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 $A_1, A_2, \ldots, A_n$ and let *X* and *Y* be subsets of $\{A_1, A_2, \ldots, A_n\}$.
>
> 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: $\mathbf{A} \rightarrow \mathbf{BC}$ FD2: $\mathbf{C} \rightarrow \mathbf{AD}$ FD3: $\mathbf{DE} \rightarrow \mathbf{F}$

- Is a state of R that contains the two tuples (3, 8, 1, 2, 3, 4) and (3, 8, 1, 7, 3, 9) valid?

  A) Yes B) No

A	B	C	D	E	F
3	8	1	2	3	4
3	8	1	7	3	9

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

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 $WX \rightarrow Z$

- 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**

Exercises: Armstrong Rules

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 $WX \rightarrow Z$

- 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)

Exercises: Armstrong Rules

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 $WX \rightarrow Z$

- 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**

Exercises: Armstrong Rules

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 $WX \rightarrow Z$

- 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)

Exercises: Armstrong Rules

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 $WX \rightarrow Z$

- 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**

Exercises: Armstrong Rules

Reflexivity: If *Y* is a subset of *X*, then *X* → *Y*

Augmentation: If *X* → *Y*, then *XZ* → *YZ*

Transitivity: If *X* → *Y* and *Y* → *Z*, then *X* → *Z*

Decomposition: If *X* → *YZ*, then *X* → *Y*

Union: If *X* → *Y* and *X* → *Z*, then *X* → *YZ*

Pseudo-transitivity: If *X* → *Y* and *WY* → *Z*, then *WX* → *Z*

- 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)

Exercise

- 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**
- Consider the following state of relation R:

A	B	C	D	E	F
3	8	1	2	3	4

- Add another tuple to this state such that the resulting extended state is valid with respect to the new FD (i.e., does not violate the FD)

Exercise

- 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**
- Consider the following state of relation R:

A	B	C	D	E	F
3	8	1	2	3	4
4	..	..	..	3	..
3	..	..	..	3	..

- Add another tuple to this state such that the resulting extended state is valid with respect to the new FD (i.e., does not violate the FD)
  - any tuple in which the value for A or E (or both) is different from the corresponding values of the given tuple

Computing (Super)Keys

```
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+ U Z;
    end while
    return X+;
end
```

Warmup (cont'd)

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

- 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 = { **A,E** } w.r.t. F = {FD1, FD2, FD3} to show that we have: **AE** → **ABCDEF**
  - Initially: X^+ = { **A,E** }
  - By using FD1: X^+ = { **A,E,B,C** }
  - By using FD2: X^+ = { **A,E,B,C,D** }
  - By using FD3: X^+ = { **A,E,B,C,D,F** }

Your Turn

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

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

Your Turn

```
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+ U Z;
    end while
    return X+;
end
```

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

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**

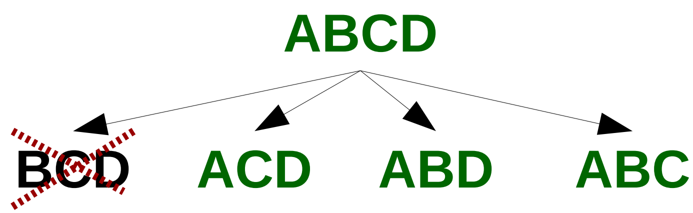
- A set *X* of attributes of *R* is a superkey if X^+ contains all the attributes of *R*

  – $\{ B,C,D \}^+ = \{ B,C,D \}$

  – $\{ A,C,D \}^+ = \{ A,C,D,B \}$

  – $\{ A,B,D \}^+ = \{ A,B,D,C \}$

  – $\{ A,B,C \}^+ = \{ A,B,C,D \}$

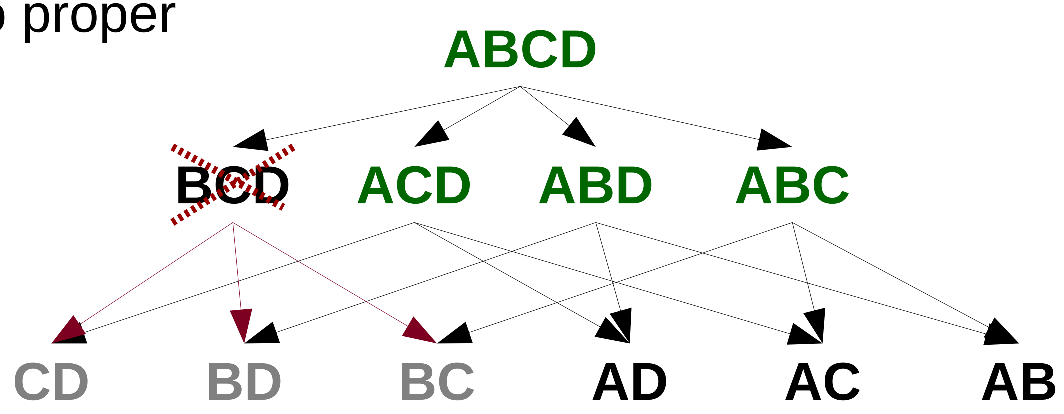


Superkeys and Candidate Keys

- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:
  FD1: **AB** → **C**
  FD2: **BC** → **D**
  FD3: **D** → **B**

