Imperial College London

Functional dependencies

Tutorial II - Normalisations

Valentin CLÉMENT - vec16@ic.ac.uk

Department of Computing

Reminders 1/2

A ***Decomposition*** of a relation $R(A, B, C, \dots)$ means finding a set of relations $R_1(\dots)$, $R_2(\dots)$, ... such that $\cup_i \mathrm{attr}(R_i) = \mathrm{attr}(R)$. It is interesting if it is:

Lossless The original relation can be retrieved by joining the decomposed relations, i.e. if we decompose $(R, \mathcal{F})$ into S and T:

$$\text{We need either} \begin{cases} \mathrm{attr}(S) \cap \mathrm{attr}(T) \to_{\mathcal{F}} \mathrm{attr}(S) \\ \mathrm{attr}(S) \cap \mathrm{attr}(T) \to_{\mathcal{F}} \mathrm{attr}(T) \end{cases}$$

Dependency preserving If we can check the functional dependencies $\mathcal{F}$ of R ***without joining*** S and T, the decomposition is *dependency preserving*

Note that lossless decomposition is much more important in practice: dependency preservation is more of a 'nice to have'. Decomposition can help reduce ***anomalies***.

Reminders (2/2)

A useful way of classifying decomposition is on properties they should have:

BCNF The Boyce-Codd Normal Form requires that for every non trivial FD, ***the LHS is a superkey***. We can decompose *any* relation into BCNF by ***splitting it on each offending FD***. It is lossless, but *not necessarily* dependency preserving

3NF The Third Normal Form: a slightly weaker form, where a FD is valid if its LHS is a superkey OR ***every attribute on its RHS is prime*** (i.e. it is part of at least one candidate key). We can obtain 3NF by adding every non-redundant relationship generated by a ***canonical cover***, plus every ***missing candidate key***.

We will see the details in the examples

Example 1
3NF decomposition

Decompose relation $R(A, B, C, D, E)$ in ***3NF***, given the following functional dependencies:

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

Reminder: One finds the 3NF decomposition by starting from an empty set of relations, and:

- For each FD $LHS \rightarrow RHS$ in a canonical cover, add relation $LHS \cup RHS$, if it is not a subset of an existing one
- If none of the resulting relations include a key, add a new relation for that key.

Solution 1

3NF decomposition

FDs $BC \rightarrow D$, $D \rightarrow E$, $A \rightarrow C$, $E \rightarrow B$

- Canonical cover: everything is minimal.
  $\mathcal{C} = \{A \rightarrow C, BC \rightarrow D, D \rightarrow E, E \rightarrow B\}$
- Starting from $\{\}$, for $a \rightarrow b$ in $\mathcal{C}$:
  - From $A \rightarrow C$, add $R_1(A, C)$: $\{R_1(A, C)\}$
  - From $BC \rightarrow D$, add $R_2(B, C, D)$: $\{R_1(A, C), R_2(B, C, D)\}$
  - From $D \rightarrow E$, add $R_3(D, E)$: $\{R_1(A, C), R_2(B, C, D), R_3(D, E)\}$
  - From $E \rightarrow B$, add $R_4(B, E)$: $\{R_1(A, C), R_2(B, C, D), R_3(D, E), R_4(B, E)\}$
- Candidate Keys: $\mathcal{K} = \{AB, AD, AE\}$. No relation contains any, so add $R_5(A, B)$, $R_5(A, D)$, or $R_5(A, E)$.

3NF decomposition: $\{R_5(A, B), R_1(A, C), R_2(B, C, D), R_4(B, E), R_3(D, E)\}$

Example 2
3NF decomposition

Decompose relation $R(A, B, C, D, E, F)$ in ***3NF***, given the following functional dependencies:

- $AB \rightarrow CD$
- $CE \rightarrow BD$
- $BC \rightarrow AD$
- $AF \rightarrow B$

Reminder: One finds the 3NF decomposition by starting from an empty set of relations, and:

- For each FD $LHS \rightarrow RHS$ in a canonical cover, add relation $LHS \cup RHS$, if it is not a subset of an existing one
- If none of the resulting relations include a key, add a new relation for that key.

