# Introduction to Database Systems

Radim Bača

Department of Computer Science, FEECS

radim.baca@vsb.cz

dbedu.cs.vsb.cz

# Content

- Motivation
- Functional dependencies
  - Armstrong's axioms
  - Closure of a set of attributes
  - Key
  - Minimal non-redundant functional dependencies

# Database Scheme Design

- There is a lot of ways how to design a database scheme corresponding to a particular assignment
- Some solutions are comparably good, others are considerably worse
- There exists an elegant theory for the database design

# Example

- We want to store this information:
    - name of customer and his/her email, which products he/she bought and how much they cost
- `Purchase(cName, email, pID, pCathegory, pLabel, when, price)`

# Example

- We want to store this information:
  - name of customer and his/her email, which products he/she bought and how much they cost
- `Purchase(cName, email, pID, pCathegory, pLabel, when, price)`

# Example

- `Purchase(cName, email, pID, pCathegory, pLabel, when, price)`

| cName | email | pID | pKat. | pLabel | when | price |
|---|---|---|---|---|---|---|
| Radim | Radim.B@vsb.cz | 1 | cleaner | Electrolux | 1.8.2012 | 520 |
| Jack | jack@theripper.cz | 1 | cleaner | Electrolux | 3.9.2012 | 500 |
| Radim | Radim.B@vsb.cz | 5 | toothpick | GlobalWood | 2.11.2012 | 6 |

- When designing a scheme, so-called anomalies can emerge:
  - anomaly during an update
  - anomaly during a deletion

    $\Downarrow$
  - Anomalies can result in an inconsistent database
- Anomalies are caused mainly by a relation redundancy

# Example

- `Purchase(cName, email, pID, pCathegory, pLabel, when, price)`

| cName | email | pID | pKat. | pLabel | when | price |
|---|---|---|---|---|---|---|
| Radim | Radim.B@vsb.cz | 1 | cleaner | Electrolux | 1.8.2012 | 520 |
| Jack | jack@theripper.cz | 1 | cleaner | Electrolux | 3.9.2012 | 500 |
| Radim | Radim.B@vsb.cz | 5 | toothpick | GlobalWood | 2.11.2012 | 6 |

- When designing a scheme, so-called anomalies can emerge:
  - anomaly during an update
  - anomaly during a deletion

    $\Downarrow$
  - Anomalies can result in an inconsistent database
- Anomalies are caused mainly by a relation redundancy

# Example - A Good Design

- Linear notation of the scheme:
  - `Customer(cName, email)`
  - `Purchase(cName, pID, price, when)`
  - `Product(pID, pCathegory, pLabel)`
- Each customer and each product are only once in the databse

# Example - A Good Design

- Linear notation of the scheme:
  - `Customer(cName, email)`
  - `Purchase(cName, pID, price, when)`
  - `Product(pID, pCathegory, pLabel)`
- Each customer and each product are only once in the databse
- Redundancy can be noticed in repeating foreign keys (for different records), but the consistency of keys is checked by a database system

# Example

- `Purchase(cName, email, pID, pCathegory, pLabel, when, price)`

- Values in the relation have certain relationship: `cName`, `email`
- If two different records in the `Purchase` relation have the same email, they both correspond the same customer
- We denote: `email` $\rightarrow$ `cName`

  and we say that the `cName` attribute is functionally dependent on the `email` attribute

# Example

- `Movie(name, year, length, director)`
- Not only pairs of attributes can be functionally dependent
- Generally, a movie is uniquely detemined by its name and year (this has been observed on IMDB's real-world data)
- So we can write: `name`, `year` $\rightarrow$ `length`, `director`

| Name | Year | Length | Director |
|---|---|---|---|
| Happiness | 1965 | 79 | Agnes Varda |
| Happiness | 1998 | 140 | Todd Solondz |
| American History X | 1998 | 119 | Tony Kaye |

# Functional Dependency (FD) - Definition

- A formal definition of a FD:
  $\forall u, v \in R$ :
  $u[A_1, \dots, A_n] = v[A_1, \dots, A_n] \Rightarrow$

