Mouhammadou Dabo (mod20)

Prof. Constantinos Costa

CS 1555

April 1, 2020

Assignment #8: Normalization

1. Consider the following set of functional dependencies:

FD1: ItemId → ItemDescription, ItemPrice, StockQuantity

FD2: PurchaseId → PurchaseDate, VendorCode

FD3: VendorCode → VendorName, VendorAddress

FD4: ItemId, PurchaceId → OrderQuantity

* 1. Using universal relational approach (top-down process), construct a set of 3NF/BCNF relations from the above functional dependencies. Indicate the primary keys for the result relations. Please show all steps clearly as mentioned in the lecture slides.
* Relation R (A, B, C, D, E, F, G, H, I, J)
* Transform all FDs to canonical form (i.e, one attribute on the right):

ItemId (A) → ItemDescription (B)

ItemId (A) → ItemPrice (C)

ItemId (A) → StockQuantity (D)

PurchaseId (E) → PurchaseDate (F)

PurchaseId (E) → VendorCode (G)

VendorCode (G) → VendorName (H)

VendorCode (G) → VendorAddress (I)

ItemId, PurchaseId (AE) → OrderQuantity (J)

which is the canonical cover of the relation

* Finding the keys of the relation:
  + A and E do not appear in the right hand side of any FDs, so they have to appear in all keys of R.
  + AE+: AE → AEJ (because AE → J) → AEJB (because A → B) → AEJBC (because A → C) → AEJBCD (because A → D) → AEJBCDF (because E → F) → AEJBCDFG (because E → G) → AEJBCDFGH (because G → H). So AE (ItemID, PurchaseId) is a key of the relation.
* Using universal relational approach:

Using A → B to decompose R, we get:

R1 (A, C, D, E, F, G, H, I, J)

R2 (A, B)

Using A → C to decompose R1, we get:

R11 (A, D, E, F, G, H, I, J)

R12 (A, C)

Using A → D to decompose R11, we get:

R111 (A, E, F, G, H, I, J)

R112 (A, D)

Using E → F to decompose R111, we get:

R1111 (A, E, G, H, I, J)

R1112 (E, F)

Using G → H to decompose R1111, we get:

R11111 (A, E, G, I, J)

R11112 (G, H)

Using G → I to decompose R11111, we get:

R111111 (A, E, G, J)

R111112 (G, I)

Using E → G to decompose R111111, we get:

R1111111 (A, E, J)

R1111112 (E, G)

Using AE → J to decompose R1111111, we get:

R11111111 (A, E)

R11111112 (AE, J)

Group the relations with the same key:

R1 (A, B, C, D)

R2 (E, F, G)

R3 (G, H, I)

R4 (AE, J)

R1, R2, R3 and R4 are in 3NF and in BCNF.

* 1. Using the table method, check whether the constructed set of relations is lossless or not. Also, state if your decomposition is good, bad or ugly. You must show all steps. Hint: Bad decomposition is a lossy one, while ugly decomposition is lossless but does not preserve some dependencies.

Initially the table looks like this:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F | G | H | I | J |
| R1(A,B,C,D) | a1 | a2 | a3 | a4 | U15 | U16 | U17 | U18 | U19 | U20 |
| R2(E,F,G) | U21 | U22 | U23 | U24 | a5 | a6 | a7 | U28 | U29 | U30 |
| R3(G,H,I) | U31 | U32 | U33 | U34 | U35 | U36 | a7 | a8 | a9 | U40 |
| R4(AE, J) | a1 | U42 | U43 | U44 | a5 | U46 | U47 | U48 | U49 | a10 |

Using A → B: we can replace U24, U by a2

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F | G | H | I | J |
| R1(A,B,C,D) | a1 | a2 | a3 | a4 | U15 | U16 | U17 | U18 | U19 | U20 |
| R2(E,F,G) | U21 | U22 | U23 | U24 | a5 | a6 | a7 | U28 | U29 | U30 |
| R3(G,H,I) | U31 | U32 | U33 | U34 | U35 | U36 | a7 | a8 | a9 | U40 |
| R4(AE, J) | a1 | **a2** | U43 | U44 | a5 | U46 | U47 | U48 | U49 | a10 |

