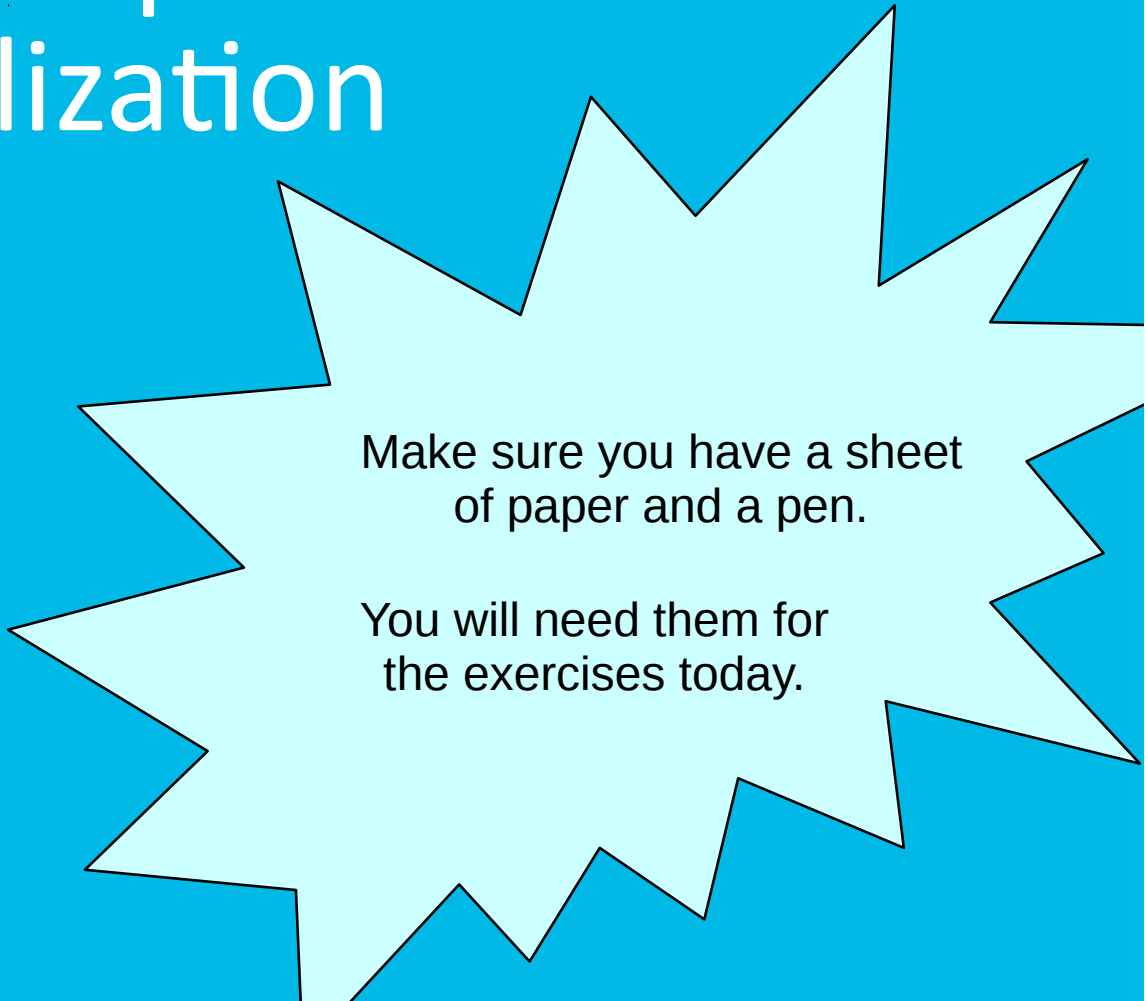


Database Technology

Functional Dependencies and Normalization

Olaf Hartig

olaf.hartig@liu.se



Make sure you have a sheet
of paper and a pen.

You will need them for
the exercises today.

Quiz

- Constraint between two sets of attributes from a relation

Let R be a relational schema with the attributes $A_1, A_2, \dots, A_n$ and let X and Y be subsets of $\{A_1, A_2, \dots, A_n\}$.

Then, the functional dependency $X \rightarrow Y$ specifies the following constraint on *any* valid relation state r of R .

For *any* two tuples t_1 and t_2 in state r we have that:

if $t_1[X] = t_2[X]$, then $t_1[Y] = t_2[Y]$.