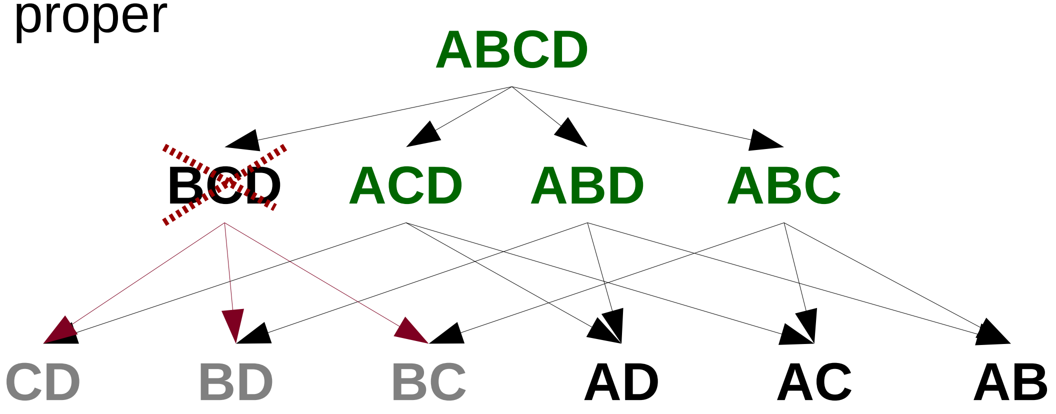
- A set *X* of attributes of *R* is a superkey if X^+ contains all the attributes of *R*

- *X* is a candidate key (CK) if no proper subset *X'* of *X* is a superkey
  - $\{A,D\}^+ =$?
  - $\{A,C\}^+ =$?
  - $\{A,B\}^+ =$?