# Functional Dependency (FD) - Definition

- A formal definition of a FD:
  $\forall u, v \in R:$
  $u[A_1, \ldots, A_n] = v[A_1, \ldots, A_n] \Rightarrow u[B_1, \ldots, B_n] = v[B_1, \ldots, B_n]$
- We write: $A_1, \ldots, A_n \rightarrow B_1, \ldots, B_n$, abbreviated as $\overline{A} \rightarrow \overline{B}$
- Functional dependencies represent a concept enabling us to correctly define database schemes
- They can also have other importances

# Functional Dependency (FD) - Definition

- A formal definition of a FD:
  $\forall u, v \in R$ :
  $u[A_1, \dots, A_n] = v[A_1, \dots, A_n] \Rightarrow u[B_1, \dots, B_n] = v[B_1, \dots, B_n]$
- We write: $A_1, \dots, A_n \to B_1, \dots, B_n$, abbreviated as $\overline{A} \to \overline{B}$
- Functional dependencies represent a concept enabling us to correctly define database schemes
- They can also have other importances

# Functional Dependency (FD) - Other Concepts

- A dependency $\overline{A} \rightarrow \overline{B}$ is said to be
  - trivial if $\overline{B} \subset \overline{A}$
  - non-trivial if $\overline{B} \not\subset \overline{A}$
  - totally non-trivial if $\overline{B} \cap \overline{A} = \emptyset$

# Armstrong's Axioms

- There are certain deriving rules for functional dependencies
- These are often called Armstrong's axioms:
  - decomposition
  - union
  - transitivity
  - augmentation

# Decomposition of a FD

- Consider $\overline{A} \rightarrow B_1, \ldots, B_n$

$$\Downarrow$$

$$\begin{array}{l} \overline{A} \rightarrow B_1 \\ \overline{A} \rightarrow B_2 \\ \quad\vdots \\ \overline{A} \rightarrow B_n \end{array}$$

- We say that the FD $\overline{A} \rightarrow B_1, \ldots, B_n$ is decomposed into elementary FDs, i.e., those having only one attribute on the right hand side

- *Can we decompose the left side of a FD?*

# Decomposition of a FD

- Consider $\overline{A} \rightarrow B_1, \ldots, B_n$

$$\Downarrow$$

$$\begin{array}{l} \overline{A} \rightarrow B_1 \\ \overline{A} \rightarrow B_2 \\ \quad\vdots \\ \overline{A} \rightarrow B_n \end{array}$$

- We say that the FD $\overline{A} \rightarrow B_1, \ldots, B_n$ is decomposed into elementary FDs, i.e., those having only one attribute on the right hand side

- *Can we decompose the left side of a FD?*

# Union of FDs

- Consider $\overline{A} \rightarrow B_1$
  $\overline{A} \rightarrow B_2$
  $\vdots$
  $\overline{A} \rightarrow B_n$

$$\Downarrow$$

$$\overline{A} \rightarrow B_1, \ldots, B_n$$

# Union of FDs

- Consider
$$\begin{aligned} \overline{A} &\to B_1 \\ \overline{A} &\to B_2 \\ &\vdots \\ \overline{A} &\to B_n \end{aligned}$$
$$\Downarrow$$
$$\overline{A} \to B_1, \ldots, B_n$$

# Augmentation of a FD

- Consider $\overline{A} \rightarrow \overline{B}$

$$\Downarrow$$

$$\overline{AZ} \rightarrow \overline{BZ} \text{ for any set } Z$$

# Augmentation of a FD

- Consider $\overline{A} \rightarrow \overline{B}$

$$\Downarrow$$

$$\overline{AZ} \rightarrow \overline{BZ} \text{ for any set } Z$$

# Transitivity of FDs

- Consider $\overline{A} \rightarrow \overline{B}$ and

$$\overline{B} \rightarrow \overline{C}$$

$$\Downarrow$$

$$\overline{A} \rightarrow \overline{C}$$

# Transitivity of FDs

- Consider $\overline{A} \to \overline{B}$ and

$$\overline{B} \to \overline{C}$$

$$\Downarrow$$

$$\overline{A} \to \overline{C}$$

# Closure of a Set of Attributes

- Consider a scheme $R$, a set of FDs, and attributes $\overline{A} \subset R$
- Find a set of all attributes $\overline{B} \subset R$ satisfying $\overline{A} \rightarrow \overline{B}$
- The set $\overline{B}$ is called a closure of $\overline{A}$ and is denoted by $\overline{A}+$

