

3.10. Boiso-Kodo norminė forma

1-25

3NF, kaip ir 1NF bei 2NF, pirmasis apibrėžė Kodas (E.F.Codd).

3NF yra dažnai pakankama praktiniuose taikymuose.

3NF ne visiškai gerai tinka lentelėms su dviem ir daugiau raktų, kai bent du iš jų yra sudėtiniai raktai ir turi bendrų atributų.

2-25

Lentelė yra **Boiso-Kodo NF (BKNF)**, jei ji yra 1NF ir kiekvienoje galiojančioje netrivialioje FP $X \rightarrow A$ atributai X yra **lentelės viršraktis**.

Kitaip tariant, lentelė yra KKNF, jei ji yra 1NF ir kiekvienoje galiojančioje FP $X \rightarrow A$:

- $A \subset X$ arba
- X yra lentelės viršraktis.

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

3-25

Tarkime, projektai yra dalinami etapais:

Projektų_Etapai (Nr, Pavadinimas, Etapas, Pabaiga)

{Nr, Etapas} – (pirminis) raktas.

Tarus, kad negali būti dviejų projektų su vienodais pavadinimais, lentelė **turi ir antrą raktą**

{Pavadinimas, Etapas}

$\{Nr, Etapas\} \cap \{Pavadinimas, Etapas\} \equiv Etapas$

$\{Pavadinimas, Etapas\} \rightarrow \{Nr, Pabaiga\}$

$\{Nr, Etapas\} \rightarrow \{Pavadinimas, Pabaiga\}$

$Nr \rightarrow Pavadinimas$

$Pavadinimas \rightarrow Nr$

Pabaiga – nepirminis. *Projektų_Etapai* yra 3NF.

4-25

Projektų_Etapai yra 3NF, bet yra duomenų **dubliavimas**.

Nr	Pavadinimas	Etapas	Pabaiga
1	Studentų apskaita	1	2001.06.30
1	Studentų apskaita	2	2001.12.31
2	Buhalterinė apskaita	1	2001.05.31
2	Buhalterinė apskaita	2	2001.12.31
3	WWW svetainė	1	2001.07.31

Dviejose (visiškai) netrivialiose FP:

$Nr \rightarrow Pavadinimas$

$Pavadinimas \rightarrow Nr$

determinantai nėra raktai (yra rakto dalys) ir juo labiau viršrakčiai \Rightarrow **Projektų_Etapai nėra BKNF**

5-25

Galimi 2 lentelės *Projektų_Etapai* skaidymai

1) *Pavadinimai_Etapai* (Pavadinimas, Etapas, Pabaiga)
Nr_Pavadinimai (Nr, Pavadinimas)

2) *Nr_Etapai* (Nr, Etapas, Pabaiga)
Nr_Pavadinimai (Nr, Pavadinimas)

Abu šie lentelių rinkiniai yra BKNF.

Teoriškai abu skaidiniai - lygiaverčiai.

Pavadinimas reikšmės yra ilgesnės negu *Nr* reikšmės.

\Rightarrow *Nr_Etapai*, praktiniais sumetimais, yra pranašesnė už *Pavadinimai_Etapai*.

6-25

Schematiškai: tarkime, $L(A, B, C)$:

$A \rightarrow B$ – visiškai netriviali bei nesuprastinama
(determinante nėra perteklinių atributų) ir
 A nėra šios lentelės raktas,

t.y. **L nėra BKNF**.

$L(A, B, C) \Rightarrow L_1(A, C)$ ir $L_2(\underline{A}, B)$.

išorinis raktas: $L_1.A \Rightarrow L_2$

7-25

Teiginys. Jei lentelė yra BKNF, tai ji yra ir 3NF.

Irodymas. Tarkime, $L(R)$ yra BKNF, bet nėra 3NF.

Jei $L(R)$ nėra 3NF, tai $\exists K, A, B \subset R$, K – raktas, A – nepirminiai, $A \not\subset KB$, $K \rightarrow B$, $B \rightarrow A$, ne $B \rightarrow K$.

$B \rightarrow K$ – negalioja $\Rightarrow B$ ne viršraktis.

$A \not\subset B \Rightarrow B \rightarrow A$ yra visiškai netriviali

$\Rightarrow B \rightarrow A$ netenkina BKNF reikalavimų

$\Rightarrow L(R)$ nėra BKNF - **prieštara**.

Tik, kai lentelėje galioja $X \rightarrow A$, kurioje X nėra viršraktis ir A yra pirminis atributas, lentelė yra 3NF, bet nėra BKNF.

