Capitolul 7

CALCULUL RELAȚIONAL

Cu toate că algebra relațională serveşte drept fundament al unor limbaje de interpelări, majoritatea limbajelor relaționale sunt bazate pe calculul relațional sau tablouri. Cauza principală constă în faptul că algebra relațională este un sistem procedural de operații, adică expresia algebrei relaționale determină o serie de operații asupra relațiilor şi ordinea lor de executare (cu exactitatea unor reguli asociative prestabilite).

Calculul relațional reprezintă o adaptare a calculului cu predicate în domeniul bazelor de date relaționale. Ideea de bază a calculului relațional este de a identifica o relație ca un predicat. Deci el eliberează utilizatorul de obligația de a defini cum să obțină rezultatul. Pe baza unor predicate (relații) inițiale, prin aplicarea unor operatori ai calculului predicatelor se pot defini noi predicate (relații).

Sunt cunoscute două variante ale calculului relațional. O variantă utilizează în calitate de valori ale variabilelor asupra relațiilor tupluri de relație. Din această cauză variabilele au fost denumite *variabile tuplu*, iar calculul relațional a primit denumirea *de calcul relațional orientat pe tuplu*. Altă variantă presupune că variabilele sunt definite asupra domeniilor. Aceste variabile se numesc *variabile domeniu*, iar calculul relațional bazat pe acest tip de variabile e cunoscut sub numele de *calcul relațional orientat pe domeniu*.

7.1. Calculul relațional orientat pe tuplu

La început vom considera calculul relațional ce permite definirea relațiilor infinite. Apoi vom introduce modificările necesare ce vor garanta că orice formulă în calculul relațional notează o relație finită.

Formulele în calculul relațional au forma $\{t \mid f(t)\}$, unde t este variabila tuplu, adică variabila ce denotă un tuplu de o lungime fixată, iar f este formula construită din atomi şi operatori.

Definiția 7.1. Fie mulțimea universală U de atribute şi pentru orice $A \in U$, dom(A) e mulțimea de valori. Fie mulțimea de operatori de comparație $\Theta = \{=, \neq, <, \leq, >, \geq\}$, schema bazei de date $Db = \{R_1, ..., R_m\}$, unde $R_i \subseteq U$, $1 \leq i \leq m$ şi baza de date $db = \{r_1, ..., r_m\}$.

Atomii în formula f sunt definiți astfel:

(1) Valorile de veridicitate, notate cu *true* sau *false* sunt atomi;

(2) Variabila tuplu t, ce reprezintă un tuplu al relației $r_j(R_j)$, notat $r_j(t)$, este atom, unde r_j este o relație cu schema R_j în baza de date db;

(3) $t(A_j)\theta s(A_k)$ este atom, unde t şi s sunt variabile tuplu (nu numaidecât distincte), A_j şi A_k sunt atribute (nu numaidecât distincte) compatibile din U,

θ este operația de comparație, şi $t(A_j)$ şi $s(A_k)$ sunt respectiv A_j-componenta lui t şi A_k–componenta lui s.

(4) cθt(A) şi t(A)θc sunt atomi, unde c este o constantă în dom(A), t(A) este A-componenta a variabilei tuplu t, şi θ este operator de comparație din Θ.

Exemplul 7.1. Fie schema bazei de date constă din schemele relaționale
funcționari(NUME SALARIU MANAGER DEPARTAMENT),
vânzări(DEPARTAMENT ARTICOL),
furnizări(ARTICOL FURNIZOR),
culori(ARTICOL CULOARE CANTITATE)
şi fie expresiile

furnizări(t), (7.1)
vânzări(s), (7.2)
t(ARTICOL) = s(ARTICOL), (7.3)
u(DEPARTAMENT)=t(DEPARTAMENT) (7.4)
s(FURNIZOR) = ”Microsoft”. (7.5)

Expresiile (7.1), (7.2), (7.3), (7.4), (7.5) sunt atomi de tipul (2), (2), (3), (3), (4), respectiv. Iar t(ARTICOL) ≠ s(CULOARE) nu este atom fiindcă atributele ARTICOL şi CULOARE nu sunt compatibile.

Pentru definirea operațiilor calculului relațional sunt utile noțiunile de variabile tuplu libere şi legate. Noțiunile acestea au acelaşi sens ca şi în calculul predicatelor. Neformal vom spune că *variabila tuplu* este *legată* într-o formulă, dacă este calificată existențial sau universal. *Variabila tuplu* se numeşte *liberă,* dacă nu e calificată.