# Closure - Algorithm

- Consider a scheme $R$, a set of FDs, and attributes $\overline{A} \subset R$
- Find $\overline{A}+$ (i.e., a closure of the $\overline{A}$ set)
- Algorithm:

$\overline{X} = \overline{A}$;
**while** $\overline{X}$ *is modified* **do**
  **if** *there is a dependency* $\overline{Y} \rightarrow \overline{B}$, *where* $\overline{Y} \subset \overline{X}$ **then**
    add $\overline{B}$ into $X$;
  **end**
  $\overline{A}+ = \overline{X}$;
**end**

# Closure - Algorithm

- Consider a scheme $R$, a set of FDs, and attributes $\overline{A} \subset R$
- Find $\overline{A}+$ (i.e., a closure of the $\overline{A}$ set)
- Algorithm:

$\overline{X} = \overline{A}$;
**while** $\overline{X}$ *is modified* **do**
  **if** *there is a dependency* $\overline{Y} \to \overline{B}$, *where* $\overline{Y} \subset \overline{X}$ **then**
    add $\overline{B}$ into $X$;
  **end**
  $\overline{A}+ = \overline{X}$;
**end**

# Closure - Example

$\overline{X} = \overline{A}$;
**while** $\overline{X}$ *is modified* **do**
  **if** *there is a dependency* $\overline{Y} \to \overline{B}$, *where* $\overline{Y} \subset \overline{X}$ **then**
    add $\overline{B}$ into $X$;
  **end**
  $\overline{A}+ = \overline{X}$;
**end**

- *Find $A+$ for the scheme $R(A, B, C, D, E)$ and this set of FDs:* $\{A \to D, AC \to B, D \to C, B \to E\}$
  1) $\overline{X} = \{A\}$
  2) $\overline{X} = \{A, D\}, (A \to D)$
  3) $\overline{X} = \{A, D, C\}, (D \to C)$
  4) $\overline{X} = \{A, D, C, B\}, (AC \to B)$
  5) $\overline{X} = \{A, D, C, B, E\}, (B \to E)$
  6) $A+ = \overline{X}$

# Closure - Example

$\overline{X} = \overline{A}$;
**while** $\overline{X}$ *is modified* **do**
  **if** *there is a dependency* $\overline{Y} \to \overline{B}$, *where* $\overline{Y} \subset \overline{X}$ **then**
    add $\overline{B}$ into $X$;
  **end**
  $\overline{A}+ = \overline{X}$;
**end**

- *Find* $A+$ *for the scheme* $R(A, B, C, D, E)$ *and this set of FDs:* $\{A \to D, AC \to B, D \to C, B \to E\}$
  1) $\overline{X} = \{A\}$
  2) $\overline{X} = \{A, D\}, (A \to D)$
  3) $\overline{X} = \{A, D, C\}, (D \to C)$
  4) $\overline{X} = \{A, D, C, B\}, (AC \to B)$
  5) $\overline{X} = \{A, D, C, B, E\}, (B \to E)$
  6) $A+ = \overline{X}$

# Closure - Example

$\overline{X}$ = $\overline{A}$;
**while** $\overline{X}$ *is modified* **do**
  **if** *there is a dependency* $\overline{Y} \to \overline{B}$*, where* $\overline{Y} \subset \overline{X}$ **then**
    add $\overline{B}$ into $X$;
  **end**
  $\overline{A}+ = \overline{X}$;
**end**

- *Find* $A+$ *for the scheme* $R(A, B, C, D, E)$ *and this set of FDs:* $\{A \to D, AC \to B, D \to C, B \to E\}$
  1) $\overline{X} = \{A\}$
  2) $\overline{X} = \{A, D\}, (A \to D)$
  3) $\overline{X} = \{A, D, C\}, (D \to C)$
  4) $\overline{X} = \{A, D, C, B\}, (AC \to B)$
  5) $\overline{X} = \{A, D, C, B, E\}, (B \to E)$
  6) $A+ = \overline{X}$

# Closure - Example

$\overline{X} = \overline{A}$;
**while** $\overline{X}$ *is modified* **do**
  **if** *there is a dependency* $\overline{Y} \rightarrow \overline{B}$, *where* $\overline{Y} \subset \overline{X}$ **then**
    add $\overline{B}$ into $X$;
  **end**
  $\overline{A}+ = \overline{X}$;
**end**

