Higher Normalization - 5NF

Overview

- Join Dependencies [JD]
- Fifth Normal Form – 5NF

Join Dependency

- There exists relvars that can be nonloss-decomposed into three or more – *n-decomposable*.
- Consider relvar SHIPMENTS from the supplier-parts-projects.
- The relvar is all key and involves no nontrivial FDs or MVDs, hence in 4NF.

SHIPMENTS

SUPPLIER_NUMBER	PART_NUMBER	PROJECT_NUMBER
S1	P1	J2
S1	P2	J1
S2	P1	J1
S1	P1	J1

Join Dependency

- SP

SUPPLIER_NUMBER	PART_NUMBER
S1	P1
S1	P2
S2	P1

PJ

PART_NUMBER	PROJECT_NUMBER
P1	J2
P2	J1
P1	J1

JS

PROJECT_NUMBER	SUPPLIER_NUMBER
J2	S1
J1	S1
J1	S2

Join Dependency

- SP Join PJ over PART_NUMBER

SUPPLIER_NUMBER	PART_NUMBER	PROJECT_NUMBER
S1	P1	J2
S1	P2	J1
S2	P1	J1
S2	P1	J2
S1	P1	J1

Spurious tuple

JS

PROJECT_NUMBER	SUPPLIER_NUMBER
J2	S1
J1	S1
J1	S2

Join over PROJECT_NUMBER, SUPPLIER_NUMBER

SHIPMENTS

Join Dependency

- SHIPMENTS is equal to the join of its three projections SP, PJ, and JS is equivalent to:

> If the pair (s1,p1) appears in SP
> and the pair (p1,j1) appears in PJ
> and the pair (j1,s1) appears in JS
> then the triple (s1,p1,j1) certainly appears in
> join of {SP,PJ,JS} = SHIPMENTS

- A relvar will be n-decomposable for some n>2 if and only if it satisfies some such (n-way) cyclic constraint.

Join Dependency

- JD is equal to the following constraints. If,
  - <s1, p1, z> : a supplier s1 supplies part p1 to some project, and
  - <x, p1, j1> : a project j1 uses part p1 supplied by some supplier, and
  - <s1, y, j1> : the supplier s1 supplies *at least one* part to project j1, then
  - <s1, p1, j1>: supplier s1 will also be supplying part p1 to project j1
- Constraint 3-decomposable [3D] is satisfied if and only if the relvar is equal to the join of certain of its projections – refer to that constraint as a *join-dependency* (JD)

Join Dependency

- Let R be a relvar, and let A,B, ..., Z be subsets of the attributes of R. Then we say that R satisfies the JD

  * { A, B, ..., Z } [star A, B, ..., Z]

  if and only if every legal value of R is equal to the join of its projections on A, B, ..., Z.
- For example, the set of attributes of SP, PJ and JS, then relvar SHIPMENTS(S,P,J) satisfies the JD *{SP, PJ, JS}

JDs and Fifth NF

Figure 11.4
Fourth and fifth normal forms.

(a) SUPPLY in 4NF, with JD(R1,R2,R3) – not in 5NF

(b) Decomposing into 5NF relations R1, R2, R3

(c) SUPPLY

Sname	Part_name	Proj_name
Smith	Bolt	ProjX
Smith	Nut	ProjY
Adamsky	Bolt	ProjY
Walton	Nut	ProjZ
Adamsky	Nail	ProjX
Adamsky	Bolt	ProjX
Smith	Bolt	ProjY

(d) R_1

Sname	Part_name
Smith	Bolt
Smith	Nut
Adamsky	Bolt
Walton	Nut
Adamsky	Nail

R_2

Sname	Proj_name
Smith	ProjX
Smith	ProjY
Adamsky	ProjY
Walton	ProjZ
Adamsky	ProjX

R_3

Part_name	Proj_name
Bolt	ProjX
Nut	ProjY
Bolt	ProjY
Nut	ProjZ
Nail	ProjX

Fifth Normal Form – 5NF

Fifth Normal Form:

Relvar R is in 5NF – also called projection-join normal form – if and only if every non-trivial join dependency that is satisfied by R is implied by the candidate key(s) of R, where

a) the JD * {A, B, ..., Z} on R is **trivial** if and only if at least one of A, B, ..., Z is the set of all attributes of R.

b) the JD * {A, B, ..., Z} on R is **implied by the candidate key(s)** of R if and only if each of A, B, ..., Z is a superkey for R.

References

- *Chapter 11: Functional Dependencies*
  An introduction to database systems, *CJ. Date*