Noțiunea de variabilă liberă e analogică noțiunii de variabilă globală din limbajele de programare, adică variabilă definită în afara procedurii curente. Variabila legată e similară valorii locale, ce e definită în procedura curentă.

Definiția 7.2. Formulele şi variabilele libere şi legate în formule se definesc recursiv.

(1) Orice atom este formulă. Orice variabilă tuplu în cadrul unui atom trebuie să fie liberă.

(2) Dacă f este formulă, atunci negația lui f, notată ⌉f, este formulă. Orice variabilă tuplu este liberă sau legată în ⌉f, dacă este liberă sau legată în f.

(3) Dacă f_1 şi f_2 sunt formule, atunci conjuncția şi disjuncția formulelor f_1 şi f_2, notate corespunzător $f_1 \wedge f_2$ şi $f_1 \vee f_2$, sunt formule. Orice variabilă tuplu liberă (legată) apărută în f_1 şi f_2 sau în ambele formule va rămâne la fel în $f_1 \wedge f_2$ sau $f_1 \vee f_2$. Orice variabilă tuplu, liberă într-o formulă şi legată în alta, este liberă sau legată în $f_1 \wedge f_2$ sau $f_1 \vee f_2$ în dependență unde ele apar.

(4) Dacă variabila tuplu t cu schema R este liberă în formula f, atunci ∀t(R)f(t) şi ∃t(R)f(t) sunt formule, unde t este calificată universal şi existențial, respectiv. Variabila tuplu t ce e liberă în f devine legată în ∀t(R)f(t) şi ∃t(R)f(t). Orice altă variabilă tuplu s, unde s≠t, este liberă sau legată în ∀t(R)f(t) sau ∃t(R)f(t) în dependență cum este în f.

(5) Dacă f e formulă, atunci (f) e formulă.

(6) Nimic altceva nu e formulă.

Se presupune următoarea ordine descrescătoare de precedenţă: operatorii de comparaţie, calificativele existenţial şi universal şi în sfârşit $\neg$, $\wedge$ şi $\vee$.

Definiţia 7.3. Se numeşte *expresie al calculului relaţional orientat pe tuplu* o construcţie de forma $\{t(R) \mid f(t)\}$, unde f este o formulă şi t este singura variabilă tuplu liberă cu schema R.

Exemplul 7.2. În calculul relaţional orientat pe tuplu uniunea relaţiilor r(R) şi s(R) se exprimă prin $\{t(R) \mid r(t) \vee s(t)\}$. Această expresie se citeşte: "mulţimea de tupluri t, ce aparţin relaţiei r sau relaţiei s". Să ne amintim că uniunea poate fi realizată, dacă r şi s au aceeaşi aritate. Similar formula $r(t) \vee s(t)$ are sens în aceleaşi condiţii, fiindcă variabila t are aceeaşi lungime.

Diferenţa r(R) \ s(R) poate fi prezentată de expresia $\{t(R) \mid r(t) \wedge \neg s(t)\}$.

Dacă r(R) şi s(S) sunt relaţii, unde $R = A_1 \ldots A_k$, şi $S = B_1 \ldots, B_m$, atunci produsul cartezian r×s în calculul relaţional orientat pe tuplu se scrie

$$\{t(A_1 \ldots A_kB_1 \ldots B_m \mid \exists t_r(R)\ \exists t_s(S)\ (r(t_r) \wedge s(t_s) \wedge (t(A_1)=t_r(A_1)) \wedge \ldots \wedge (t(A_k)=t_r(A_k)) \wedge (t(B_1)=t_s(B_1)) \wedge \ldots \wedge (t(B_l)=t_s(B_l)))\}.$$

Expresia de mai sus ne spune că r×s este o mulţime de tupluri t de aritatea |R|+|S|, pentru care există t_r, ce aparţine relaţiei r şi t_s, ce aparţine relaţiei s, şi primele componente ale lui t formează t_r, iar celelalte componente formează t_s.

Fie relaţia r(R) şi $X \subseteq R$, unde $X=B_1 \ldots B_k$, atunci proiecţia $\pi_{B1\ldots Bk}(r)$ se exprimă astfel:

$$\{t(B_1 \ldots B_k) \mid \exists t_r(R)\ (r(t_r) \wedge (t[B_1]=t_r[B_1]) \wedge \ldots \wedge (t[B_k]=t_r[B_k])\}.$$

Selecţia $\sigma_F(r)$ este expresia de forma $\{t(R) \mid r(t) \wedge F^1\}$, unde F^1 este formula F, în care fiecare operând i, ce denotă componenta i e substituită de t(i).