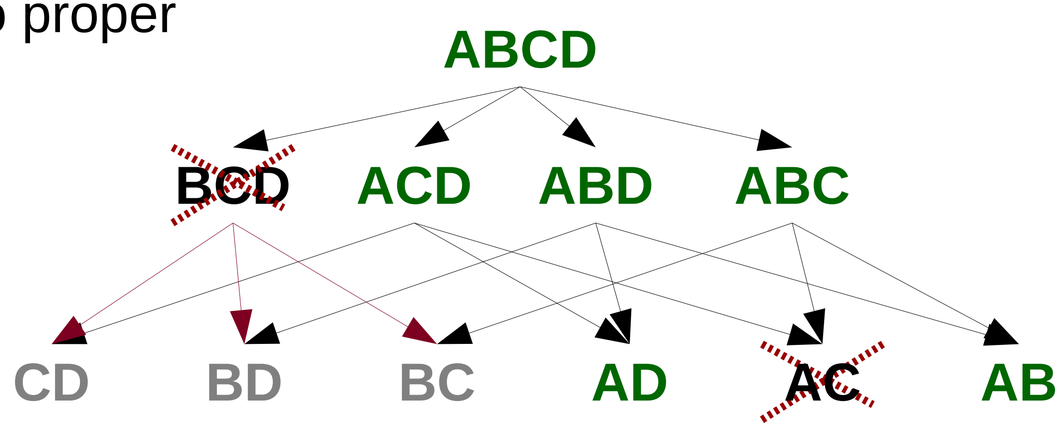
Superkeys and Candidate Keys

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



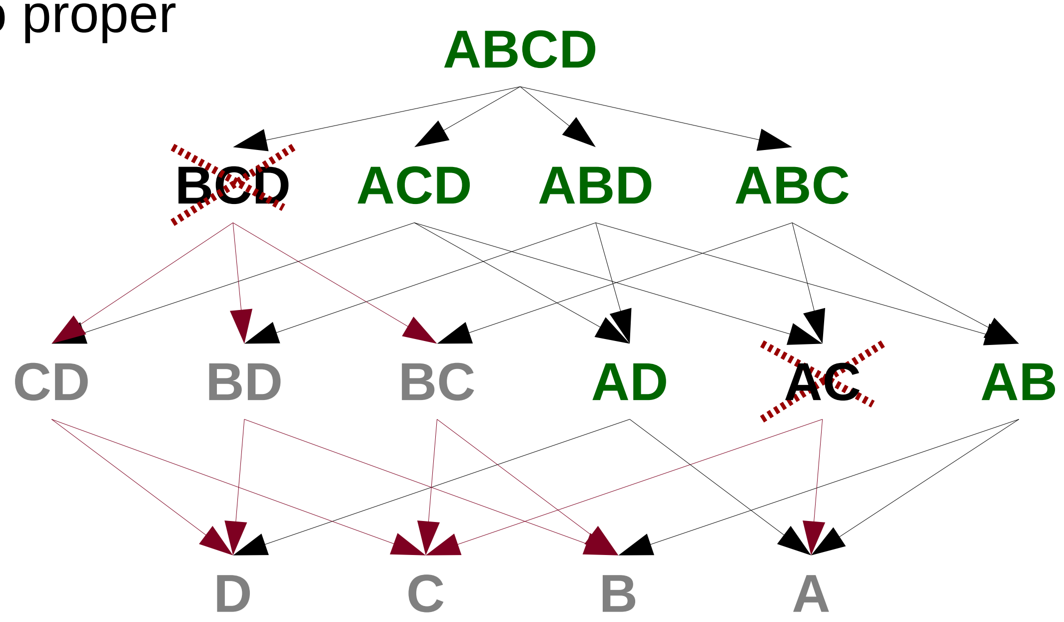
Superkeys and Candidate Keys

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



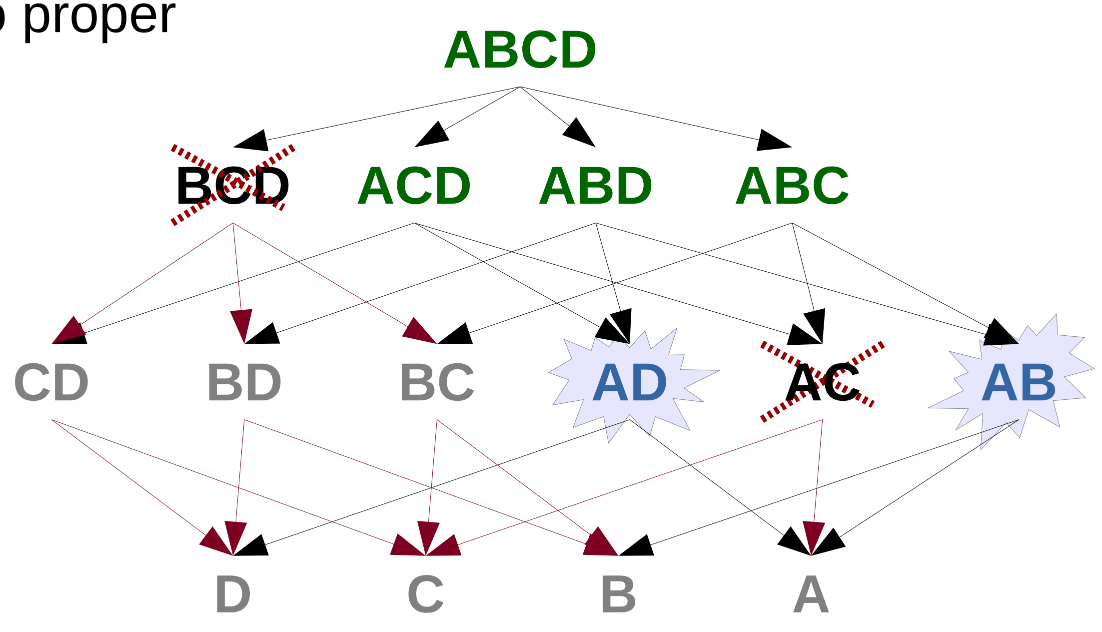
Superkeys and Candidate Keys

- 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



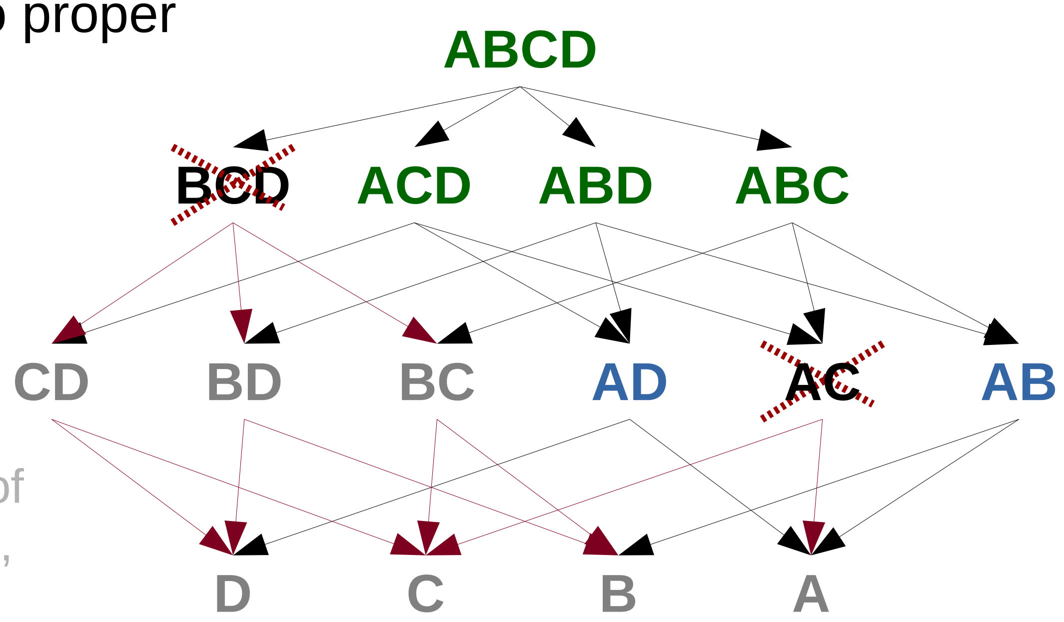
Superkeys and Candidate Keys

- 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



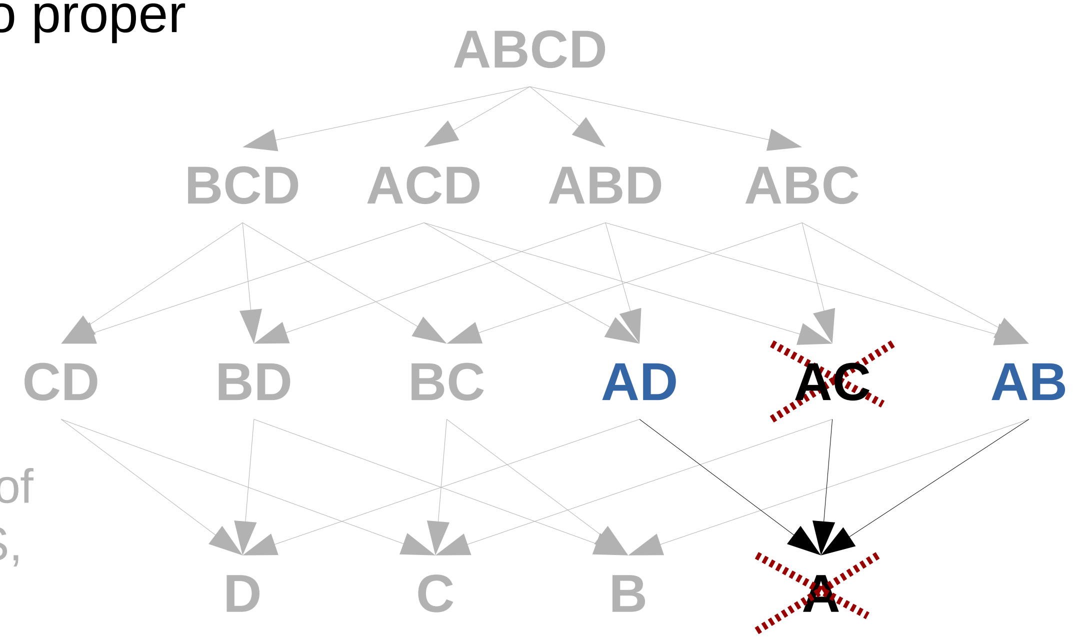
