

ALGEBRA RELATIONALĂ

1. Obiective

- Studiul conceptelor și operatorilor algebrei relationale
- Exemple de operații pe seturi de date modelate cu operatorii algebrei relaționale

2. Consideratii teoretice

Interogarea (query) este operația prin care se obțin datele dorite dintr-o bază de date, selectate conform unor criterii. Deoarece această operație este cea mai importantă operație, limbajele de manipulare a datelor sunt denumite și *limbaje de interogare*. Pentru formularea conceptuală a interogărilor în bazele de date relaționale s-au dezvoltat două limbaje abstracte de interogare: algebra relațională și calculul relațional.

Algebra relațională (relational algebra), introdusă de Codd în 1970, definește cadrul formal al limbajelor relaționale pentru baze de date. *Algebra relațională* introduce o colecție de operatori algebrici care se aplică relațiilor (tabelelor). Fiecare operație din algebra relațională are drept operanzi una sau mai multe relații, iar rezultatul este tot o relație. Această uniformitate (proprietatea algebrică de închidere) permite aplicarea de combinații de operatori relațiilor.

Prin analogie cu un *compiler* care, plecând de la un *program sursă* produce un *program executabil*, rezultatul compilării unei interogări (cereri) de către un SGBD relațional este o expresie algebrică care va fi evaluată. În cadrul modelului relațional se consideră *limbaje relaționale complete* numai acele limbaje de interogare care permit implementarea tuturor operațiilor prevăzute de unul din limbajele abstracte de interogare. Limbajele de interogare reale implementate în sistemele de baze de date relaționale sunt limbaje definite pe baza limbajelor abstracte de interogare.

Cunoașterea limbajului abstract de interogare bazat pe *algebra relațională* este obligatorie pentru înțelegerea aprofundată a modului de execuție a interogărilor.

Algebra relațională conține două tipuri de operații: operații pe mulțimi: reuniunea, intersecția, diferența și produsul cartezian. Pentru a determina reuniunea, intersecția și diferența a două relații, acestea trebuie să fie compatibile cu reuniunea (trebuie să aibă același număr de atribute și atributele corespondente să fie definite pe domenii compatibile, adică să aibă formatul datelor identic). operații relaționale: selecția, proiecția, joncțiunea și diviziunea.

Operații de bază - reprezintă un ansamblu minimal de operații, în sensul că niciuna din operații nu poate fi exprimată ca o combinație a celorlalte operații. Pentru algebra relațională există mai multe ansambluri minimale. În continuare prezentarea va considera ca operații de bază: trei operații pe mulțimi: *reuniunea*, *diferența*, *produsul cartezian* și două operații relaționale unare: *selecția* și *proiecția*. Operatorii de *intersecție*, *joncțiune* și *diviziune* pot fi obținuți din cei cinci operatori de bază.

În definirea operatorilor *algebrei relaționale* vom nota cu:

t – un tuplu din relația R $t(A)$ – subtuplu din R relativ la atributul A .

2.1. OPERATORII ALGEBREI RELAȚIONALE

1. **UNION** – reuniunea a două relații R_1 și R_2 , compatibile cu reuniunea, este dată de mulțimea tuplurilor care aparțin fie relației R_1 , fie relației R_2 , fie ambelor relații. Tuplurile care aparțin ambelor relații se introduc în reuniune o singură dată, adică nu se duplică.

$$\begin{aligned} & \text{Relatie} \times \text{Relatie} \rightarrow \text{Relatie} \\ R_3 = R_1 \cup R_2 = \{t \mid t \in R_1 \text{ sau } t \in R_2\} \end{aligned}$$

Sintaxa: **UNION (R₁, R₂)** - rezultatul este o relație.

2. **DIFF** – diferența a două relații R_1 și R_2 , compatibile cu reuniunea, este dată de mulțimea tuplurilor care aparțin relației R_1 și nu aparțin relației R_2 .

$$\begin{aligned} & \text{Relatie} \times \text{Relatie} \rightarrow \text{Relatie} \\ R_3 = R_1 - R_2 = \{t \mid t \in R_1 \text{ si } t \notin R_2\} \end{aligned}$$

Exemplu.

<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 10%;">R₁</th> <th style="width: 20%;">A</th> <th style="width: 10%;">B</th> </tr> </thead> <tbody> <tr><td></td><td>a</td><td>b</td></tr> <tr><td></td><td>c</td><td>d</td></tr> <tr><td></td><td>x</td><td>y</td></tr> </tbody> </table>	R₁	A	B		a	b		c	d		x	y	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 10%;">R₂</th> <th style="width: 20%;">A</th> <th style="width: 10%;">B</th> </tr> </thead> <tbody> <tr><td></td><td>d</td><td>f</td></tr> <tr><td></td><td>x</td><td>y</td></tr> <tr><td></td><td>h</td><td>r</td></tr> <tr><td></td><td>c</td><td>d</td></tr> </tbody> </table>	R₂	A	B		d	f		x	y		h	r		c	d	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 10%;">R₃</th> <th style="width: 20%;">A</th> <th style="width: 10%;">B</th> </tr> </thead> <tbody> <tr><td></td><td>a</td><td>b</td></tr> </tbody> </table>	R₃	A	B		a	b
R₁	A	B																																	
	a	b																																	
	c	d																																	
	x	y																																	
R₂	A	B																																	
	d	f																																	
	x	y																																	
	h	r																																	
	c	d																																	
R₃	A	B																																	
	a	b																																	

Sintaxa: **DIFF(R₁, R₂)** - rezultatul este o relație.

