Teorija baza podataka
Poopćene relacijske baze podataka

Izv. prof. dr. sc. Markus Schatten

Fakultet organizacije i informatike,
Sveučilište u Zagrebu
Pavlinska 2, 42000 Varaždin
markus.schatten@foi.hr

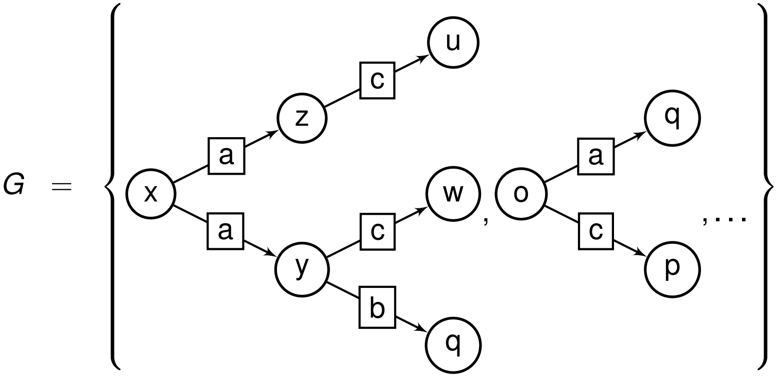
Uvod

- U konvencionalnom relacijskom modelu pretpostavljamo da su relacijske sheme u $1NF$.
- Navedeno znači da su svi atributi sheme relacijske baze podataka jednostavni, tj. domene atributa se sastoje od jednostavnih objekata, koje još nazivamo elementarni ili atomarni objekti.
- Zahtjevom za boljom reprezentacijom dolazimo do relacijskih shema koje su u višim normalnim formama ($2NF$, $3NF$, $BCNF$, $4NF$, $5NF$).
- Razne aplikacije za prirodne i tehničke znanosti i neki logički modeli (na primjer temporalni model, objektno/relacijski model) zahtijevaju relacije čija se shema sastoji i od složenih atributa.
- Koristeći razne konstrukte, na primjer: konstrukte za skupove, liste i relacije, dobivamo skupove složenih ili neelemenatarnih objekata (neatomarni objekti).

Skupovi objekata

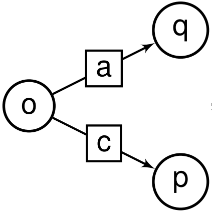
$$
\begin{aligned}
A &= Alf \cup I \cup R \cup \mathrm{Dat} \cup \mathrm{Novac};\\
S &= \{\{a, b\}, \varnothing, \ldots\};\\
L &= \{\langle a, b, a, c\rangle, \langle\rangle, \ldots\};\\
R &= \left\{ \begin{array}{c|cc} r_1 & A & B \\ \hline & 1 & 2 \\ & 3 & 1 \end{array}, \begin{array}{c|ccc} r_2 & B & C & D \\ \hline & 4 & 1 & 1 \\ & 7 & 5 & 1 \end{array}, \ldots \right\};
\end{aligned}
$$

Skupovi objekata



Skupovi objekata

$$K = \left\{ \left\langle \begin{array}{c|cc} r_1 & A & B \\ \hline & 1 & 2 \\ & 3 & 1 \end{array}, \{a, b\}, \right. \right.$$



$$\left. \left. , \right\rangle, \ldots \right\}$$

Skupovi objekata

A je skup atomskih objekata (sadrži alfanumerički tip, cjelobrojni tip, realne brojeve, datum i td.). S, L, R, i K se sastoje od složenih objekata: S je skup čiji elementi su skupovi, L je skup čiji elementi su liste, R je skup koji se sastoji od relacija, G, je skup koji se sastoji od grafova, a K je skup koji se sastoji od elemenata koji su izgrađeni kombinacijom elemenata iz skupova S, L, R i G.

Jednostavni i složeni atributi

Definicija

Jednostavni i složeni atributi
Atribut A je jednostavan atribut ako se domena atributa A, $Dom(A)$, sastoji samo od jednostavnih objekata; u protivnome, A je složen atribut.

Istaknimo da $Dom(A)$ složenog atributa A sadrži barem jedan složeni objekt. Podsjetimo se da je relacijska shema (R, F) u $1NF$ ako je svaki atribut iz R jednostavan.

Poopćena normalna forma

Definicija

Poopćena normalna forma *PONF*
(R, F) *je u poopćenoj normalnoj formi, PONF, ako je svaki atribut iz R jednostavan ili složen.*

Iz definicije *PONF* vidi se da je *PONF* poopćenje $1NF$.

U daljnjem pretpostavljamo da su vrijednosti složenih atributa relacije. Dakle imamo, u tabličnom prikazu relacije, pojavu tablica unutar tablica.

PONF relacije

U nastavku ćemo skup atributa $R = \{A_1, A_2, \dots, A_k\}$ pisati u obliku $R(A_1, A_2, \dots, A_k)$.

Definicija

Ako je (R, F) u X normalnoj formi (XNF), onda za relaciju r nad R kažemo da je XNF relacija.

Primjer I

Neka je zadana relacija r nad shemom R(Radnik#, Mentor), gdje je R u $1NF$.

r	Radnik#	Mentor
	R_1	Murn
	R_2	Murn
	R_3	Murn
	R_4	Lovrić
	R_5	Lovrić

Relacija r je $1NF$ relacija.

Primjer II

Informacijski sadržaj relacije r moguće je prikazati kompaktnije koristeći relaciju r_1 nad relacijskom shemom R_1(Rad(Radnik#),Mentor), gdje je R_1 u *PONF*. Uočite da je atribut Rad složen.

r	Rad(Radnik#)	Mentor
	R_1, R_2, R_3	Murn
	R_4, R_5	Lovrić

Relacija r_1 je *PONF* relacija.

