDs $\mathcal{F}$ implies an FD $\alpha \rightarrow \beta$ if and only if $\beta \subseteq \alpha^+_{\mathcal{F}}$.

Implication of Functional Dependencies

Cover computation

Input: α (set of attributes)
$\alpha_1 \to \beta_1, \ldots, \alpha_n \to \beta_n$ (set of FDs $\mathcal{F}$)

Output: α^+ (the cover of α)

```
x = α;
while x did change do
    for all given FD α_i → β_i do
        if α_i ⊆ x then
            x = x ∪ β_i;    (add attributes in β_i to x)
        end if
    end for
end while
return x;
```

Implication of Functional Dependencies

Compute the cover $\{\texttt{ISBN}\}^+$ for the following FDs:

$$\texttt{ISBN} \rightarrow \texttt{TITLE}, \texttt{PUBLISHER}$$
$$\texttt{ISBN}, \texttt{NO} \rightarrow \texttt{AUTHOR}$$
$$\texttt{PUBLISHER} \rightarrow \texttt{PUB\_URL}$$

1. We start with $x = \{\texttt{ISBN}\}$.
2. $\texttt{ISBN} \rightarrow \texttt{TITLE}, \texttt{PUBLISHER}$ is applicable.

   The left-hand side is completely contained in x.

   We get $x = \{\texttt{ISBN}, \texttt{TITLE}, \texttt{PUBLISHER}\}$.
3. $\texttt{PUBLISHER} \rightarrow \texttt{PUB\_URL}$ is applicable.

   We get $x = \{\texttt{ISBN}, \texttt{TITLE}, \texttt{PUBLISHER}, \texttt{PUB\_URL}\}$.
4. No further way to extend set x, the algorithm returns

   $$\{\texttt{ISBN}\}^+ = \{\texttt{ISBN}, \texttt{TITLE}, \texttt{PUBLISHER}, \texttt{PUB\_URL}\}$$
5. We may now conclude, e.g., $\texttt{ISBN} \rightarrow \texttt{PUB\_URL}$.

How to Determine Keys

Given a set of FDs and the set of $\mathcal{A}$ all attributes of a relation *R*:

$$\alpha \subseteq \mathcal{A} \text{ is key of } R \iff \alpha^+ = \mathcal{A}$$

That is α is a key if the cover α^+ contains all attributes.

We can use FDs to determine all possible keys of *R*.

Normally, we are interested in **minimal keys** only.

A key α is **minimal** if every $A \in \alpha$ is **vital**, that is

$$(\alpha - \{A\})^+ \neq \mathcal{A}$$

How to Determine Keys

Finding a Minimal Key

Input: $\mathcal{A}$ (set of all attributes of R)
$\alpha_1 \rightarrow \beta_1, \ldots, \alpha_n \rightarrow \beta_n$ (set of FDs $\mathcal{F}$)
Output: α (a minimal key of R)

$x = \mathcal{A}$;
for all attributes $A \in X$ **do**
  if $A \in \{x - A\}^+_{\mathcal{F}}$ **then**
    $x = x - A$; (remove A from x)
  end if
end for
return x;

We might get different keys depending on the order in **for all.**

How to Determine Keys

Finding all Minimal Keys

Input: $A_1, A_2, \ldots, A_n$ (all attributes of R) and $\mathcal{F}$ (set of FDs)

$Results = \varnothing$;
$Candidates = \{ \{ A_i \mid A_i$ is not part of any right-hand side in $\mathcal{F} \} \}$;
while $Candidates \neq \varnothing$ **do**
  choose and remove a smallest $\kappa \in Candidates$;
  if $\kappa^+_{\mathcal{F}} = \{A_1, A_2, \ldots, A_n\}$ **then**
    if κ contains no key in *Results* **then**
      $Results = Results \cup \{\kappa\}$;
    end if
  else
    for all $A_i \notin \kappa^+_{\mathcal{F}}$ **do**
      $\kappa_i = \kappa \cup \{A_i\}$;
      $Candidates = Candidates \cup \{\kappa_i\}$;
    end for
  end if
end while
return *Results*;

How to Determine Keys: Examples

Find **all** minimal keys the relation R

R				
A	**B**	**C**	**D**	**E**

with the functional dependencies

$$A \rightarrow D \qquad B \rightarrow C \qquad B \rightarrow D \qquad D \rightarrow E$$

We get

1. $\mathit{Candidates} = \{\,\{\,A, B\,\}\,\}$
   since A, B do not occur in any right-hand side
2. $\{\,A, B\}^+ = \{\,A, B, C, D, E\,\}$
   So $\{\,A, B\}$ is a key.
3. $\mathit{Candidates} = \{\ \ \}$
   No more candidate keys to check, we terminate.