3. **INTERSECT** – intersecția a două relații R_1 și R_2 , compatibile cu reuniunea, este dată de mulțimea tuplurilor care aparțin atât relației R_1 , cât și relației R_2 .

$$\begin{aligned} & \text{Relatie} \times \text{Relatie} \rightarrow \text{Relatie} \\ R_3 = R_1 \cap R_2 = \{t \mid t \in R_1 \text{ si } t \in R_2\} \end{aligned}$$

Exemplu.

<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 10%;">R₁</th> <th style="width: 20%;">A</th> <th style="width: 10%;">B</th> </tr> </thead> <tbody> <tr><td></td><td>a</td><td>b</td></tr> <tr><td></td><td>c</td><td>d</td></tr> <tr><td></td><td>x</td><td>y</td></tr> </tbody> </table>	R₁	A	B		a	b		c	d		x	y	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 10%;">R₂</th> <th style="width: 20%;">A</th> <th style="width: 10%;">B</th> </tr> </thead> <tbody> <tr><td></td><td>d</td><td>f</td></tr> <tr><td></td><td>x</td><td>r</td></tr> <tr><td></td><td>h</td><td>d</td></tr> </tbody> </table>	R₂	A	B		d	f		x	r		h	d	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 10%;">R₃</th> <th style="width: 20%;">A</th> <th style="width: 10%;">B</th> </tr> </thead> <tbody> <tr><td></td><td>c</td><td>d</td></tr> <tr><td></td><td>x</td><td>y</td></tr> </tbody> </table>	R₃	A	B		c	d		x	y
R₁	A	B																																	
	a	b																																	
	c	d																																	
	x	y																																	
R₂	A	B																																	
	d	f																																	
	x	r																																	
	h	d																																	
R₃	A	B																																	
	c	d																																	
	x	y																																	

Sintaxa: **INTERSECT (R₁, R₂)** - rezultatul este o relație.

Observație. Intersecția poate fi exprimată cu ajutorul operației de *diferență* (operație de bază):

$$R \cap S = R - (R - S) = S - (S - R)$$

4. **PRODUCT** – produsul cartezian a două relații R_1 și R_2 , produce o nouă relație care are ca atribute, reuniunea atributelor din cele două relații (atributele comune vor fi luate separat, calificările fiind făcute cu numele relației), iar fiecare element din R_1 se combină (concatenează) cu fiecare element din R_2 .

$$\begin{aligned} & \text{Relatie} \times \text{Relatie} \rightarrow \text{Relatie} \\ R_3 &= R_1 \times R_2 = \{(t_1, t_2) \mid t_1 \in R_1 \text{ si } t_2 \in R_2\} \end{aligned}$$

Exemplu.

<table border="1" style="border-collapse: collapse; width: 100%; text-align: center;"> <thead> <tr> <th style="border: none;">R_1</th> <th style="border: none;">A</th> <th style="border: none;">B</th> </tr> </thead> <tbody> <tr> <td style="border: none;"></td> <td style="border: none;">a</td> <td style="border: none;">b</td> </tr> <tr> <td style="border: none;"></td> <td style="border: none;">c</td> <td style="border: none;">d</td> </tr> <tr> <td style="border: none;"></td> <td style="border: none;">x</td> <td style="border: none;">y</td> </tr> </tbody> </table>	R_1	A	B		a	b		c	d		x	y	<table border="1" style="border-collapse: collapse; width: 100%; text-align: center;"> <thead> <tr> <th style="border: none;">R_2</th> <th style="border: none;">A</th> </tr> </thead> <tbody> <tr> <td style="border: none;"></td> <td style="border: none;">a d</td> </tr> </tbody> </table>	R_2	A		a d	<table border="1" style="border-collapse: collapse; width: 100%; text-align: center;"> <thead> <tr> <th style="border: none;">R_3</th> <th style="border: none;">$R_1.A$</th> <th style="border: none;">B</th> <th style="border: none;">$R_2.A$</th> </tr> </thead> <tbody> <tr> <td style="border: none;"></td> <td style="border: none;">a</td> <td style="border: none;">b</td> <td style="border: none;">a</td> </tr> <tr> <td style="border: none;"></td> <td style="border: none;">c</td> <td style="border: none;">d</td> <td style="border: none;">a</td> </tr> <tr> <td style="border: none;"></td> <td style="border: none;">x</td> <td style="border: none;">y</td> <td style="border: none;">a</td> </tr> <tr> <td style="border: none;"></td> <td style="border: none;">a</td> <td style="border: none;">b</td> <td style="border: none;">d</td> </tr> <tr> <td style="border: none;"></td> <td style="border: none;">c</td> <td style="border: none;">d</td> <td style="border: none;">d</td> </tr> <tr> <td style="border: none;"></td> <td style="border: none;">x</td> <td style="border: none;">y</td> <td style="border: none;">d</td> </tr> </tbody> </table>	R_3	$R_1.A$	B	$R_2.A$		a	b	a		c	d	a		x	y	a		a	b	d		c	d	d		x	y	d
R_1	A	B																																												
	a	b																																												
	c	d																																												
	x	y																																												
R_2	A																																													
	a d																																													
R_3	$R_1.A$	B	$R_2.A$																																											
	a	b	a																																											
	c	d	a																																											
	x	y	a																																											
	a	b	d																																											
	c	d	d																																											
	x	y	d																																											

Sintaxa: **PRODUCT** (R_1, R_2) - rezultatul este o relație.