Primjer

Razmotrimo *PONF* relaciju r_s

$$
\begin{array}{c|cc}
r_s & A & D(B\ C) \\
\hline
 & 1 & \boxed{\begin{array}{cc} 2 & 2 \\ 1 & 2 \end{array}} \\
 & 3 & \boxed{\begin{array}{cc} 1 & 1 \end{array}}
\end{array}
$$

Semantika relacije r_s je kao što slijedi.
Shema relacije r je $sh(r) = R(A, D(B, C))$; atribut A je jednostavan, a atribut D je složen; vrijednosti atributa D su relacije nad (B, C). Neka je t_1 prvi red relacije r_s, a t_2 drugi red relacije r_s. Tada vrijedi

$$
t_1[A] = 1,\ t_1[D] = \begin{array}{cc} B & C \\ \hline 2 & 2 \\ 1 & 2 \end{array},\ t_2[A] = 3,\ t_2[D] = \begin{array}{cc} B & C \\ \hline 1 & 1 \end{array},\ t_1[D] \neq t_2[D].
$$

Particijska normalna forma

Dodatni zahtjev za jezgrovitijom tj. kompaktnijom reprezentacijom ostvaruje se particijskom normalnom formom.

Uvodni primjer I

Neka je zadana relacija

r	$S\#$	Prezime	Predmet(P-ime Ocjena))
	S_1	Mrak	Mat 3 Fiz 4
	S_1	Mrak	BP 3

U r ne vrijedi funkcijska zavisnost $S\#$, Prezime $\rightarrow$ Predmet. Dakle jednostavni atributi $S\#$, *Prezime* ne određuju funkcijski složeni atribut *Predmet*.

Uvodni primjer II

Kompaktnija reprezentacija relacije r dana je u relaciji r_1.

r_1	$S\#$	Prezime	Predmet(P-ime Ocjena))
	S_1	Mrak	Mat 3 Fiz 4 BP 3

U relaciji r_1 vrijedi $S\#$, Prezime $\to$ Predmet.

Particijska normalna forma

Definicija

PNF

Neka je zadana PONF relacija $r(X\ Y)$*, gdje su u* X *jednostavni atributi, a u* Y *su složeni atributi. Kažemo da je* r *u PNF ako* $X \rightarrow Y$ *vrijedi u* r*, i* $(\forall t \in R)(\forall A \in Y)[$*relacija* $t[A]$ *je u PNF*$]$*.*

Dakle, *PNF* zahtijeva da jednostavni atributi funkcijski određuju složene atribute, a taj zahtjev je na snazi i za svaku ugniježđenu relaciju u relaciji r. Slobodnije kazano: *PNF* uvjet vrijedi na putu iz vana prema unutra do kraja ugniježđenja.

Relacija r_1 iz uvodnog primjera je u *PNF*, jer jednostavni atributi funkcijski određuju složene atribute tj. funkcijska zavisnost $S\#$, Prezime $\rightarrow$ Predmet vrijedi u r_1.

Primjer I

Razmotrimo slijedeći prikaz relacije r_1 iz uvodnog primjera, koji je označen sa r_2. Neka relacija r_2 ima ovakav oblik:

r_2	$S\#$	Prezime	Predmet(Pr(Ime)Ocjena))
	S_1	Mrak	Mat 3 Fiz 4 BP 3

Primjer II

Ovdje su atributi $S\#$, Prezime i Ocjena jednostavni atributi, a atributi Predmet i Pr su složeni. Dakle, unutar složenog atributa Predmet imamo složen atribut Pr i jednostavan atribut Ocjena.

Iako vrijedi, na vanjskoj razini, funkcijska zavisnost $S\#$, Prezime $\rightarrow$ Predmet, na unutarnjoj razini u relaciji nad Predmet ne vrijedi funkcijska zavisnost Ocjena $\rightarrow$ Pr. Zaključujemo da r_1 nije u PNF.

Primjer III

Relaciju r_2 iz prošlog primjera možemo transformirati u oblik relacije, koja je u PNF. Prikaz označimo sa r_3 .

r_3	$S\#$	Prezime	Predmet(Pr(Ime)Ocjena))
	S_1	Mrak	$\boxed{\text{Mat BP}}$ 3 $\boxed{\text{Fiz}}$ 4

U r_3 vrijedi $S\#$, Prezime $\rightarrow$ Predmet i Ocjena $\rightarrow$ Pr. Zaključujemo da je r_2 u PNF .

Propozicija ($1NF$, PNF)

Propozicija

($1NF$, PNF)
Neka je (R, F) u $1NF$. Tada je (R, F) u PNF.

Dokaz.

Iz pretpostavke da je (R, F) u $1NF$ slijedi da su svi atributi iz R jednostavni. Dakle, skup jednostavnih atributa $X = R$ i skup složenih atributa $Y = \cup$. Kako uvijek vrijedi $X \to \varnothing$ zaključujemo da je (R, F) u PNF. $\square$

Kao posljedicu propozicije ($1NF$, PNF) imamo: ako je $r(R)$ $1NF$ relacija, onda je $r(R)$ PNF relacija.

Relacijski operatori za *PONF* relacije

- Nakon razmatranja *PONF* relacija tj. strukturne komponente poopćenih relacijskih baza podataka (PORBP), u nastavku opisuje se operativna komponenta. Dakle, bit će razmatrani relacijski operatori za *PONF* relacije.
- Razmatranje počinjemo opisom operatora grupiranja, Gr, i operatora rastavljanja, Ra.
- Nakon toga, posvetit ćemo pažnju proširenju konvencionalnih relacijskih operatora (operatori na 1*NF* relacijama) na *PONF* relacije.
- Proširenja će biti označena indeksiranjem slovom p standardnih oznaka za konvencionalne operatore: $\overset{p}{\cup}$, $\overset{p}{\cap}$, $\overset{p}{-}$, $\overset{p}{\bowtie}$, $\overset{p}{\otimes}$, $\overset{p}{\Pi}$, $\overset{p}{\sigma}$, $\overset{p}{\delta}$.

