KU Leuven

Tutorial: Normaalvormen

Gegevensbanken [H01O9A]

Gestart: 02 juni 2014
Cecompileerd: 3 juni 2014

Auteur:
Tom Sydney Kerckhove

Professor:
BettinaBerendt

Inhoudsopgave

Voorkennis

- Eerste orde formele logica.

Hoofdstuk 1

Theorie

1.1 Informele richtlijnen

1. Ontwerp een relatieschema zo dat zijn betekenis gemakkelijk verklaard kan worden.
2. Ontwerp een relatieschema zo dat redundantie vermeden wordt en geen toevoeg-, weglaat- of wijziging-anomalieën kunnen voorkomen.
3. Vermijd zoveel mogelijk attributen waarvan de waarden nul kunnen zijn.
4. Ontwerp relatieschema's zo dat ze na een equi-join op attributen die primaire of verwijssleutels zijn, geen onechte tupels opleveren.

1.2 Functionele afhankelijkheden (functional dependencies)

Definitie 1.2.1. Zij X en Y attributenverzamelingen. Y is **functioneel afhankelijk** X als vanuit de waarden van X de waarden van Y deterministisch bepaald kunnen worden.

$$X \to Y$$

Definitie 1.2.2. Een functionele afhankelijkheid $X \to Y$ is **partieel** als er een kleinere deelverzameling Z bestaat van X ($Z \subset X$) zodat Y functioneel afhankelijk is van Z. ($Z \to Y$)

Definitie 1.2.3. Een functionele afhankelijkheid noemen we **volledig** wanneer ze niet partieel is.

Definitie 1.2.4. Een functionele afhankelijkheid $X \to Y$ is **triviaal** wanneer Y een deel is van X.

$$Y \subseteq X$$

Definitie 1.2.5. De **sluiting van een verzameling van attributen** X onder een verzameling functionele afhankelijkheden X_F^+ is de verzameling van alle attribuutverzamelingen die functioneel afhankelijk zijn van X

$$X_F^+ = \{ \ Y \mid X \to Y \ \}$$

Definitie 1.2.6. Een verzameling functionele afhankelijkheden E **overdekt** een andere verzameling functionele afhankelijkheden F als voor elke $e = X \to Y$ geldt dat $Y \subseteq X_F^+$

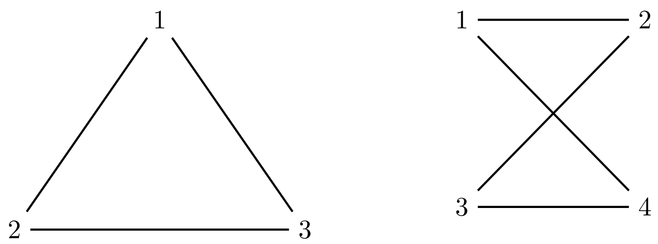
Definitie 1.2.7. We spreken van **equivaltente verzamelingen** E en F van **functionele afhankelijkheden** als zowel E F overdekt, als F E overdekt.

Definitie 1.2.8. Een verzameling functionele afhankelijkheden is minimaal als en slechts als er geen equivalente verzameling G te vormen valt door ...

- een afhankelijkheid uit F.
- een attribuut uit de rechterkant van een afhankelijkheid uit F.
- een attribuut uit de linkerkant van een afhankelijkheid uit F.

Definitie 1.2.9. Een functionele afhankelijkheid $X \to Y$ in een relatieschema $S_R = (U, F)$ is **transitief** als en slechts als er een Z bestaat zodat aan volgende voorwaarden voldaan is.

- Z is volledig en niet-triviaal afhankelijk van X
- Z is geen deelverzameling van een kandidaatsleutel voor R
- Y is niet-triviaal functioneel afhankelijk van Z.



1.2.1 Afleidingsregels

Regel 1. Reflexiviteit

$$Y \subseteq X \Rightarrow X \to Y$$

Regel 2. Uitbreiding

$$\{X \to Y\} \models XZ \to YZ$$

Regel 3. Transitiviteit

$$\{X \to Y, Y \to Z\} \models X \to Z$$

Regel 4. Decompositie

$$\{X \to YZ\} \models X \to Y$$

Regel 5. Vereniging

$$\{X \to Y, X \to Z\} \models X \to YZ$$

Regel 6. Pseudo-Transitiviteit

$$\{X \to Y, WY \to Z\} \models WX \to Z$$

1.3 Sleutels

Definitie 1.3.1. Een verzameling attributen K is een **supersleutel** voor een relatie R met schema $S_R = (U, F)$ als en slechts als K elk attribuut in U determineert.

$$K_F^+ = U$$

Definitie 1.3.2. Een verzameling attributen K is een **(kandidaat)sleutel** voor een relatie R met schema $S_R = (U, F)$ als en slechts als K een supersleutel is en er geen kleinere deelverzameling van K een supersleutel is voor R.