- *Find $A+$ for the scheme $R(A, B, C, D, E)$ and this set of FDs:* $\{A \rightarrow D, AC \rightarrow B, D \rightarrow C, B \rightarrow E\}$
  1) $\overline{X} = \{A\}$
  2) $\overline{X} = \{A, D\}, (A \rightarrow D)$
  3) $\overline{X} = \{A, D, C\}, (D \rightarrow C)$
  4) $\overline{X} = \{A, D, C, B\}, (AC \rightarrow B)$
  5) $\overline{X} = \{A, D, C, B, E\}, (B \rightarrow E)$
  6) $A+ = \overline{X}$

# Closure - Example

$\overline{X} = \overline{A}$;
**while** $\overline{X}$ *is modified* **do**
  **if** *there is a dependency* $\overline{Y} \rightarrow \overline{B}$, *where* $\overline{Y} \subset \overline{X}$ **then**
    add $\overline{B}$ into $X$;
  **end**
  $\overline{A}+ = \overline{X}$;
**end**

- *Find* $A+$ *for the scheme* $R(A, B, C, D, E)$ *and this set of FDs:* $\{A \rightarrow D, AC \rightarrow B, D \rightarrow C, B \rightarrow E\}$
  1) $\overline{X} = \{A\}$
  2) $\overline{X} = \{A, D\}, (A \rightarrow D)$
  3) $\overline{X} = \{A, D, C\}, (D \rightarrow C)$
  4) $\overline{X} = \{A, D, C, B\}, (AC \rightarrow B)$
  5) $\overline{X} = \{A, D, C, B, E\}, (B \rightarrow E)$
  6) $A+ = \overline{X}$

# Closure - Example

$\overline{X} = \overline{A}$;
**while** $\overline{X}$ *is modified* **do**
  **if** *there is a dependency* $\overline{Y} \to \overline{B}$, *where* $\overline{Y} \subset \overline{X}$ **then**
    add $\overline{B}$ into $X$;
  **end**
  $\overline{A}+ = \overline{X}$;
**end**

- *Find* $A+$ *for the scheme* $R(A, B, C, D, E)$ *and this set of FDs:* $\{A \to D, AC \to B, D \to C, B \to E\}$
  1) $\overline{X} = \{A\}$
  2) $\overline{X} = \{A, D\}, (A \to D)$
  3) $\overline{X} = \{A, D, C\}, (D \to C)$
  4) $\overline{X} = \{A, D, C, B\}, (AC \to B)$
  5) $\overline{X} = \{A, D, C, B, E\}, (B \to E)$
  6) $A+ = \overline{X}$

# Closure - Example

$\overline{X} = \overline{A}$;
**while** $\overline{X}$ *is modified* **do**
  **if** *there is a dependency* $\overline{Y} \rightarrow \overline{B}$, *where* $\overline{Y} \subset \overline{X}$ **then**
    add $\overline{B}$ into $X$;
  **end**
  $\overline{A}+ = \overline{X}$;
**end**

- *Find $A+$ for the scheme $R(A, B, C, D, E)$ and this set of FDs:* $\{A \rightarrow D, AC \rightarrow B, D \rightarrow C, B \rightarrow E\}$
  1) $\overline{X} = \{A\}$
  2) $\overline{X} = \{A, D\}, (A \rightarrow D)$
  3) $\overline{X} = \{A, D, C\}, (D \rightarrow C)$
  4) $\overline{X} = \{A, D, C, B\}, (AC \rightarrow B)$
  5) $\overline{X} = \{A, D, C, B, E\}, (B \rightarrow E)$
  6) $A+ = \overline{X}$

# Key

- A set of attributes $K \subset R$ is a key of $R$ if all attributes of the scheme are functionally dependent on $K$
- So if $K \rightarrow$ all attributes of $R$
- Usually there is, moreover, stated that there is no subset of $K$ (different from $K$) which is a key of $R$

# Key and Closure

- Consider $\overline{A} \subset R$ and find out if this is a key of $R$
- We solve this problem by finding a closure of $\overline{A}$
- If the closure $\overline{A}+$ involves all attributes of $R$, then $\overline{A}$ is a key of $R$