How to Determine Keys: Examples

Find **all** minimal keys the relation `R(A,B,C,D,E)` with

$$A, D \rightarrow B, D \qquad B, D \rightarrow C \qquad A \rightarrow E$$

We get

1. $\mathit{Candidates} = \{\,\{\,A\,\}\,\}$ since A not in any right-hand side
2. $\{\,A\,\}^+ = \{\,A, E\,\}$, so we extend with B, C, D:
   $\mathit{Candidates} = \{\,\{\,A, B\,\}, \{\,A, C\,\}, \{\,A, D\,\}\,\}$
3. $\{\,A, D\,\}^+ = \{\,A, B, C, D, E\,\}$. So $\{\,A, D\,\}$ is a **key**.
4. $\{\,A, B\,\}^+ = \{\,A, B, E\,\}$, so we extend with C, D:
   $\mathit{Candidates} = \{\,\{\,A, B, C\,\}, \{\,A, B, D\,\}, \{\,A, C\,\}\,\}$
5. $\{\,A, C\,\}^+ = \{\,A, C, E\,\}$, so we extend with B, D:
   $\mathit{Candidates} = \{\,\{\,A, B, C\,\}, \{\,A, B, D\,\}, \{\,A, C, D\,\}\,\}$
6. Remove $\{\,A, B, D\,\}$ and $\{\,A, C, D\,\}$ since they contain a key.
7. $\{\,A, B, C\,\}^+ = \{\,A, B, C, E\,\}$ is not a key!
   Extension with D again contains a key.

Determinants

Determinants (Non-trivial, minimal FDs)

$\{ A_1, \ldots, A_n \}$ is a **determinant** for $\{ B_1, \ldots, B_m \}$ if

- the FD $A_1, \ldots, A_n \rightarrow B_1, \ldots B_m$ holds, and
- the **left-hand side is minimal**, that is, if any A_i is removed then $A_1, \ldots, A_{i-1}, A_{i+1}, A_n \rightarrow B_1, \ldots B_m$ does *not* hold, and
- it is **not trivial**, that is, $\{B_1, \ldots, B_m\} \not\subseteq \{A_1, \ldots, A_n\}$.

$$\mathcal{F} = \left\{ \begin{array}{rcl} \texttt{STUD\_ID}, \texttt{EX\_NO} & \rightarrow & \texttt{POINTS} \\ \texttt{EX\_NO} & \rightarrow & \texttt{MAX\_POINTS} \end{array} \right\}$$

Are the following determinants?

- `POINTS`, `MAX_POINTS` for `POINTS`, `MAX_POINTS` ? No
- `EX_NO` for `POINTS`, `MAX_POINTS` ? No
- `STUD_ID`, `EX_NO` for `POINTS`, `MAX_POINTS` ? Yes
- `EX_NO`, `POINTS` for `POINTS`, `MAX_POINTS` ? Yes

Relational Normal Forms

Overview

1. Functional Dependencies (FDs)
2. Anomalies, FD-based Normal Forms
3. BCNF and 3NF Synthesis
4. Multivalued Dependencies (MVDs) and 4NF
5. Normal Forms and ER Design
6. Denormalization

Consequences of Bad DB Design

Usually a severe sign of **bad DB design** if a table contains an FD (encodes a partial function) that is **not implied by a key.**

INAME $\longrightarrow$ PHONE

COURSES			
CRN	**TITLE**	**INAME**	**PHONE**
22268	Databases I	Grust	7111
42232	Functional Programming	Grust	7111
31822	Graph Theory	Klotz	2418

This leads to

- **redundant storage of certains facts**
  (here, phone numbers)
- **insert, update, deletion anomalies**

Consequences of Bad DB Design

Redundant storage is bad for several reasons:

- it **wastes storage space**
- difficult to ensure **integrity** when updating the database
  - all redundant copies need to be updated
  - **wastes time**, inefficient
- need for **additional constraints** to guarantee integrity
  - ensure that the redundant copies indeed agree
  - e.g. the constraint `INAME` $\longrightarrow$ `PHONE`

Problem

General FDs are **not** supported by **relational databases.**

The solution is to transform FDs into **key constraints.**
This is what **DB normalization** tries to do.

Consequences of Bad DB Design

Update anomalies

- When a single value needs to be changed (e.g., a phone number), **multiple tuples** must be updated. This **complicates programs and updates takes longer**.
- Redundant copies potentially get **out of sync** and it is impossible/hard to identify the correct information.

Insertion anomalies

- The phone number of a new instructor cannot be inserted into the DB until it is known what course she/he will teach.
- Insertion anomalies arise when **unrelated concepts** are **stored together in a single table**.

Deletion anomalies

- When the last course of an instructor is deleted, his/her phone number is lost.

Boyce-Codd Normal Form

A relation R is in **Boyce-Codd Normal Form (BCNF)**
if all its FDs are implied by its key constraints.

That is, for every FD $A_1, \ldots, A_n \rightarrow B_1, \ldots, B_m$ of R we have:

- $\{B_1, \ldots, B_m\} \subseteq \{A_1, \ldots, A_n\}$ (the FD is trivial), or
- $\{A_1, \ldots, A_n\}$ contains a key of R.

The relation

COURSES (<u>CRN</u>, TITLE, INAME, PHONE)

with the FDs

$$\begin{aligned} \texttt{CRN} &\rightarrow \texttt{TITLE}, \texttt{INAME}, \texttt{PHONE} \\ \texttt{INAME} &\rightarrow \texttt{PHONE} \end{aligned}$$

is **not in BCNF** because of the FD INAME $\rightarrow$ PHONE:

- the FD is not trivial, and
- INAME is not a key

However, the relation COURSES (<u>CRN</u>, TITLE, INAME) without the attribute PHONE is in BCNF.