5. **SELECT** – este o operație unară de restricție care selectează din tuplurile relației R , acele tupluri care satisfac o condiție specificată. Condiția este o expresie logică (predicat) specificată asupra atributelor relației R . Condiția poate cuprinde nume de atribute, constante, operatori logici (and, or, not), operatori aritmetici de comparare ($<, >, =, \leq, \geq, \neq$).

$$\text{Relatie} \times \text{Expresie logica} \rightarrow \text{Relatie}$$

$S = \sigma_{\text{conditie}}(R) = \{t \mid t \in R \text{ si } t(A)\theta\alpha\}$ unde: $t(A) \theta \alpha$ definește condiția de selecție (θ este un operator aritmetic de comparare, iar α un tuplu care poate fi înlocuit de un atribut sau valoarea unui atribut).

Exemplu.

$$S = \sigma_{B='x'}(R)$$

<table border="1"> <thead> <tr> <th>R</th> <th>A</th> <th>B</th> <th>C</th> </tr> </thead> <tbody> <tr> <td></td> <td>x</td> <td>y x</td> <td>z a</td> </tr> <tr> <td></td> <td>t</td> <td>x u</td> <td>b</td> </tr> <tr> <td></td> <td>z</td> <td></td> <td>w</td> </tr> <tr> <td></td> <td>c</td> <td></td> <td></td> </tr> </tbody> </table>				R	A	B	C		x	y x	z a		t	x u	b		z		w		c			<table border="1"> <thead> <tr> <th>S</th> <th>A</th> <th>B</th> <th>C</th> </tr> </thead> <tbody> <tr> <td></td> <td>t</td> <td>x</td> <td>a</td> </tr> <tr> <td></td> <td>z</td> <td>x</td> <td>b</td> </tr> </tbody> </table>				S	A	B	C		t	x	a		z	x	b
R	A	B	C																																				
	x	y x	z a																																				
	t	x u	b																																				
	z		w																																				
	c																																						
S	A	B	C																																				
	t	x	a																																				
	z	x	b																																				

Sintaxa: **SELECT (R; condiție)** - rezultatul este o relație.

Observații.

1. Această operație nu trebuie confundată cu instrucțiunea *SELECT*, care este instrucțiunea generală de interogare din limbajele de manipulare a datelor.
2. În termenii limbajului de interogare *SQL*, operația de selecție realizează o *decupare pe orizontală a tabelului operand R*.
3. Cardinalul relației rezultat *S* este mai mic sau egal decât cardinalul relației *R*. Egalitatea poate apare în situația în care condiția este adevărată pentru toate tuplurile din relație.

6. **PROJECT** - este o operație unară de restricție prin care se selectează din relația *R*, numai acele atribute specificate explicit în cadrul operației. Relația rezultată *P* va avea ca atribute submulțimea selectată.

Signatura: $Relatie \times Lista\ atribute \rightarrow Relatie$

Fie $R = [A_1, A_2, \dots, A_n]$ și $\alpha = \{A_1, \dots, A_k\} \subset \{A_1, A_2, \dots, A_n\}$

$$P = \Pi_{\alpha}(R) = \{t(a_1, \dots, a_k)\}$$

Exemplu.

$$P = \Pi_{A,C}(R)$$

<table border="1"> <thead> <tr> <th>R</th> <th>A</th> <th>B</th> <th>C</th> </tr> </thead> <tbody> <tr> <td></td> <td>x</td> <td>u</td> <td>x</td> </tr> <tr> <td></td> <td>z</td> <td>x</td> <td>y</td> </tr> <tr> <td></td> <td>x</td> <td>z</td> <td>x</td> </tr> <tr> <td></td> <td>z</td> <td>y</td> <td>y</td> </tr> <tr> <td></td> <td>x</td> <td>t</td> <td>x</td> </tr> </tbody> </table>				R	A	B	C		x	u	x		z	x	y		x	z	x		z	y	y		x	t	x	<table border="1"> <thead> <tr> <th>P'</th> <th>A</th> <th>C</th> </tr> </thead> <tbody> <tr> <td></td> <td>x</td> <td>x</td> </tr> <tr> <td></td> <td>z</td> <td>y</td> </tr> <tr> <td></td> <td>x</td> <td>x</td> </tr> <tr> <td></td> <td>z</td> <td>y</td> </tr> <tr> <td></td> <td>x</td> <td>x</td> </tr> </tbody> </table>			P'	A	C		x	x		z	y		x	x		z	y		x	x	<table border="1"> <thead> <tr> <th>P</th> <th>A</th> <th>C</th> </tr> </thead> <tbody> <tr> <td></td> <td>x</td> <td>x</td> </tr> <tr> <td></td> <td>z</td> <td>y</td> </tr> </tbody> </table>		P	A	C		x	x		z	y
R	A	B	C																																																								
	x	u	x																																																								
	z	x	y																																																								
	x	z	x																																																								
	z	y	y																																																								
	x	t	x																																																								
P'	A	C																																																									
	x	x																																																									
	z	y																																																									
	x	x																																																									
	z	y																																																									
	x	x																																																									
P	A	C																																																									
	x	x																																																									
	z	y																																																									

Sintaxa: **PROJECT (R; lista atribute)** - rezultatul este o relație.