- Consider the relation **R(A,B,C,D,E,F)** with the following FDs:
FD1: **A** $\rightarrow$ **BC**
FD2: **C** $\rightarrow$ **AD**
FD3: **DE** $\rightarrow$ **F**
- Is a state that contains the tuples (3,8,1,2,3,4) and (3,7,1,2,3,4) valid?
1) Yes 2) No

Reasoning About FDs

- Logical implications can be derived by using inference rules called **Armstrong's rules**:
 - *Reflexivity*: If Y is a subset of X , then $X \rightarrow Y$
 - *Augmentation*: If $X \rightarrow Y$, then $XZ \rightarrow YZ$
(we use XY as a short form for $X \cup Y$)
 - *Transitivity*: If $X \rightarrow Y$ and $Y \rightarrow Z$, then $X \rightarrow Z$
- Additional rules can be derived:
 - *Decomposition*: If $X \rightarrow YZ$, then $X \rightarrow Y$
 - *Union*: If $X \rightarrow Y$ and $X \rightarrow Z$, then $X \rightarrow YZ$
 - *Pseudo-transitivity*: If $X \rightarrow Y$ and $WY \rightarrow Z$, then $WX \rightarrow Z$

Exercises: Armstrong Rules

- Consider the relation **R(A,B,C,D,E,F)** with the following FDs:

FD1: **A** → **BC**

FD2: **C** → **AD**

FD3: **DE** → **F**

- Use the Armstrong rules to derive the following FD: **AC** → **D**

Reflexivity: If Y is a subset of X , then $X \rightarrow Y$

Augmentation: If $X \rightarrow Y$, then $XZ \rightarrow YZ$

Transitivity: If $X \rightarrow Y$ and $Y \rightarrow Z$, then $X \rightarrow Z$

Decomposition: If $X \rightarrow YZ$, then $X \rightarrow Y$

Union: If $X \rightarrow Y$ and $X \rightarrow Z$, then $X \rightarrow YZ$

Pseudo-transitivity: If $X \rightarrow Y$ and $WY \rightarrow Z$,
then $WX \rightarrow Z$

Exercises: Armstrong Rules

- Consider the relation **R(A,B,C,D,E,F)** with the following FDs:

FD1: **A** → **BC**

FD2: **C** → **AD**

FD3: **DE** → **F**

- Use the Armstrong rules to derive the following FD: **AC** → **D**
 - **FD4: AC** → **AD** (Augmentation of FD2 with A)
 - **FD5: AC** → **D** (Decomposition of FD4)

Reflexivity: If Y is a subset of X , then $X \rightarrow Y$

Augmentation: If $X \rightarrow Y$, then $XZ \rightarrow YZ$

Transitivity: If $X \rightarrow Y$ and $Y \rightarrow Z$, then $X \rightarrow Z$

Decomposition: If $X \rightarrow YZ$, then $X \rightarrow Y$

Union: If $X \rightarrow Y$ and $X \rightarrow Z$, then $X \rightarrow YZ$

Pseudo-transitivity: If $X \rightarrow Y$ and $WY \rightarrow Z$, then $WX \rightarrow Z$

Exercises: Armstrong Rules

- Consider the relation **R(A,B,C,D,E,F)** with the following FDs:

FD1: **A** → **BC**

FD2: **C** → **AD**

FD3: **DE** → **F**

- Use the Armstrong rules to derive the following FD: **AC** → **D**
 - **FD4: AC** → **AD** (Augmentation of FD2 with A)
 - **FD5: AC** → **D** (Decomposition of FD4)
- Use the Armstrong rules to derive the following FD: **A** → **D**

Reflexivity: If Y is a subset of X , then $X \rightarrow Y$

Augmentation: If $X \rightarrow Y$, then $XZ \rightarrow YZ$

Transitivity: If $X \rightarrow Y$ and $Y \rightarrow Z$, then $X \rightarrow Z$

Decomposition: If $X \rightarrow YZ$, then $X \rightarrow Y$

Union: If $X \rightarrow Y$ and $X \rightarrow Z$, then $X \rightarrow YZ$

Pseudo-transitivity: If $X \rightarrow Y$ and $WY \rightarrow Z$, then $WX \rightarrow Z$

Exercises: Armstrong Rules

- Consider the relation **R(A,B,C,D,E,F)** with the following FDs:

FD1: **A** → **BC**

FD2: **C** → **AD**

FD3: **DE** → **F**

- Use the Armstrong rules to derive the following FD: **AC** → **D**
 - FD4: **AC** → **AD** (Augmentation of FD2 with A)
 - FD5: **AC** → **D** (Decomposition of FD4)
- Use the Armstrong rules to derive the following FD: **A** → **D**
 - FD6: **A** → **C** (Decomposition of FD1)
 - FD7: **A** → **AD** (Transitivity of FD6 and FD2)
 - FD8: **A** → **D** (Decomposition of FD7)