Superkeys and Candidate Keys

- 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:
  - *D* cannot be in any CK
  - *A* and *B* must be in every CK
  - and, since $\{A,B\}^+ = \{A,B,C,D\}$, we see that {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: $\mathbf{A} \rightarrow \mathbf{BC}$

  FD2: $\mathbf{C} \rightarrow \mathbf{AD}$

  FD3: $\mathbf{DE} \rightarrow \mathbf{F}$

- What are the candidate keys?

Your Turn

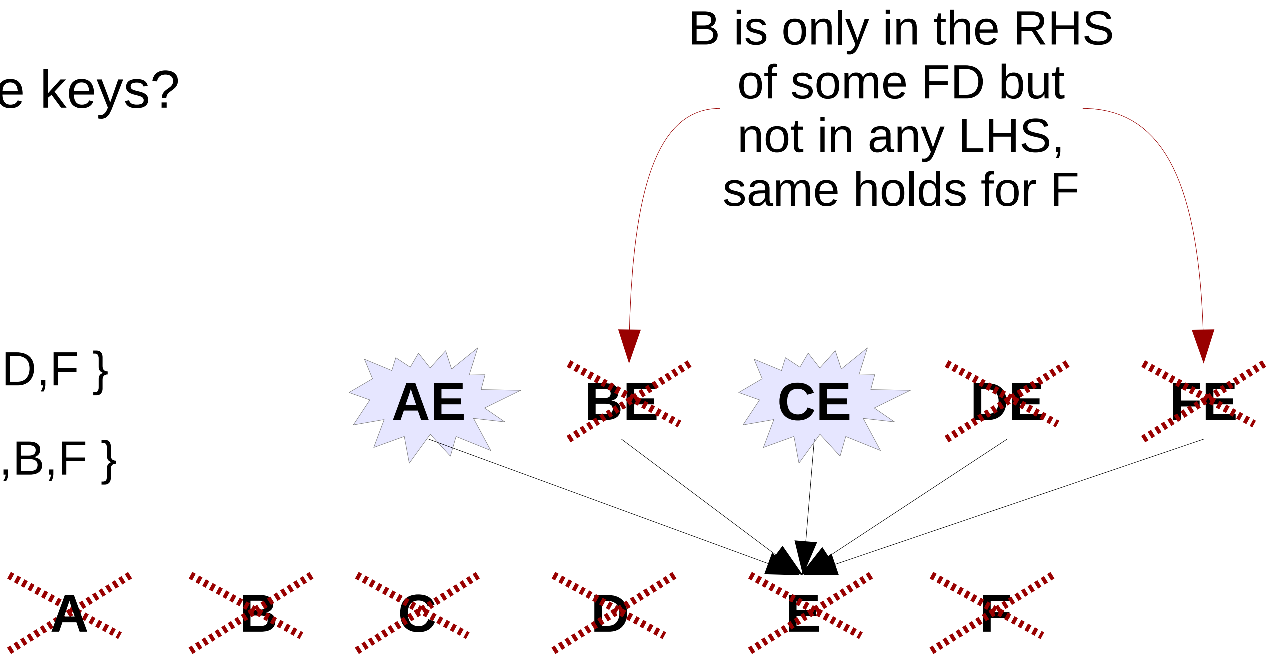
- Consider the relation **R(A,B,C,D,E,F)** with the following FDs:

