

3.8. Antroji norminė forma

Atributai B **visiškai priklauso** nuo atributų A FP aibės F atžvilgiu, jei B - f -priklauso nuo visos aibės A , bet nėra f -priklausantys nei nuo jokio aibės A poaibio, t.y.

$$A \rightarrow B \in F^+ \Rightarrow \forall C \subset A : C \rightarrow B \notin F^+$$

$L(R)$ atributas a ($a \in R$) - **pirminis** FP aibės F atžvilgiu, jei a priklauso L raktui, kitaip, a - **nepirminis atributas**.

$L(R)$ yra 2NF F atžvilgiu,

jei ji yra 1NF ir visi nepirminiai atributai visiškai priklauso nuo kiekvieno L rakto.

DB yra 2NF FP aibės F atžvilgiu tada ir tik tada, kai visos lentelės yra 2NF atžvilgiu F .

Projektai_Vykdymas

| Projektas | Pavadinimas | Svarba | Trukmė | Vykdytojas | ... |
|-----------|----------------------|----------|--------|------------|-----|
| 1 | Studentų apskaita | Maža | 12 | 1 | |
| 1 | Studentų apskaita | Maža | 12 | 2 | |
| 1 | Studentų apskaita | Maža | 12 | 3 | |
| 1 | Studentų apskaita | Maža | 12 | 4 | |
| 2 | Buhalterinė apskaita | Vidutinė | 10 | 1 | |
| 2 | Buhalterinė apskaita | Vidutinė | 10 | 2 | |
| 2 | Buhalterinė apskaita | Vidutinė | 10 | 4 | |
| 3 | WWW svetainė | Didelė | 6 | 1 | |
| 3 | WWW svetainė | Didelė | 6 | 2 | |
| 3 | WWW svetainė | Didelė | 6 | 3 | |

Projektai_Vykdymas (Projektas, Pavadinimas, Svarba, Trukmė, Pradžia, Vykdytojas, Statusas, Valandos)

Projektas \rightarrow {Pavadinimas, Svarba, Trukmė, Pradžia}
 {Projektas, Vykdytojas} \rightarrow {Statusas, Valandos}

{Projektas, Vykdytojas} - lentelės raktas
 Projektas, Vykdytojas - pirminiai atributai

Nepirminiai Pavadinimas, Svarba, Trukmė, Pradžia nevisiškai f -priklauso nuo rakto, jie f -priklauso tik nuo Projektas

\Rightarrow Projektai_Vykdymas nėra 2NF.

Projektai ir Vykdymas yra 2NF.

Projektai (Projektas, Pavadinimas, Svarba, Trukmė, Pradžia)

Projektas \rightarrow {Pavadinimas, Svarba, Trukmė, Pradžia}

Vykdymas (Projektas, Vykdytojas, Statusas, Valandos)
 {Projektas, Vykdytojas} \rightarrow {Statusas, Valandos}

2NF sumažina duomenų perteklių ir pašalina anomalijas.

Darbas su 2 lentelėmis yra sudėtingesnis, nei su 1.

Tačiau:

- duomenų paieškose galima jungti kelias lenteles
- SQL leidžia apibrėžti virtualiąsias lenteles, apjungiančias kelias lenteles.

Lentelės skaidymas į kelias lenteles, esančias 2NF:

- 1) kuriama nauja lentelė, kurios atributais yra pradinės lentelės atributai, įeinantys į FP tarp nepirminių atributų (atributo) ir rakto dalies. Šios FP determinantas - naujos lentelės raktas;
- 2) FP dešinėje esantys atributai pašalinami iš pradinės lentelės;
- 3) žingsnius 1 ir 2 kartoti kiekvienam tokiai FP.

Schematiškai **lentelę, nesančią 2NF, galima pavaizduoti taip:**

$L(\underline{A}, \underline{B}, C, D)$, kurioje galioja: $AB \rightarrow CD$ ir $A \rightarrow D$
 D f -priklauso nuo rakto dalies A .

L skaidome į dvi: $L_1(\underline{A}, \underline{B}, C)$ ir $L_2(\underline{A}, D)$.

L_1 galioja FP $AB \rightarrow C$,

L_2 galioja FP $A \rightarrow D$.