Fie relaţia r(AB) de aritatea 2. Atunci

$$\{t(AB) \mid \exists t_r(AB)\ (r(t) \wedge r(t_r) \wedge (t(A) \neq t_r(A) \vee t(B) \neq t_r(B))\}$$

este o expresie a calculului relaţional, ce defineşte relaţia r, dacă r are două sau mai multe tupluri, şi o relaţie vidă, dacă r e vidă sau constă dintr-un singur tuplu.

7.2. Expresii bine formate

Ca şi expresiile algebrei relaţionale, expresiile din calculul relaţional orientat pe tuplu reprezintă definiţii ale unor relaţii. În forma prezentată mai sus, aceste expresii permit definirea unor relaţii cu un număr infinit de tupluri. De exemplu, expresia $\{t(R) \mid \neg r(t)\}$ este mulţimea de tupluri de lungime egală cu aritatea relaţiei r ce nu aparţin r.

Deoarece e greu de precizat "toate tuplurile posibile" se impune excludere unor expresii absurde de tipul celei de mai sus. Aceasta se poate atinge, dacă ne vom limita doar la expresiile bine formate. În astfel de expresii $\{t(R) \mid f(t)\}$, fiecare componentă a lui t ce satisface f trebuie să fie element al lui dom(f). Mulţimea dom(f) se defineşte ca o mulţime de simboluri, care, fie apar explicit în f, fie sunt componentele unor tupluri din relaţia r, citată în f. Astfel definit dom(f) este finit întotdeauna.

Exemplul 7.3. Fie f(t) este $t(a) = a \vee r(t)$, unde r(AB) e o relaţie. Atunci dom(f) poate fi definită de formula algebrei relaţionale $\{a\} \cup \pi_A(r) \cup \pi_B(r)$.

Definiția 7.4. Expresia $\{t(R) \mid f(t)\}$ a calculului relațional orientat pe tuplu este *bine formată*, dacă sunt satisfăcute condițiile:

(1) Fiecare componentă a lui t, ce satisface f, aparține dom(f).

(2) În orice subexpresie de forma $\exists t_1(R)f_1(t_1)$, dacă orice componență a lui t_1 aparține dom (f_1), atunci t_1 satisface f_1.

(3) În orice subexpresie de forma $\forall t_1(R)f_1(t_1)$, dacă orice componentă a lui t_1 nu aparține lui dom (f_1), atunci t_1 satisface f_1.

Condițiile (2) și (3) stabilesc veridicitatea formulelor calificate $\exists t_1(R)f_1(t_1)$ și $\exists t_1(R)f_1(t_1)$, considerând numai t_1, constituit din simboluri ce aparțin lui dom(f_1). De exemplu, orice formulă $\exists t_1(R)(r(t_1)\vee...)$ satisface condiția (2) și orice formulă $\forall t_1(R)(\rceil r(t_1)\vee...)$ satisface condiția (3).

Chiar dacă condiția (3) pare neobișnuită, trebuie să observăm că formula $\forall t_1(R)f_1(t_1)$ este echivalentă formulei $\rceil\exists t_1(R)\rceil f_1(t_1)$. Ultima expresie nu e bine formată, dacă și numai dacă există un t_1^0 pentru care e adevărată $\rceil f_1(t_1^0)$ și t_1^0 nu aparține domeniului formulei $\rceil f_1$. Întrucât domeniile lui f_1 și $\rceil f_1$ sunt aceleași, condiția (3) afirmă că formula $\forall t_1(R)f_1(t_1)$ e bine formată, dacă e bine formată formula $\rceil\exists t_1(R)\rceil f_1(t_1)$.

Exemplul 7.4. Fie f(t) e o formulă și fie că orice subformulă a ei de forma $\exists t_1(R)f_1(t_1)$ sau $\forall t_1(R)f_1(t_1)$ este bine formată. Atunci sunt bine formate expresiile de forma $\{t(R) \mid r(t)\wedge f(t)\}$, fiindcă orice tuplu ce satisface $r(t)\wedge f(t)$ aparține relației r, prin urmare, orice componentă a lui t aparține lui dom$(r(t)\wedge f(t))$. În calitate de exemplu al acestui tip de formule poate fi formula diferenței a două relații $\{t(R) \mid r(t)\wedge\rceil s(t)\}$, unde $f(t)=\rceil s(t)$. Dacă $f(t)=F^1$, atunci formula exprimă selecția.