  FD1: **A** → **BC**

  FD2: **C** → **AD**

  FD3: **DE** → **F**

- What are the candidate keys?
  - {A,E} and {C,E}
  - $\{E\}^{+} = \{E\}$
  - $\{A,E\}^{+} = \{A,E,B,C,D,F\}$
  - $\{C,E\}^{+} = \{C,E,A,D,B,F\}$
  - $\{D,E\}^{+} = \{D,E,F\}$

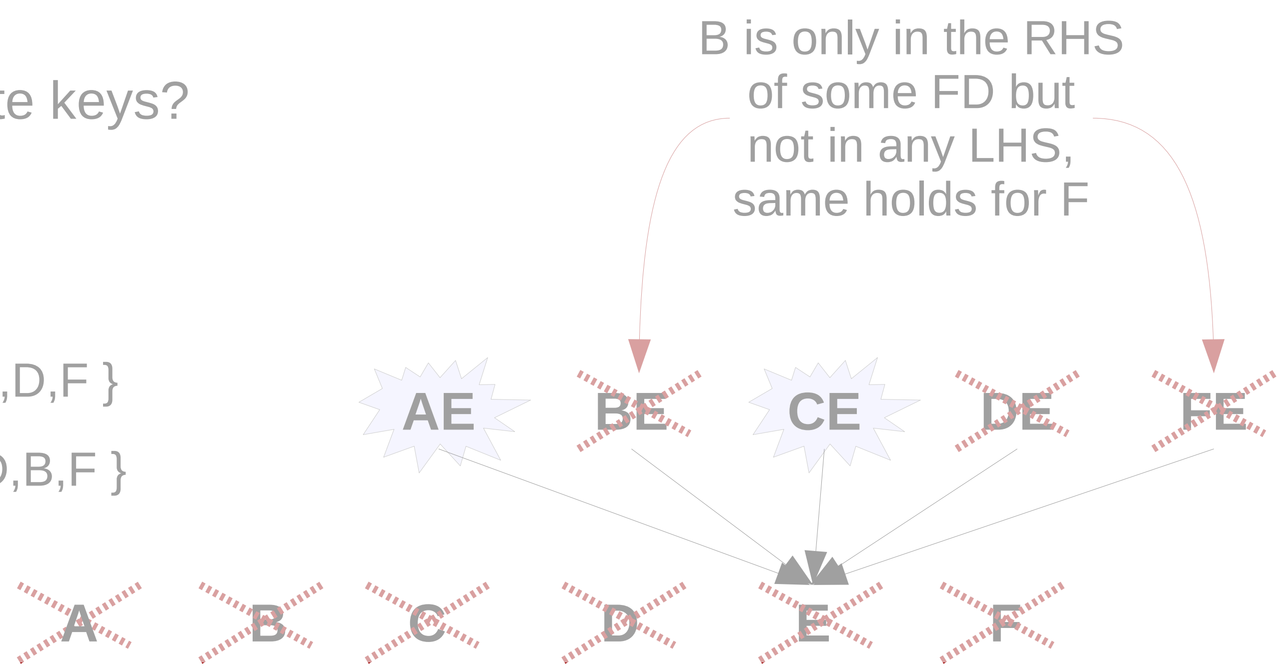


Boyce-Codd Normal Form (BCNF)

- Relation schema *R* with a set *F* of functional dependencies is in BCNF if for **every** non-trivial **FD $X \rightarrow Y$ in F^+** we have that **X is a superkey**

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



BCNF Example

- Relation schema *R* with a set *F* of functional dependencies is in BCNF if for **every** non-trivial **FD** $X \rightarrow Y$ **in** F^+ we have that ***X* is a superkey**
- Consider the relation **R(A,B,C,D,E,F)** with the following FDs:

