

מסדי נתונים – 89-281

תרגול 6 - אלג' יחסים 2

מבוסס על מצגות של עמיעד רוזנברג

Computers are like air conditioners.
They work fine until you start opening windows

תזכורת

- בתרגול הקודם ראינו 4 פעולות :
- Select – מאפשר לבחור רשומות ספציפיות מתוך היחס.
- Project – מאפשר לבחור עמודות ספציפיות מתוך היחס.
- Union – מאחד רשומות בין 2 יחסים תואמים.
- Difference – מחזיר רשומות המופיעות ביחס אחד מתוך רשומות המופיעות ביחס אחר. (היחסים צריכים להיות תואמים).

× – Cartesian Product מכפלה קרטזית

- פעולה בינארית המאפשרת לשלב מידע משני יחסים.

- סימון: $R_1 \times R_2$

- אם R_1 יחס מדרגה n (n תכונות) ו R_2 יחס מדרגה m , אזי $R_1 \times R_2$ היא יחס מדרגה $n+m$.

- $R_1 \times R_2 = \{t, s \mid t \in R_1, s \in R_2\}$

- מס' השורות ביחס המתקבל מהפעולה $R_1 \times R_2$ = מס' השורות ב- R_1 כפול מס' השורות ב- R_2 .

- אם בשני היחסים יש עמודות עם שם זהה אזי בתוצאת המכפלה הקרטזית שמות העמודות המתאימות יהיו `relationName.ColumnName`

- לדוגמה, הסכימה של היחס המתקבל כתוצאה מ- $\text{borrower} \times \text{loan}$ היא `customer-name, borrower.loan-number, branch-name, loan.loan-number, ammount`

× – Cartesian Product מכפלה קרטזית

• דוגמא:

R1		
<i>Name</i>	<i>ID</i>	<i>Phone</i>
Avi	012	345
Bar	013	456
Gal	014	567

R2	
<i>Course</i>	<i>Number</i>
DB	281
Complexity	320

$$R_1 \times R_2 =$$

<i>Name</i>	<i>ID</i>	<i>Phone</i>	<i>Course</i>	<i>Number</i>
Avi	012	345	DB	281
Avi	012	345	Complexity	320
Bar	013	456	DB	281
Bar	013	456	Complexity	320
Gal	014	567	DB	281
Gal	014	567	Complexity	320

כינוי $\rho_X(R)$ – Rename

- פעולה אונארית המעניקה את השם X ליחס נתון R .

- דוגמא:

- $\rho_{\text{no-borrow}}(\pi_{\text{customer-name}}(\text{customer}) - \pi_{\text{customer-name}}(\text{borrower}))$

customer
customer-name
Jones
Smith
Hayes
Curry
Lindsay
Turner
Williams
Adams
Johnson
Glenn
Brooks
Green

=

borrower
customer-name
Jones
Smith
Hayes
Jackson
Curry
Williams
Adams

=

no-borrow
customer-name
Lindsay
Turner
Johnson
Glenn
Brooks
Green

תרגיל ביניים 1

- מציאת שמות כל הלקוחות שלקחו הלוואה בסניף 'Downtown':
- המידע הדרוש נמצא ביחסים *loan* ו *borrower*, לכן בשלב ראשון נבצע מכפלה קרטזית בין שני היחסים.

loan		
branch-name	loan-number	amount
Downtown	L-17	1000
Redwood	L-23	2000
Perryridge	L-15	1500
Downtown	L-14	1500
Mianus	L-93	500
Round Hill	L-11	900
Perryridge	L-16	1300

X

borrower	
customer-name	loan-number
Jones	L-17
Smith	L-23
Hayes	L-15
Jackson	L-14
Curry	L-93
Smith	L-11
Williams	L-17
Adams	L-16

=

branch-name	loan.loan-number	amount	customer-name	borrower.loan-number
Downtown	L-17	1000	Jones	L-17
Downtown	L-17	1000	Smith	L-23
Downtown	L-17	1000	Hayes	L-15
Downtown	L-17	1000	Jackson	L-14
Downtown	L-17	1000	Curry	L-93
Downtown	L-17	1000	Smith	L-11
Downtown	L-17	1000	Williams	L-17
Redwood	L-23	2000	Jones	L-17
Redwood	L-23	2000	Smith	L-23
Redwood	L-23	2000	Hayes	L-15
...

תרגיל ביניים 1

- מציאת שמות כל הלקוחות שלקחו הלוואה בסניף 'Downtown':
- שלב 2: קיבלנו הרבה שורות לא רלוונטיות. נבחר את השורות הרלוונטיות לנו, כלומר אילו שבהן יש התאמה במספר החשבון.