# Key and Closure

- Consider $\overline{A} \subset R$ and find out if this is a key of $R$
- We solve this problem by finding a closure of $\overline{A}$
- If the closure $\overline{A}+$ involves all attributes of $R$, then $\overline{A}$ is a key of $R$

# How to Find a Key?

- We want to find all keys for a given set of FDs
- Theoretically, we should determine a closure of every subset of attributes
- Practically, we start with the shortest subsets and proceed to longer ones
- After we find some key, we do not have to test supersets of this key since they will be keys too

# How to Find a Key?

- We want to find all keys for a given set of FDs
- Theoretically, we should determine a closure of every subset of attributes
- Practically, we start with the shortest subsets and proceed to longer ones
- After we find some key, we do not have to test supersets of this key since they will be keys too

# Minimal Non-redundant FDs

- The goal is:
  to find a minimal set of totally non-trivial and non-redundant FDs such that all FDs for the relational scheme are implied by this set
- When determining a set of FDs for some scheme, we usually intuitively create a set satisfying this condition
- In the following slides, we introduce a technique how to find this set

# Minimal Non-redundant FDs

- The goal is:
  to find a minimal set of totally non-trivial and non-redundant FDs such that all FDs for the relational scheme are implied by this set
- When determining a set of FDs for some scheme, we usually intuitively create a set satisfying this condition
- In the following slides, we introduce a technique how to find this set

# Redundant FDs

- Having a set $F$ of FDs, we want to determine if $\overline{A} \rightarrow B$ is implied by $F$ (i.e., if $\overline{A} \rightarrow B$ is a redundant FD)
- Note that $B$ is a single attribute (we deal with an elementary FD) - every set of FDs can be easily decomposed by using Armstrong's decomposition rule into a set of elementary FDs
- Basically, we have two options how to resolve this problem:
  - to determine a closure of $\overline{A}$ by using the rules from $F$; if the closure involves $B$, then the dependency $\overline{A} \rightarrow B$ is redundant
  - to derive $\overline{A} \rightarrow B$ directly from $F$ by using Armstrong's axioms

# Redundant FDs

- Having a set $F$ of FDs, we want to determine if $\overline{A} \rightarrow B$ is implied by $F$ (i.e., if $\overline{A} \rightarrow B$ is a redundant FD)
- Note that $B$ is a single attribute (we deal with an elementary FD) - every set of FDs can be easily decomposed by using Armstrong's decomposition rule into a set of elementary FDs
- Basically, we have two options how to resolve this problem:
  - to determine a closure of $\overline{A}$ by using the rules from $F$; if the closure involves $B$, then the dependency $\overline{A} \rightarrow B$ is redundant
  - to derive $\overline{A} \rightarrow B$ directly from $F$ by using Armstrong's axioms

# Redundant FDs - Example

- *Consider $R(X, Y, Z)$ and this set of FDs: $\{X \rightarrow YZ, Y \rightarrow XZ\}$. Determine non-redundant set of FDs.*
- We create a set of elementary FDs: $\{X \rightarrow Y, X \rightarrow Z, Y \rightarrow X, Y \rightarrow Z\}$
- We can pick every elementary FD and try to find out if it is redundant

# Redundant FDs - Example

- *Consider $R(X, Y, Z)$ and this set of FDs: $\{X \rightarrow YZ, Y \rightarrow XZ\}$. Determine non-redundant set of FDs.*
- We create a set of elementary FDs: $\{\mathbf{X} \rightarrow \mathbf{Y}, X \rightarrow Z, Y \rightarrow X, Y \rightarrow Z\}$
- We can pick every elementary FD and try to find out if it is redundant
- Let is start with $X \rightarrow Y$:
  - The remaining FDs are $\{X \rightarrow Z, Y \rightarrow X, Y \rightarrow Z\}$
  - $\overline{X}+ = \{X, Z\}$, which does not contain $Y$, so that the rule is not redundant

# Redundant FDs - Example

- *Consider $R(X, Y, Z)$ and this set of FDs: $\{X \rightarrow YZ, Y \rightarrow XZ\}$. Determine non-redundant set of FDs.*
- We create a set of elementary FDs: $\{\mathbf{X} \rightarrow \mathbf{Y}, X \rightarrow Z, Y \rightarrow X, Y \rightarrow Z\}$
- We can pick every elementary FD and try to find out if it is redundant
- Let is start with $X \rightarrow Y$:
  - The remaining FDs are $\{X \rightarrow Z, Y \rightarrow X, Y \rightarrow Z\}$
  - $\overline{X}+ = \{X, Z\}$, which does not contain $Y$, so that the rule is not redundant