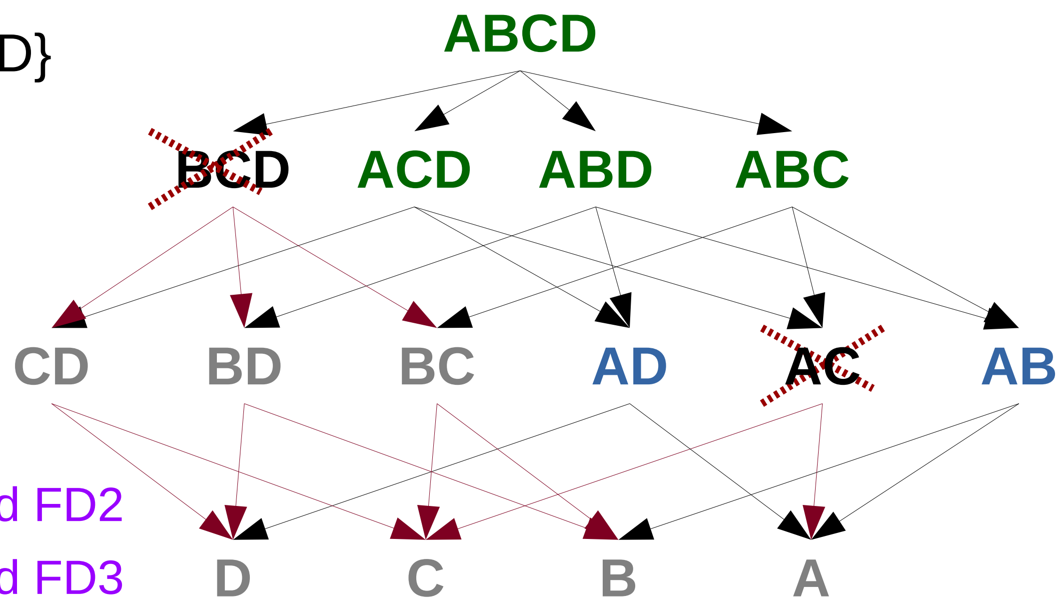
  FD1: **A** → **BC**

  FD2: **C** → **AD**

  FD3: **DE** → **F**
- What are the candidate keys?
  - {A,E} and {C,E}
- Is the given relation in BCNF?
  - If not, identify the FDs that violate the BCNF condition.

Your Turn

- Relation schema *R* with a set *F* of functional dependencies is in BCNF if for **every** non-trivial **FD *X*→*Y* in *F*⁺** we have that ***X* is a superkey**
- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:
  FD1: **AB → C**
  FD2: **BC → D**
  FD3: **D → B**
- Candidate keys: {A,B} and {A,D}
- Is the given relation in BCNF? If not, identify the FDs that violate the BCNF condition.
  1) Yes, BCNF
  2) Not BCNF, because of FD1
  3) Not BCNF, because of FD1 and FD2
  4) Not BCNF, because of FD2 and FD3



BCNF Decomposition Step

- By using an FD $X \rightarrow Y$ that violates BCNF, decompose R into
  - $R1$ with all the attributes in X and in Y, and
  - $R2$ with all attributes from R except those that are in Y and not in X
- Consider the relation **R(A,B,C,D)** with the following set F of FDs:

  FD1: **AB** → **C**

  FD2: **BC** → **D**

  FD3: **D** → **B**
- Let's decompose based on FD2

BCNF Decomposition Step

- By using an FD $X \rightarrow Y$ that violates BCNF, decompose *R* into
  - *R1* with all the attributes in *X* and in *Y*, and
  - *R2* with all attributes from *R* except those that are in *Y* and not in *X*
- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:

  FD1: **AB** → **C**

  FD2: **BC** → **D**

  FD3: **D** → **B**
- Let's decompose based on FD2
  - R1(B,C,D)
  - R2(A,B,C)

BCNF Decomposition Step

- By using an FD $X \rightarrow Y$ that violates BCNF, decompose *R* into
  - *R1* with all the attributes in *X* and in *Y*, and
  - *R2* with all attributes from *R* except those that are in *Y* and not in *X*
- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:

  FD1: **AB** → **C**

  FD2: **BC** → **D**

  FD3: **D** → **B**

- Let's decompose based on FD2 – and don't forget to determine the FDs of the resulting relation schemas
  - R1(B,C,D)
  - R2(A,B,C)

BCNF Decomposition Step

- By using an FD $X \rightarrow Y$ that violates BCNF, decompose R into
  - $R1$ with all the attributes in X and in Y, and
  - $R2$ with all attributes from R except those that are in Y and not in X
- Consider the relation **R(A,B,C,D)** with the following set F of FDs:
  FD1: **AB** → **C**
  FD2: **BC** → **D**
  FD3: **D** → **B**
- Let's decompose based on FD2 – and don't forget to determine the FDs of the resulting relation schemas
  - R1(B,C,D) with FDs: FD2 and FD3
  - R2(A,B,C) with FDs: FD1

BCNF Decomposition Step

- By using an FD *X* → *Y* that violates BCNF, decompose *R* into
  - *R1* with all the attributes in *X* and in *Y*, and
  - *R2* with all attributes from *R* except those that are in *Y* and not in *X*
- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:

  FD1: **AB** → **C**

  FD2: **BC** → **D**

  FD3: **D** → **B**
- Let's decompose based on FD2 – and don't forget to determine the FDs of the resulting relation schemas, and the CKs
  - R1(B,C,D) with FDs: FD2 and FD3
  - R2(A,B,C) with FDs: FD1