Grupiranje *Gr*

Definicija

Grupiranje *Gr*

Neka je zadana relacija $r(B_1, \dots, B_k, B_{k+1}, \dots, B_n)$*; neka je C atribut koji se ne pojavljuje u* $sh(r)$*. Grupiranje relacije r po atributima* $B_{k+1}, \dots, B_n$ *agregiranim u C, je relacija* $Gr_{C(B_{k+1},\dots,B_n)}(r)$*:*

(a) $sh(Gr_{C(B_{k+1},\dots,B_n)}(r)) = (B_1, \dots, B_k, C(B_{k+1}, \dots, B_n))$

(b) *Slogovi u Gr dobiju se iz slogova iz r agregiranjem slogova koji imaju jednake vrijednosti na* $B_1, \dots, B_k$*.*

Formalno, $t \in Gr_{C(B_{k+1},\dots,B_n)}(r)$ *ako*

(1) $\exists u \in r(t[B_1, \dots, B_k] = u[B_1, \dots, B_k])$, *i*

(2) $t[C] = \{v[B_{k+1}, \dots, B_n] \mid v \in r \text{ i } v[B_1, \dots, B_k] = t[B_1, \dots, B_k]\}$

Primjer I

Neka je zadana relacija

r	A	B	C
	2	1	1
	2	0	1
	3	1	2

Tada je

$Gr_{D(B,C)}(r)$	A	$D(B\ C)$
	2	$\begin{array}{\|cc\|} \hline 1 & 1 \\ 0 & 1 \\ \hline \end{array}$
	3	$\begin{array}{\|cc\|} \hline 1 & 2 \\ \hline \end{array}$

Primjer II

Komentar: agregirani su jednostavni atributi B, C u novi, složeni atribut D. Vrijednosti složenog atributa D su relacije nad BC. Za sve redove iz r, koji su jednaki na atributu A, obavlja se grupiranje njihovih BC vrijednosti iz r u redove relacije nad BC. Primijetite da je zadnji red $(3, 1, 2)$ iz relacije r grupiran u red $(3, \boxed{1 \quad 2})$ u relaciji $Gr_{D(B,C)}(r)$. Iz ovog primjera vidi se da se postupkom grupiranja postiže jezgrovitiji prikaz podataka.

U izvjesnom smislu, koji će biti poslije objašnjen, inverzna operacija operacije grupiranja je sljedeća operacija rastavljanja *Ra*.

Rastavljanje *Ra*

Definicija

Rastavljanje *Ra*

Neka je zadana relacija $r(B_1, \dots, B_m, B(A_1, \dots, A_k))$, gdje je B složeni atribut. Rastavljanje od r s obzirom na B je relacija $Ra_B(r)$ definirana ovako:

(1) $sh(Ra_B(r)) = (B_1, \dots, B_m, A_1, \dots, A_k)$

(2) $t \in Ra_B(r)$ *ako* $\exists u \in r(t[B_1, \dots, B_m] = u[B_1, \dots, B_m]$ *i* $t[A_1, \dots, A_k] \in u[B])$

Primjer

Neka je zadana relacija $s = Gr_{D(B,C)}(r)(A \quad D(BC))$ iz prošlog primjera.

$s = Gr_{D(B,C)}(r)$	A	$D(B\ C)$
	2	$\boxed{\begin{array}{cc} 1 & 1 \\ 0 & 1 \end{array}}$
	3	$\boxed{\begin{array}{cc} 1 & 2 \end{array}}$

Tada $Ra_D(s)$ ima oblik:

$Ra_D(s)$	A	B	C
	2	1	1
	2	0	1
	3	1	2

Uočite da smo dobili početnu relaciju r iz prethodnog primjera.

Primjer I

Za relaciju $r(A\ B\ C(DE(F\ G)))$ odredimo $Ra_{E(F,G)}(r)(A\ B\ C(D\ F\ G))$.

$$
\begin{array}{c|cc|c|}
r & A & B & C(DE(F\ G)) \\
\hline
 & 1 & 2 & \boxed{\begin{array}{cc} 3 & \boxed{\begin{array}{cc} 2 & 2 \\ 1 & 0 \end{array}} \end{array}} \\
 & 1 & 2 & \boxed{\begin{array}{cc} 4 & \boxed{\begin{array}{cc} 2 & 2 \\ 1 & 1 \\ 0 & 1 \end{array}} \end{array}}
\end{array}
$$

Primjer II

Shema relacije r sastoji se redom od jednostavnih atributa A i B, složenog atributa C čija se shema sastoji od jednostavnog atributa D i složenog atributa E čija se shema sastoji od jednostavnih atributa F i G. Rastavljanje je potrebno izvršiti po složenom atributu E. Dobivamo relaciju $Ra_{E(F,G)}(r)(A\ B\ C(D\ F\ G))$

$$
\begin{array}{c|ccc}
Ra_{E(F,G)}(r) & A & B & C(DE(F\ G)) \\
\hline
 & 1 & 2 & \boxed{\begin{array}{ccc} 3 & 2 & 2 \\ 3 & 1 & 0 \end{array}} \\
 & 1 & 2 & \boxed{\begin{array}{ccc} 4 & 2 & 2 \\ 4 & 1 & 1 \\ 4 & 0 & 1 \end{array}}
\end{array}
$$

Propozicije

U sljedeće dvije propozicije utvrđujemo odnos između operatora grupiranja *Gr* i operatora rastavljanja *Ra*.

Propozicija

(Ra je inverz Gr)
Operator Ra je inverz od Gr, tj. $r = Ra_A(Gr_{A(Y)})(r))$

Propozicija

(Gr nije inverz od Ra)
Operator Gr nije inverz od Ra, tj. postoji relacija r takva da je

$$r \neq Gr_{A(Y)}(RaA(Y)(r)).$$

Dokaz I

Neka je zadana relacija

r	Izlet	Posjeta(Grad Dan)
	l_1	Split 1
	l_1	Rijeka 1 Pula 2
	l_2	Zadar 3

Izračunajmo $Gr_{\text{Posjeta(Grad Dan)}}(Ra_{\text{Posjeta}}(r))$. Dobivamo

Dokaz II

$Ra_{\text{Posjeta(Grad Dan)}}(r)$	Izlet	Grad	Dan
	I_1	Split	1
	I_1	Rijeka	1
	I_1	Pula	2
	I_2	Zadar	3

$Gr_{\text{Posjeta(Grad Dan)}}(Ra_{\text{Posjeta}}(r))$	Izlet	Posjeta(Grad Dan)	
	I_1	Split 1 Rijeka 1 Pula 2	$\neq r$
	I_2	Zadar 3	

Propozicija

Slabost relacije r, koju smo koristili u dokazu prošle propozicije, ogleda se u činjenici da relacija r nije u particijskoj normalnoj formi (PNF). Naime, zavisnost $I1 \rightarrow Posjeta$ ne vrijedi u relaciji r . U slučaju da je relacija u PNF, onda je Gr inverz od Ra. O tome se govori u sljedećoj propoziciji.