Observații.

1. Dacă în lista atributelor de proiecție există o cheie a relației operand *R*, atunci relația rezultat are toate tuplurile distincte, adică relațiile *R* și *P* vor avea același cardinal.

2. Dacă în lista atributelor de proiecție nu există o cheie a relației operand R, atunci în relația rezultat P pot apare tupluri duplicate care vor fi eliminate. În exemplul prezentat după eliminarea tuplurilor duplicate din relația intermediară P', s-a obținut relația P care conține două tupluri distincte.
3. Relația rezultat P are gradul k, dat de numărul atributelor din listă.
4. În termenii limbajului de interogare SQL, operația de proiecție realizează o decupare pe verticală a tablei operand R.

7. **JOIN** – este o operație definită pe două relații R₁ și R₂. Relația rezultat R₃ va fi construită prin concatenarea unor tupluri din R₁ cu tupluri din R₂ care satisfac o anumită condiție (*condiția de joncțiune* - θ) specificată explicit în cadrul relației. Condiția de joncțiune - θ este o expresie logică (predicat) specificată asupra atributelor relațiilor R₁ și R₂. Condiția de joncțiune - θ poate cuprinde nume de atribute, constante, operatori logici (and, or, not), operatori aritmetici de comparare (<, >, =, ≤, ≥, ≠).

$$Relatie \times Relatie \times \theta \text{ expresie} \rightarrow Relatie$$

Fie relațiile R₁ = [A, B₁] și R₂ = [B₂, C], unde B₁ și B₂ sunt atribute definite pe același domeniu. Atunci:

$$R_3 = R_1 \otimes_{\theta} R_2 = \{t \mid t \in R_1 \times R_2 \text{ si } t(B_1)\theta t(B_2)\}$$

Observație. Joncțiunea se poate exprima în funcție de operațiile de bază: *produs cartezian* și *selecție* astfel:

$$R_1 \otimes_{\theta} R_2 = \sigma_{\theta}(R_1 \times R_2)$$

Exemplu.

$$R_3 = R_1 \otimes_{\theta} R_2 \text{ unde } \theta : R_1.A > R_2.A$$

<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">R1</th> <th style="text-align: center;">A</th> <th style="text-align: center;">B</th> <th style="text-align: center;">C</th> </tr> </thead> <tbody> <tr><td></td><td style="text-align: center;">a</td><td style="text-align: center;">x</td><td style="text-align: center;">c</td></tr> <tr><td></td><td style="text-align: center;">b</td><td style="text-align: center;">y</td><td style="text-align: center;">c</td></tr> <tr><td></td><td style="text-align: center;">d</td><td style="text-align: center;">z</td><td style="text-align: center;">g</td></tr> </tbody> </table>	R1	A	B	C		a	x	c		b	y	c		d	z	g	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">R2</th> <th style="text-align: center;">D</th> <th style="text-align: center;">E</th> <th style="text-align: center;">A</th> </tr> </thead> <tbody> <tr><td></td><td style="text-align: center;">0</td><td style="text-align: center;">11</td><td style="text-align: center;">a</td></tr> <tr><td></td><td style="text-align: center;">1</td><td style="text-align: center;">13</td><td style="text-align: center;">a</td></tr> <tr><td></td><td style="text-align: center;">3</td><td style="text-align: center;">11</td><td style="text-align: center;">a</td></tr> <tr><td></td><td style="text-align: center;">4</td><td style="text-align: center;">11</td><td style="text-align: center;">d</td></tr> <tr><td></td><td style="text-align: center;">6</td><td style="text-align: center;">12</td><td style="text-align: center;">d</td></tr> <tr><td></td><td style="text-align: center;">7</td><td style="text-align: center;">13</td><td style="text-align: center;">c</td></tr> </tbody> </table>	R2	D	E	A		0	11	a		1	13	a		3	11	a		4	11	d		6	12	d		7	13	c	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">R3</th> <th style="text-align: center;">R₁. A</th> <th style="text-align: center;">B</th> <th style="text-align: center;">C</th> <th style="text-align: center;">D</th> <th style="text-align: center;">E</th> <th style="text-align: center;">R₂. A</th> </tr> </thead> <tbody> <tr><td></td><td style="text-align: center;">b b</td><td style="text-align: center;">y y</td><td style="text-align: center;">c c</td><td style="text-align: center;">0</td><td style="text-align: center;">11</td><td style="text-align: center;">a a</td></tr> <tr><td></td><td style="text-align: center;">b d</td><td style="text-align: center;">y z</td><td style="text-align: center;">c g</td><td style="text-align: center;">1</td><td style="text-align: center;">13</td><td style="text-align: center;">a a</td></tr> <tr><td></td><td style="text-align: center;">d d</td><td style="text-align: center;">z z</td><td style="text-align: center;">g g</td><td style="text-align: center;">3</td><td style="text-align: center;">11</td><td style="text-align: center;">a a</td></tr> <tr><td></td><td></td><td></td><td></td><td style="text-align: center;">0</td><td style="text-align: center;">11</td><td></td></tr> <tr><td></td><td></td><td></td><td></td><td style="text-align: center;">1</td><td style="text-align: center;">13</td><td></td></tr> <tr><td></td><td></td><td></td><td></td><td style="text-align: center;">3</td><td style="text-align: center;">11</td><td></td></tr> <tr><td></td><td style="text-align: center;">d</td><td style="text-align: center;">z</td><td style="text-align: center;">g</td><td style="text-align: center;">7</td><td style="text-align: center;">13</td><td style="text-align: center;">c</td></tr> </tbody> </table>	R3	R ₁ . A	B	C	D	E	R ₂ . A		b b	y y	c c	0	11	a a		b d	y z	c g	1	13	a a		d d	z z	g g	3	11	a a					0	11						1	13						3	11			d	z	g	7	13	c
R1	A	B	C																																																																																																			
	a	x	c																																																																																																			
	b	y	c																																																																																																			
	d	z	g																																																																																																			
R2	D	E	A																																																																																																			
	0	11	a																																																																																																			
	1	13	a																																																																																																			
	3	11	a																																																																																																			
	4	11	d																																																																																																			
	6	12	d																																																																																																			
	7	13	c																																																																																																			
R3	R ₁ . A	B	C	D	E	R ₂ . A																																																																																																
	b b	y y	c c	0	11	a a																																																																																																
	b d	y z	c g	1	13	a a																																																																																																
	d d	z z	g g	3	11	a a																																																																																																
				0	11																																																																																																	
				1	13																																																																																																	
				3	11																																																																																																	
	d	z	g	7	13	c																																																																																																