Boyce-Codd Normal Form: Examples

Each course meets once per week in a dedicated room:

CLASS (CRN, TITLE, WEEKDAY, TIME, ROOM)

The relation thus satisfies the following FDs (plus implied ones):

$$\begin{aligned} \mathtt{CRN} &\rightarrow \mathtt{TITLE}, \mathtt{WEEKDAY}, \mathtt{TIME}, \mathtt{ROOM} \\ \mathtt{WEEKDAY}, \mathtt{TIME}, \mathtt{ROOM} &\rightarrow \mathtt{CRN} \end{aligned}$$

The minimal keys of CLASS are

- $\{ \mathtt{CRN} \}$
- $\{ \mathtt{WEEKDAY}, \mathtt{TIME}, \mathtt{ROOM} \}$

Is the relation in BCNF?

- both FDs are implied by keys
  (their left-hand sides even coincide with the keys)

Thus CLASS **is in BCNF.**

Boyce-Codd Normal Form: Examples

Consider the relation

`PRODUCT (NO, NAME, PRICE)`

and the following FDs:

$$\begin{array}{lclrcl} \texttt{NO} & \longrightarrow & \texttt{NAME} & \texttt{PRICE}, \texttt{NAME} & \longrightarrow & \texttt{NAME} \\ \texttt{NO} & \longrightarrow & \texttt{PRICE} & \texttt{NO}, \texttt{PRICE} & \longrightarrow & \texttt{NAME} \end{array}$$

Is this relation in BCNF?

- The two left FDs indicate that `NO` is a key.
  Both FDs are thus implied by a key.
- The third FD is trivial (and may be ignored).
- The left-hand side of the last FD contains a key.

Thus the relation `PRODUCT` **is in BCNF**.

Boyce-Codd Normal Form

Advantages of Boyce-Codd Normal Form

If a relation R is in BCNF, then...

- Ensuring its key constraints automatically satisfies all FDs. Hence, no additional constraints are needed!
- The **anomalies** (udpate/insertion/deletion) **do not occur.**

Boyce-Codd Normal Form: Quiz

BCNF Quiz

1. Consider the relation

   RESULTS (STUD_ID, EX_NO, POINTS, MAX_POINTS)

   with the following FDs

   $$\begin{aligned} \text{STUD\_ID, EX\_NO} &\longrightarrow \text{POINTS} \\ \text{EX\_NO} &\longrightarrow \text{MAX\_POINTS} \end{aligned}$$

   Is this relation in BCNF?

2. Consider the relation

   INVOICE (INV_NO, DATE, AMOUNT, CUST_NO, CUST_NAME)

   with the following FDs

   $$\begin{aligned} \text{INV\_NO} &\longrightarrow \text{DATE, AMOUNT, CUST\_NO} \\ \text{INV\_NO, DATE} &\longrightarrow \text{CUST\_NAME} \\ \text{CUST\_NO} &\longrightarrow \text{CUST\_NAME} \\ \text{DATE, AMOUNT} &\longrightarrow \text{DATE} \end{aligned}$$

   Is this relation in in BCNF?

Third Normal Form

A **key attribute** is an attribute that appears in a minimal key.
Minimality is important, otherwise all attributes are key attributes.

Assume that FDs with multiple attributes on rhs have been expanded. That is, every FD has a single attribute on the right-hand side.

Third Normal Form (3NF)

A relation R is in **Third Normal Form (3NF)** if and only if every FD $A_1, \ldots, A_n \to B$ satisfies at least one of the conditions:

- $B \in \{A_1, \ldots, A_n\}$ (the FD is trivial), or
- $\{A_1, \ldots, A_n\}$ contains a key of R, or
- B is a **key attribute** of R.

The only difference with BCNF is the last condition.

Third Normal Form (3NF) is slightly weaker than BCNF:
If a relation is in BCNF, it is automatically in 3NF.

Third Normal Form

In short, we can say:

BCNF $\iff$ for every non-trivial FD:

- the left-hand side contains a key

3NF $\iff$ for every non-trivial FD:

- the left-hand side contains a key, or
- the right-hand side is an attribute of a minimal key

Third Normal Form Quiz

3NF vs BCNF

BOOKINGS			
COURT	**START_TIME**	**END_TIME**	**RATE**
1	9:30	11:00	SAVER
2	9:30	12:00	PREMIMUM-A
1	12:00	14:00	STANDARD

The table contains bookings for one day at a tennis club:

- there are courts 1 (hard court) and 2 (grass court)
- the rates are
  - SAVER for member bookings of court 1
  - STANDARD for non-member bookings of court 1
  - PREMIMUM-A for member bookings of court 2
  - PREMIMUM-B for non-member bookings of court 2

Quiz:

- Find a representative FDs set.
- Is the table in BCNF? Is the table in 3NF?

Splitting Relations

If a table R is not in BCNF, we can **split** it into two tables.

The violating FD determines how to split:

Table Decomposition

If the FD $A_1, \ldots, A_n \rightarrow B_1, \ldots B_m$ violates BCNF:

- create a new relation $S(\underline{A_1}, \ldots, \underline{A_n}, B_1, \ldots, B_m)$ and
- remove $B_1, \ldots, B_m$ from the original relation R.