Generalizând cele spuse, observăm că este bine formată și formula de forma

$$\{t \mid r_1(t) \vee r_2(t)\vee ... \vee r_k(t) \wedge f(t)\}.$$

Aici componenta t(i) trebuie să fie un simbol ce apare în componenta i al unui tuplu dintr-o relație r_j. Această formă are, de exemplu, formula ce exprimă uniunea a două relații, adică $\{t(R) \mid r(t) \vee s(t)\}$. În ea lipsește f, fiindcă aici f este identic adevărată cum ar fi, spre exemplu, t(1)=t(1).

O altă expresie bine formată este

$$\{t \mid \exists t_1(R_1)... \exists t_k(R_k)\ (r_1(t_1)\wedge ... \wedge r_k(t_k) \wedge (t(1)=t_{i1}(j_1))\wedge ... \wedge(t(m)=t_{im}(m)) \wedge f(t, t_1, ..., t_{k3}))\}.$$

Asupra componentei t(p) se aplică restricția că ea trebuie să fie un simbol ce apare în componenta jp al unui tuplu din R_{ip}. Formula produsului cartezian din exemplul 7.2 are această formă.

7.3. Reducerea algebrei relaționale la calculul relațional orientat pe tuplu

Vom arăta că orice expresie a algebrei relaționale se poate reduce la o expresie a calculului relațional orientat pe tuplu.

Teorema 7.1. Dacă Ea este o expresie a algebrei relaționale, atunci există o expresie, Et, a calculului orientat pe tuplu echivalentă expresiei Ea.

Demonstrare. Teorema se demonstrează prin inducție după numărul operațiilor în Ea.

Baza inducţiei: Considerăm zero operatori în Ea. În acest caz, Ea sau este o relaţie constantă asupra R, adică $\{t_1, ..., t_n\}$, sau o variabilă r, ce denotă o relaţie.

În primul caz expresia algebrică Ea e echivalentă expresiei $Et=\{t(R) \mid t=t_1 \vee ... \vee t=t_n\}$, unde $t=t_i$ este notaţia prescurtată pentru $t(A_1)=t_i(A_1) \wedge ... \wedge t(A_k)=t_i(A_k)$. Este clar că $t(A_i)$ este un simbol ce se prezintă explicit în calitate de componentă a unui tuplu constantă t_j.

În al doilea caz, Ea e echivalentă expresiei $Et=\{t(R) \mid r(t)\}$, care, precum e arătat în exemplul 7.4, este bine formată.

Inducţia: Presupunem că teorema a validă pentru expresii algebrice cu mai puţin de k operatori. Fie Ea are k operatori. Atunci avem de considerat cinci cazuri (fiindcă sunt cinci operaţii de bază ale algebrei relaţionale, celelalte deducându-se din ele).

(1) Uniunea: $Ea = Ea_1 \cup Ea_2$. Din ipoteza inductivă, există $Et_1 = \{t_1(R) \mid f_1(t_1)\}$ şi $Et_2 = \{t_2(R) \mid f_2(t_2)\}$ echivalente cu Ea_1 şi Ea_2, respectiv, unde Ea_1 şi Ea_2 au mai puţin de k operatori. Atunci Ea este echivalentă expresiei $Et=\{t(R) \mid f_1(t) \vee f_2(t)\}$. Dacă t satisface $f_1(t) \vee f_2(t)$, atunci orice componentă a lui t aparţine $dom(f_1)$ sau $dom(f_2)$. Întrucât $dom(f_1(t) \vee f_2(t)) = dom(f_1) \cup dom(f_2)$ rezultă că Et e expresie bine formată.

(2) Diferenţa: $Ea = Ea_1 \setminus Ea_2$. Fie Ea_1 şi Ea_2 au mai puţin de k operatori. Atunci Ea este echivalentă expresiei $Et = \{t(R) \mid f_1(t) \wedge \neg f_2(t)\}$, unde $f_1(t)$ şi $f_2(t)$ sunt cele din cazul (1), iar $\neg f_2(t)$ este negaţia lui $f_2(t)$. Întrucât $dom(f_1(t) \wedge \neg f_2(t)) = dom(f_1) \cup dom(f_2)$ expresia Et este expresie bine formată.