L_1 turi išorinį raktą A , nukreipiantį į L_2 .

Suskaidydami lentelę, nepraradome savybių, nes $AB \rightarrow CD$ yra išvedama iš $AB \rightarrow C$ ir $A \rightarrow D$

Lentelė, nesanti 2NF, mažiausiai turi 3 stulpelius:

$T(\underline{A}, B, C)$, kurioje galioja FP: $AB \rightarrow C$ ir $A \rightarrow C$

T skaidome į dvi: $T_1(\underline{A}, B)$ ir $T_2(\underline{A}, C)$.

T_1 galioja tik triviali FP $AB \rightarrow AB$,

T_2 galioja FP $A \rightarrow C$.

T_1 turi išorinį raktą A , nukreipiantį į L_2 .

Suskaidydami lentelę, nepraradome savybių, nes

$AB \rightarrow C$ yra išvedama iš $A \rightarrow C$, papildant determinantą stulpeliu B .

Skaidymas (dekompozicija) išsaugant FP – lentelės skaidymas, kai neprarandamos jokios FP.

Skaidymas (dekompozicija) be praradimo – lentelės skaidymas, kai jungiant naująsias lenteles gaunami pradinės lentelės duomenys.

Skaidant lentelę $L(A, B, C)$ į $L_1(A, B)$ ir $L_2(A, C)$ duomenys yra neprarandami, jeigu $L(A, B, C)$ duomenis visada galima gauti jungiant L_1 ir L_2 :

SELECT A, B, C **FROM** L_1, L_2
WHERE $L_1.A = L_2.A$

Hezo (I.J. Heath) teorema. Tarkime, lentelėje $L(A, B, C)$ galioja $A \rightarrow B$ arba $A \rightarrow C$, kur A, B ir C – lentelės atributų aibės poaibiai. **Tuomet** lentelę L galima gauti jungiant jos projekcijas $L_1(A, B)$ ir $L_2(A, C)$.

Irodymas. Tarkime, $A \rightarrow B$ arba $A \rightarrow C$ ir $L \neq L_1 \bowtie L_2$.

Jei L_1 ir L_2 yra L projekcijos, tai $L \subseteq L_1 \bowtie L_2$

Jei $L \neq L_1 \bowtie L_2$, tai $\langle a, b, c \rangle \in L_1 \bowtie L_2$, kurios **nėra** L .

Jei $\langle a, b, c \rangle \in L_1 \bowtie L_2$, tai $\langle a, b \rangle \in L_1$ ir $\langle a, c \rangle \in L_2$.

Vadinasi, $\langle a, b, c' \rangle \in L$ ir $\langle a, b', c \rangle \in L$, kuriose $b' \neq b$ ir $c' \neq c$. Todėl, negalioja nei $A \rightarrow B$ nei $A \rightarrow C$. **Prieštara.**

Jei $L(A, B, C)$ galioja $A \rightarrow B$ arba $A \rightarrow C$, **tai** jungiant L skaidinį: $L_1(A, B)$ ir $L_2(A, C)$ per bendrą stulpelį A , gaunami L duomenys, t. y.

SELECT A, B, C **FROM** L_1, L_2
WHERE $L_1.A = L_2.A$

rezultatas sutaps su L .

Formaliai: $L = \pi_{AB}(L) \bowtie \pi_{AC}(L) = L_1 \bowtie L_2$.

Kitaip tariant,

Jei $A \rightarrow B$ arba $A \rightarrow C$,

tai $\pi_{AB}(L)$ ir $\pi_{AC}(L)$ yra L skaidymas be praradimo.

Atvirkščias teiginys nėra teisingas:

tai, kad $\pi_{AB}(L)$ ir $\pi_{AC}(L)$ yra L skaidymas be praradimo **nereiškia**, kad $A \rightarrow B$ arba $A \rightarrow C$

| L | | | L_1 | | L_2 | |
|-----|------|------|-------|------|-------|------|
| A | B | C | A | B | A | C |
| a | b | c | a | b | a | c |
| a | b' | c | a | b' | a | c' |
| a | b | c' | | | | |
| a | b' | c' | | | | |

$L_1 = \pi_{AB}(L)$, $L_2 = \pi_{AC}(L)$ ir $L = L_1 \bowtie L_2$,

bet negalioja $A \rightarrow B$, $A \rightarrow C$.