Reflexivity: If Y is a subset of X , then $X \rightarrow Y$

Augmentation: If $X \rightarrow Y$, then $XZ \rightarrow YZ$

Transitivity: If $X \rightarrow Y$ and $Y \rightarrow Z$, then $X \rightarrow Z$

Decomposition: If $X \rightarrow YZ$, then $X \rightarrow Y$

Union: If $X \rightarrow Y$ and $X \rightarrow Z$, then $X \rightarrow YZ$

Pseudo-transitivity: If $X \rightarrow Y$ and $WY \rightarrow Z$, then $WX \rightarrow Z$

Exercises: Armstrong Rules

- Consider the relation **R(A,B,C,D,E,F)** with the following FDs:

FD1: **A** → **BC**

FD2: **C** → **AD**

FD3: **DE** → **F**

- Use the Armstrong rules to derive the following FD: **AE** → **ABCDEF**

Reflexivity: If Y is a subset of X , then $X \rightarrow Y$

Augmentation: If $X \rightarrow Y$, then $XZ \rightarrow YZ$

Transitivity: If $X \rightarrow Y$ and $Y \rightarrow Z$, then $X \rightarrow Z$

Decomposition: If $X \rightarrow YZ$, then $X \rightarrow Y$

Union: If $X \rightarrow Y$ and $X \rightarrow Z$, then $X \rightarrow YZ$

Pseudo-transitivity: If $X \rightarrow Y$ and $WY \rightarrow Z$,
then $WX \rightarrow Z$

Exercises: Armstrong Rules

- Consider the relation **R(A,B,C,D,E,F)** with the following FDs:

FD1: **A** → **BC**

FD2: **C** → **AD**

FD3: **DE** → **F**

- Use the Armstrong rules to derive the following FD: **AE** → **ABCDEF**
 - FD9: **AE** → **BCE** (Augmentation of FD1 with E)
 - FD10: **AE** → **C** (Decomposition of FD9)
 - FD11: **AE** → **AD** (Transitivity of FD10 and FD2)
 - FD12: **AE** → **ADE** (Augmentation of FD11 with E)
 - FD13: **AE** → **DE** (Decomposition of FD12)
 - FD14: **AE** → **F** (Transitivity of FD13 and FD3)
 - FD15: **AE** → **ABCDEF** (Union of FD9, FD11, and FD14)

Reflexivity: If Y is a subset of X , then $X \rightarrow Y$

Augmentation: If $X \rightarrow Y$, then $XZ \rightarrow YZ$

Transitivity: If $X \rightarrow Y$ and $Y \rightarrow Z$, then $X \rightarrow Z$

Decomposition: If $X \rightarrow YZ$, then $X \rightarrow Y$

Union: If $X \rightarrow Y$ and $X \rightarrow Z$, then $X \rightarrow YZ$

Pseudo-transitivity: If $X \rightarrow Y$ and $WY \rightarrow Z$, then $WX \rightarrow Z$

Computing (Super)Keys

```
function ComputeAttrClosure(  $X$ ,  $F$  )  
begin  
     $X^+ := X$ ;  
    while  $F$  contains an FD  $Y \rightarrow Z$  such that  
        (i)  $Y$  is a subset of  $X^+$ , and  
        (ii)  $Z$  is not a subset of  $X^+$  do  
         $X^+ := X^+ \cup Z$ ;  
    end while  
    return  $X^+$ ;  
end
```

Warmup (cont'd)

- Consider the relation **R(A,B,C,D,E,F)** with the following FDs:

FD1: **A** → **BC**

FD2: **C** → **AD**

FD3: **DE** → **F**

- Compute the **attribute closure** of $X = \{ \mathbf{A, E} \}$ w.r.t. $F = \{ \text{FD1, FD2, FD3} \}$ to show that we have: **AE** → **ABCDEF**
 - Initially: $X^+ = \{ \mathbf{A, E} \}$
 - By using FD1: $X^+ = \{ \mathbf{A, E, B, C} \}$
 - By using FD2: $X^+ = \{ \mathbf{A, E, B, C, D} \}$
 - By using FD3: $X^+ = \{ \mathbf{A, E, B, C, D, F} \}$

```
function ComputeAttrClosure( X, F )
begin
    X+ := X;
    while F contains an FD Y → Z such that
        (i) Y is a subset of X+, and
        (ii) Z is not a subset of X+ do
        X+ := X+ ∪ Z;
    end while
    return X+;
end
```

Your Turn

- Consider the relation **R(A,B,C,D)** with the following set F of FDs:

FD1: **AB** $\rightarrow$ **C**

FD2: **BC** $\rightarrow$ **D**

FD3: **D** $\rightarrow$ **B**

- Compute the following attribute closures w.r.t. $F = \{\text{FD1, FD2, FD3}\}$
 - $\{B, C, D\}^+ = ?$
 - $\{A, C, D\}^+ = ?$
 - $\{A, B, D\}^+ = ?$
 - $\{A, B, C\}^+ = ?$

```
function ComputeAttrClosure( X, F )
begin
    X+ := X;
    while F contains an FD Y  $\rightarrow$  Z such that
        (i) Y is a subset of X+, and
        (ii) Z is not a subset of X+ do
        X+ := X+  $\cup$  Z;
    end while
    return X+;
end
```

Your Turn

- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:

FD1: **AB** → **C**

FD2: **BC** → **D**

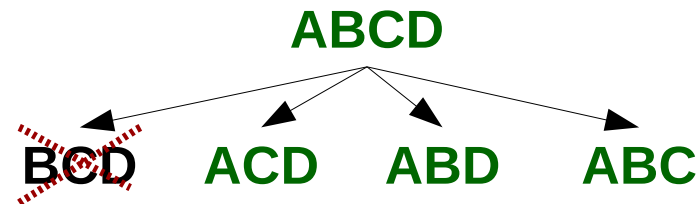
FD3: **D** → **B**

- Compute the following attribute closures w.r.t. $F = \{FD1, FD2, FD3\}$
 - $\{B, C, D\}^+ = \{B, C, D\}$
 - $\{A, C, D\}^+ = \{A, C, D, B\}$
 - $\{A, B, D\}^+ = \{A, B, D, C\}$
 - $\{A, B, C\}^+ = \{A, B, C, D\}$

```
function ComputeAttrClosure( X, F )
begin
    X+ := X;
    while F contains an FD Y → Z such that
        (i) Y is a subset of X+, and
        (ii) Z is not a subset of X+ do
        X+ := X+ ∪ Z;
    end while
    return X+;
end
```

Superkeys and Candidate Keys

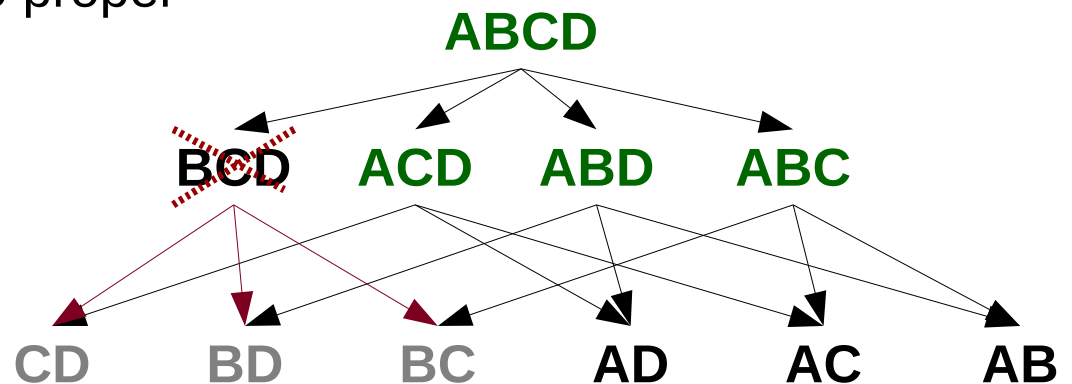
- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:
FD1: **AB** → **C**
FD2: **BC** → **D**
FD3: **D** → **B**
- A set *X* of attributes of *R* is a **superkey** if X^+ contains all the attributes of *R*
 - $\{B,C,D\}^+ = \{B,C,D\}$
 - $\{A,C,D\}^+ = \{A,C,D,B\}$
 - $\{A,B,D\}^+ = \{A,B,D,C\}$
 - $\{A,B,C\}^+ = \{A,B,C,D\}$