Propozicija

(PNF, Gr, Ra)
Neka je $r(X\ A(Y))$ u PNF, gdje je $X \cap Y = \varnothing$ i $A \notin XY$. Tada vrijedi $r = Gr_{A(Y)}(Ra_{A(Y)}(r))$.

Zatvorenost obzirom na *Ra* i *Gr*

Dodatno dobro svojstvo *PNF* relacija sadržano je u činjenici da je klasa *PNF* relacija zatvorena s obzirom na rastavljanje *Ra* .

Propozicija

(zatvorenost PNF s obzirom na Ra)
Klasa PNF relacija je zatvorena s obzirom na Ra.

Prema tome, ako je r *PNF* relacija, onda je i $Ra(r)$ *PNF* relacija. Međutim, klasa *PNF* relacija nije zatvorena s obzirom na grupiranje *Gr* .

Propozicija

(nezatvorenost s obzirom na Gr)
Klasa PNF relacija nije zatvorena s obzirom na Gr.

Dokaz I

Razmotrimo *PNF* relaciju r:

r	A	B	$C(D\ E)$
	1	2	2 2
	1	3	1 0

Da je r *PNF* relacija proizlazi iz činjenice da u relaciji r vrijedi funkcijska zavisnost $AB \to C$.

Dokaz II

Izvršimo grupiranje relacije r po atributu B agregiranjem u atribut B_1 . Dobivamo relaciju

$Gr_{B_1}(B)](r)$	A	$B_1(B)$	$C(D\ E)$
	1	$\boxed{2}$	$\boxed{2\quad 2}$
	1	$\boxed{3}$	$\boxed{1\quad 0}$

Budući da funkcijska zavisnost $A \to B_1 C$ ne vrijedi u $Gr(r)$, zaključujemo da $Gr(r)$ nije u PNF. $\square$

Primjer I

Promotrimo malo detaljnije relaciju r , korištenu u dokazu propozicije.

r	A	B	$C(D\ E)$
	1	2	$\boxed{2 \quad 2}$
	1	3	$\boxed{1 \quad 0}$

Iako je r *PNF* relacija, slabost relacije r proizlazi iz činjenice da funkcijska zavisnost $A \to C$ ne vrijedi u r; jasno je onda da grupiranjem po atributu B u atribut B_1 dolazimo u situaciju da $A \to B_1 C$ također ne vrijedi u r , a time je narušeno svojstvo *PNF*. Izbjegavanjem spomenute slabosti dolazi se u situaciju zatvorenosti *PNF* s obzirom na grupiranje. Navedeno je iskazano u sljedećoj propoziciji:

Propozicija

Propozicija

(PNF, Gr)
Neka je $r(X\ Y\ Z)$ *takva relacija gdje je* X *skup jednostavnih atributa,* Y *je skup složenih atributa, te* $Z = R – XY$. *Tada vrijedi:* $Gr[A(Z)](r)$ *je u PNF akko* $X \rightarrow Y$ *vrijedi u* r.

Dokaz.

Agregacijom atributa iz Z u složeni atribut $A(Z)$ dobiva se $sh(Gr_{A(Z)}(r) = XYA(Z)$, gdje su u X svi jednostavni atributi, a $YA(Z)$ su svi složeni attributi. Iz pretpostavke da $X \rightarrow Y$ vrijedi u relaciji r , proizlazi da $X \rightarrow YA(Z)$ vrijedi u $Gr_{A(Z)}(r)$; prema tome, $Gr_{A(Z)}(r)$ je u *PNF*. Obratno, ako je $Gr_{A(Z)}(r)$ u *PNF* onda u $Gr_{A(Z)}(r)$ vrijedi $X \rightarrow YA(Z)$. Zbog toga, $X \rightarrow Y$ vrijedi u r. $\square$

Relacijski operatori za PONF relacije

Prelazimo na proširenje konvencionalnih relacijskih operatora tako da dobijemo relacijske operatore za *PONF* relacije uz ispunjenje svojstva da je klasa *PNF* relacija zatvorena s obzirom na definirana proširenja.

Primjer - Unija I

Neka su zadane relacije $r_1(A, B(C, D))$, $r_2(A, B(C, D))$

r_1	A	$B(CD)$
	1	2 2 1 0
	2	1 1

r_2	A	$B(CD)$
	1	2 1 1 0
	3	2 2

Relacije r_1 i r_2 su PNF.

Primjer - Unija II

Ako izračunamo $r_1 \cup r_2$, dobivamo

$r_1 \cup r_2$	A	$B(CD)$
	1	$\boxed{\begin{array}{cc} 2 & 2 \\ 1 & 0 \end{array}}$
	2	$\boxed{\begin{array}{cc} 1 & 1 \end{array}}$
	1	$\boxed{\begin{array}{cc} 2 & 1 \\ 1 & 0 \end{array}}$
	3	$\boxed{\begin{array}{cc} 2 & 2 \end{array}}$

Primjer - Unija III

Budući da funkcijska zavisnost $A \to B$ ne vrijedi u $r_1 \cup r_2$, zaključujemo da $r_1 \cup r_2$ nije u PNF.

Prema tome, klasa PNF relacija nije zatvorena s obzirom na konvencionalnu uniju $\cup$.

Navedenu slabost konvencionalne unije možemo otkloniti tako da za redove iz $r_1 \cup r_2$ koji su jednaki na atributu A , izvršimo uniranje pripadnih B vrijednosti. Na taj način, dobivamo relaciju $r_1 \overset{p}{\cup} r_2(AB(CD))$, koja je u PNF.