Definitie 1.3.3. Een attribuut A is een sleutelattribuut voor een relatie R met schema $S_R = (U, F)$ als en slecht als er een sleutel K voor R bestaat waar A in zit.

1.4 Normaalvormen

Definitie 1.4.1. Elk relatieschema is in **nulde normaalvorm (NF0)**. Er zijn geen voorwaarden opgelegd aan de attributen of functionele afhankelijkheden.

Definitie 1.4.2. Een relatieschema $S_R = (U, F)$ is in **eerste normaalvorm (NF1)** als en slechts als het domein van elk attribuut in U enkelvoudig is.

Definitie 1.4.3. Een relatieschema $S_R = (U, F)$ is in **tweede normaalvorm (NF2)** als en slechts als voor elk niet-sleutelattribuut $A \in U$ geldt dat A partieel functioneel afhankgelijk is van een kandidaatsleutel van R. *In mensentaal: "Voor elk niet-sleutel-attribuut moet de hele primaire sleutel nodig zijn om het te determineren."*

1.4.1 Derde normaalvorm: NF3

1.4.2 Boyce-Codd normaalvorm: BNCF

1.4.3 Vierde normaalvorm: NF4

1.4.4 Vijfde normaalvorm: NF5

Hoofdstuk 2

Generisch

Vraag 1 Bepaal de sluiting van X onder F: X_F^+

V1 Opgave

Gegeven zijn een attribuutverzameling $X = \{a_1, a_2, \cdot, a_n\}$ en een verzameling functionele afhankelijkheden $F = \{f_1, f_2, \cdot, f_n\}$. Bepaal X_F^+

V1 Antwoord

Ga één voor één de functionele afhankelijkheden in F af. voeg steeds de rechterkant toe als de linkerkant al in je verzameling zit. Herhaal dit tot X_F^+ niet meer verandert.

$X_F^+ \leftarrow X$
repeat
  $OLD \leftarrow X_F^+$
  foreach $f = Y \rightarrow Z \in F$ **do**
    if $y \subseteq X_F^+$ **then**
      $X_F^+ = X_F^+ \cup Z$
    end
  end
until $X_F^+ = OLD$;

Vraag 2 Bespreek de equivalentie van E en F.

V2 Opgave

Gegeven zijn twee verzamelingen functionele afhankelijkheden E en F. Bereken of E F overdekt, Bereken of F E overdekt en concludeer of E en F equivalent zijn.

V2 Antwoord

Ga simpelweg alle functionele afhankelijkheden $X \rightarrow Y \in E$ af en controleer of $Y \subseteq X_F^+$.

foreach $e = X \rightarrow Y \in E$ **do**
  If $Y \not\subseteq X_F^+$ **return** *False*
end
return *True*

Vraag 3 Bereken een minimale overdekking G van F.

V3 Opgave

Gegeven een verzameling functionele afhankelijkheden $F = \{f_1, f_2, \cdots, f_n$. Bepaal een minimale overdekking G van F.

V3 Antwoord

Ontdubbel eerst de rechterkant van elke functionele afhankelijkheid in F. Ontdubbel daarna zoveel mogelijk de linkerkant van elke functionele afhankelijkheid. Let op, dit mag alleen als er dan geen informatie verloren gaat. Elimineer ten slotte alle redundante functionele afhankelijkheden.

```
G = ∅
/* Ontdubbel de rechterkanten.                                  */
foreach f = X → A1 A2 ··· An inF do
    G ← G ∪ X → A1 ∪ X → A2 ∪ ··· ∪ X → An
end
/* Ontdubbel zoveel mogelijk de linkerkanten.                   */
foreach g = X1 X2 ··· Xn → A ∈ G do
    for i = 1..n do
        X' ← X1 X2 ··· X(i−1) X(i+1) ··· Xn
        G' ← G \ {X → A} ∪ {X' → A}
        /* Enkel wanneer er geen informatie verloren gaat!       */
        if Xi ∈ X'+_G' then
            G ← G'
        end
    end
end
/* Elimineer redundante functionele afhankelijkheden.           */
foreach g = X → A ∈ G do
    G' ← G \ {X → A}
    if A ∈ X+_G' then
        G ← G'
    end
end
```

Hoofdstuk 3

Voorbeeld

3.1 Bepaal de sluiting van X onder F: X_F^+

3.1.1 Opgave

$$X_1 = \{A\},\ X_2 = \{C\} \text{ en } X_3 = \{AC\}$$
$$F = \{A \to B, C \to DE, AC \to F\}$$