Splitting "along an FD"

The FD `INAME` $\rightarrow$ `PHONE` is the reason why table

```
COURSES (CRN, TITLE, INAME, PHONE)
```

violates BCNF because of `INAME` $\rightarrow$ `PHONE`. We split into:

```
INSTRUCTORS (CRN, TITLE, INAME)
PHONEBOOK (INAME, PHONE)
```

Splitting Relations

It is important that this splitting transformation is **lossless**, i.e., that the original relation can be reconstructed by a join.

Reconstruction after split

Recall that we have split

```
COURSES (CRN, TITLE, INAME, PHONE)
```

into tables

```
INSTRUCTORS (CRN, TITLE, INAME)
PHONEBOOK (INAME, PHONE)
```

We can reconstruct the original table as follows:

```
CREATE VIEW COURSES (CRN, TITLE, INAME, PHONE)
    AS
SELECT I.CRN, I.TITLE, I.INAME, P.PHONE
  FROM INTSTRUCTORS I, PHONEBOOK P
 WHERE I.INAME = P.INAME
```

Splitting Relations

When is a split lossless?

Decomposition Theorem

The split of relations is **guaranteed to be lossless** if the intersection (the shared set) of the attributes of the new tables is a key of at least one of them.

The join $\bowtie$ connects tuples depending on the attribute (values) in the intersection. If these values uniquely identify tuples in one relation we do not lose information.

"Lossy" decomposition

Original table (key A, B, C)

A	B	C
a_{11}	b_{11}	c_{11}
a_{11}	b_{11}	c_{12}
a_{11}	b_{12}	c_{11}

Decomposition

R_1

A	B
a_{11}	b_{11}
a_{11}	b_{12}

R_2

A	C
a_{11}	c_{11}
a_{11}	c_{12}

"Reconstruction" $R_1 \bowtie R_2$

A	B	C
a_{11}	b_{11}	c_{11}
a_{11}	b_{11}	c_{12}
a_{11}	b_{12}	c_{11}
a_{11}	b_{12}	c_{12}

Splitting Relations

Lossless split condition satisfied

Recall that we have split

COURSES (CRN, TITLE, INAME, PHONE)

into tables

INSTRUCTORS (CRN, TITLE, INAME)

PHONEBOOK (INAME, PHONE)

The lossless split condition is satisfied since

$$\{\texttt{CRN}, \texttt{TITLE}, \texttt{INAME}\} \cap \{\texttt{INAME}, \texttt{PHONE}\} = \{\texttt{INAME}\}$$

and INAME is a key of the table PHONEBOOK.

All splits initiated by the **table decomposition method** for transforming relations into BCNF satisfy the condition of the decomposition theorem.

It is **always possible** to transform a relation into BCNF by lossless splitting.

Splitting Relations

Lossless split guarantees that the schema after splitting can represent all DB states that were possible before.

- we can translate states from the old into the new schema
- old schema can be “simulated” via views

Lossless splits can lead to **more general schemas**!

- the new schema allows states which do not correspond to the state in the old schema

Recall that we have split

`COURSES (CRN, TITLE, INAME, PHONE)`

into tables

`INSTRUCTORS (CRN, TITLE, INAME)`

`PHONEBOOK (INAME, PHONE)`

We may now store instructors and phone numbers without any affiliation to courses.

Splitting Relations

Not every lossless split is reasonable!

STUDENTS		
SSN	FIRST_NAME	LAST_NAME
111-22-3333	John	Smith
123-45-6789	Maria	Brown

Splitting STUDENTS into

STUD_FIRST (SSN, FIRST_NAME)
STUD_LAST (SSN, LAST_NAME)

is lossless, but

- the split is **not** necessary to enforce a normal form,
- only requires costly joins in subsequent queries

Splitting Relations: Computable Columns

Although **computable columns** lead to violations of BCNF, splitting the relation is **not** the right solution.

E.g. `AGE` which is derivable from `BIRTDATE`.

As a consequence we have a functional dependency:

$$\texttt{BIRTDATE} \longrightarrow \texttt{AGE}$$

A split would yield a relation:

`R(BIRTHDAY, AGE)`

which would try to materialise the computable function.

The **correct solution** is to **eliminate `AGE`** from the table and to **define a view** which contains all columns plus the **computed** column `AGE`.

Preservation of Functional Dependencies

Besides losslessness, a property which a good decomposition of a relation should guarantee is the **preservation of FDs**:

- An FD can refer only to attributes of a single relation.
- When splitting a relation into two, there might be FDs that can no longer be expressed (these FDs are not preserved).

FD gets lost during decomposition

ADRESSES (STREET_ADDR, CITY, STATE, ZIP)

with functional dependencies

$$\begin{aligned} \texttt{STREE\_ADDR}, \texttt{CITY}, \texttt{STATE} &\rightarrow \texttt{ZIP} \\ \texttt{ZIP} &\rightarrow \texttt{STATE} \end{aligned}$$

The second FD violates BCNF and would lead to the split:

- ADDRESSES1 (STREET_ADDR, CITY, ZIP) and
- ADDRESSES2 (ZIP, STATE).

But now the first FD can no longer be expressed.