Superkeys and Candidate Keys

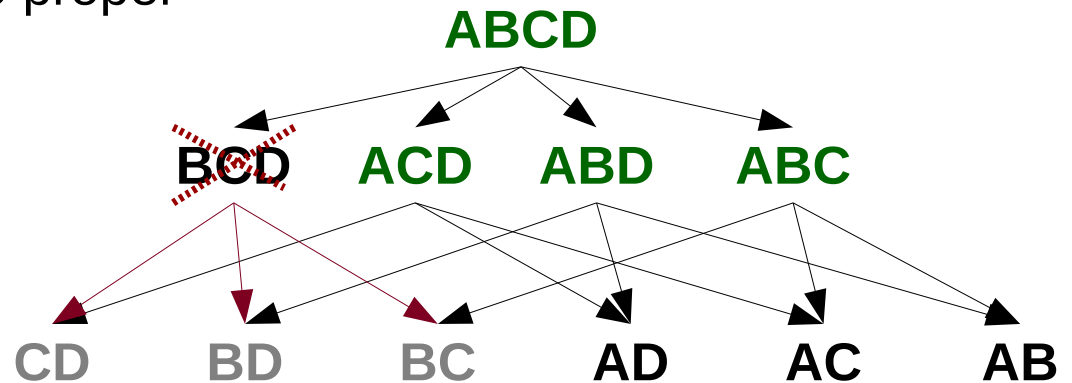
- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:
FD1: **AB** → **C**
FD2: **BC** → **D**
FD3: **D** → **B**
- A set *X* of attributes of *R* is a **superkey** if X^+ contains all the attributes of *R*
- X* is a **candidate key** (CK) if no proper subset *X'* of *X* is a **superkey**

- $\{A,D\}^+ = ?$
- $\{A,C\}^+ = ?$
- $\{A,B\}^+ = ?$



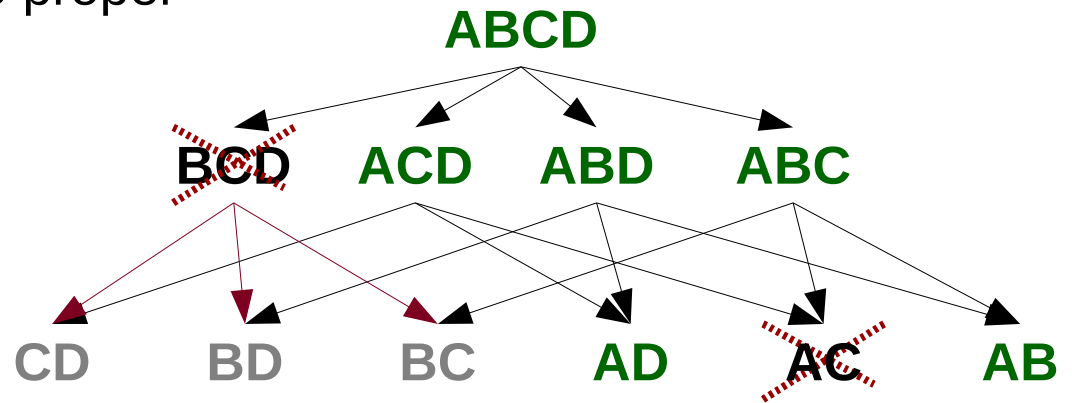
Superkeys and Candidate Keys

- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:
FD1: **AB** → **C**
FD2: **BC** → **D**
FD3: **D** → **B**
- A set *X* of attributes of *R* is a **superkey** if X^+ contains all the attributes of *R*
- X* is a **candidate key** (CK) if no proper subset *X'* of *X* is a **superkey**
 - $\{A,D\}^+ = \{A,D,B,C\}$
 - $\{A,C\}^+ = \{A,C\}$
 - $\{A,B\}^+ = \{A,B,C,D\}$