Primjer - Unija IV

$r_1 \overset{p}{\cup} r_2$	A	$B(CD)$
	1	$\boxed{\begin{array}{cc} 2 & 2 \\ 1 & 0 \\ 2 & 1 \end{array}}$
	2	$\boxed{\begin{array}{cc} 1 & 1 \end{array}}$
	3	$\boxed{\begin{array}{cc} 2 & 2 \end{array}}$

Unija $\overset{p}{\cup}$

Definicija

Unija $\overset{p}{\cup}$

Neka su zadane relacije $r_1(X\ Y)$, $r_2(X\ Y)$, gdje su u X jednostavni, a u Y složeni atributi.

1. $sh(r_1 \overset{p}{\cup} r_2) = SH(r_1) = SH(r_2)$
2. $t \in r_1 \overset{p}{\cup} r_2$ *ako i samo ako:*
    1. $t \in r_1$ *i* $(\forall t_1 \in r_2)(\exists A_i \in X)(t[A_i] \neq t_1[A_i])$, *ili*
    2. $t \in r_2$ *i* $(\forall t_1 \in r_1)(\exists A_i \in X)(t[A_i] \neq t_1[A_i])$, *ili*
    3. $(\exists t_1 \in r_1)(\exists t_2 \in r_2)(\forall A_i \in X)(\forall B_j \in Y):$
       $t[A_i] = t_1[A_i] = t_2[A_i]$ *i*
       $t[B_j] = t_1[B_j] \overset{p}{\cup} t_2[B_j]$

Pojašnjenje

Definicijom unije $\overset{p}{\cup}$ se prvo utvrđuje u točki 1. što je ime i shema operatora $\overset{p}{\cup}$, a zatim se u točki 2. navodi kako se dobiju pripadni redovi: svi redovi iz r_1 i r_2 , koji su različiti na skupu svih jednostavnih atributa X , uključuje se u $r_1 \overset{p}{\cup} r_2$ (točke (a) i (b)), a za sve one redove koji su jednaki na skupu X , treba izvršiti uniju $\overset{p}{\cup}$ pripadnih relacija nad svim složenim atributima iz skupa složenih atributa Y .

Slobodnije kazano operator $\overset{p}{\cup}$ ‘radi tako da se ide od jednostavnih atributa prema složenim atributima’.

Primjer I

Neka su opet zadane relacije $r_1(A, B(C, D))$, $r_2(A, B(C, D))$ iz prethodnog primjera.

$$
\begin{array}{c|cc}
r_1 & A & B(CD) \\
\hline
 & 1 & \boxed{\begin{array}{cc} 2 & 2 \\ 1 & 0 \end{array}} \\
 & 2 & \boxed{\begin{array}{cc} 1 & 1 \end{array}}
\end{array}
\qquad
\begin{array}{c|cc}
r_2 & A & B(CD) \\
\hline
 & 1 & \boxed{\begin{array}{cc} 2 & 1 \\ 1 & 0 \end{array}} \\
 & 3 & \boxed{\begin{array}{cc} 2 & 2 \end{array}}
\end{array}
$$

Računajući uobičajeni presjek $r_1 \cup r_2$, dobivamo praznu relaciju

$$
\begin{array}{c|c}
r_1 \cap r_2 & A\ B(CD) \\
\hline
 & \varnothing
\end{array}
$$

Primjer II

U dobivenoj relaciji $r_1 \cup r_2$ izgubljena je informacija da prvi redovi u relaciji r_1 i r_2 imaju iste vrijednosti na jednostavnom atributu A, te da na složenom atributu B njima pripadne relacije imaju neprazan presjek. Poželjan rezultat dan je relacijom $r_1 \overset{p}{\cup} r_2$, koja izgleda ovako:

$r1 \overset{p}{\cap} r2$	A	$B(C\ D)$
	1	$\boxed{1 \quad 0}$

Prema tome, sve redove t iz $r_1 \overset{p}{\cup} r_2$ dobijemo tako da pronađemo sve redove t_1 iz r_1, t_2 iz r_2 takve da vrijedi: $t_1[A] = t_2[A]$ i $t_1[B] \overset{p}{\cup} t_2[B] \neq \emptyset$, a zatim stavimo: $t[A] = t_1[A] = t_2[A]$ i $t[B] = t_1[B] \overset{p}{\cup} t_2[B]$.

Presjek $\overset{p}{\cap}$

Definicija

Presjek $\overset{p}{\cap}$

Neka su zadane relacije $r_1(XY)$, $r_2(XY)$, *gdje je* X *skup svih jednostavnih atributa iz* R, *a* Y *skup svih složenih atributa iz* R.

1. $sh(r_1 \overset{p}{\cap} r_2) = SH(r_1) = SH(r_2)$
2. $t \in (r_1 \overset{p}{\cap} r_2)$ *ako i samo ako* $(\exists t_1 \in r_1)(\exists t_2 \in r_2)(\forall A_i \in X)(\forall B_j \in Y)$*:* $t[A_i] = t_1[A_i] = t_2[A_i]$ *i* $t[B_j] = t_1[B_j] \overset{p}{\cap} t_2[B_j]$ *i* $t[B_j] \neq \varnothing$.

Dakle, redovi t_1 i t_2 daju u presjeku t ako se t_1 i t_2 podudaraju na svim jednostavnim atributima X, i imaju neprazan prošireni presjek na svakom složenom atributu iz Y.