• $\sigma_{\text{borrower.loan-number} = \text{loan.loan-number}} (\text{borrower} \times \text{loan})$

branch-name	loan.loan-number	amount	customer-name	borrower.loan-number
Downtown	L-17	1000	Jones	L-17
Downtown	L-17	1000	Williams	L-17
Redwood	L-23	2000	Smith	L-23
Perryridge	L-15	1500	Hayes	L-15
Downtown	L-14	1500	Jackson	L-14
Mianus	L-93	500	Curry	L-93
Round Hill	L-11	900	Williams	L-17
Perryridge	L-16	1300	Adams	L-16

תרגיל ביניים 1

- מציאת שמות כל הלקוחות שלקחו הלוואה בסניף 'Downtown':
- נבחר את השורות המדברות על ההלוואה בסניף הרצוי.

- $\sigma_{\text{branch-name} = \text{'Downtown'}} (\sigma_{\text{borrower.loan-number} = \text{loan.loan-number}} (\text{borrower} \times \text{loan}))$

branch-name	loan.loan-number	amount	customer-name	borrower.loan-number
Downtown	L-17	1000	Jones	L-17
Downtown	L-17	1000	Williams	L-17
Downtown	L-14	1500	Jackson	L-14

תרגיל ביניים 1

- מציאת שמות כל הלקוחות שלקחו הלוואה בסניף 'Downtown':
- נבחר את המידע הרצוי.

- $\pi_{\text{customer-name}} ($
 $\sigma_{\text{branch-name} = \text{'Downtown'}} ($
 $\sigma_{\text{borrower.loan-number} = \text{loan.loan-number}} (\text{borrower} \times \text{loan}))$

customer-name
Jones
Williams
Jackson

תרגיל ביניים 2

- מציאת שמות הלקוחות הגרים באותה עיר כמו 'Green'.
- שלב ראשון – נבצע מכפלה קרטזית בין טבלאת לקוחות לעצמה.
מכיוון שלא נוכל לדעת איזו עמודה שייכת לאיזה יחס – ניתן לאחד היחסים שם אחר לפני ביצוע ההכפלה.
- $\text{customer} \times \rho_s(\text{customer})$
- שלב שני – נבחר את השורות שבהן מופיע Green
- $\sigma_{\text{customer.customer-name}='Green'}(\text{customer} \times \rho_s(\text{customer}))$

תרגיל ביניים 2

- מציאת שמות הלקוחות הגרים באותה עיר כמו 'Green'.

- שלב שלישי – נבחר את השורות שבהן שם העיר זהה לשם העיר של Green.

- $\sigma_{\text{customer.customer-city} = \text{s. customer-city}} (\sigma_{\text{customer.customer-name} = \text{'Green'}} (\text{customer} \times \rho_s(\text{customer})))$

- שלב רביעי – נציג את השמות המופיעים באותן רשומות

- $\pi_{\text{s.customer-name}} (\sigma_{\text{customer.customer-city} = \text{s. customer-city}} (\sigma_{\text{customer.customer-name} = \text{'Green'}} (\text{customer} \times \rho_s(\text{customer}))))$

חיתוך – Intersection – \cap

- פעולה בינארית היוצרת יחס המכיל רק n -יות המופיעות ביחס הראשון והשני.

- סימון: $R_1 \cap R_2$

- R_1 ו- R_2 צריכים להיות תואמים (מבחינת דרגה ותחומים)

- ליחס המוחזר אין סכמה עבור עמודות עם שם שונה!

- חיתוך ע"י פעולות בסיסיות:

$$R_1 \cap R_2 = R_1 - (R_1 - R_2) = R_2 - (R_2 - R_1)$$

חיתוך – Intersection – \cap

• דוגמא:

מצא את שמות כל הלקוחות שהפקידו כסף ולקחו הלוואה.

• $\pi_{c-name}(depositor) \cap \pi_{c-name}(borrower)$

צירוף טטה – Theta Join – \bowtie_{θ}

- פעולה בינארית המשלבת את המכפלה הקרטזית ופעולת הבחירה לפעולה אחת.

- סימון: $R_1 \bowtie_{\theta} R_2 = \sigma_{\theta}(R_1 \times R_2)$

- הפרדיקט θ :

- יכול להכיל יחס שוויון: $=, \neq, <, >, \leq, \geq, \wedge, \vee, \neg$,

- משווה בין תכונה ב- R_1 לתכונה ב- R_2 .

- היחס החדש שנוצר יכיל את התכונות של שני היחסים.

צירוף טטה – Theta Join – \bowtie_{θ}

• דוגמא:

מספרי כל הלוואות שנלקחו בעיר Brooklyn.