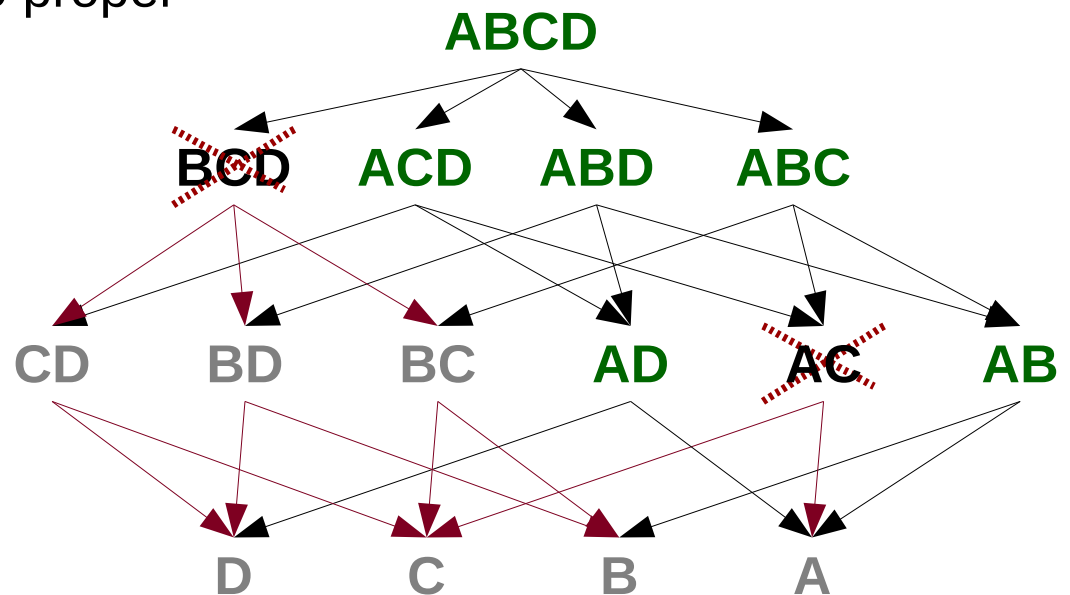
Superkeys and Candidate Keys

- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:
FD1: **AB** → **C**
FD2: **BC** → **D**
FD3: **D** → **B**
- A set *X* of attributes of *R* is a **superkey** if X^+ contains all the attributes of *R*
- X* is a **candidate key** (CK) if no proper subset *X'* of *X* is a **superkey**
 - $\{A,D\}^+ = \{A,D,B,C\}$
 - $\{A,C\}^+ = \{A,C\}$
 - $\{A,B\}^+ = \{A,B,C,D\}$



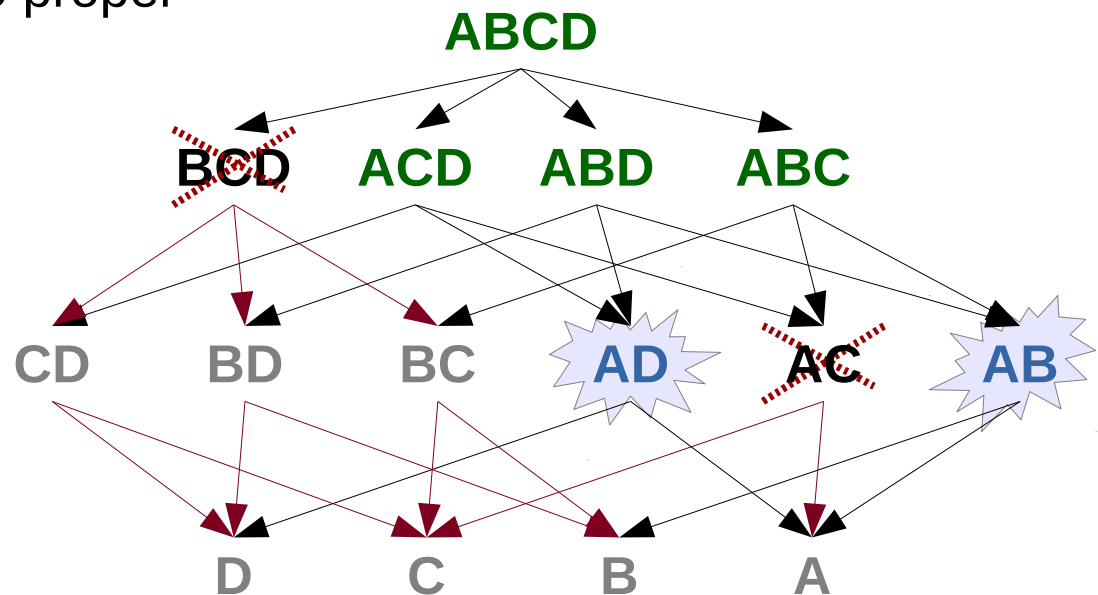
Superkeys and Candidate Keys

- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:
FD1: **AB** → **C**
FD2: **BC** → **D**
FD3: **D** → **B**
- A set *X* of attributes of *R* is a **superkey** if X^+ contains all the attributes of *R*
- X* is a **candidate key** (CK) if no proper subset *X'* of *X* is a **superkey**