# Redundant FDs - Example

- *Consider $R(X, Y, Z)$ and this set of FDs: $\{X \rightarrow YZ, Y \rightarrow XZ\}$. Determine non-redundant set of FDs.*
- We create a set of elementary FDs: $\{X \rightarrow Y, \mathbf{X} \rightarrow \mathbf{Z}, Y \rightarrow X, Y \rightarrow Z\}$
- We proceed with the rule $X \rightarrow Z$:
  - The remaining FDs are $\{X \rightarrow Y, Y \rightarrow X, Y \rightarrow Z\}$
  - $\overline{X}+ = \{X, Y, Z\}$, which contains $Z$, so that the rule is redundant
  - It can be noticed that the rule $X \rightarrow Z$ can be derived from $X \rightarrow Y$ and $Y \rightarrow Z$ by using transitivity
  - The set of FDs without $X \rightarrow Z$ is already non-redundant (it can be shown analogously)

# Redundant FDs - Example

- *Consider $R(X, Y, Z)$ and this set of FDs: $\{X \rightarrow YZ, Y \rightarrow XZ\}$. Determine non-redundant set of FDs.*
- We create a set of elementary FDs: $\{X \rightarrow Y, \mathbf{X} \rightarrow \mathbf{Z}, Y \rightarrow X, Y \rightarrow Z\}$
- We proceed with the rule $X \rightarrow Z$:
  - The remaining FDs are $\{X \rightarrow Y, Y \rightarrow X, Y \rightarrow Z\}$
  - $\overline{X}+ = \{X, Y, Z\}$, which contains $Z$, so that the rule is redundant
  - It can be noticed that the rule $X \rightarrow Z$ can be derived from $X \rightarrow Y$ and $Y \rightarrow Z$ by using transitivity
  - The set of FDs without $X \rightarrow Z$ is already non-redundant (it can be shown analogously)

# Redundant FDs - Example

- *Consider $R(X, Y, Z)$ and this set of FDs: $\{X \rightarrow YZ, Y \rightarrow XZ\}$. Determine non-redundant set of FDs.*
- We create a set of elementary FDs: $\{X \rightarrow Y$, $\mathbf{X} \rightarrow \mathbf{Z}$, $Y \rightarrow X$, $Y \rightarrow Z\}$
- We proceed with the rule $X \rightarrow Z$:
  - The remaining FDs are $\{X \rightarrow Y, Y \rightarrow X, Y \rightarrow Z\}$
  - $\overline{X}+ = \{X, Y, Z\}$, which contains $Z$, so that the rule is redundant
  - It can be noticed that the rule $X \rightarrow Z$ can be derived from $X \rightarrow Y$ and $Y \rightarrow Z$ by using transitivity
  - The set of FDs without $X \rightarrow Z$ is already non-redundant (it can be shown analogously)

# Redundant FDs - Example

- *Consider $R(X, Y, Z)$ and this set of FDs: $\{X \to YZ, Y \to XZ\}$. Determine non-redundant set of FDs.*
- We create a set of elementary FDs: $\{X \to Y, X \to Z, \mathbf{Y} \to \mathbf{X}, Y \to Z\}$
- Let us examine the rule $Y \to X$:
  - The remaining FDs are $\{X \to Y, X \to Z, Y \to Z\}$
  - $\overline{Y}+ = \{Y, Z\}$, which does not contain $X$, so that the rule is not redundant

# Redundant FDs - Example

- *Consider $R(X, Y, Z)$ and this set of FDs: $\{X \rightarrow YZ, Y \rightarrow XZ\}$. Determine non-redundant set of FDs.*
- We create a set of elementary FDs: $\{X \rightarrow Y, X \rightarrow Z,$ **$\mathbf{Y} \rightarrow \mathbf{X}$**, $Y \rightarrow Z\}$
- Let us examine the rule $Y \rightarrow X$:
  - The remaining FDs are $\{X \rightarrow Y, X \rightarrow Z, Y \rightarrow Z\}$
  - $\overline{Y}+ = \{Y, Z\}$, which does not contain $X$, so that the rule is not redundant