Projektai_Vykdymas (Projektas, Pavadinimas, Svarba, Trukmė, Pradžia, Vykdytojas, Statusas, Valandos)

skaidant į

Projektai (Projektas, Pavadinimas, Svarba, Trukmė, Pradžia)

Vykdymas (Projektas, Vykdytojas, Statusas, Valandos)

išsaugomos FP ir neprarandami duomenys, nes **Projektai** sudaryta pagal Hezo teoremą, naudojant FP:

$\text{Projektas} \rightarrow \{\text{Pavadinimas}, \text{Svarba}, \text{Trukmė}, \text{Pradžia}\}$

Pvz., lentelę

Projektai_Vykdymas(Projektas, Pavadinimas, Svarba, Trukmė, Pradžia, Vykdytojas, Statusas, Valandos)

išskaidykime kitaip:

Projektai_Vykdymas1(Projektas, Pavadinimas, Svarba, Trukmė, Pradžia, Vykdytojas)

Projektai_Vykdymas2(Vykdytojas, Statusas, Valandos)

Projektai_Vykdymas

| <u>Projektas</u> | <u>Pavadinimas</u> | <u>Svarba</u> | <u>Trukmė</u> | <u>Vykdytojas</u> | ... |
|------------------|----------------------|---------------|---------------|-------------------|-----|
| 1 | Studentų apskaita | Maža | 12 | 1 | |
| 1 | Studentų apskaita | Maža | 12 | 2 | |
| 1 | Studentų apskaita | Maža | 12 | 3 | |
| 1 | Studentų apskaita | Maža | 12 | 4 | |
| 2 | Buhalterinė apskaita | Vidutinė | 10 | 1 | |
| 2 | Buhalterinė apskaita | Vidutinė | 10 | 2 | |
| 2 | Buhalterinė apskaita | Vidutinė | 10 | 4 | |
| 3 | WWW svetainė | Didelė | 6 | 1 | |
| 3 | WWW svetainė | Didelė | 6 | 2 | |
| 3 | WWW svetainė | Didelė | 6 | 3 | |

Projektai_Vykdymas1

| Projektas | Pavadinimas | Svarba | Trukmė | Vykdytojas |
|-----------|----------------------|----------|--------|------------|
| 1 | Studentų apskaita | Maža | 12 | 1 |
| 1 | Studentų apskaita | Maža | 12 | 2 |
| 1 | Studentų apskaita | Maža | 12 | 3 |
| 1 | Studentų apskaita | Maža | 12 | 4 |
| 2 | Buhalterinė apskaita | Vidutinė | 10 | 1 |
| 2 | Buhalterinė apskaita | Vidutinė | 10 | 2 |
| 2 | Buhalterinė apskaita | Vidutinė | 10 | 4 |
| 3 | WWW svetainė | Didelė | 6 | 1 |
| 3 | WWW svetainė | Didelė | 6 | 2 |
| 3 | WWW svetainė | Didelė | 6 | 3 |

Projektai_Vykdymas2

| Vykdytojas | Statusas | Valandos |
|------------|-----------------|----------|
| 1 | Programuotojas | 30 |
| 2 | Dokumentuotojas | 100 |
| 3 | Testuotojas | 100 |
| 4 | Vadovas | 100 |
| 1 | Programuotojas | 300 |
| 2 | Analitikas | 250 |
| 4 | Vadovas | 100 |
| 1 | Programuotojas | 250 |
| 2 | Vadovas | 400 |
| 3 | Dizaineris | 150 |

Lentelei Projektai_Vykdymas2

negalima nustatyti rakto ir prarandame dalį informacijos.

Lentelės Projektai_Vykdymas skaidinio

Projektai_Vykdymas1 ir Projektai_Vykdymas2 **junginį**

Projektai_Vykdymas1 \bowtie Projektai_Vykdymas2

sudaro 26 eilutės, nors pradinėje lentelėje tėra 10.