Sintaxa: **THETA - JOIN (R₁, R₂; θ - expresie)** - rezultatul este o relație.

În continuare se prezintă patru forme ale operației de joncțiune.

7.1. **EQUI - JOIN** – este un caz particular al lui **THETA - JOIN**, când θ este egalitate.

Exemplu.

$$R_3 = R_1 \otimes_{\theta} R_2 \text{ unde } \theta : R_1.A = R_2.A$$

<table border="1"> <thead> <tr> <th>R1</th> <th>A</th> <th>B</th> <th>C</th> </tr> </thead> <tbody> <tr> <td></td> <td>a</td> <td>x</td> <td>c</td> </tr> <tr> <td></td> <td>b</td> <td>y</td> <td>c</td> </tr> <tr> <td></td> <td>d</td> <td>z</td> <td>a</td> </tr> </tbody> </table>	R1	A	B	C		a	x	c		b	y	c		d	z	a	<table border="1"> <thead> <tr> <th colspan="3">R2</th> </tr> <tr> <th>D</th> <th>E</th> <th>A</th> </tr> </thead> <tbody> <tr> <td>0</td> <td>11</td> <td>a a</td> </tr> <tr> <td>1</td> <td>13</td> <td>a</td> </tr> <tr> <td>3</td> <td>11</td> <td>d</td> </tr> <tr> <td>4</td> <td>11</td> <td>d</td> </tr> <tr> <td>6</td> <td>12</td> <td>c</td> </tr> <tr> <td>7</td> <td>13</td> <td></td> </tr> </tbody> </table>	R2			D	E	A	0	11	a a	1	13	a	3	11	d	4	11	d	6	12	c	7	13		<table border="1"> <thead> <tr> <th>R3</th> <th>R1. A</th> <th>B</th> <th>C</th> <th>D</th> <th>E</th> <th>R2. A</th> </tr> </thead> <tbody> <tr> <td></td> <td>a</td> <td>x</td> <td>c</td> <td>0</td> <td>11</td> <td>a</td> </tr> <tr> <td></td> <td>a</td> <td>x</td> <td>c</td> <td>1</td> <td>13</td> <td>a</td> </tr> <tr> <td></td> <td>a</td> <td>x</td> <td>c</td> <td>3</td> <td>11</td> <td>a</td> </tr> <tr> <td></td> <td>d</td> <td>z</td> <td>a</td> <td>4</td> <td>11</td> <td>d</td> </tr> <tr> <td></td> <td>d</td> <td>z</td> <td>a</td> <td>6</td> <td>12</td> <td>d</td> </tr> </tbody> </table>	R3	R1. A	B	C	D	E	R2. A		a	x	c	0	11	a		a	x	c	1	13	a		a	x	c	3	11	a		d	z	a	4	11	d		d	z	a	6	12	d
R1	A	B	C																																																																																	
	a	x	c																																																																																	
	b	y	c																																																																																	
	d	z	a																																																																																	
R2																																																																																				
D	E	A																																																																																		
0	11	a a																																																																																		
1	13	a																																																																																		
3	11	d																																																																																		
4	11	d																																																																																		
6	12	c																																																																																		
7	13																																																																																			
R3	R1. A	B	C	D	E	R2. A																																																																														
	a	x	c	0	11	a																																																																														
	a	x	c	1	13	a																																																																														
	a	x	c	3	11	a																																																																														
	d	z	a	4	11	d																																																																														
	d	z	a	6	12	d																																																																														

Sintaxa: **EQUI - JOIN** (R₁, R₂; θ - expresie) - rezultatul este o relație.

