# Tutorial – Functional Depedency

CSCC43 - Introduction to Databases
Summer 2009

# Anomalie

- Redundancy leads to anomalies. Consider the relation schema: *Rents(CustomerID, Title, Price, Rating, Date)*, and following instance of the schema (which records customer renting movies at some dates). Describe update, deletion and insertion anomalies caused by this design.

| CustomerID | Title | Price | Rating | Date |
|---|---|---|---|---|
| 0001 | Braveheart | 4.20 | PG13 | 2003-01-08 |
| 0001 | The Patriot | 3.30 | R | 2003-01-08 |
| 0002 | The Patriot | 3.30 | R | 2003-03-03 |
| 0003 | Ransom | 3.50 | PG13 | 2003-04-01 |

# Functional Dependency

- FD: given schema **R(X)** and non-empty subsets **Y** and **Z** of the attributes **X**, we say that there is a *functional dependency* between **Y** and **Z** (**Y**→**Z**), iff for every relation instance **r** of **R(X)** and every pair of tuples **t1, t2** of **r**, if **t1.Y = t2.Y**, then **t1.Z = t2.Z**
- …, or, … the values of **Y** uniquely determines the values of **Z** in all possible instances of **R.** (***FD on non-keys causes redundancy***)
- Examples:
  - A film has a unique title, rental price and distributor.

    $$\text{FilmID} \rightarrow \text{Title, RentalPrice, Distributor}$$

  - The customerID uniquely identifies the customer and his/her address

    $$\text{CustomerID} \rightarrow \text{Name, Street, City, State}$$

  - On any particular day, a film tape can be rented to at most one customer.

    $$\text{Date, FilmID, TapeNum} \rightarrow \text{CustomerID}$$

  - A performer can have only one role in a particular movie.

    $$\text{PerformerID, FilmID} \rightarrow \text{Role}$$

# Armstrong's Axioms

- **Armstrong**
  - ***Reflexivity***: $Y \subseteq X$ **|-** $X \rightarrow Y$
  - ***Augmentation***: $X \rightarrow Y$ **|-** $XZ \rightarrow YZ$
  - ***Transitivity***: $X \rightarrow Y$, $Y \rightarrow Z$ **|-** $X \rightarrow Z$
- **Derived Rules**
  - ***Decomposition*****:** $X \rightarrow YZ$ **|-** $X \rightarrow Y$, $X \rightarrow Z$
  - ***Union***: $X \rightarrow Y$ and $X \rightarrow Z$ **|-** $X \rightarrow YZ$

# Armstrong's Axioms

- Proof following rules using Armstrong's Axioms
  - ***Left Augmentation***
    $X \rightarrow Y$ **|-** $XZ \rightarrow Y$
  - ***Pseudotransitivity***
    $X \rightarrow Y, YZ \rightarrow W$ **|-** $XZ$ -> $W$
  - ***Addition***
    $X \rightarrow Y, Z \rightarrow W$ |- $XZ$ -> $YW$

# Armstrong's Axioms

- Let ***F*** be the following set of functional dependencies: {$AB \rightarrow CD$, $B \rightarrow DE$, $C \rightarrow F$, $E \rightarrow G$, $A \rightarrow B$}. Use Armstrong's axioms to show that {$A \rightarrow FG$} is logically implied by ***F***

| | | |
|---|---|---|
| 1 | $A \rightarrow B$ | Given |
| 2 | $A \rightarrow AB$ | 1, Augmentation |
| 3 | $AB \rightarrow CD$ | Given |
| 4 | $A \rightarrow CD$ | 2, 3, Transitivity |
| 5 | $B \rightarrow DE$ | Given |
| 6 | $A \rightarrow DE$ | 1, 5, Transitivity |
| 7 | $A \rightarrow ACD$ | 4, Augmentation |
| 8 | $ACD \rightarrow CDE$ | 6, Augmentation twice |
| 9 | $A \rightarrow CDE$ | 7, 8, Transitivity |
| 10 | $A \rightarrow CE$ | 9, Trivial dependency |
| 11 | $C \rightarrow F$ | Given |
| 12 | $CE \rightarrow FE$ | 11, Augmentation |
| 13 | $E \rightarrow G$ | Given |
| 14 | $FE \rightarrow FG$ | 13, Augmentation |
| 15 | $CE \rightarrow FG$ | 12, 14, Transitivity |
| 16 | $A \rightarrow FG$ | 10, 15, Transitivity |

# FD Closure

- A FD **f**: **Y→Z** on schema **R(X)** is a constraint on *all* allowable instances of **R**
- **F** *entails* **f** if every instance of R that *satisfies* **F** also *satisfies* **f**.
- The *closure* of **F**, denoted $\mathbf{F^+}$, is the set of all FDs entailed by **F**.
- Given a set of FDs **F**, the attribute closure of a set of attributes X is $X^+_F$ = set of all attributes A such that X → A (entailed by **F**)

```
closure := X;            // since X ⊆ X+F
repeat
  old := closure;
  if there is an FD Z → V in F such that
     Z ⊆ closure and V ⊈ closure
  then closure := closure ∪ V
until old = closure
```