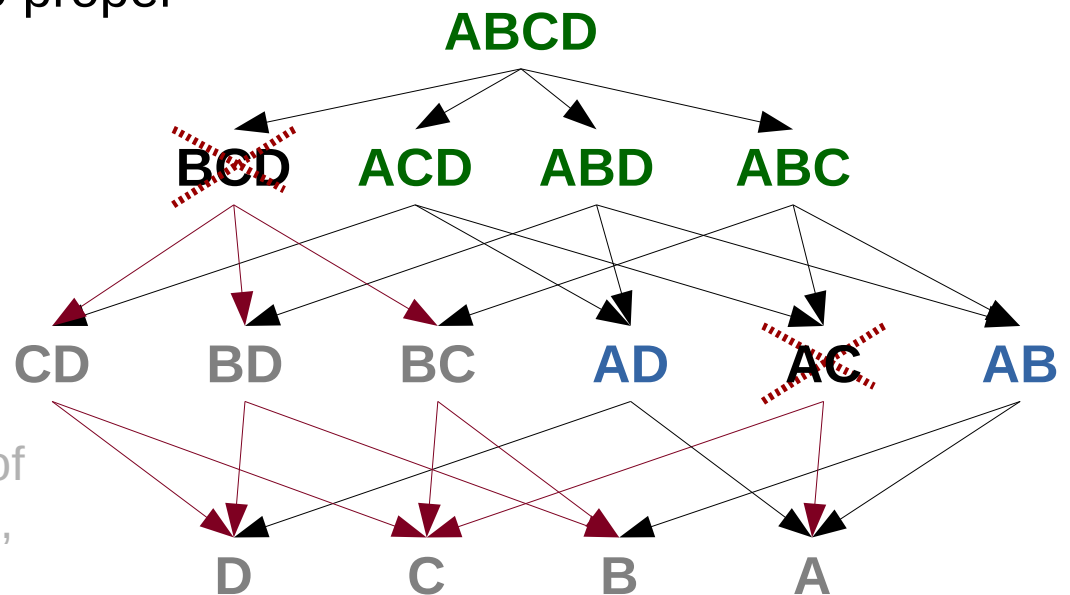
Superkeys and Candidate Keys

- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:
FD1: **AB** → **C**
FD2: **BC** → **D**
FD3: **D** → **B**
- A set *X* of attributes of *R* is a **superkey** if X^+ contains all the attributes of *R*
- X* is a **candidate key** (CK) if no proper subset *X'* of *X* is a **superkey**



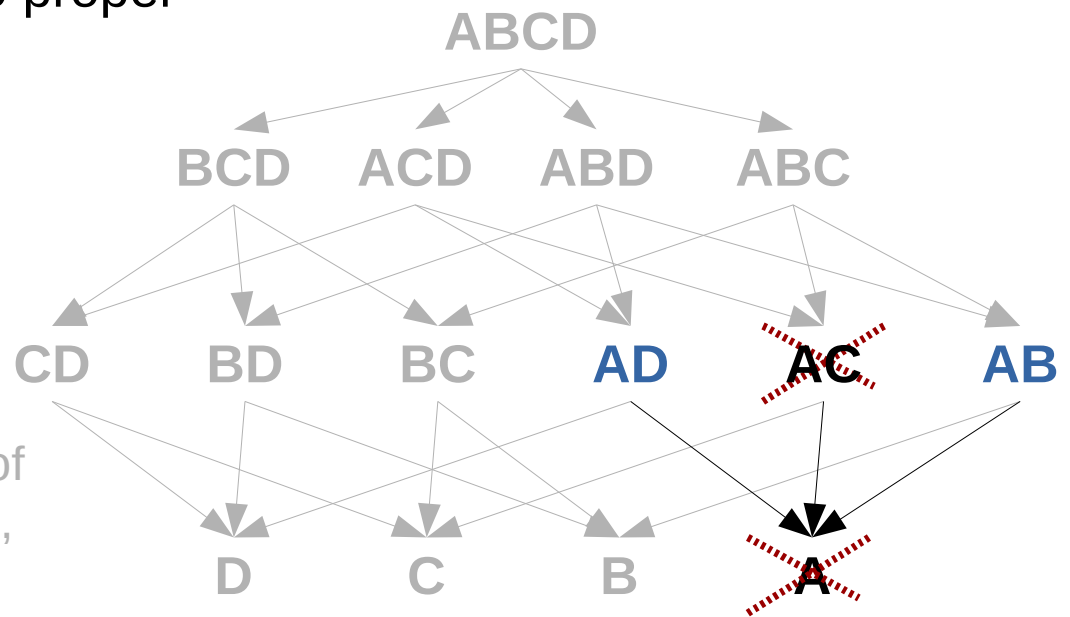
Superkeys and Candidate Keys

- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:
 - FD1: **AB** → **C**
 - FD2: **BC** → **D**
 - FD3: **D** → **B**
- A set *X* of attributes of *R* is a **superkey** if X^+ contains all the attributes of *R*
- X* is a **candidate key** (CK) if no proper subset *X'* of *X* is a **superkey**
- Some steps may be skipped
 - If an attribute is nowhere in the RHS, it must be part of every candidate key (such as *A* in the example)
 - If an attribute is in the RHS of FDs but nowhere in the LHS, it cannot be part of any CK



Superkeys and Candidate Keys

- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:
 - FD1: **AB** → **C**
 - FD2: **BC** → **D**
 - FD3: **D** → **B**
- A set *X* of attributes of *R* is a **superkey** if X^+ contains all the attributes of *R*
- X* is a **candidate key** (CK) if no proper subset *X'* of *X* is a **superkey**
- Some steps may be skipped
 - If an attribute is nowhere in the RHS, it must be part of every candidate key (such as *A* in the example)
 - If an attribute is in the RHS of FDs but nowhere in the LHS, it cannot be part of any CK