Solution 2

3NF decomposition

FDs $AB \rightarrow CD$, $CE \rightarrow BD$, $BC \rightarrow AD$, $AF \rightarrow B$

- Canonical cover: $\{AB \rightarrow CD, CE \rightarrow BD, BC \rightarrow AD, AF \rightarrow B\}$
- Add $R_1(A, B, C, D)$
- Add $R_2(B, C, D, E)$
- Add $R_3(A, B, C, D)$? No, because $R_3 \subseteq R_1$
- Add $R_4(A, B, F)$
- Is any $\mathrm{attr}(R_i)^+ = \mathrm{attr}(R)$? No: we need to add a key.
- $ABCDEF \mapsto ABEF \mapsto AEF$: Add $R_5(A, E, F)$

3NF decomposition:
$\{R_1(A, B, C, D), R_2(B, C, D, E), R_4(A, B, F), R_5(A, E, F)\}$

Example 3
BCNF decomposition

Decompose relation R(A, B, C, D, E, F) in ***BCNF***, given the following functional dependencies:

- $AB \rightarrow CD$
- $CE \rightarrow BD$
- $BC \rightarrow AD$
- $AF \rightarrow B$

Reminder: One finds the BCNF decomposition by recursively:

- Splitting the relation on each offending FD, i.e. each FD $LHS \rightarrow RHS$ such that LHS is *not* a superkey into $R_1(LHS \cup RHS)$ and $R_2(ALL - RHS)$
- Allocating to each child relation only the relations on its attributes

Imperial College London

Solution 3

BCNF decomposition

FDs $AB \to CD$, $CE \to BD$, $BC \to AD$, $AF \to B$

- AB is *not* a superkey: splitting R in $R_1(A, B, C, D)$ and $R_2(\{A, B, C, D, E, F\} - \{C, D\}) = R_2(A, B, E, F)$
- $AB \to CD$ and $BC \to AD$ match BCNF criterion on R_1. Its keys are AB and BC, and no FD can be derived from A, B, C, D, AD, AC or BD, so R_1 is BCNF.
- $AF \to B$ applies to R_2, but AF is not a superkey there: splitting into $R_{21}(A, B, F)$ and $R_{22}(\{A, B, E, F\} - \{B\}) = R_{22}(A, E, F)$
- R_{21} is BCNF: AF is a superkey, and no FD can be derived from A, B, F, AB or BF which applies there.
- R_{22} is BCNF: there is no way to derive E or F from anything, and all derived FDs that apply boil down to $A \to A$, which is trivial

BCNF decomposition: $\{R_1(A, B, C, D), R_{21}(A, B, F), R_{22}(A, E, F)\}$

Example 4
BCNF decomposition

Find a ***BCNF decomposition*** of $R(A, B, C, D)$ with:

- $AB \rightarrow D$
- $D \rightarrow C$

Solution 4
BCNF decomposition

FDs $AB \rightarrow D$, $D \rightarrow C$

- $D \rightarrow C$ not BCNF: $R_1(C, D)$, $R_2(A, B, D)$
- R_1 is BCNF
- R_2 is BCNF

BCNF decomposition: $\{R_1(C, D), R_2(A, B, D)\}$

Example 5
BCNF decomposition

Find a ***BCNF decomposition*** of $R(A, B, C, D)$

- $AB \to C$
- $C \to D$
- $D \to A$

Solution 5

BCNF decomposition

FDs $AB \to C$, $C \to D$, $D \to A$

- $AB \to C$ *matches* BCNF: $\{A, B\}^+ = \{A, B, C, D\}$
- $C \to D$ is not BCNF: $\{C\}^+ = \{C, D\}$. Splitting as $R_1(C, D)$, $R_2(A, B, C)$
- R_1 is BCNF: $C \to D$ is the only relevant FD
- R_2 is *not* BCNF: derived FD $C \to D \to A \Rightarrow C \to A$ does not match BCNF: $\{C\}^+ = \{A, C\}$. Split $R_{21}(A, C)$, $R_{22}(A, B)$
- R_{21} is BCNF
- R_{22} is BCNF

BCNF decomposition: $\{R_1(C, D), R_{21}(A, C), R_{22}(A, B)\}$

Conclusion

3NF : Add relations one by one from the canonical cover, purging redundant ones. Add keys eventually, if needed.

BCNF : For each offending FD, split relation until you've reached BCNF. Do not forget ***derived FDs*** !

Questions ?