# Redundant FDs - Example

- *Consider $R(X, Y, Z)$ and this set of FDs: $\{X \to YZ, Y \to XZ\}$. Determine non-redundant set of FDs.*
- We create a set of elementary FDs: $\{X \to Y, X \to Z, Y \to X,$ **$Y \to Z$**$\}$
- The last rule is $Y \to Z$:
  - The remaining FDs are $\{X \to Y, X \to Z, Y \to X\}$
  - $\overline{Y}+ = \{X, Y, Z\}$, which contains $Z$, so that the rule is redundant
  - This rule can be again derived by using transitivity

# Redundant FDs - Example

- *Consider $R(X, Y, Z)$ and this set of FDs: $\{X \rightarrow YZ, Y \rightarrow XZ\}$. Determine non-redundant set of FDs.*
- We create a set of elementary FDs: $\{X \rightarrow Y, X \rightarrow Z, Y \rightarrow X,$ **Y** $\rightarrow$ **Z**$\}$
- The last rule is $Y \rightarrow Z$:
  - The remaining FDs are $\{X \rightarrow Y, X \rightarrow Z, Y \rightarrow X\}$
  - $\overline{Y}+ = \{X, Y, Z\}$, which contains $Z$, so that the rule is redundant
  - This rule can be again derived by using transitivity

# Redundant FDs - Example

- *Consider $R(X, Y, Z)$ and this set of FDs: $\{X \rightarrow YZ, Y \rightarrow XZ\}$. Determine non-redundant set of FDs.*
- We create a set of elementary FDs: $\{X \rightarrow Y, X \rightarrow Z, Y \rightarrow X,$ $\mathbf{Y} \rightarrow \mathbf{Z}\}$
- The last rule is $Y \rightarrow Z$:
  - The remaining FDs are $\{X \rightarrow Y, X \rightarrow Z, Y \rightarrow X\}$
  - $\overline{Y}+ = \{X, Y, Z\}$, which contains $Z$, so that the rule is redundant
  - This rule can be again derived by using transitivity

# Redundant FDs - Example

- *Consider $R(X, Y, Z)$ and this set of FDs: $\{X \rightarrow YZ, Y \rightarrow XZ\}$. Determine non-redundant set of FDs.*
- We create a set of elementary FDs: $\{X \rightarrow Y, X \rightarrow Z, Y \rightarrow X, Y \rightarrow Z\}$
- So the result is that we have two non-redundant sets of FDs:
  1. $\{X \rightarrow Y, Y \rightarrow X, Y \rightarrow Z\}$
  2. $\{X \rightarrow Y, X \rightarrow Z, Y \rightarrow X\}$

# Removal of Redundant Attributes

- In the previous example, we have shown how to remove FDs
- To obtain a set of FDs as small as possible, it is necessary to remove redundant attributes on the left hand side of FDs
- If $\overline{A} \rightarrow \overline{B}$ and for a $C \in \overline{A}$ it holds that $(\overline{A} - C)+ = \overline{A}+$, then the $C$ attribute is redundant for this FD

# Removal of Redundant Attributes

- In the previous example, we have shown how to remove FDs
- To obtain a set of FDs as small as possible, it is necessary to remove redundant attributes on the left hand side of FDs
- If $\overline{A} \rightarrow \overline{B}$ and for a $C \in \overline{A}$ it holds that $(\overline{A} - C)+ = \overline{A}+$, then the $C$ attribute is redundant for this FD

# Removal of Redundant Attributes

- In the previous example, we have shown how to remove FDs
- To obtain a set of FDs as small as possible, it is necessary to remove redundant attributes on the left hand side of FDs
- If $\overline{A} \rightarrow \overline{B}$ and for a $C \in \overline{A}$ it holds that $(\overline{A} - C)+ = \overline{A}+$, then the $C$ attribute is redundant for this FD

# Removal of Redundant Attributes - Example

- *Consider $R(A, B, C, D, E)$ and this set of FDs:*
  $\{ABC \rightarrow D, E \rightarrow C, AB \rightarrow E, C \rightarrow D\}$.
  *Remove redundant attributes.*