Iš junginio \Rightarrow vykdytojas Nr. 2 projektą Nr. 1 vykdo: ir dokumentuotojo, ir analitiko, ir vadovo statusuose - **netiesa**. Be to, taip skaidydami **praradome FP**

$\{Projektas, Vykdytojas\} \rightarrow \{Statusas, Valandos\}$

3.9. Trečioji norminė forma

$L(R)$ atributų aibės R poaibis C ($C \subset R$) **tranzityviai priklauso** nuo atributų aibės A FP aibės F atžvilgiu, jei egzistuoja toks atributų aibės R poaibis B , kad A f-apibrėžia B , bet **neatvirkščiai**, bei B f-apibrėžia C FP aibės F atžvilgiu, t.y. $\exists B \subset R$:

$A \rightarrow B \in F^+, B \rightarrow A \notin F^+, B \rightarrow C \in F^+$ ir $C \not\subset A \cup B$

$L(R)$ yra **3NF FP** aibės F atžvilgiu, jei ji yra 1NF ir nėra nepirminių atributų, tranzityviai priklausančių nuo rakto.

DB yra 3NF FP aibės F atžvilgiu tada ir tik tada, kai visos jos lentelės yra 3NF F atžvilgiu.

Vykdytojai_AM (Nr, Pavardė, Kvalifikacija, Kategorija, Išsilavinimas, AM_Adresas)

| Nr | Pavardė | Išsilavinimas | AM_Adresas |
|----|------------|---------------|-------------------------|
| 1 | Jonaitis | VU | Universiteto 3, Vilnius |
| 2 | Petraitis | VU | Universiteto 3, Vilnius |
| 3 | Gražulytė | NULL | NULL |
| 4 | Onaitytė | VDU | Donelaičio 58, Kaunas |
| 5 | Antanaitis | VU | Universiteto 3, Vilnius |

Nr - vienintelis lentelės raktas. Lentelėje **galioja FP**:

$Nr \rightarrow \{Pavardė, Kvalifikacija, Kategorija, Išsilavinimas, AM_Adresas\}$

$Išsilavinimas \rightarrow AM_Adresas$

Lentelės raktas tik iš vieno atributo \Rightarrow ji yra 2NF.

Tačiau, jai yra būdingas duomenų perteklius.

Lentelėje Vykdytojai_AM galioja FP

$Nr \rightarrow Išsilavinimas$,

kuri nėra apverčiama, t.y. negalioja

$Išsilavinimas \rightarrow Nr$,

be to $Išsilavinimas$ f-apibrėžia $AM_Adresas$

\Rightarrow šioje lentelėje nepirminis atributas $AM_Adresas$ tranzityviai priklauso nuo rakto Nr

\Rightarrow lentelė Vykdytojai_AM nėra 3NF.

Lentelės Vykdytojai_AM blogybės išnyksta, suskaidžius ją į lenteles:

Vykdytojai ir AM_Adresai(Pavadinimas, Adresas)

Išorinis raktas: Vykdytojai.Išsilavinimas nukreipia į

AM_Adresai

| Pavadinimas | Adresas |
|-------------|-------------------------|
| VU | Universiteto 3, Vilnius |
| VDU | Kaunas |

Vykdytojai ir AM_Adresai yra 3NF.

3NF sumažino duomenų perteklių ir panaikino anomalijas.

Schematiškai: tarkime, $L(\underline{A}, B, C)$ - galioja:

$A \rightarrow B, B \rightarrow C$, ir negalioja $B \rightarrow A$

t.y. L nėra 3NF.

L skaidome į dvi: $L_1(\underline{A}, B)$ ir $L_2(\underline{B}, C)$.

$L_1 : A \rightarrow B$

$L_2 : B \rightarrow C$

$L_1.B - L_1$ išorinis raktas į L_2

Heath teorema užtikrina, kad neprarandami duomenys.

Be to, išsaugomos FP

$\{A \rightarrow B, B \rightarrow C\} \equiv \{A \rightarrow BC, B \rightarrow C\}$

Teorema. Lentelė, esanti 3NF FP aibės F atžvilgiu, yra ir 2NF F atžvilgiu.