8-25

Teiginys. Jei lentelė yra 3NF ir nėra BKNF FP aibės F atžvilgiu, **tai** $\exists X \rightarrow A \in F^+$, kurioje X nėra viršraktis ir A yra pirminis atributas.

Irodymas. Tarkime, $L(R)$ yra 3NF, bet nėra BKNF.

Ne BKNF $\Rightarrow \exists$ netrivialioji $X \rightarrow A$, kurioje X nėra viršraktis. Tarkime, A – nepirminis.

Jei X – ne viršraktis, tai raktui K negalioja FP $X \rightarrow K$.

K – raktas $\Rightarrow K \rightarrow X$.

$X \rightarrow A$ yra visiškai netrivialioji $\Rightarrow A \notin X$.

\Rightarrow nepirminis A tranzityviai priklauso nuo rakto K

($K \rightarrow X \in F^+$, $X \rightarrow K \notin F^+$, $X \rightarrow A$, ir $A \notin KX$)

$\Rightarrow L$ nėra 3NF. **Prieštara**.

3.11. Nedalomosios lentelės

9-25

Tarkime,

- darbuotojas gali turėti kelias kvalifikacijas, baigdamas skirtingas mokymo programas;
- konkrečią kvalifikaciją galima įgyti tik vieną kartą;
- kvalifikacija gali būti suteikiama pagal kelias programas, bet kiekvienai jų - tik 1 kvalifikacija.

Studijos (Nr, Kvalifikacija, Programa)

Nr	Kvalifikacija	Programa
1	Informatikas	Kompiuterių mokslas
1	Statistikas	Ekonometrika
5	Informatikas	Programų sistemos
5	Statistikas	Matematinė statistika

Studijos (Nr, Kvalifikacija, Programa)

10-25

Antras raktas: Nr, Programa

Nr	Kvalifikacija	Programa
1	Informatikas	Kompiuterių mokslas
1	Statistikas	Ekonometrika
5	Informatikas	Programų sistemos
5	Statistikas	Matematinė statistika

Programa → Kvalifikacija

{Nr, Programa} → Kvalifikacija

{Nr, Kvalifikacija} → Programa

Visi atributai – pirminiai ⇒ *Studijos* yra 3NF.

Tačiau nėra BKNF.

Studijos skaidome (be duomenų praradimo):

Vykdytojų studijos (Nr, Programa)

Nr	Programa
1	Kompiuterių mokslas
1	Ekonometrija
5	Programų sistemos
5	Matematinė statistika

Programos (Programa, Kvalifikacija)

Programa	Kvalifikacija
Kompiuterių mokslas	Informatikas
Ekonometrija	Statistikas
Programų sistemos	Informatikas
Matematinė statistika	Statistikas

11-25

Vykdytojų studijos (Nr, Programa):

{Nr, Programa} → {Nr, Programa}

– trivialioji,

Programos (Programa, Kvalifikacija):

Programa → Kvalifikacija

Negalioja {Nr, Kvalifikacija} → Programa

– ši FP neišsaugoma

Studijos yra nedalomoji. Lentelė - **nedalomoji**, jei jos negalima suskaidyti, išsaugant visas FP.

Jei *Studijos* neskaidome, tai yra duomenų perteklius

Jei skaidome, prarandame FP – reliacinio modelio ribotumas.

12-25

Schematiškai nedalomoji lentelė: $L(A, B, C)$

13-25

Galioja FP: $AB \rightarrow C$ ir $C \rightarrow B$

L turi 2 raktus: AB ir AC

Visi atributai (A, B, C) yra pirminiai ⇒ L yra 3NF

$C \rightarrow B$ netriviali priklausomybė, kurios determinantas C nėra viršraktis ⇒ L nėra BKNF

Jei skaidome pagal Hezo teoremą:

$L(A, B, C) \Rightarrow L_1(\underline{A}, \underline{C})$ ir $L_2(\underline{C}, B)$.

išorinis raktas: $L_1.C \Rightarrow L_2$

tai neišsaugome FP: $AB \rightarrow C$

3.11. Ketvirtoji norminė forma

14-25

Lentelėje $L(A, B, C)$ yra **daugiareikšmė** (angl. *multi-valued*) priklausomybė, $A \twoheadrightarrow B \Leftrightarrow B$ reikšmių aibė, atitinkanti bet kurias A reikšmes, nepriklauso nuo atributų C reikšmių.

DR – priklausomybėje (DRP), vienai atributo reikšmei gali atitikti kelios kito atributo reikšmės. DRP yra FP apibendrinimas, nes kiekviena FP kartu yra ir DRP.

DRP $A \twoheadrightarrow B$ yra **netriviali** lentelėje $L(R)$,

jei nei vienas atributas iš B neįeina į A ir be atributų A ir B lentelėje yra dar ir kitų atributų,

t.y. jei $A \cap B = \emptyset$ ir $A \cup B \subset R$.