Using A → C: we can replace U43 by a3

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F | G | H | I | J |
| R1(A,B,C,D) | a1 | a2 | a3 | a4 | U15 | U16 | U17 | U18 | U19 | U20 |
| R2(E,F,G) | U21 | U22 | U23 | U24 | a5 | a6 | a7 | U28 | U29 | U30 |
| R3(G,H,I) | U31 | U32 | U33 | U34 | U35 | U36 | a7 | a8 | a9 | U40 |
| R4(AE, J) | a1 | **a2** | **a3** | U44 | a5 | U46 | U47 | U48 | U49 | a10 |

Using A → D: we can replace U44 by a4

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F | G | H | I | J |
| R1(A,B,C,D) | a1 | a2 | a3 | a4 | U15 | U16 | U17 | U18 | U19 | U20 |
| R2(E,F,G) | U21 | U22 | U23 | U24 | a5 | a6 | a7 | U28 | U29 | U30 |
| R3(G,H,I) | U31 | U32 | U33 | U34 | U35 | U36 | a7 | a8 | a9 | U40 |
| R4(AE, J) | a1 | **a2** | **a3** | **a4** | a5 | U46 | U47 | U48 | U49 | a10 |

Using E → F: we can replace U46 by a6

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F | G | H | I | J |
| R1(A,B,C,D) | a1 | a2 | a3 | a4 | U15 | U16 | U17 | U18 | U19 | U20 |
| R2(E,F,G) | U21 | U22 | U23 | U24 | a5 | a6 | a7 | U28 | U29 | U30 |
| R3(G,H,I) | U31 | U32 | U33 | U34 | U35 | U36 | a7 | a8 | a9 | U40 |
| R4(AE, J) | a1 | **a2** | **a3** | **a4** | a5 | **a6** | U47 | U48 | U49 | a10 |

Using E → G: we can replace U47 by a7

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F | G | H | I | J |
| R1(A,B,C,D) | a1 | a2 | a3 | a4 | U15 | U16 | U17 | U18 | U19 | U20 |
| R2(E,F,G) | U21 | U22 | U23 | U24 | a5 | a6 | a7 | U28 | U29 | U30 |
| R3(G,H,I) | U31 | U32 | U33 | U34 | U35 | U36 | a7 | a8 | a9 | U40 |
| R4(AE, J) | a1 | **a2** | **a3** | **a4** | a5 | **a6** | **a7** | U48 | U49 | a10 |

Using G → H: we can replace U28 and U48 by a8

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F | G | H | I | J |
| R1(A,B,C,D) | a1 | a2 | a3 | a4 | U15 | U16 | U17 | U18 | U19 | U20 |
| R2(E,F,G) | U21 | U22 | U23 | U24 | a5 | a6 | a7 | **a8** | U29 | U30 |
| R3(G,H,I) | U31 | U32 | U33 | U34 | U35 | U36 | a7 | a8 | a9 | U40 |
| R4(AE, J) | a1 | **a2** | **a3** | **a4** | a5 | **a6** | **a7** | **a8** | U49 | a10 |

Using G → I: we can replace U29 and U49 by a9

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F | G | H | I | J |
| R1(A,B,C,D) | a1 | a2 | a3 | a4 | U15 | U16 | U17 | U18 | U19 | U20 |
| R2(E,F,G) | U21 | U22 | U23 | U24 | a5 | a6 | a7 | **a8** | **a9** | U30 |
| R3(G,H,I) | U31 | U32 | U33 | U34 | U35 | U36 | a7 | a8 | a9 | U40 |
| R4(AE, J) | a1 | **a2** | **a3** | **a4** | a5 | **a6** | **a7** | **a8** | **a9** | a10 |

We cannot proceed, but there is a row of all known values, so the decomposition is not lossy and is a good decomposition.