- Let us check only this FD: $ABC \rightarrow D$

- First we obtain that $ABC+ = \{A, B, C, D, E\}$

- Then we determine the closures $BC+ = \{B, C, D\}$, $AC+ = \{A, C, D\}$, and $AB+ = \{A, B, C, D, E\}$

- Evidently, the $C$ attribute is redundant since $ABC+ = AB+$

- So the result is this set of FDs:
  $\{AB \rightarrow DE, E \rightarrow C, C \rightarrow D\}$

# Removal of Redundant Attributes - Example

- *Consider $R(A, B, C, D, E)$ and this set of FDs:*
  $\{ABC \rightarrow D, E \rightarrow C, AB \rightarrow E, C \rightarrow D\}$.
  *Remove redundant attributes.*
- Let us check only this FD: $ABC \rightarrow D$
- First we obtain that $ABC+ = \{A, B, C, D, E\}$
- Then we determine the closures $BC+ = \{B, C, D\}$,
  $AC+ = \{A, C, D\}$, and $AB+ = \{A, B, C, D, E\}$
- Evidently, the $C$ attribute is redundant since $ABC+ = AB+$
- So the result is this set of FDs:
  $\{AB \rightarrow DE, E \rightarrow C, C \rightarrow D\}$

# Removal of Redundant Attributes - Example

- *Consider $R(A, B, C, D, E)$ and this set of FDs:*
  $\{ABC \to D, E \to C, AB \to E, C \to D\}$.
  *Remove redundant attributes.*

- Let us check only this FD: $ABC \to D$
- First we obtain that $ABC+ = \{A, B, C, D, E\}$
- Then we determine the closures $BC+ = \{B, C, D\}$, $AC+ = \{A, C, D\}$, and $AB+ = \{A, B, C, D, E\}$
- Evidently, the $C$ attribute is redundant since $ABC+ = AB+$
- So the result is this set of FDs:
  $\{AB \to DE, E \to C, C \to D\}$

# Removal of Redundant Attributes - Example

- *Consider $R(A, B, C, D, E)$ and this set of FDs:*
  $\{ABC \rightarrow D, E \rightarrow C, AB \rightarrow E, C \rightarrow D\}$.
  *Remove redundant attributes.*

- Let us check only this FD: $ABC \rightarrow D$
- First we obtain that $ABC+ = \{A, B, C, D, E\}$
- Then we determine the closures $BC+ = \{B, C, D\}$, $AC+ = \{A, C, D\}$, and $AB+ = \{A, B, C, D, E\}$
- Evidently, the $C$ attribute is redundant since $ABC+ = AB+$
- So the result is this set of FDs:
  $\{AB \rightarrow DE, E \rightarrow C, C \rightarrow D\}$

# Removal of Redundant Attributes - Example

- *Consider $R(A, B, C, D, E)$ and this set of FDs:*
  $\{ABC \rightarrow D, E \rightarrow C, AB \rightarrow E, C \rightarrow D\}$.
  *Remove redundant attributes.*

- Let us check only this FD: $ABC \rightarrow D$
- First we obtain that $ABC+ = \{A, B, C, D, E\}$
- Then we determine the closures $BC+ = \{B, C, D\}$,
  $AC+ = \{A, C, D\}$, and $AB+ = \{A, B, C, D, E\}$
- Evidently, the $C$ attribute is redundant since $ABC+ = AB+$
- So the result is this set of FDs:
  $\{AB \rightarrow DE, E \rightarrow C, C \rightarrow D\}$

# Removal of Redundant Attributes - Example

- *Consider $R(A, B, C, D, E)$ and this set of FDs:*
  $\{ABC \to D, E \to C, AB \to E, C \to D\}$.
  *Remove redundant attributes.*

- Let us check only this FD: $ABC \to D$
- First we obtain that $ABC+ = \{A, B, C, D, E\}$
- Then we determine the closures $BC+ = \{B, C, D\}$, $AC+ = \{A, C, D\}$, and $AB+ = \{A, B, C, D, E\}$
- Evidently, the $C$ attribute is redundant since $ABC+ = AB+$
- So the result is this set of FDs:
  $\{AB \to DE, E \to C, C \to D\}$

# References


- Jennifer Widom. Introduction to Databases. `https://www.coursera.org/course/db`
- Course home pages `http://dbedu.cs.vsb.cz`