BCNF Decomposition Step

- By using an FD $X \rightarrow Y$ that violates BCNF, decompose R into
  - *R1* with all the attributes in *X* and in *Y*, and
  - *R2* with all attributes from *R* except those that are in *Y* and not in *X*
- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:
  FD1: **AB** → **C**
  FD2: **BC** → **D**
  FD3: **D** → **B**
- Let's decompose based on FD2 – and don't forget to determine the FDs of the resulting relation schemas, and the CKs
  - R1(B,C,D) with FDs: FD2 and FD3, CKs: {B,C}, {C,D}
  - R2(A,B,C) with FDs: FD1, CK: {A,B}

BCNF Decomposition Step

- By using an FD $X \rightarrow Y$ that violates BCNF, decompose R into
  - $R1$ with all the attributes in X and in Y, and
  - $R2$ with all attributes from R except those that are in Y and not in X
- Consider the relation **R(A,B,C,D)** with the following set F of FDs:
  FD1: **AB** → **C**
  FD2: **BC** → **D**
  FD3: **D** → **B**
- Let's decompose based on FD2 – and don't forget to determine the FDs of the resulting relation schemas, and the CKs
  - R1(B,C,D) with FDs: FD2 and FD3, CKs: {B,C}, {C,D}
  - R2(A,B,C) with FDs: FD1, CK: {A,B}
  - Are they in BCNF?

BCNF Decomposition Step

- By using an FD *X* → *Y* that violates BCNF, decompose *R* into
  - *R1* with all the attributes in *X* and in *Y*, and
  - *R2* with all attributes from *R* except those that are in *Y* and not in *X*
- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:

  FD1: **AB** → **C**

  FD2: **BC** → **D**

  FD3: **D** → **B**
- Let's decompose based on FD2 – and don't forget to determine the FDs of the resulting relation schemas, and the CKs
  - R1(B,C,D) with FDs: FD2 and FD3, CKs: {B,C}, {C,D}
  - R2(A,B,C) with FDs: FD1, CK: {A,B}
  - Are they in BCNF? No, R1 is not because of FD3 (but R2 is)

BCNF Decomposition Step

- By using an FD $X \rightarrow Y$ that violates BCNF, decompose R into
  - $R1$ with all the attributes in X and in Y, and
  - $R2$ with all attributes from R except those that are in Y and not in X
- Consider the relation **R(A,B,C,D)** with the following set F of FDs:
  FD1: **AB** → **C**
  FD2: **BC** → **D**
  FD3: **D** → **B**
- Let's decompose based on FD2 – and don't forget to determine the FDs of the resulting relation schemas, and the CKs
  - R1(B,C,D) with FDs: FD2 and FD3, CKs: {B,C}, {C,D}
  - R2(A,B,C) with FDs: FD1, CK: {A,B}
  - Your turn: decompose R1 based on FD3 (and don't forget ...)

BCNF Decomposition Step

- By using an FD $X \rightarrow Y$ that violates BCNF, decompose *R* into
  - *R1* with all the attributes in *X* and in *Y*, and
  - *R2* with all attributes from *R* except those that are in *Y* and not in *X*
- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:

  FD1: **AB** → **C**

  FD2: **BC** → **D**

  FD3: **D** → **B**
- Let's decompose based on FD2 – and don't forget to determine the FDs of the resulting relation schemas, and the CKs
  - ~~R1(B,C,D) with FDs: FD2 and FD3, CKs: {B,C}, {C,D}~~
  - R2(A,B,C) with FDs: FD1, CK: {A,B}
  - R3(D,B) with FD3, CK {D} – R4(C,D) only trivial FDs, CK: {C,D}

Your Turn

- By using an FD $X \rightarrow Y$ that violates BCNF, decompose R into
  - $R1$ with all the attributes in X and in Y, and
  - $R2$ with all attributes from R except those that are in Y and not in X
- Consider the relation **R(A,B,C,D)** with the following set F of FDs:

  FD1: **AB** → **C**

  FD2: **BC** → **D**

  FD3: **D** → **B**

- Let's decompose based on FD3 – and don't forget to determine the FDs of the resulting relation schemas, and the CKs

Your Turn

- By using an FD *X* → *Y* that violates BCNF, decompose *R* into
  - *R1* with all the attributes in *X* and in *Y*, and
  - *R2* with all attributes from *R* except those that are in *Y* and not in *X*
- Consider the relation **R(A,B,C,D)** with the following set *F* of FDs:

  FD1: **AB** → **C**

  FD2: **BC** → **D**

  FD3: **D** → **B**

- Let's decompose based on FD3 – and don't forget to determine the FDs of the resulting relation schemas, and the CKs
  - R1(B,D) with FD3, CK: {D}
  - R2(A,C,D) with FD4: AD → C, CK: {A,D}

  R1 and R2 are in BCNF

can be derived from FD3 and FD1 by using the augmentation rule and the transitivity rule