7.2. NATURAL – JOIN – este o joncțiune pe egalitate (*EQUI – JOIN*) pentru toate atributele cu același nume din cele două relații, urmată de o proiecție pe reuniunea atributelor celor două relații. În cazul *EQUI – JOIN* schema relației rezultat conține toate atributele celor doi operanzi și rezultă că în fiecare tuplu al relației rezultat vor exista cel puțin două valori egale. Introducerea *joncțiunii naturale* va elimina această redundanță.

Schema relației rezultat R₃ se obține prin reuniunea atributelor celor două relații R₁ și R₂ (atributele cu același nume se iau o singură dată), iar extensia relației R₃ va conține tuplurile obținute prin concatenarea tuplurilor din R₁ cu tupluri din R₂, care au aceleași valori pentru atributele cu același nume.

Joncțiunea naturală este joncțiunea cea mai utilizată în practică și poate fi definită cu ajutorul operațiilor de bază: *proiecție, selecție și produs cartezian*.

Dacă se notează cu: θ - condiția de egalitate între valorile atributelor din intersecția schemelor relațiilor R₁ și R₂ (coloanele comune), atr - reuniunea atributelor celor două scheme (atributele cu același nume se iau o singură dată), atunci:

$$R_3 = R_1 \otimes R_2 = \pi_{atr}(\sigma_{\theta}(R_1 \times R_2))$$

Exemplu

A. Se dau relațiile: R₁[A, B, C] și R₂[B, C, D] θ = (R₁.B = R₂.B) and (R₁.C = R₂.C) atr = A, B, C, D

<table border="1"> <thead> <tr> <th>R1</th> <th>A</th> <th>B</th> <th>C</th> </tr> </thead> <tbody> <tr> <td></td> <td>a</td> <td>b</td> <td>c</td> </tr> <tr> <td></td> <td>d</td> <td>b</td> <td>c</td> </tr> <tr> <td></td> <td>b</td> <td>b</td> <td>f</td> </tr> <tr> <td></td> <td>c</td> <td>a</td> <td>d</td> </tr> </tbody> </table>	R1	A	B	C		a	b	c		d	b	c		b	b	f		c	a	d	<table border="1"> <thead> <tr> <th colspan="3">R2</th> </tr> <tr> <th>B</th> <th>C</th> <th>D</th> </tr> </thead> <tbody> <tr> <td>b</td> <td>c</td> <td>d</td> </tr> <tr> <td>b</td> <td>c</td> <td>e</td> </tr> <tr> <td>a</td> <td>d</td> <td>b</td> </tr> </tbody> </table>	R2			B	C	D	b	c	d	b	c	e	a	d	b	<table border="1"> <thead> <tr> <th>R3</th> <th>A</th> <th>B</th> <th>C</th> <th>D</th> </tr> </thead> <tbody> <tr> <td></td> <td>a a</td> <td>b</td> <td>c c</td> <td>d</td> </tr> <tr> <td></td> <td>d</td> <td>b</td> <td>c c</td> <td>e</td> </tr> <tr> <td></td> <td>d</td> <td>b</td> <td></td> <td>d</td> </tr> <tr> <td></td> <td></td> <td>b</td> <td></td> <td>e</td> </tr> <tr> <td></td> <td>c</td> <td>a</td> <td>d</td> <td>b</td> </tr> </tbody> </table>	R3	A	B	C	D		a a	b	c c	d		d	b	c c	e		d	b		d			b		e		c	a	d	b
R1	A	B	C																																																																
	a	b	c																																																																
	d	b	c																																																																
	b	b	f																																																																
	c	a	d																																																																
R2																																																																			
B	C	D																																																																	
b	c	d																																																																	
b	c	e																																																																	
a	d	b																																																																	
R3	A	B	C	D																																																															
	a a	b	c c	d																																																															
	d	b	c c	e																																																															
	d	b		d																																																															
		b		e																																																															
	c	a	d	b																																																															

B. Se dau relațiile: $R_1 [A, B, C]$ și $R_2 [D, E, A]$ $\theta = R_1.A = R_2.A$
 $atr = A, B, C, D, E$

R_1	A	B	C
	a	x	c
	b	y	c
	d	z	a

R_2	D	E	A
	0	11	a
	1	13	a
	3	11	a
	4	11	d
	6	12	d
	7	13	c

R_3	A	B	C	D	E
	a	x	c	0	11
	a	x	c	1	13
	a	x	c	3	11
	d	z	a	4	11
	d	z	a	6	12

Sintaxa: **NATURAL - JOIN (R_1, R_2 ; atribut(e) joncțiune*)** - rezultatul este o relație.

* pentru mărirea clarității va apare atributul / atributele de joncțiune.

Observație.

Selecția este un caz particular de *joncțiune naturală* a unei relații cu o *relație constantă*. Înțelegem prin *relație constantă* o relație care are un singur tuplu, eventual redus la o singură valoare.

7.3. SEMI – JOIN - este joncțiunea dintre două relații R_1 și R_2 , urmată de o proiecție pe atributele relației R_1 .

Semi – joncțiunea conservă atributele unei relații participante la joncțiune (R_1). *Semi – joncțiunea* mai poate fi privită ca o generalizare a operației de *selecție*, rezultatul fiind o selecție asupra relației R_1 , realizată pe baza valorilor din R_2 ale atributului de joncțiune.