Primjer I

Neka su opet zadane relacije $r_1(A, B(C, D))$, $r_2(A, B(C, D))$ iz naša dva prošla primjera.

$$
\begin{array}{c|cc}
r_1 & A & B(CD) \\
\hline
 & 1 & \boxed{\begin{array}{cc} 2 & 2 \\ 1 & 0 \end{array}} \\
 & 2 & \boxed{\begin{array}{cc} 1 & 1 \end{array}}
\end{array}
\qquad
\begin{array}{c|cc}
r_2 & A & B(CD) \\
\hline
 & 1 & \boxed{\begin{array}{cc} 2 & 1 \\ 1 & 0 \end{array}} \\
 & 3 & \boxed{\begin{array}{cc} 2 & 2 \end{array}}
\end{array}
$$

Računajući uobičajenu razliku $r_1 - r_2$, dobivamo relaciju r_1:

Primjer II

$$
\begin{array}{c|cc}
r_1 - r_2 & A & B(C\ D) \\
\hline
 & 1 & \boxed{\begin{array}{cc} 2 & 2 \\ 1 & 0 \end{array}} \\
 & 2 & \boxed{\begin{array}{cc} 1 & 1 \end{array}} \\
\end{array}
$$

I dok je drugi red u rezultatu $r_1 - r_2$ očekivan, prvi red na atributu B ima vrijednost relaciju s_1 čiji se drugi red $(1, 0)$ nalazi i u relaciji $s2$ koja je komponenta prvog reda u relaciji r_2 :

$$
s_1 = \boxed{\begin{array}{cc} 2 & 2 \\ 1 & 0 \end{array}} \qquad s_2 = \boxed{\begin{array}{cc} 2 & 1 \\ 1 & 0 \end{array}}
$$

Primjer III

Razlika relacija $s_1 - s_2 = \boxed{2 \quad 2}$ treba biti druga komponenta prvog reda, tako da je korektan rezultat dan relacijom $r_1 \overset{p}{-} r_2$

$r_1 \overset{p}{-} r_2$	A	$B(C\ D)$
	1	$\boxed{2 \quad 2}$
	2	$\boxed{1 \quad 1}$

Razlika $\overset{p}{-}$

Definicija

Razlika $\overset{p}{-}$

Neka su zadane relacije $r_1(X\ Y)$, $r_2(X\ Y)$, gdje je X skup svih jednostavnih atributa iz R, a Y skup svih složenih atributa iz R.

1. $sh(r_1 \overset{p}{-} r_2) = SH(r_1) = SH(r_2)$
2. $t \in (r_1 \overset{p}{-} r_2)$ *ako i samo ako*
    1. $t \in r_1$ *i* $(\forall t_1 \in r_2)(\exists A_i \in X) : t[A_i] \neq t_1[A_i]$, *ili*
    2. $(\exists t_1 \in r_1)(\exists t_2 \in r_2)(\forall A_i \in X)(\forall B_j \in Y)$:

       $t[A_i] = t_1[A_i] = t_2[A_i]$ *i* $t[B_j] = t_1[B_j] \overset{p}{-} t_2[B_j]$ *i* $t[B_j] \neq \emptyset$.

Prema tome, u $\overset{p}{-}$ su oni redovi iz r_1 koji se ne podudaraju niti s jednim slogom iz r_2 na skupu jednostavnih atributa X, ili ako se podudaraju, onda imaju nepraznu, proširenu razliku na svakom složenom atributu iz Y.

Primjer I

Neka su zadane relacije $r_1(A\ B\ C(D\ E))$ i $r_2(F\ B\ C(D\ E))$.

$$\begin{array}{c|ccc} r_1 & A & B & C(D\ E) \\ \hline & a_1 & b_1 & \boxed{\begin{array}{cc} d_1 & e_1 \\ d_2 & e_2 \end{array}} \\ & a_2 & b_2 & \boxed{\begin{array}{cc} d_3 & e_1 \end{array}} \end{array} \qquad \begin{array}{c|ccc} r_2 & F & B & C(D\ E) \\ \hline & f_1 & b_1 & \boxed{\begin{array}{cc} d_1 & e_1 \\ d_1 & e_2 \end{array}} \\ & f_2 & b_2 & \boxed{\begin{array}{cc} d_3 & e_2 \\ d_1 & e_2 \end{array}} \end{array}$$

Primjer II

Uobičajeni prirodni spoj $r_1 \bowtie r_2$, gdje se spajaju svi redovi iz r_1 i r_2 koji se podudaraju na zajedničkim atributima relacija r_1 i r_2 , izgleda ovako:

$r_1 \bowtie r_2$	$A\ B\ F\ C(D\ E)$
	$\varnothing$

Naime, kako nema redova iz r_1 i r_2 koji se podudaraju na zajedničkim atributima B i C, rezultat je prazna relacija nad $ABFC(D,E)$. Međutim, u ovom primjeru imamo pojavu gubljenja informacije: prvi redovi iz r_1 i r_2 podudaraju se na jednostavnom atributu B, a imaju neprazan presjek pripadnih relacija nad složenim atributom C.

Primjer III

Zbog toga, korektan rezultat je dan relacijom $r_1 \overset{p}{\bowtie} r_2$, gdje je otklonjena spomenuta slabost.

$r_1 \overset{p}{\bowtie} r_2$	A	B	F	$C(D\ E)$
	a_1	b_1	f_1	$\boxed{d_1 \quad e_1}$

Prirodni spoj $\overset{p}{\bowtie}$

Definicija

Prirodni spoj $\overset{p}{\bowtie}$

Za relacije $r_1(R_1), r_2(R2), X = \{A \in R_1 \cap R_2 \mid A \text{ je složen atribut}\}$, $Y = R_1 - X$, $Z = R_2 - X$ *, operator prirodni spoj* $\overset{p}{\bowtie}$ *definiran je kao što slijedi.*

1. $sh(r_1 \overset{p}{\bowtie} r_2) = R_1 \cup R_2 = (Y\ X\ Z)$
2. $t \in (r_1 \overset{p}{\bowtie} r_2)$ *ako i samo ako*

   $(\exists t_1 \in r_1)(\exists t_2 \in r_2)[t[Y] = t_1[Y] \text{ i } t[Z] = t_2[Z] \text{ i } t[X] = t_1[X] \overset{p}{\cap} t_2[X] \neq \varnothing$

Primjer I

Neka je zadana relacija $r(A\ B\ C(D, E))$

$$
\begin{array}{c|ccc}
r & A & B & C(D\ E) \\
\hline
 & 0 & 0 & \boxed{\begin{array}{cc} 1 & 1 \end{array}} \\
 & 0 & 1 & \boxed{\begin{array}{cc} 1 & 0 \\ 2 & 1 \end{array}} \\
 & 1 & 1 & \boxed{\begin{array}{cc} 1 & 1 \end{array}}
\end{array}
$$

Relacija r je u *PNF*, jer funkcijska zavisnost $AB \to C$ vrijedi u r.