(3) Produsul cartezian: $Ea = Ea_1 \times Ea_2$. Fie Ea_1 şi Ea_2 sunt expresii ale algebrei relaţionale cu mai puţin de k operatori. Conform ipotezei inducţiei există expresii ale calculului orientat pe tuplu $Et_1 = \{t_1(R) \mid f_1(t_1)\}$ şi $Et_2 = \{t_2(S) \mid f_2(t_2)\}$ echivalente cu Ea_1 şi Ea_2, respectiv. Atunci Ea este echivalentă expresiei

$$Et = \{t(RS) \mid \exists t_1(R)\ \exists t_2(S)\ (f_1(t_1) \wedge f_2(t_2) \wedge (t(R)=t_1(R)) \wedge (t(S)=t_2(S)))\},$$

unde $t(R)=t_1(R)$ este scrierea scurtă pentru $t(A_1)=t_1(A_1) \wedge ... \wedge t(A_n)=t_1(A_m)$, $R=A_1...A_m$. Este evident că Et este o expresie bine formată.

(4) Proiecţia: $Ea = \pi_{Ai1\ Ai2\ ...\ Aij}(Ea_1)$. Fie Ea_1 reprezintă o relaţie cu schema R, iar $A_{i1},...,A_{ij}$, sunt atribute din R. Atunci Ea este echivalentă expresiei

$$Er = \{t(A_{i1}...A_{ij}) \mid \exists t_1(R)\ (f_1(t_1) \wedge (t(A_{i1})=t_1(A_{i1})) \wedge \dots \wedge (t(A_{ij})=t_1(A_{ij})))\}.$$

Expresia este bine formată din aceleaşi considerente că şi expresia din cazul (3).

(5) Selecţia: $Ea = \sigma_F(Ea_1)$. Fie Ea_1 reprezintă o relaţie cu schema R şi formula F este aplicabilă. Atunci Ea e echivalentă expresiei $Et = \{t(R) \mid f_1(t) \wedge F^1\}$, unde F^1 este obţinută din F substituind orice atribut A_i ce apare în F cu A_i–componenta variabilei tuplu t, $t(A_i)$.

Celelalte operaţii ale algebrei relaţionale se deduc din aceste cinci operaţii de bază.

Exemplul 7.5. Fie expresia algebrei relaţionale

$$Ea = r(AB) \setminus (s(A) \times \pi_B(r(AB))).$$

Expresia calculului orientat pe tuplu echivalentă ei este

$$Et = \{t(AB) \mid r(t) \wedge \neg(\exists t_1(A)\ \exists t_2(B)\ s(t_1) \wedge \exists t_3(AB)\ (r(t_3) \wedge (t_2(B)=t_3(B) \wedge (t(A)=t_1(A)) \wedge (t(B)=t_2(B)))\}.$$

Exemplul 7.6. Fie r(AB) şi s(CD) două relaţii. Să se găsească expresia calculului orientat pe tuplu echivalentă expresiei $Ea = \pi_{AD}(\sigma_{B=C}(r\times s))$ a algebrei relaţionale.
Folosind teorema 7.1, expresia calculului orientat pe tuplu echivalentă expresiei $r\times s$ este

$$\{t_3(ABCD) \mid \exists t_1(AB)\ \exists t_2(CD)\ (r(t_1) \wedge s(t_2) \wedge (t_3(A)=t_1(A)) \wedge (t_3(B)=t_1(B)) \wedge (t_3(C)=t_2(C)) \wedge (t_3(D)=t_2(D)))\}.$$

Pentru expresia $\sigma_{B=C}(r\times s)$ la formula de mai sus se mai adaugă $t_3(B)=t_3(C)$. Atunci expresia algebrică $\pi_{AD}(\sigma_{B=C}(r\times s)$ este echivalenta următoarei expresii a calculului relaţional orientat pe tuplu

$$Et = \{t(AD) \mid \exists t_3(ABCD)\ \exists t_1(AB)\ \exists t_2(CD)\ (r(t_1) \wedge s(t_2) \wedge (t_3(A)=t_1(A)) \wedge (t_3(B)=t_1(B)) \wedge (t_3(C)=t_2(C)) \wedge (t_3(D)=t_2(D)) \wedge (t_3(B)=t_3(C)) \wedge (t(A)=t_3(A)) \wedge (t(D)=t_3(D)))\}.$$

Această expresie nu e cea mai scurtă. Ea poate fi simplificată, dacă se înlătură t_3 şi se substituie toate componentele lui t_3 cu componentele lui t_1şi t_2. Atunci obţinem

$$Et = \{t(AD) \mid \exists t_1(AB)\ \exists t_2(CD)\ (r(t_1) \wedge s(t_2) \wedge (t_1(B)=t_2(C)) \wedge (t(A)=t_1(A)) \wedge (t(D)=t_2(D)))\}.$$

7.4. Calculul relaţional orientat pe domeniu

Calculul relaţional orientat pe domeniu utilizează în construcţiile sale aceiaşi operatori ca şi în calculul relaţional orientat pe tuplu, dar variabilele, care apar în aceste construcţii, sunt definite asupra domeniilor.