1. Consider the following set of functional dependencies:

A → B,

B → CD,

A → D,

B → C,

AB → CD,

A → C,

E → F

* 1. Using synthesis approach (bottom-up process), construct a set of 3NF/BCNF relations from the above functional dependencies. Indicate the primary keys for the result relations and whether or not they are in 3NF or BCNF. Please show all steps clearly as mentioned in the lecture slides.
     + Transform all FDs to canonical form (i.e, one attribute on the right):

A → B

B → CD becomes B → C and B → D

A → D

B → C

AB → CD becomes AB → C and AB → D

A → C

E → F

* + - Remove redundant dependencies:

A → B

B → C

B → D

A → D

AB → C

AB → D

A → C

E → F

* + - Drop extraneous attributes:

AB → C can be removed because we have A → C so B is redundant

AB → D can be removed because we have A → D so B is redundant

A → B

B → C

B → D

A → D

A → C

E → F

* + - Drop redundant FDs:

A → D can be deduced from A → B and B → D so we remove A → D

A → C can be deduced from A → B and B → C so we remove A → C

The set of FDs becomes:

A → B

B → C

B → D

E → F

which is the canonical cover of the relation.

* + - Finding the keys of the relation:
      * A and E do not appear in the right hand side of any FDs, so they have to appear in all the keys of the relation.
      * AE+: AE → AEB (because A → B) → AEBF (because E → F) → AEBFC (because B → C) → AEBFCD (because B → D). So AE is a key of R. In this case, we do not need to consider any other combination, because any other combination containing AE (e.g., AEBF) is a super key and not minimal.
    - Using synthesis approach:
      * Group FDs with same determinant:

A → B

B → CD

E → F

* + - * Construct a relation for each group:

R1 (A, B)

R2 (B, C, D)

R3 (E, F)

* + - * If none of the relations contain the key for the original relation (AE), add a relation with the key

R4 (A, E)

R1, R2, R3 and R4 are in 3NF and in BCNF

* 1. Using the table method, check whether the constructed set of relations is lossless or not. If not, correct them.

Initially the table looks like this:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F |
| R1(A,B) | a1 | a2 | U13 | U14 | U15 | U16 |
| R2(B,C,D) | U21 | a2 | a3 | a4 | U25 | U26 |
| R3(E,F) | U31 | U32 | U33 | U34 | a5 | a6 |
| R4(A,E) | a1 | U42 | U43 | U44 | a5 | U46 |

Using A → B: we can replace U42 by a2

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F |
| R1(A,B) | a1 | a2 | U13 | U14 | U15 | U16 |
| R2(B,C,D) | U21 | a2 | a3 | a4 | U25 | U26 |
| R3(E,F) | U31 | U32 | U33 | U34 | a5 | a6 |
| R4(A,E) | a1 | **a2** | U43 | U44 | a5 | U46 |

Using B → C: we can replace U13 and U43 by a3

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F |
| R1(A,B) | a1 | a2 | **a3** | U14 | U15 | U16 |
| R2(B,C,D) | U21 | a2 | a3 | a4 | U25 | U26 |
| R3(E,F) | U31 | U32 | U33 | U34 | a5 | a6 |
| R4(A,E) | a1 | **a2** | **a3** | U44 | a5 | U46 |

Using B → D: we can replace U14 and U44 by a4

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F |
| R1(A,B) | a1 | a2 | **a3** | **a4** | U15 | U16 |
| R2(B,C,D) | U21 | a2 | a3 | a4 | U25 | U26 |
| R3(E,F) | U31 | U32 | U33 | U34 | a5 | a6 |
| R4(A,E) | a1 | **a2** | **a3** | **a4** | a5 | U46 |

Using E → F: we can replace U46 by a6

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F |
| R1(A,B) | a1 | a2 | **a3** | **a4** | U15 | U16 |
| R2(B,C,D) | U21 | a2 | a3 | a4 | U25 | U26 |
| R3(E,F) | U31 | U32 | U33 | U34 | a5 | a6 |
| R4(A,E) | a1 | **a2** | **a3** | **a4** | a5 | **a6** |

We cannot proceed, but there is a row of all known values, so the decomposition is not lossy.