Primjer II

Izračunajmo uobičajenu projekciju $\Pi_{AC}(r)$. Imamo

$\Pi_{AC}(r)$	A	$C(D\ E)$
	0	1 1
	0	1 0 2 1
	1	1 1

Relacija $\Pi_{AC}(r)$ nije u *PNF*, jer funkcijska zavisnost $A \to C$ ne vrijedi u $\Pi_{AC}(r)$.

Primjer III

Prikažimo sad svaki red iz $\Pi_{AC}(r)$ u obliku odgovarajuće relacije nad $AC(D, E)$. Tako dobivamo relacije r_1, r_2, r_3:

$$
\begin{array}{c|cc}
r_1 & A & C(D\ E) \\
\hline
 & 0 & \boxed{\begin{array}{cc} 1 & 1 \end{array}}
\end{array}
\qquad
\begin{array}{c|cc}
r_2 & A & C(D\ E) \\
\hline
 & 0 & \boxed{\begin{array}{cc} 1 & 0 \\ 2 & 1 \end{array}}
\end{array}
\qquad
\begin{array}{c|cc}
r_3 & A & C(D\ E) \\
\hline
 & 1 & \boxed{\begin{array}{cc} 1 & 1 \end{array}}
\end{array}
$$

Ako izračunamo $r_1 \overset{p}{\cup} r_2 \overset{p}{\cup} r_3$, onda dobivamo relaciju $\overset{p}{\Pi}_{AC}(r)$

Primjer IV

$\overset{p}{\Pi}_{AC}(r)$	A	$C(D\ E)$
	0	1 1 1 0 2 1
	1	1 1

Budući da funkcijska zavisnost $A \to C$ vrijedi u $\overset{p}{\Pi}_{AC}(r)$, zaključujemo da je $\overset{p}{\Pi}_{AC}(r)$ u PNF.

Primjer V

Zaključimo da klasa *PNF* relacija nije zatvorena s obzirom na uobičajenu projekciju Π.

Navedenu slabost projekcije Π otklanjamo tako da za sve redove iz $\Pi_{AC}(r)$ uvedemo prikaz u obliku odgovarajuće relacije nad AC, a zatim izvršimo proširenu uniju $\overset{p}{\cup}$ dobivenih relacija. Na taj način, dobivamo relaciju $\overset{p}{\Pi} AC(r)$, koja je u *PNF*.

Projekcija $\overset{p}{\Pi}$

Definicija

Projekcija $\overset{p}{\Pi}$

Za relaciju $r(R)$, $X \subseteq R$, $X \neq \varnothing$, *operator* $\overset{p}{\Pi}$ *definira se na sljedeći način:*

1. $sh(\overset{p}{\Pi}_X(r)) = X$
2. *svaki red iz* $\Pi_X(r)$ *prikažimo u obliku relacije* $r_t(X)$, *koja sadrži samo red* t; *tada je* $\overset{p}{\Pi}_X(r) = \overset{p}{\cup}\{r_t(X)\}$, $t \in \Pi_X(r)$ $\overset{p}{\Pi}_V(r) = \overset{p}{\cup}\{t\}$, *gdje* $t \in \Pi_V(r)$

Dakle, u koraku (1) dano je ime i shema za $\overset{p}{\Pi}_X(r)$; korak (2) kaže da se prvo računa uobičajena projekcija $\Pi_X(r)$, te se, nakon toga, svaki red "t" iz $\Pi_X(r)$ prikazuje u obliku relacije $r_t(X)$, koja sadrži samo red t , a na kraju se $\overset{p}{\Pi}_X(r)$ uvedimo oznake $\overset{p}{\cup}$ dobivenih relacija $r_t(X)$.

Primjer I

Neka imamo relacije $r(A\ B\ C(D, E))$ i $r_1(C(D, E))$.

r	A	B	$C(D\ E)$
	a_1	1	$d_1\ \ e_1$ $d_2\ \ e_2$
	a_2	2	$d_3\ \ e_1$

r_1	$C(D\ E)$
	$d_1\ \ e_1$

Također, neka je zadan uvjet selekcije $F_p = (B \geq 2) \wedge [(C \overset{p}{\cap} r_1) \neq \emptyset]$. Odredimo $\overset{p}{\sigma}_{F_p}(r)$.

Primjer II

Neka je t_1 prvi, a t_2 drugi red relacije r . Trebamo izračunati $F_p(t_1)$ i $F_p(t_2)$. Uvedimo oznake

$$s_1 = \boxed{\begin{array}{cc} d_1 & e_1 \\ d_2 & e_2 \end{array}} \qquad s_2 = \boxed{\begin{array}{cc} d_3 & e_1 \end{array}}$$

Sad računamo:

$$F_p(t_1) = (1 \geq 2) \wedge [(s_1 \overset{p}{\cap} r_1) = \varnothing] = \bot \wedge \bot = \bot$$

$$F_p(t_2) = (2 \geq 2) \wedge [(s_2 \overset{p}{\cap} r_1) = \varnothing] = \top \wedge \top = \top$$

Primjer III

Zaključujemo da je $\overset{p}{\sigma}_{F_p}(r)$ sljedeća relacija

$$\begin{array}{c|ccc} \overset{p}{\sigma}_{F_p}(r) & A & B & C(D\ E) \\ \hline & a_2 & 2 & \boxed{d_3 \quad e_1} \end{array}$$

Ovaj primjer pokazuje da uvjet selekcioniranja, formula F_p, može sadržavati elementarne (atomarne) formule koje izražavaju uvjete između složenog atributa i relacija (u našem primjeru, $(C \overset{p}{\cap} r_1) \neq \emptyset$), te uvjete između složenih atributa (na primjer, $H \subseteq G$).