Atomul, ca construcţie elementară din calculul orientat pe domeniu, se defineşte recursiv în felul următor.

Definiţia 7.5. Atomul în calculul orientat pe domeniu poate avea una din formele:

(1) Dacă $r_j(A_{i1}\ A_{i2}\ ...\ A_{ik})$ este o relaţie în baza de date db, atunci $r_j(a_1\ a_2\ ...\ a_k)$ este formulă atomică, unde orice a_p, $1 \le p \le k$, este sau variabilă domeniu cu schema A_{ip}, sau constantă în $dom(A_{ip})$.
(2) Dacă a şi b sunt variabile domeniu cu aceeaşi schemă şi θ este operator de comparaţie şi c este o constantă cu schema identică lui a, atunci $a\theta b$, $a\theta c$ şi $c\theta a$ sunt formule atomice.
(3) Valorile de veridicitate *true* şi *false* sunt atomi.
(4) Nici o altă formulă nu e atom

Noţiunile de variabile legate şi libere în calculul orientat pe domeniu se definesc ca şi în calculul orientat pe tuplu.

Expresiile în calculul relaţional orientat pe domeniu au forma $\{d_1(A_1)...d_k(A_k) \mid f(d_1,...,d_k)\}$

Definiţia 7.6. Expresia calculului orientat pe domeniu este bine formată, dacă:

(1) din veridicitatea funcţiei $f(d_1, ..., d_k)$ urmează că $d_i \in dom(f)$;
(2) $\exists a(A)\ f_1(a)$ e o subformulă a lui f şi din veridicitatea $f_1(a)$ urmează că $a \in dom(f_1)$;
(3) $\forall a(A)\ f_1(a)$ e o subformulă a lui f şi din veridicitatea $\neg f_1(a)$ urmează că $a \in dom(f_1)$.

r	A	B
	a_1	b_2
	a_1	b_3

Fig.7.1(a).

r	A	B
	a_1	b_2

Fig.7.1(b).

Exemplul 7.7. Apelăm din nou la ultima parte a exemplului 7.2, unde trebuia scrisă expresia în termenii calculului orientat pe tuplu, care ar exprima relaţia r(AB), dacă are două sau mai multe tupluri şi o relaţie vidă în caz contrar. În calculul orientat pe domeniu expresia ar avea forma

$$Ed = \{w(A)x(B) \mid \exists y(A)\ \exists z(B)\ (r(wx) \wedge r(yz) \wedge (w\neq y \wedge x\neq z))\}.$$

Într-adevăr, fie $f(w,x,y,z,) = r(wx) \wedge r(yz) \wedge (w\neq y\wedge x\neq z)$ şi r e relaţia din fig.7.1(a). Dacă presupunem că $w=a_1$ şi $x=b_2$, atunci $\exists y(A)\ \exists z(B)\ (f(a_1b_2,\ y_1\ z))$ este validă, fiindcă prin alegerea lui $y=a_1$ şi $z=b_3$ va fi validă formula f. În acelaşi mod această formulă e validă pentru $w=a_1$ şi $x=b_3$, fiindcă putem selecta $y=a_1$ şi $z=b_2$. Prin urmare, ambele tupluri $<a_1\ b_2>$ şi $<a_1\ b_3>$ aparţin relaţiei reprezentată de expresia calculului orientat pe domeniul de mai sus. Dacă, însă, se iau alte valori pentru w şi x, atunci subformula r(w, x) din f este falsă. Prin urmare, şi formula $\exists y(A)\ \exists z(B)\ (f(w, x, y, z))$ e falsă. Deci, dacă în r sunt două sau mai multe tupluri expresia Ed reprezintă relaţia r.

Fie, acum, r este relaţia cu un singur tuplu din fig.7.1(b). În acest caz, nici o valoare a lui w şi x nu satisface formulele $\exists y(A)\ \exists z(B)\ (f(w, x, y, z))$.