Tam, kad lentelėje būtų netriviali DRP reikia, kad joje būtų bent 3 atributai.

15-25

Tarkime, firmos darbuotojai gali įgyti kelias kvalifikacijas ir mokėti kelias užsienio kalbas:

Išsilavinimas (Nr, Kvalifikacija, Kalba)

Nr	Kvalifikacija	Kalba
1	Informatikas	Anglų
1	Informatikas	Vokiečių
1	Fizikas	Anglų
1	Fizikas	Vokiečių

Lentelėje *Išsilavinimas* yra dvi DRP:

$Nr \twoheadrightarrow Kvalifikacija$

$Nr \twoheadrightarrow Kalba$

16-25

Kitas DRP apibrėžimas: $L(A, B, C)$ galioja $A \rightarrow\rightarrow B$, jei bet kurioms 2 eilutėms: $\langle a, b, c \rangle$, $\langle a, b', c' \rangle$ yra dar 2 eilutės: $\langle a, b', c \rangle$, $\langle a, b, c' \rangle$

Teiginys. Jei $L(A, B, C)$ galioja DRP $A \rightarrow\rightarrow B$, tai galioja ir $A \rightarrow\rightarrow C$

Irodymas. Tarkime, $A \rightarrow\rightarrow B$, bet NE $A \rightarrow\rightarrow C$. Jei negalioja $A \rightarrow\rightarrow C$, tai A reikšmei a yra 2 eilutės: $\langle a, b, c \rangle$, $\langle a, b', c' \rangle$, bet nėra $\langle a, b', c \rangle$ arba $\langle a, b, c' \rangle$, t.y. L negalioja $A \rightarrow\rightarrow B$. **Prieštara.**

Netrivialios DRP visuomet sudaro poras, kurios dažnai vaizduojamos taip:

$$A \rightarrow\rightarrow B \mid C$$

Išsilavinimas galioja

$$Nr \rightarrow\rightarrow Kvalifikacija \mid Kalba$$

$L(R)$ yra 4NF FP ir DRP aibės F atžvilgiu



\forall DRP $A \rightarrow\rightarrow B \in F^+$, $A \subset R$, $B \subset R$, arba yra triviali, arba determinantas A yra L viršraktis.

Jei DRP $A \rightarrow\rightarrow B$ determinantas A yra viršraktis, tai ši DRP yra FP.

4NF apibrėžime galima reikalauti, kad \forall netriviali DRP būtų FP.

Kad būtų 4NF, turi būti tik FP, su determinantais – viršrakčiais.

\Rightarrow lentelė, esanti 4NF yra ir BKNF.

Išsilavinimas(Nr , Kvalifikacija, Kalba)

Nr	<u>Kvalifikacija</u>	<u>Kalba</u>
1	Informatikas	Anglų
1	Informatikas	Vokiečių
1	Fizikas	Anglų
1	Fizikas	Vokiečių

Galioja dvi DRP: $Nr \rightarrow\rightarrow Kvalifikacija \mid Kalba$
– netrivialios DRP ir ne FP.
– nėra 4NF

Kadangi lentelėje tarp atributų nėra jokių FP (galioja tik trivialis FP), tai ji yra BKNF.

Teorema (R. Fagin). Tarkime, A , B ir C yra lentelės $L(A, B, C)$ atributų aibės. Tuomet lentelė L gali būti gaunama sujungiant jos projekcijas $L_1(A, B)$ ir $L_2(A, C)$ tada ir tik tada, kai $A \rightarrow\rightarrow B \mid C$.

Kadangi DRP yra FP apibendrinimas, tai Feigino teorema yra Hezo teoremos apibendrinimas.

Išsilavinimas galima suskaidyti į dvi projekcijas:

Kvalifikacijos(Nr , Kvalifikacija)
Kalbos(Nr , Kalba)

Abiejose šiose lentelėse yra po vieną DRP, tačiau abi jos yra trivialis, t.y. juose dalyvauja visi atitinkamos lentelės atributai. Todėl abi šios dvi lentelės yra 4NF.

Papildykime lentelę Kalbos stulpeliu Lygmuo.

Kalbų mokėjimas

Nr	<u>Kalba</u>	<u>Lygmuo</u>
1	Anglų	Puikiai
1	Vokiečių	Gerai
2	Anglų	Gerai

yra tik 1 FP $\{Nr, Kalba\} \rightarrow Lygmuo$, determinantas – raktas – 4NF.

3.13. Norminių formų tarpusavio ryšys

Lentelės 1NF (normalizuotos lentelės)