Preservation of Functional Dependencies

ADRESSES (STREET_ADDR, CITY, STATE, ZIP)

with functional dependencies

$$\text{STREE\_ADDR}, \text{CITY}, \text{STATE} \rightarrow \text{ZIP}$$
$$\text{ZIP} \rightarrow \text{STATE}$$

Is the table in 3NF? Yes

- Most designers would not split the table since it is in 3NF.
- **Pro split**: if there are many addresses with the same ZIP code, there will be significant redundancy.
- **Contra split**: queries will involve more joins.

Whether or not to split depends on the intended application:

- A table of ZIP codes might be of interest on its own.
  E.g. for the database of a mailing company.

Overview

1. Functional Dependencies (FDs)
2. Anomalies, FD-based Normal Forms
3. BCNF and 3NF Synthesis
4. Multivalued Dependencies (MVDs) and 4NF
5. Normal Forms and ER Design
6. Denormalization

Canonical Set of Functional Dependencies

Determine a **minimal** (**canonical**) set of FDs that is equivalent to the given FDs $\mathcal{F}$ as follows:

1. **Make the right-hand sides singular**

   Replace every FD $\alpha \rightarrow B_1, \ldots, B_m$ by $\alpha \rightarrow B_i$, $1 \leqslant i \leqslant m$.

2. **Minimise left-hand sides**

   For each $A_1, \ldots, A_n \rightarrow B$ and each $i = 1, \ldots, n$:
   - If the cover $\{A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_n\}^+_{\mathcal{F}}$ contains B, then
     - drop $A_1, \ldots, A_n \rightarrow B$ from $\mathcal{F}$, and
     - add $A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_n \rightarrow B$ to $\mathcal{F}$.

   Keep repeating until all left-hand sides are minimal.

3. **Remove implied FDs**

   For each FD $\alpha \rightarrow B$:
   - If the cover $\alpha^+_{\mathcal{F}'}$ for $\mathcal{F}' = \mathcal{F} - \{\alpha \rightarrow B\}$ contains B, then drop $\alpha \rightarrow B$ from $\mathcal{F}$.

Canonical Set of Functional Dependencies

Consider the relation $R = (A, B, C, D, E)$ with FDs

$$A \rightarrow DE \qquad B \rightarrow C \qquad BC \rightarrow D \qquad D \rightarrow E$$

1. Make the right-hand sides singular

$$A \rightarrow D \quad A \rightarrow E \quad B \rightarrow C \quad BC \rightarrow D \quad D \rightarrow E$$

2. Minimise left-hand sides

$$A \rightarrow D \quad A \rightarrow E \quad B \rightarrow C \quad B \rightarrow D \quad D \rightarrow E$$

   We can drop C from $BC \rightarrow D$ because $B \rightarrow C$.

3. Remove implied FDs

$$A \rightarrow D \qquad B \rightarrow C \qquad B \rightarrow D \qquad D \rightarrow E$$

   $A \rightarrow E$ can still be derived from $A \rightarrow D$ and $D \rightarrow E$.

Canonical Set of Functional Dependencies

Compute the canonical set of FDs for

$$A, B, C \longrightarrow D, E$$
$$B \longrightarrow C$$
$$B \longrightarrow E$$
$$C \longrightarrow E$$
$$C, D \longrightarrow D, F$$

BCNF Synthesis Algorithm

BNCF Synthesis Algorithm

Input: relation R and a set of FDs for R.

1. Compute a canonical (minimal) set of FDs $\mathcal{F}$.
2. **Maximise the right-hand sides of the FDs:**
   Replace every functional dependency $X \rightarrow Y \in \mathcal{F}$ by

$$X \rightarrow X^{+} - X$$

3. **Split off violating FD's one by one:**
   Start with $\mathcal{S} = \{ R \}$. For every $R_i \in \mathcal{S}$ and $X \rightarrow Y \in \mathcal{F}$: if
   - $X \subseteq R_i$, and (X is contained in R_i)
   - $R_i \not\subseteq X^{+}$, and (X is not a key of R_i)
   - $Y \cap R_i \neq \emptyset$, ($Y$ overlaps with R_i)

   then, let $Z = Y \cap R_i$ and
   - remove attributes Z from the relation R_i, and
   - add a relation with attributes XZ to $\mathcal{S}$.

BCNF Synthesis Algorithm: Example

Consider $R = (A, B, C, D, E)$ with the canonical set of FDs

$$A \to D \qquad B \to C \qquad B \to D \qquad D \to E$$

Here $\{A, B\}$ is the only minimal key. Is R in BCNF? No.

1. Maximise the right-hand sides of the FDs:

$$A \to D, E \qquad B \to C, D, E \qquad D \to E$$

2. Split off violating FD's one by one:
   - $\mathcal{S} = \{ R_0(\underline{A}, \underline{B}, C, D, E) \}$
   - $A \to D, E$ violates BCNF of R_0
   - $\mathcal{S} = \{ R_0(\underline{A}, \underline{B}, C),\ R_1(\underline{A}, D, E) \}$
   - $B \to C, D, E$ violates BCNF of R_0
   - $\mathcal{S} = \{ R_0(\underline{A}, \underline{B}),\ R_1(\underline{A}, D, E),\ R_2(\underline{B}, C) \}$
   - $D \to E$ violates BCNF of R_1
   - $\mathcal{S} = \{ R_0(\underline{A}, \underline{B}),\ R_1(\underline{A}, D),\ R_2(\underline{B}, C),\ R_3(\underline{D}, E) \}$ - done!