3.1.2 Antwoord

$$X_1^+ = \{A, B\}$$
$$X_2^+ = \{C, D, E\}$$
$$X_3^+ = \{A, B, C, D, E, F\}$$

3.2 Bespreek de equivalentie van E en F

3.2.1 Opgave

$$E = \{A \to BC, D \to AE\}$$
$$F = \{A \to B, AB \to C, D \to AC, D \to E$$

3.2.2 Antwoord

- Overdekt E F?
  - $A \to BC$
    $BC \subseteq A_F^+ = \{A, B, C\}$ OK
  - $D \to AE$
    $AE \subseteq D_F^+ = \{A, B, C, D, E\}$ OK

  Dus E overdekt F.
- Overdekt F E?
  - $A \to B$
    $B \subseteq A_E^+ = \{A, B, C\}$ OK
  - $AB \to C$
    $C \subseteq AB_E^+ = \{A, B, C\}$ OK
  - $D \to AC$
    $AC \subseteq D_E^+ = \{A, B, C, D, E\}$ OK
  - $D \to E$
    $E \subseteq D_E^+ = \{A, B, C, D, E\}$ OK

  Dus F overdekt E.
- Zijn E en F equivalent?
  E overdekt F en F overdekt E dus E en F zijn equivalent.

3.3 Bereken een minimale overdekking G van F.

3.3.1 Opgave

Zij F een verzameling functionele afhankelijkheden. Geef een minimale overdekking G van F.

$$F = \{B \rightarrow A, D \rightarrow A, AB \rightarrow D, H \rightarrow IJ\}$$

3.3.2 Antwoord

We voeren het beschreven algoritme uit, stap voor stap.

1. Ontdubbel de rechterkanten.
   Enkel $G \rightarrow GH$ kan ontdubbelt worden.

$$G = \{B \rightarrow A, D \rightarrow A, AB \rightarrow D, H \rightarrow I, H \rightarrow J\}$$

2. Ontdubbel zo veel mogelijk de llinker kanten.
   Enkel $AB \rightarrow D$ kan ontdubbelt worden.

   - Kunnen we $AB \rightarrow D$ vervangen door $A \rightarrow D$? Noem G' de verzameling die we zouden bekomen.

$$G' = \{B \rightarrow A, D \rightarrow A, A \rightarrow D, H \rightarrow I, H \rightarrow J\}$$

$$B \notin A^+_{G'} \rightarrow \text{ vervanging niet mogelijk.}$$

   - Kunnen we $AB \rightarrow D$ vervangen door $B \rightarrow D$? Noem G' de verzameling die we zouden bekomen.

$$G' = \{B \rightarrow A, D \rightarrow A, B \rightarrow D, H \rightarrow I, H \rightarrow J\}$$

$$A \in B^+_{G'} \rightarrow \text{ vervanging mogelijk.}$$

   We vervangen $AB \rightarrow D$ dus door $B \rightarrow D$.

$$G = \{B \rightarrow A, D \rightarrow A, B \rightarrow D, H \rightarrow I, H \rightarrow J\}$$

3. Ga alle $g \in G$ af en kijk na of ze redundant zijn. Verwijder indien mogelijk.

   - $B \rightarrow A$
     Zonder $B \rightarrow A$ zou G G' zijn.

$$G' = \{D \rightarrow A, B \rightarrow D, H \rightarrow I, H \rightarrow J$$

$$A \in B^+_{G'} = \{A, B, D\} \rightarrow \text{ Weglating OK}$$

   - $D \rightarrow A$

$$G' = \{B \rightarrow D, H \rightarrow I, H \rightarrow J$$

$$A \notin D^+_{G'} = \{D\} \rightarrow \text{ Weglating niet OK}$$

   - $B \rightarrow D$

$$G' = \{D \rightarrow A, H \rightarrow I, H \rightarrow J$$

$$D \notin B^+_{G'} = \{B\} \rightarrow \text{ Weglating niet OK}$$

   - $H \rightarrow I$

$$G = \{D \rightarrow A, B \rightarrow D, H \rightarrow J\}$$

$$I \notin H^+_{G'} = \{H, J\} \rightarrow \text{ Weglating niet OK}$$

   - $H \rightarrow J$

$$G = \{D \rightarrow A, B \rightarrow D, H \rightarrow I\}$$

$$J \notin H^+_{G'} = \{H, I\} \rightarrow \text{ Weglating niet OK}$$

Het resultaat is volgende verzameling G.

$$G = \{D \rightarrow A, B \rightarrow D, H \rightarrow I, H \rightarrow J\}$$