Proširenje formule F_p

Definicija

Proširenje formule F_p
Formula se definirana rekurzivno kao što slijedi: Atom je izraz oblika:

- $A_t \ \alpha \ k$, $k \ \alpha \ A_t$, $A_t \ \alpha \ B_t$, *gdje su* A_t *i* B_t *bilo koji jednostavni atributi,* k *je konstanta i* $\alpha \in \{=, <, >, \leqslant, \geqslant, \neq\}$ *je aritmetički operator uspoređivanja; ili*
- $C_t \ \rho \ r$, $r \ \rho \ C_t$, $C_t \ \rho \ D_t$, *gdje su* C_t *i* D_t *bilo koji složeni atributi,* r *je relacija i* $\rho \in \{=, \subset, \subseteq, \supset, \supseteq, \neq\}$ *je relacijski operator uspoređivanja.*

1. *Svaki atom je formula. Naziva se atomarna ili elementarna formula.*
2. *Formula je svaki izraz oblika* $\neg F$, $F \vee G$, $F \wedge G$, $F \Rightarrow G$, $F \Leftrightarrow G$, *gdje su* F *i* G *formule.*

Selekcija $\overset{p}{\sigma}$

Definicija

Selekcija $\overset{p}{\sigma}$

Neka je zadana formula F_p primjenjiva na relaciju $r(R)$. Tada je

$$\overset{p}{\sigma}_{F_p}(r) = \{t \in r \| F_p(t) = T\};\ \overset{p}{\sigma}N_{F_p}(r) = \{t \in r \| F_p(t) = N\}$$

Dakle, $\overset{p}{\sigma}_{F_p}(r)$ sadrži sve one redove t iz r za koje vrijedi $F_p(t)$, tj za koje je $F_p(t)$ istinito (oznaka T*); $\overset{p}{\sigma}N_{F_p}(r)$ sadrži sve one redove t iz r za koje se ne zna da li vrijedi $F_p(t)$, tj. za koje je $F_p(t)$ nepoznato (oznaka* N*).*

Produkt $\overset{p}{\otimes}$ i preimenovanje $\overset{p}{\delta}$

Definicija

Produkt $\overset{p}{\otimes}$

Definicija $\overset{p}{\otimes}$ *ostaje nepromijenjena u odnosu na konvencionalni produkt* $\otimes$*; prema tome,* $\overset{p}{\otimes} = \otimes$.

Isto vrijedi i za preimenovanje $\overset{p}{\delta}$

Definicija

Preimenovanje $\overset{p}{\delta}$

$\overset{p}{\delta} = \delta$*, gdje je* δ *konvencionalno preimenovanje.*

Aktivni komplement $\overset{p}{AC}$

Definicija

Aktivni komplement $\overset{p}{AC}$

U definiciji konvencionalnog aktivnog komplementa AC potrebno je umjesto operatora projekcije Π *koristiti operator proširene projekcije* $\overset{p}{\Pi}$*, umjesto operatora spoja* $\bowtie$ *uvrstiti operator proširenog spoja* $\overset{p}{\bowtie}$*, te umjesto operatora razlike* $-$ *uvrstiti operator proširene razlike* $\overset{p}{-}$*. Tako dobivamo*

$$\overset{p}{AC}(r) = (\overset{p}{\Pi}_{A_1}(r)\overset{p}{\bowtie}..\overset{p}{\bowtie}\overset{p}{\Pi}_{A_n}(r))\overset{p}{-}r.$$

Kvocijent $\overset{p}{\div}$

Definicija

Kvocijent $\overset{p}{\div}$

Postupak je analogan postupku u definiranju $\overset{p}{\mathrm{AC}}$.
Za relacije $r(R)$ *i* $s(S)$ *takve da je* $S \subset R$ *, prošireni kvocijent relacija r i s je relacija za koju vrijedi:*

1. *Ime relacije je* $r\overset{p}{\div}s$
2. $sh(r\overset{p}{\div}s) = T$, *gdje je* $T = R - S$
3. *redovi su dani jednakošću* $r\overset{p}{\div}s = \overset{p}{\Pi}_T(r)\overset{p}{-}\overset{p}{\Pi}_T((\overset{p}{\Pi}_T(r)\overset{p}{\bowtie}s)\overset{p}{-}r)$

Zadaci

Neka su zadane relacije $r_0(R_0)$, $r_1(R_1)$, $r_2(R_1)$

r_0	C	D
	1	1
	2	2
	2	1

<table>
<tr><th>r_1</th><th>A</th><th>$B(C\ D)$</th></tr>
<tr><td></td><td>2</td><td><table><tr><td>2</td><td>2</td></tr><tr><td>1</td><td>0</td></tr></table></td></tr>
<tr><td></td><td>1</td><td><table><tr><td>1</td><td>1</td></tr><tr><td>2</td><td>1</td></tr></table></td></tr>
</table>

<table>
<tr><th>r_2</th><th>A</th><th>$B(C\ D)$</th></tr>
<tr><td></td><td>1</td><td><table><tr><td>2</td><td>1</td></tr><tr><td>1</td><td>0</td></tr></table></td></tr>
<tr><td></td><td>2</td><td><table><tr><td>2</td><td>2</td></tr><tr><td>1</td><td>0</td></tr></table></td></tr>
</table>

i formule: $F = (B \subseteq r_0) \wedge (A \neq 0)$; $G = [(B \overset{p}{-} r_0) = \phi]$.

Zadaci

Odredite:

(a) $r_1 \overset{p}{\cup} r_2$

(b) $r_1 \overset{p}{\cap} r_2$

(c) $r_1 \overset{p}{-} r_2$

(d) $r_2 \overset{p}{-} r_1$

(e) $r_1 \overset{p}{\bowtie} r_2$

(f) $\overset{p}{\Pi}_A(r_1)$

(g) $\overset{p}{\Pi}_B(r_2)$

(h) $r_1 \overset{p}{\otimes} r_2$

(i) $\overset{p}{\underset{B}{\Pi}}(\overset{p}{\sigma}_F(r_1))$

(j) $\overset{p}{\sigma}_G(r_2)$

(k) $(r_1 \overset{p}{\bowtie} r_2) \overset{p}{-} \overset{p}{\sigma}_G(r_2)$

Pitanja?

Izvor

Maleković, M., Schatten, M. (2017) Teorija i primjena baza podataka, Fakultet organizacije i informatike, Varaždin.