Different Example

- By using an FD $X \rightarrow Y$ that violates BCNF, decompose *R* into
  - *R1* with all the attributes in *X* and in *Y*, and
  - *R2* with all attributes from *R* except those that are in *Y* and not in *X*
- Consider the relation schema **R(A,B,C,D,E,F)** with the following FDs:
  FD1: **AB** → **CDEF**
  FD2: **E** → **F**
- Your turn:
  - Determine candidate key(s)
  - Is *R* in BCNF?
  - If not, normalize into a set of BCNF relation schemas
    (and don't forget to determine FDs and CKs along the way)

Different Example

- By using an FD $X \rightarrow Y$ that violates BCNF, decompose R into
  - *R1* with all the attributes in *X* and in *Y*, and
  - *R2* with all attributes from *R* except those that are in *Y* and not in *X*
- Consider the relation schema **R(A,B,C,D,E,F)** with the following FDs:

  FD1: **AB** → **CDEF**

  FD2: **E** → **F**
- Your turn:
  - Determine candidate key(s) *{A,B}*
  - Is *R* in BCNF? *No, FD2 violates the BCNF condition.*
  - If not, normalize into a set of BCNF relation schemas

    We decompose R based on FD2:

    - R1(E,F) with FD2; candidate key is {E}

    - R2(A,B,C,D,E) with a new FD: AB→CDE; candidate key is {A,B}

    R1 and R2 are in BCNF

can be derived from FD1 by using the decomposition rule

One More

- By using an FD $X \rightarrow Y$ that violates BCNF, decompose *R* into
  - *R1* with all the attributes in *X* and in *Y*, and
  - *R2* with all attributes from *R* except those that are in *Y* and not in *X*
- Consider the relation schema **R(A,B,C,D,E,F)** with the following FDs:
  FD1: **AB** → **CDEF**
  FD2: **E** → **F**
  FD3: **A** → **D**
- Your turn:
  - Determine candidate key(s)
  - Is *R* in BCNF?
  - If not, normalize into a set of BCNF relation schemas
    (and don't forget to determine FDs and CKs along the way)

One More

- By using an FD *X* → *Y* that violates BCNF, decompose *R* into
  - *R1* with all the attributes in *X* and in *Y*, and
  - *R2* with all attributes from *R* except those that are in *Y* and not in *X*
- Consider the relation schema **R(A,B,C,D,E,F)** with the following FDs:

  FD1: **AB** → **CDEF**

  FD2: **E** → **F**

  FD3: **A** → **D**

- Solution: *CK is {A,B}; R is not in BCNF because of FD2 and FD3.*

  We decompose R based on FD2:

  - R1(E,F) with FD2; candidate key is {E}

  - R2(A,B,C,D,E) with FD3 and a new FD: AB→CDE; candidate key is {A,B}

  R1 is in BCNF, but R2 is not because of FD3. So, we have to decompose R2 using FD3:

  - R3(A,D) with FD3; candidate key is {A}

  - R4(A,B,C,E) with new FD: AB→CE; candidate key is {AB}

  R3 and R4 are in BCNF. Hence, the result of normalizing R consists of R1, R3, and R4.

Back to the Earlier Running Example

- By using an FD $X \rightarrow Y$ that violates BCNF, decompose R into
  - $R1$ with all the attributes in X and in Y, and
  - $R2$ with all attributes from R except those that are in Y and not in X
- Consider the relation schema **R(A,B,C,D,E,F)** with the following FDs:
  FD1: **A → BC** FD2: **C → AD** FD3: **DE → F**
- Recall: CKs are {A,E}, {C,E}; all three FDs violate the BCNF condition
- Your turn: Normalize into a set of BCNF relation schemas
  (and don't forget to determine FDs and CKs along the way)

Back to the Earlier Running Example

- By using an FD $X \rightarrow Y$ that violates BCNF, decompose R into
  - *R1* with all the attributes in X and in Y, and
  - *R2* with all attributes from R except those that are in Y and not in X
- Consider the relation schema **R(A,B,C,D,E,F)** with the following FDs:
  FD1: **A → BC** FD2: **C → AD** FD3: **DE → F**
- Recall: CKs are {A,E}, {C,E}; all three FDs violate the BCNF condition
- A possible solution: *We decompose R based on FD1:*
  - R1(A,B,C) with FD1 and a new FD: C→A; candidate keys are {A} and {C}
  - R2(A,D,E,F) with FD3 and a new FD: A→D; candidate key is {A,E}
  R1 is in BCNF, but R2 is not because of FD3 and A→D. Let's decompose R2 based on FD3:
  - R3(D,E,F) with FD3; candidate key is {D,E}
  - R4(A,D,E) with A→D; candidate key is {A,E}
  R3 is in BCNF, but R4 is not because of A→D. Let's decompose R4 based on A→D
  - R5(A,D) with A→D; candidate key is {A}
  - R6(A,E) with only trivial FDs; candidate key is {A,E}
  R5 and R6 are in BCNF. Hence, the result of the normalization consists of R1, R3, R5, R6.

www.liu.se

LINKÖPING UNIVERSITY