Irodymas. Tarkime, $L(R)$ yra 3NF, bet nėra 2NF FP aibės F atžvilgiu. Tuomet $\exists a \in R : a$ - nepirminis ir a f-priklauso nuo kurio nors rakto K dalies K' , $K' \subset K \subseteq R$. Tai reiškia, kad $K' \rightarrow a \in F^+$, $K \rightarrow K' \in F^+$ ir $K' \rightarrow K \notin F^+$, nes kitaip K' būtų L raktas, o ne K . Be to, $a \notin K$, kadangi K – raktas ir a - nepirminis.

Turime: $K \rightarrow K'$, ne $K' \rightarrow K$, $K' \rightarrow a$ bei $a \notin K$ ir juo labiau $a \notin K'$, t.y. $a \notin K \cup K'$. Tokiu būdu, nepirminis a tranzityviai priklauso nuo rakto $K \Rightarrow L(R)$ nėra 3NF - **prieštara** prielaidai.

25/30

Dalinė priklausomybė reiškia ir tranzityviąją.

Jei lentelė nėra 2F, tai joje egzistuoja nepirminis atributas, tranzityviai priklausantis nuo rakto.

Tarkime, $L(\underline{A}, \underline{B}, C)$ galioja: $AB \rightarrow C$ ir $A \rightarrow C$.
 $\Rightarrow L$ nėra 2NF.

C tranzityviai priklauso nuo rakto AB

$\Rightarrow L$ nėra 3NF:

Galioja: $AB \rightarrow A$, $A \rightarrow C$

Negalioja: $A \rightarrow AB$, $C \not\rightarrow AB$

26/30

Tarkime $L(\underline{A}, B, C)$ - lentelė, kurioje galioja FP

$A \rightarrow B$ ir $B \rightarrow C$

Tuomet L skaidome į dvi: $L_1(\underline{A}, B)$ ir $L_2(\underline{B}, C)$.

Pagal Hezo teoremą galimas ir toks **skaidinys**:

$L_3(\underline{A}, B)$, kurioje $A \rightarrow B$

$L_4(\underline{A}, C)$, kurioje $A \rightarrow C$

Pagal Hezo teoremą, tai dekompozicija **be praradimo**, bet šis skaidymas **neišsaugo FP**, nes

$B \rightarrow C$ **neišvedama** iš $A \rightarrow B$, $A \rightarrow C$

\Rightarrow taip nedera skaidyti.

27/30

Skaidymas gali būti teisingas, bet **neefektyvus**:

$L(\underline{A}, B, C, D) : FP : A \rightarrow B, B \rightarrow C, B \rightarrow D$

arba : $A \rightarrow B, B \rightarrow CD$ – **nėra 3NF**

Naikinam CD tranzityvumą nuo rakto A :

$L_1(\underline{A}, B)$ ir $L_2(\underline{B}, C, D)$

Jei iš pradžių likviduojame vieną (C) tranzityvumą, o po to kitą, tai gauname 3 lenteles:

$L_1(\underline{A}, B)$, $L_3(\underline{B}, C)$ ir $L_4(\underline{B}, D)$

Abu skaidymai yra be praradimo ir išsaugo FP.

1-asis – efektyvesnis, nes reikalauja mažiau atminties.

28/30

Algoritmas sudaryti 3NF lenteles FP aibės F atžvilgiu.

$G := F_{\min}$;

$i := 0$;

for each $X : \exists(X \rightarrow Y) \in G$

$R := \emptyset$;

for each $Y : \exists(X \rightarrow Y) \in G$

$R := R \cup Y$;

endfor

$i := i + 1$;

Į DB reliacinę schemą įtraukti $L_i(\underline{X}, R)$;

endfor

Lentelių skaičius – determinantų skaičius aibėje F_{\min}

29/30

Pvz.: $L(\underline{A}, B, C, D)$

$F: A \rightarrow B, B \rightarrow C, B \rightarrow D$

F – minimalioji aibė

$F_{\min} = F$

Aibėje F_{\min} yra 2 skirtingi determinantai (A ir B).

Išorinį ciklą atliekame 2 kartus.

1-as $X := A$, vykdome **vidinį ciklą 1 kartą**, gauname $L_1(\underline{A}, B)$.

2-as $X := B$, **vidinį – 2 kartus** ($Y := C, Y := D$) gauname $L_2(\underline{B}, C, D)$

30/30