Într-adevăr, prima subformulă r(w x) a formulei f e satisfăcută numai pentru $w=a_1$ şi $x=b_2$. A doua subformulă r(yz) e satisfăcută numai pentru $y=a_1$ şi $z=b_2$, dar, în acest caz, nu e satisfăcută subformula $(w\neq y\wedge x\neq z)$.

Exemplul 7.8. Expresia calculului relaţional pe domeniu pentru interpelarea "Care sunt funcţionarii care-şi au serviciul în departamentul "Soft"?" are forma:

Ed = {n(NUME) | ∃s(SALARIU) ∃m(MANAGER) ∃d(DEPARTAMENT)
(*funcţionari*(n s m d) ∧ d="Soft")}.

Exemplul 7.9. Fie interpelarea: "Care sunt departamentele ce vând articole "imprimantă EPSON" şi "imprimantă HP"?"
Expresia în termenii calculului orientat pe domeniu ce corespunde acestei interpelări este

Ed = {d(DEPARTAMENT) | ∃a(DEPARTAMENT) ∃b(DEPARTAMENT) (*vânzări*(a "imprimantă EPSON") ∧ d=a ∧ *vânzări*(b "imprimantă HP") ∧ d=b)}.

Exemplul 7.10. Fie interpelarea "Care sunt departamentele ce vând articole "imprimantă EPSON" sau "imprimantă HP" ?" Atunci

Ed = {d(DEPARTAMENT) | ∃a(DEPARTAMENT) ∃b(DEPARTAMENT) ((*vânzări*(a "imprimanta EPSON") ∧ d=a) ∨ *vânzări*(b "imprimanta HP") ∧ d=b))}.

7.5. Reducerea calculului orientat pe tuplu la calculul orientat pe domeniu

Expresia Et a calculului orientat pe tuplu se transformă destul de uşor într-o expresie Ed a calculului orientat pe domeniu.

Fie $\{t(R) \mid f(t)\}$ expresia calculului orientat pe tuplu, unde $R=A_1...A_k$. Atunci:

(1) orice atom $r(t)$ din f este substituit de $r(d_1 \ldots d_k)$, unde d_i este variabila domeniu cu schema A_i.

(2) orice atom $t(A_i)\theta c$ sau $c\theta t(A_j)$, unde c este componentă a altei variabile tuplu sau o constantă, se substituie de $d_i\theta c$ sau $c\theta d_i$, respectiv, unde d_i este variabila domeniu ce denotă componenta A_i a variabilei tuplu t;

(3) orice atom $t_1(A_i)\theta t_2(B_j)$ se substituie de $d_{1i}\theta d_{2j}$, unde d_{1i} şi d_{2j} sunt variabile domeniu ce denotă componentele A_i şi B_j ale variabilelor tuplu t_1 şi t_2, corespunzător;

(4) subformula calificată $\exists t(R)f$ este substituită de $\exists d_1(A_1)\ \exists d_2(A_2) \ldots \exists d_k(A_k)f$;

(5) subformula calificată $\forall t(R)f$ este substituită de $\forall d_1(A_1)\ \forall d_2(A_2) \ldots \forall d_k(A_k)f$;

(6) $t(R)$ este substituită de $d_1(A_1)\ d_2(A_2) \ldots d_k(A_k)$. Deci d_i primeşte acele valori pe care le primea $t(i)$.

Este evident că dacă expresia $\{t(R) \mid f(t)\}$ este bine formată, atunci e bine formată şi expresia calculului pe domeniu. Vom formula următoarea teoremă fără a o demonstra.

Teorema 7.2. Orice expresie bine formată din calculul relațional orientat pe tuplu are o expresie bine formată echivalentă în cadrul calculului relațional orientat pe domeniu.

Exemplul 7.11. Să examinăm expresia

$$Et = \{t(AD) \mid \exists t_1(AB)\ \exists t_2(CD)\ (r(t_1) \wedge s(t_2) \wedge (t_1(B)=t_2(C)) \wedge (t(A)=t_1(A)) \wedge (t(D)=t_2(D)))\}$$

în termenii calculului relațional orientat pe tuplu din exemplul 7.6 şi să construim expresia echivalentă în termenii calculului relațional orientat pe tuplu.