Superkeys and Candidate Keys

- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:
 - FD1: **AB** → **C**
 - FD2: **BC** → **D**
 - ~~FD3: **D** → **B**~~
- A set *X* of attributes of *R* is a **superkey** if X^+ contains all the attributes of *R*
- X* is a **candidate key** (CK) if no proper subset *X'* of *X* is a **superkey**
- Some steps may be skipped
 - If an attribute is nowhere in the RHS, it must be part of *every* candidate key (such as *A* in the example)
 - If an attribute is in the RHS of FDs but nowhere in the LHS, it cannot be part of any CK
- For instance, if we had only FD1 and FD2 (i.e., not FD3), we could immediately say that:
 - A* and *B* must be in every CK
 - and, since $\{A,B\}^+ = \{A,B,C,D\}$, $\{A,B\}$ is a superkey and thus the only CK

Your Turn

- Consider the relation **R(A,B,C,D,E,F)** with the following FDs:
FD1: **A** → **BC**
FD2: **C** → **AD**
FD3: **DE** → **F**
- What are the candidate keys?

Your Turn

- Consider the relation **R(A,B,C,D,E,F)** with the following FDs:

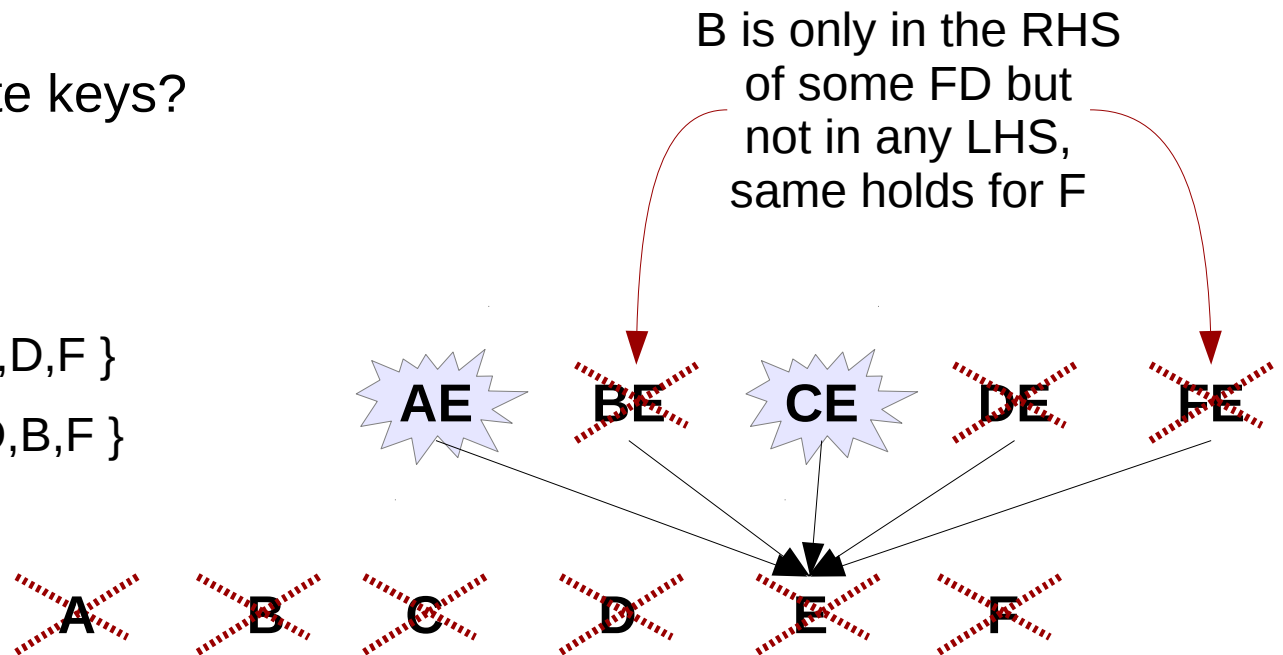
FD1: $A \rightarrow BC$

FD2: $C \rightarrow AD$

FD3: DE $\rightarrow$ F

- What are the candidate keys?

- $\{A,E\}$ and $\{C,E\}$
- $\{E\}^+ = \{E\}$
- $\{A,E\}^+ = \{A,E,B,C,D,F\}$
- $\{C,E\}^+ = \{C,E,A,D,B,F\}$
- $\{D,E\}^+ = \{D,E,F\}$



www.liu.se