Note that we lost the dependency $B \to D$!

3NF Synthesis Algorithm

The **3NF synthesis algorithm** produces a lossless decomposition of a relation into 3NF that **preserves the FDs**.

3NF Synthesis Algorithm

Input: relation R and a set of FDs for R.

1. Compute a canonical (minimal) set of FDs $\mathcal{F}$.
2. For each left-hand side α of an FD in $\mathcal{F}$ create a relation with attributes $\mathcal{A} = \alpha \cup \{B \mid \alpha \to B \in \mathcal{F}\}$.
3. If none of the relations constructed in step 2 contains a key of the original relation R, add one relation containing the attributes of a minimal key of R.
4. For any two relations R_1, R_2 constructed in steps 2,3, if the schema of R_1 is contained in the schema R_2, discard R_1.

3NF Synthesis Algorithm: Example

Use the 3NF synthesis algorithm to normalise the relation

R (A,B,C,D,E,F)

with the following canonical functional dependencies:

$$A \longrightarrow D$$

$$B \longrightarrow C$$

$$B \longrightarrow D$$

$$D \longrightarrow E$$

Efficiency Considerations: BCNF vs 3NF

BCNF does not retain all FDs, therefore 3NF is popular.

Database systems are good at checking **key** constraints, because they create an index on the key columns.

If we leave a table in 3NF (and not BCNF), we have non-key constraints. Namely those FDs that are not implied by keys.

Sometimes we can enforce non-key constraints as key constraints on **materialised views**.

Overview

Introduction

The development of BCNF/3NF has been guided by a particular type of constraint: **functional dependencies.**

The goal of normalization into BCNF/3NF is to

- **eliminate the redundant storage** of data that follows from these constraints, and to
- transform tables such that the **constraints are automatically enforced by means of keys**

However, there are **further types of constraints** which are also useful to during DB design.

Introduction

Recall the Decomposition Theorem

The split of relations is **guaranteed to be lossless** if the intersection (the shared set attributes) of the attributes of the new tables is a key of at least one of them.

The condition in the decomposition theorem is only

- **sufficient** (it guarantees losslessness),
- but **not necessary** (a decomposition might be lossless even if the condition is not satisfied).

Multivalued dependencies (MVDs) are constraints that give a **necessary and sufficient** condition for lossless decomposition

MVDs lead to the **Fourth Normal Form (4NF)**.

Multivalued Dependencies

The following table shows for each employee:

- knowledge of programming languages
- knowledge of programming DBMSs

EMP_KNOWLEDGE		
ENAME	**PROG_LANG**	**DBMS**
John Smith	C	Oracle
John Smith	C	DB2
John Smith	C++	Oracle
John Smith	C++	DB2
Maria Brown	Prolog	PostgreSQL
Maria Brown	Java	PostgreSQL

- There are no non-trivial functional dependencies.
- The table is in **BCNF**.

Nevertheless, there is **redundant information**.

Multivalued Dependencies

The table contains redundant data & must be split.

EMP_LANG

ENAME	PROG_LANG
John Smith	C
John Smith	C++
Maria Brown	Prolog
Maria Brown	Java

EMP_DBMS

ENAME	DBMS
John Smith	Oracle
John Smith	DB2
Maria Brown	PostgreSQL

Note: table may only be decomposed if PROG_LANG and DBMS are **independent**; otherwise **loss of information**.

E.g. it may not be decomposed if the semantics of the table is that the employee knows the interface between the language and the database.

Multivalued Dependencies

The **multivalued dependency (MVD)**

$$\text{ENAME} \twoheadrightarrow \text{PROG\_LANG}$$

means that the **set of values** in column PROG_LANG associated with every ENAME is **independent of all other columns.**

EMP_KNOWLEDGE		
ENAME	**PROG_LANG**	**DBMS**
John Smith	C	Oracle
John Smith	C	DB2
John Smith	C++	Oracle
John Smith	C++	DB2
Maria Brown	Prolog	PostgreSQL
Maria Brown	Java	PostgreSQL

That is, the table contains an

embedded function from ENAME to **sets of** PROG_LANG

Multivalued Dependencies

Formally, ENAME $\twoheadrightarrow$ PROG_LANG holds if: whenever two tuples agree on ENAME, one can exchange their PROG_LANG values and the resulting tupes are in the same table.

From the two table rows

ENAME	PROG_LANG	DBMS
John Smith	C	Oracle
John Smith	C++	DB2

and the MVD ENAME $\twoheadrightarrow$ PROG_LANG, we can conclude that the table must also contain the following rows:

ENAME	PROG_LANG	DBMS
John Smith	C++	Oracle
John Smith	C	DB2

This expresses the **independence** of PROG_LANG for a given ENAME from the rest of the table columns.