– If $T \subseteq closure$ then $X \to T$ is entailed by $F$

# FD Closure

- Given **$R$** = $ABCD$ and **$F$** = $\{A \to B, A \to C, CD \to A\}$. *Compute* $F^+$.
- *Solution:*
  - $A^+_F = \{ABC\}$
  - $B^+_F = \{B\}$
  - …
  - $AB^+_F = \{ABC\}$
  - $AC^+_F = \{ABC\}$
  - …
  - $ABC^+_F = \ldots$

# FD Closure

- (Previous final question) Given $\boldsymbol{R} = ABCDEGH$ and $\boldsymbol{F} = \{A \rightarrow DE, C \rightarrow ADH, BH \rightarrow GE, ABH \rightarrow C, BGH \rightarrow C\}$. *Compute* $X^+_F$ *for sets of attributes X such that X appears on the left hand side of a FD in F.*
- *Solution:*
  - $A^+_F = \{ADE\}$
  - $C^+_F = \{ACDEH\}$
  - $BH^+_F = \{BCEGH\}$
  - $ABH^+_F = \{ABCDEGH\}$
  - $BGH^+_F = \{ABCDEGH\}$
- *Use Armstrong's Axioms to prove each FD!*

# Key

- Recall a key is a minimal superkey, where superkey of a schema **R** is a set of attributes in **R** that functionally determines all attributes in **R**.
- Consider the relation schema **R**(*A, B, C, D*) with FDs: *A*⟶*C* and *B*⟶*D*. Is {*A*,*B*} a key for **R**?

Fact: $\{A, B\}$ is a superkey.

Indeed from Armstrong's Axioms we can infer:

$$A \rightarrow C \Rightarrow AB \rightarrow \text{ABC (augmentation by AB)}$$

$$B \rightarrow D \Rightarrow ABC \rightarrow \text{ABCD (augmentation by ABC)}$$

We obtain $AB \rightarrow ABCD$ (transitivity)

$\{A, B\}$ is a candidate key (minimal). We must show that neither $\{A\}$ nor $\{B\}$ alone are candidate keys.

Neither $\{A\}$ nor $\{B\}$ are superkeys since $\{A\}^+ = \{A, C\}$ , $\{B\}^+ = \{B, D\}$

(But $\{A, B\}^+ = \{A, B, C, D\}$)

# key

- Consider a schema **R**={*S,T,V,C,P,D*} and **F**= {*S* → *T, V* → *SC, SD* → *P*}. Find keys for **R**.

- **Solution**
  - *V* and *D* do not appear on the right side, they must be in the key.
  - $VD^+_F$ = {*STVCPD*}. So VD is the only key.

- Now proof: *F* |= *VD* → *STVCPD* using Armstrong's Axioms
  **Solution**

| | |
|---|---|
| *(1) V → SC* | *# given* |
| *(2) VD → SCD* | *# aug. (1)* |
| *(3) VD → VSCD* | *# aug. (2)* |
| *(4) S → T* | *# given* |
| *(5) SC → TC* | *# aug. (4)* |
| *(6) V → TC* | *# (1), (5)* |
| *(7) VD → TCD* | *# (6) aug.* |
| *(8) SD → P* | *# given* |
| *(9) VD → SD* | *# decomp. (2)* |
| *(10) VD → P* | *# trans. (8), (9)* |
| *(11) VD → STVCPD* | *# union (3),(7),(10)* |

# All together

- Consider a relation with schema R(A,B,C,D) and FD's **F** = {*AB* → *C, C* → *D, D* → *A*}.
  - a) Compute $\mathbf{F}^+$; You can restrict yourself to non-trivial FD's with single attributes on the right side.

**Solution**:

$A^+_F = \{A\}$
$B^+_F = \{B\}$
$C^+_F = \{ACD\}$ # add C → A to $\mathbf{F}^+$
$D^+_F = \{AD\}$
$\mathbf{AB^+_F = \{ABCD\}}$ # add AB → D to $\mathbf{F}^+$
$AC^+_F = \{ACD\}$ # add AC → D to $\mathbf{F}^+$
$AD^+_F = \{AD\}$
$\mathbf{BC^+_F = \{ABCD\}}$ # add BC → A, BC → D to $\mathbf{F}^+$
$\mathbf{BD^+_F = \{ABCD\}}$ # add BD → A, BD → C to $\mathbf{F}^+$
$CD^+_F = \{ACD\}$ # add CD → A to $\mathbf{F}^+$
$\mathbf{ABC^+_F = \{ABCD\}}$ # add ABC → D to $\mathbf{F}^+$
$\mathbf{ABD^+_F = \{ABCD\}}$ # add ABD → C to $\mathbf{F}^+$
$ACD^+_F = \{ACD\}$
$\mathbf{BCD^+_F = \{ABCD\}}$ # add BCD → A to $\mathbf{F}^+$
$\mathbf{ABCD^+_F = \{ABCD\}}$

$\mathbf{F}^+ = \mathbf{F} \cup$ {C → A, AB → D, AC → D, BC → A, BC → D, BD → A, BD → C, CD → A, ABC → D, ABD → C, BCD → A} + {trivial FDs}.

# All together

– What are all the keys of R?

**Solution**

- ***Superkeys*** are sets of attributes whose closures are all attributes: AB, BC, BD, ABC, ABD, BCD, ABCD
- ***Keys*** are minimal superkeys: AB, BC, BD