Exemplu

$$R_3 = R_1 \otimes_{\theta} R_2 \text{ unde } \theta : R_1.A = R_2.A$$

R_1	A	B	C
	a	x	c
	b	y	c
	d	z	a

R_2	D	E	A
	0	11	a
	1	13	a
	3	11	a
	4	11	d
	6	12	d
	7	13	c

R_3	A	B	C
	a	x	c
	a	x	c
	a	x	c
	d	z	a
	d	z	a

Sintaxa: **SEMI - JOIN (R_1, R_2 ; θ - expresie)** - rezultatul este o relație.

Observație.

Considerăm *joncțiunea naturală* dintre $R_1 = [A, B_1]$ și $R_2 = [B_2, C]$, unde B_1 și B_2 sunt atributele de joncțiune.

Dacă $B_1 = B_2 = \Phi$, atunci joncțiunea corespunde *produsului cartezian*.

Dacă $A = C = \Phi$, atunci joncțiunea corespunde *intersecției*.

Dacă $A = \Phi$ sau $C = \Phi$ (dar nu amândouă), atunci operația este o *semijoncțiune*.

7.4. OUTER – JOIN. Joncțiunea dintre două relații R_1 și R_2 poate conduce la pierdere de tupluri, dacă relațiile participante la joncțiune nu au proiecții identice pe atributul de joncțiune, adică nu au aceleași valori în relațiile care se joncționează. Relația rezultat R_3 conține *joncțiunea naturală* dintre R_1 și R_2 , la care se adaugă tuplurile din R_1 și R_2 , care nu au participat la joncțiune. În aceste tupluri se va atribui valoarea *null* pentru atributele relației corespondente.

Exemplu

$$R_3 = R_1 \otimes_{ext} R_2$$

R1	A	B	C		R2	D	E	A		R3	A	B	C	D	E
	a	x	c			0	11	a			a	x	c c	0	11
	b	y	c			1	13	a			a	x	c a	1	13
	d	z	a			3	11	a			a	x	a	3	11
						4	11	d			d	z z	c	4	11
						6	12	d			d	y		6	12
						7	13	c			b			null	null
											c	null	null	7	13

Observații. În urma *joncțiunii naturale* se pierd informațiile din tuplurile $\langle b, y, c \rangle$ din R_1 și $\langle 7, 13, c \rangle$ din R_2 . Aceste tupluri se adaugă în cazul *joncțiunii externe* și se completează cu *null* pe atributele relației corespondente.

Sintaxa: **OUTER - JOIN (R₁, R₂; atribut(e) joncțiune)** - rezultatul este o relație.

8. DIVISION - este o operație definită pe două relații care au schema

$$R_1 [A_1, A_2, \dots, A_n] \text{ și } R_2 [A_{p+1}, A_{p+2}, \dots, A_n]. \text{ Relația rezultat } R_3 = R_1 \div R_2 \text{ are schema}$$

$R_3 [A_1, A_2, \dots, A_p]$ și este formată din toate tuplurile care, concatenate cu fiecare tuplu din R_2 , dau întotdeauna un tuplu din R_1 .

Notăm: $ATR_1 = \{A_1, A_2, \dots, A_{p+1}, A_{p+2}, \dots, A_n\}$ $ATR_2 = \{A_{p+1}, A_{p+2}, \dots, A_n\}$

Definiția 1.

$$t \in R_1 \div R_2 \text{ dacă } \forall t_2 \in R_2 \exists t_1 \in R_1, \text{ astfel încât } \Pi_{ATR_1 - ATR_2}(t_1) = t \text{ și } \Pi_{ATR_2}(t_1) = t_2.$$

Definiția 2. *Diviziunea* se poate exprima în funcție de operațiile de bază: *produs cartezian*, *diferență* și *proiecție* astfel:

$$R_1 \div R_2 = \Pi_{ATR_1 - ATR_2}(R_1) - \Pi_{ATR_1 - ATR_2}((\Pi_{ATR_1 - ATR_2}(R_1) \times R_2) - R_1)$$

Problemă (Exemplu de diviziune).

Fie relația $R_1 [K, P]$ unde atributul K are ca valori codurile angajaților unui institut de cercetare, iar atributul P conține codurile proiectelor în derulare. Un cercetător poate lucra la unul sau mai multe proiecte. Să se determine codurile angajaților angrenați simultan în proiectele $P3$ și $P4$.

Rezolvare.

Construim relația $R_2 [P]$ care va conține două tupluri: $\langle P3 \rangle$ și $\langle P4 \rangle$.

Codurile angajaților care lucrează la proiectele $P3$ și $P4$ sunt date de rezultatul diviziunii

$R_1 \div R_2$.