Substituim $t(AD)$ cu $d^1(A)$ şi $d^2(D)$; $t_1(AB)$ cu $d_1^{\,1}(A)$ şi $d_1^{\,2}(B)$; $t_2(CD)$ cu $d_2^{\,1}(C)$ şi $d_2^{\,2}(D)$. Atunci

$$Ed = \{d^1(A)d^2(D) \mid \exists d_1^{\,1}(A)\ \exists d_1^{\,2}(B)\ \exists d_2^{\,1}(C)\ \exists d_2^{\,2}(D)\ (r(d_1^{\,1}\ d_1^{\,2}) \wedge s(d_2^{\,1}\ d_2^{\,2}) \wedge d_1^{\,2}= d_2^{\,1} \wedge d^1=d_1^{\,1} \wedge d^2= d_2^{\,2})\}.$$

Pentru a demonstra echivalența dintre algebra relațională şi calculul relațional, este necesar să se demonstreze faptul că, pentru orice expresie bine formată din calculul relațional (expresie pe tuplu sau pe domeniu), există o expresie echivalentă din algebra relațională.

Mai jos vom prezenta doar formularea teoremei fără a aduce demonstrarea ei.

Teorema 7.3. Orice expresie bine formată din calculul relațional orientat pe domeniu are o expresie echivalentă în cadrul algebrei relaționale.

Exemplul 7.12. Fie schema bazei de date constă din schemele relaționale
domiciliu(PERS_NUME STRADĂ ORAŞ),
serviciu(PERS_NUME COMPANIE SALARIU),
sediu(COMPANIE ORAŞ),
manageri(PERS_NUME MNG_NUME).

Să se scrie expresiile în termenii algebrei relaționale, calculului relațional orientat pe domeniu, calculului orientat pe tuplu pentru următoarea interpelare: "Să se găsească numele de persoane care-şi au serviciul la compania MSG". Expresiile sunt notate cu Ea, Ed şi Et, corespunzător.

$Ea = \pi_{PERS\_NUME}(\sigma_{COMPANIE=\text{"MSG"}}(\mathit{serviciu}))$

$Ed = \{n(PERS\_NUME) \mid \exists c(COMPANIE)\ \exists s(SALARIU)\ (\mathit{serviciu}(n, c, s) \wedge (c=\text{"MSG"}))\}$

$Et = \{t(PERS\_NUME) \mid \exists t_1(PERS\_NUME\ COMPANIE\ SALARIU)\ \mathit{serviciu}(t_1) \wedge (t_1(PERS\_NUME) = t(PERS\_NUME)) \wedge (t_1(COMPANIE)=\text{"MSG"}))\}$.

Exemplul 7.13. Fie schema bazei de date din exemplul 7.12. Să se aducă expresiile Ea, Ed şi Et pentru interpelarea "Să se aducă numele şi oraşele unde locuiesc persoanele care-şi au serviciul la compania MSG". Expresiile sunt următoarele:

$Ea=\pi_{PERS\text{-}NUME,ORAŞ}(\sigma_{COMPANIE=\text{"MSG"}}(\mathit{domiciliu}|\times|\mathit{serviciu}))$.

$Ed = \{n(PERS\_NUME), o(ORAŞ) \mid \exists st(STRADĂ)\ \exists s(SALARIU)\ (\mathit{domiciliu}(n, st, o) \wedge \mathit{serviciu}(n, \text{"MSG"}, s)\}$.

$Et = \{t(PERS\_NUME\ ORAŞ) \mid \exists t_1(PERS\_NUME\ STRADĂ\ ORAŞ)\ \exists t_2(PERS\_NUME\ COMPANIE\ SALARIU)\ (t(PERS\_NUME) = (t_1(PERS\_NUME)) \wedge (t(ORAŞ)=t_1(ORAŞ)) \wedge (t_1(PERS\_NUME)=t_2(PERS\_NUMA)) \wedge (t_2(COMPANIE)=\text{"MSG"}))\}$.

7.6. Exerciții

7.1. Fie schema bazei de date din exemplul 7.12. Să se scrie expresiile algebrei relaționale, calculului relațional orientat pe tuplu pentru interpelările:
   (a) Să se afişeze lista numelor de persoane care nu îşi au serviciul la compania MSG.
   (b) Să se afişeze lista numelor de persoane care nu au serviciu.

7.2. Să se găsească expresiile calculului orientat pe tuplu echivalente pentru operațiile algebrei relaționale:
   (a) uniunea;
   (b) diferența;
   (c) produsul cartezian;
   (d) proiecția;
   (e) selecția;
   (f) intersecția;
   (g) θ - joncțiunea;
   (h) diviziunea.

7.3. Să se găsească expresiile calculului relațional orientat pe domeniu echivalente expresiilor calculului relațional orientat pe tuplu, obținute în exercițiul 7.2.