Multivalued Dependencies

Multivalued Dependency

A **multivalued dependency (MVD)**

$$A_1, \ldots, A_n \twoheadrightarrow B_1, \ldots, B_m$$

is satisfied in a DB state I if and only if

- for all tuples t, u in $I(R)$ with $t.A_i = u.A_i, 1 \leqslant i \leqslant n$,

  there are two further tuples t', u' in $I(R)$ such that

  1. t' agrees with t except that $t'.B_i = u.B_i, 1 \leqslant i \leqslant m$, and
  2. u' agrees with u except that $u'.B_i = t.B_i, 1 \leqslant i \leqslant m$.

The condition means that the values of the B_i are swapped:

t	$a_1, \ldots, a_n,\ b_1, \ldots, b_m,\ c_1, \ldots, c_k$	t'	$a_1, \ldots, a_n,\ b'_1, \ldots, b'_m,\ c_1, \ldots, c_k$
u	$a_1, \ldots, a_n,\ b'_1, \ldots, b'_m,\ c'_1, \ldots, c'_k$	u'	$a_1, \ldots, a_n,\ b_1, \ldots, b_m,\ c'_1, \ldots, c'_k$

Multivalued Dependencies

Multivalued dependencies always **come in pairs**!

If ENAME $\twoheadrightarrow$ PROG_LANG holds, then ENAME $\twoheadrightarrow$ DBMS is automatically satisfied.

More general:

For a relation $R(A_1, \ldots, A_n, B_1, \ldots, B_m, C_1, \ldots, C_k)$,
the following multivalued dependencies are equivalent

- $A_1, \ldots, A_n \twoheadrightarrow B_1, \ldots, B_m$
- $A_1 \ldots, A_n \twoheadrightarrow C_1, \ldots, C_k$

Swapping the B_j values in two tuples is the same as swapping the values for all other columns (the A_i values are identical, so swapping them has no effect).

Multivalued Dependencies

If the FD $A_1, \ldots A_n \rightarrow B_1, \ldots B_m$ holds, the corresponding MVD

$$A_1, \ldots, A_n \twoheadrightarrow B_1, \ldots, B_m$$

is trivially satisfied.

The FD means that if tuples t, u agree on the A_i then also on the B_j. Swapping thus has no effect (yields t, u again).

Deduction rules to derive **all** implied FDs/MVDs

- The three Armstrong Axioms for FDs.
- If $\alpha \twoheadrightarrow \beta$ then $\alpha \twoheadrightarrow \gamma$, where γ are all remaining columns.
- If $\alpha_1 \twoheadrightarrow \beta_1$ and $\alpha_2 \supseteq \beta_2$ then $\alpha_1 \cup \alpha_2 \twoheadrightarrow \beta_1 \cup \beta_2$.
- If $\alpha \twoheadrightarrow \beta$ and $\beta \twoheadrightarrow \gamma$ then $\alpha \twoheadrightarrow (\gamma - \beta)$.
- If $\alpha \rightarrow \beta$, then $\alpha \twoheadrightarrow \beta$.
- If $\alpha \twoheadrightarrow \beta$ and $\beta' \subseteq \beta$ and there is γ with $\gamma \cap \beta = \emptyset$ and $\gamma \rightarrow \beta'$, then $\alpha \rightarrow \beta'$.

Fourth Normal Form

Fourth Normal Form (4NF)

A relation is in **Fourth Normal Form (4NF)** if every MVD

$$A_1, \ldots, A_n \twoheadrightarrow B_1, \ldots, B_m$$

is

- either trivial, or
- implied by a key.

Note: this definition of 4NF is very similar to BCNF but with a focus on implied MVDs (not FDs).

Since every FD is also an MVD, 4NF is stronger than BCNF.
That is, if a relation is in 4NF, it is automatically in BCNF.

However, it is not very common that 4NF is violated, but BCNF is not.

Fourth Normal Form

The relation

`EMP_KNOWLEDGE (ENAME, PROG_LANG, DBMS)`

is an example of a relation that is in BCNF, but not in 4NF.

The relation has no non-trivial FDs.

Other Constraints

Multiple choice test

The following relation encodes the correct solution to a typical multiple choice test:

ANSWERS			
QUESTION	ANSWER	TEXT	CORRECT
1	A	...	Y
1	B	...	N
1	C	...	N
2	A	...	N
2	B	...	Y
2	C	...	N

Using keys to enforce other constraints

The constraint is not an FD, MVD, or JD:

"*Each question can only have one correct answer.*"

- Can you suggest a transformation of table ANSWERS such that the above constraint is already implied by a key?

Relational Normal Forms

Overview

Introduction

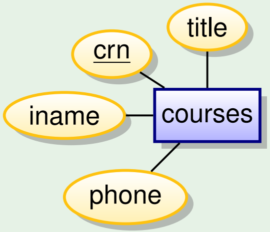
If a "good" ER schema is transformed into the relational model, the result will **satisfy all normal forms** (4NF, BCNF, 3NF).

A normal form violation detected in the generated relational schema indicates a **flaw** in the input ER schema.

This needs to be corrected on the ER level.