R_1	K	P
	17	P1
	17	P2
	17	P3
	17	P4
	29	P1
	29	P3
	53	P3
	53	P4
	80	P3

R_2	P
	P3
	P4

Calculăm *diviziunea* conform definiției 2: $R_1 \div R_2 = Q_1 - Q_2$

Pasul 1. Calculăm $Q_1 = \Pi_{ATR1} -_{ATR2} (R_1)$

Pasul 2. Calculăm $S = Q_1 \times R_2$

Q_1	K
	17
	29
	53
	80

S	K	P
	17	P3
	29	P3
	53	P3
	80	P3
	17	P4
	29	P4
	53	P4
	80	P4

Pasul 3. Calculăm $T = S - R_1$

T	K	P
	29	P4
	80	P4

Pasul 4. Calculăm $Q_2 = \prod_{ATR1} -_{ATR2} (T)$

T	K
	29
	80

Pasul 5. Calculăm $R_1 \div R_2 = Q_1 - Q_2$

R1:R2	K
	17
	53

Rezultatul interogării: angajații cu codul <17> și <53> lucrează simultan în proiectele <P3> și <P4>.

Sintaxa: **DIVISION (R₁, R₂)** - rezultatul este o relație.

2.2. OPERAȚII DE CALCUL

La operațiile descrise anterior se pot adăuga *operații de calcul pe relații*. Aceste operații sunt justificate de numeroasele interogări (cereri) care necesită operații de calcul. Operațiile de calcul sunt implementate în toate limbajele de interogare. Acești operatori de calcul formează deci o extensie a operatorilor de bază și nu pot fi exprimați cu ajutorul acestora.

1. COUNT - este o operație care permite numărarea tuplurilor dintr-o relație (liniilor dintr-o tabelă) care au aceeași valoare pe atributul considerat (sau aceleași valori pe atributele considerate). Relația rezultantă va conține numai atributul (atributele) de regrupare X_i , iar tuplurile vor fi formate din valorile distincte și numărul de apariții.

Notăm:

$$T = COUNT_{X_1, \dots, X_n} (R), \text{ unde } X_1, \dots, X_n \text{ sunt atributele de regrupare.}$$

Operatorul *COUNT*

Exemplu.

R	A	B	C	$Count_B(R)$	B	Count	$Count_{B,C}(R)$	B	C	Count
	a	n	17		n	2		n	17	2
	b	o	14		m	2		m	20	1
	c	n	17		o	1		m	10	1
	d	p	13		p	1		o	14	1
	e	m	20					p	13	1
	f	m	10							

Dacă nu este precizat niciun atribut de regrupare, operația *COUNT* va determina numărul de tupluri din relație:

Count(R)	Count
	6

Sintaxa: **COUNT (R; X₁, X₂, ..., X_n)** - rezultatul este o relație;

COUNT (R) - rezultatul este un număr (poate fi interpretat și ca o relație cu un singur atribut și un singur tuplu, care are ca valoare numărul de linii din tabelă).

2. SUM – este o operație care permite efectuarea sumei valorilor atributului Y pentru fiecare din valorile diferite ale atributelor de regrupare X_1, \dots, X_n . Atributul Y trebuie să fie numeric.

Notăm:

$T = SUM_{X_1, \dots, X_n} (R, Y)$, unde X_1, \dots, X_n sunt atributele de regrupare.

Dacă nu este precizat niciun atribut de regrupare, operația SUM va determina suma valorilor atributului Y .

Sintaxa: **SUM (R, Y; X₁, X₂, ..., X_n)** - rezultatul este o relație;

SUM (R, Y) - rezultatul este un număr (poate fi interpretat și ca o relație cu un singur atribut și un singur tuplu, care are ca valoare suma valorilor atributului Y din toate liniile tablei).

3. AVG – este o operație care permite efectuarea mediei aritmetice a valorilor atributului Y pentru fiecare din valorile diferite ale atributelor de regrupare X_1, \dots, X_n . Atributul Y trebuie să fie numeric.

Notăm:

$T = AVG_{X_1, \dots, X_n} (R, Y)$, unde X_1, \dots, X_n sunt atributele de regrupare.

Dacă nu este precizat niciun atribut de regrupare, operația AVG va determina media aritmetică a valorilor atributului Y din toată relația.

Sintaxa: **AVG (R, Y; X₁, X₂, ..., X_n)** - rezultatul este o relație;

AVG (R, Y) - rezultatul este un număr (poate fi interpretat și ca o relație cu un singur atribut și un singur tuplu, care are ca valoare media aritmetică a valorilor atributului Y din toate liniile tablei).

4. MAX și MIN - este o operație care permite determinarea valorii maxime / minime a atributului Y pentru fiecare din valorile diferite ale atributelor de regrupare X_1, \dots, X_n . Atributul Y trebuie să fie numeric.

Notăm:

$T = MAX_{X_1, \dots, X_n} (R, Y)$, unde X_1, \dots, X_n sunt atributele de regrupare.

$T = MIN_{X_1, \dots, X_n} (R, Y)$, unde X_1, \dots, X_n sunt atributele de regrupare.

Dacă nu este precizat niciun atribut de regrupare, operația MAX (MIN) va determina maximul (minimul) valorilor atributului Y din toată relația.

Sintaxa: **MAX (R, Y; X₁, X₂, ..., X_n)** - rezultatul este o relație;

MAX (R, Y) - rezultatul este un număr (poate fi interpretat și ca o relație cu un singur atribut și un singur tuplu, care are ca valoare maximul valorilor atributului Y din toate liniile tablei).

Observație. Pentru operația MIN sintaxa este similară.

3. Referințe bibliografice

https://en.wikipedia.org/wiki/Relational_algebra

https://www.cs.purdue.edu/homes/bertino/348Spring2012/relational_algebra_and_calculus.pdf