FDs in the ER model

The ER equivalent of the very first example in this chapter:



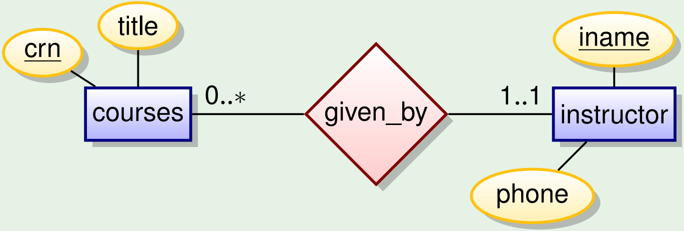
- Obviously, the FD iname $\rightarrow$ phone leads to a violation of BCNF in the resulting table for entity Course.
- Also in the ER model, FDs between attributes of an entity should be implied by a key constraint.

Examples

In the ER model, the solution is the “same” as in the relational model: we have to **split** the entity.

ER entity split

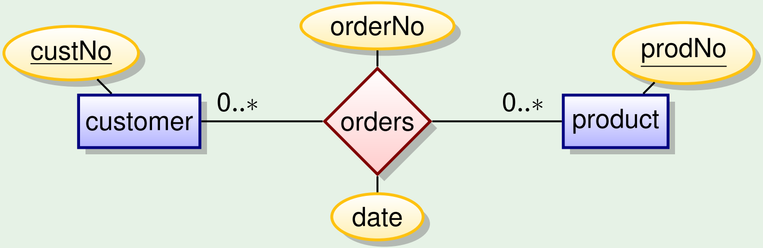
In this case, the instructor is an independent entity:



Examples

Functional dependencies between **attributes of a relationship** always violate BCNF.

Violation of BCNF on the ER level



The FD orderNo $\rightarrow$ date violates BCNF.

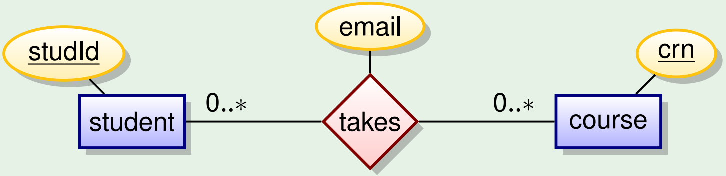
- The key of the table corresponding to the relationship “orders” consists of the attributes custNo, prodNo.

This shows that the concept “order” is an independent entity.

Examples

Violations of BCNF might also be due to the **wrong placement of an attribute.**

Questionable attribute placement



- The relationship is translated into

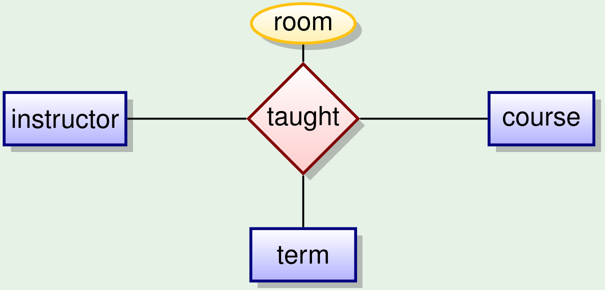
  `TAKES (STUD_ID, CRN, EMAIL)`

- Then the FD `STUD_ID` $\longrightarrow$ `EMAIL` violates BCNF.
- Obviously, email should be an attribute of Student.

Examples

If an attribute of a ternary relationship depends only on two of the entities, this violates BCNF.

Ternary relationship



If every course is taught only once per term, then attribute room depends only on term and course (but not instructor).

Then the FD TERM, COURSE $\rightarrow$ ROOM violates BCNF.

Normalization: Summary

Relational normalization is about:

- **Avoiding redundancy.**
- **Storing separate facts (functions) separately.**
- **Transforming general integrity constraints** into constraints that are supported by the DBMS: **keys.**

Overview

1. Functional Dependencies (FDs)
2. Anomalies, FD-based Normal Forms
3. BCNF and 3NF Synthesis
4. Multivalued Dependencies (MVDs) and 4NF
5. Normal Forms and ER Design
6. Denormalization

Denormalization

Denormalization is the process of **adding redundant columns** to the database in order to **improve performance**.

Redundant data storage

For example, if an application extensively access the phone number of instructors, performance-wise it may make sense to add column PHONE to table COURSES.

COURSES			
CRN	**TITLE**	**INAME**	**PHONE**

This **avoids the otherwise required joins** (on attribute INAME) between tables COURSES and PHONEBOOK.

Denormalization

- Since there is still the separate table `PHONEBOOK`, **insertion and deletion anomalies are avoided.**
- But there will be **update anomalies** (changing a single phone number requires the update of many rows).
- The performance gain is thus paid for with
  - a more complicated application logic (e.g., the need for triggers)
  - and the risk that a faulty application will turn the DB inconsistent
- Denormalization may not only be used to avoid joins:
  - Complete **separate, redundant tables** may be created (increasing the potential for parallel operations).
  - Columns may be added which **aggregate** information in other columns/rows.

Relational Normal Forms: Objectives

After completing this chapter, you should be able to

- work with **functional dependencies** (FDs),
  - define what they are
  - detect them in database schemas
  - decide implication, determine keys
- explain insert, update, and delete **anomalies,**
- understand, explain and use **BCNF**
  - test a given relation for BCNF, and
  - transform a relation into BCNF
- understand, explain and use **3NF**
  - test a given relation for 3NF, and
  - transform a relation into 3NF
- understand, explain **MVDs** and **4NF**
- detect **normal form violations** on the level of ER,
- explain when and how to **denormalize** a DB schema