- $R_1 = (branch) \bowtie_{\substack{branch.b-name = \\ loan.b-name}} (loan)$
- $R_2 = \sigma_{b-city='brooklyn'}(R_1)$
- $\pi_{loan-number}(R_2)$

צירוף טבעי – Natural Join – \bowtie

- פעולה בינארית המשלבת :

1. מכפלה קרטזית.
2. בחירת tuples בעלי ערכים זהים בעמודות המשותפות.
3. ביצוע הטלה (כך שלא יהיו עמודות כפולות).

- סימון : $R_1 \bowtie R_2$

- פורמאלית : בהינתן 2 יחסים $R_1(S_1), R_2(S_2)$, הפעלת הצירוף תחזיר :

- $R_1 \bowtie R_2 = \pi_{S_1 \cup S_2} \left(\sigma_{R_1.A_1=R_2.A_1 \wedge \dots \wedge R_1.A_n=R_2.A_n} (R_1 \times R_2) \right)$
- העמודות $A_1 \dots A_n$ הן **עמודות בעלות שם זהה** בשני הטבלאות.

צירוף טבעי – Natural Join – \bowtie

• דוגמא:

R_1	
name	Grade
Avi	89
Beni	95
Gadi	97

R_2		
name	Grade1	Grade2
Avi	94	91
Beni	87	98

• שלב ראשון – מכפלה קרטזית:

$R_1 \times R_2$				
name	Grade	name	Grade1	Grade2
Avi	89	Avi	94	91
Avi	89	Beni	87	98
Beni	95	Avi	94	91
Beni	95	Beni	87	98
Gadi	97	Avi	94	91
Gadi	97	Beni	87	98

צירוף טבעי – Natural Join – \bowtie

$R_1 \times R_2$				
name	Grade	name	Grade1	Grade2
Avi	89	Avi	94	91
Avi	89	Beni	87	98
Beni	95	Avi	94	91
Beni	95	Beni	87	98
Gadi	97	Avi	94	91
Gadi	97	Beni	87	98

- שלב שני – בחירת tuples בעלי ערכים זהים בעמודות המשותפות :

$\sigma_{R_1.name=R_2.name} R_1 \times R_2$				
name	Grade	name	Grade1	Grade2
Avi	89	Avi	94	91
Beni	95	Beni	87	98

צירוף טבעי – Natural Join – \bowtie

$\sigma_{R_1.name=R_2.name} R_1 \times R_2$				
name	Grade	name	Grade1	Grade2
Avi	89	Avi	94	91
Beni	95	Beni	87	98

- שלב שלישי – ביצוע הטלה (כך שלא יהיו עמודות כפולות):

$\pi_{name, Grade, Grade1, Grade2} \sigma_{R_1.name=R_2.name} R_1 \times R_2$			
name	Grade	Grade1	Grade2
Avi	89	94	91
Beni	95	87	98

דוגמא נוספת: צירוף טבעי (Natural Join)

הסבר נוסף: כל הצירופים של רשומה מ-S ורשומה מ-T המסכימות ביניהן על התכונות המשותפות, כאשר מכל זוג תכונות משותפות מותירים עמודה אחת בלבד.

S	צבע	חיה
	לבן	סוס
	ורוד	פיל
	שחור	חתול

T	צבע	גוון
	כחול	בהיר
	כחול	כהה
	ורוד	בהיר
	ורוד	כהה

$S \bowtie T =$		
חיה	צבע	גוון
פיל	ורוד	בהיר
פיל	ורוד	כהה

– Outer Join – צירוף עם שמירת מידע

- בצירוף טבעי ראינו שחלק מהרשומות שלנו נעלמות. בדוגמא שראינו, גדי לא הופיע בטבלה השניה ולכן הרשומה שלו נעלמה לחלוטין מהתוצאה הסופית.
- במקרה שבו נרצה לשמור על המידע – נשתמש ב-Outer Join.
- ישנם 3 סוגים של Outer Join :

● Outer Join – שמירת מידע משני הטבלאות. 

● Right Outer Join – שמירת מידע מהטבלה הימנית. 

● Left Outer Join – שמירת מידע מהטבלה השמאלית. 

דוגמה

צירוף חיצוני של S ו-T

S

צבע	חיה
לבן	סוס
ורוד	פיל
שחור	חתול

T

צבע	גוון
כחול	בהיר
כחול	כהה
ורוד	בהיר
ורוד	כהה

$S \bowtie_{\text{outer}} T =$

גוון	צבע	חיה
בהיר	ורוד	פיל
כהה	ורוד	פיל
null	לבן	סוס
null	שחור	חתול
בהיר	כחול	null
כהה	כחול	null

קיימים צירוף חיצוני ימני (right outer join) וצירוף חיצוני שמאלי (left outer join) שבהם מרפדים ב-null רק את היחס הימני או השמאלי, בהתאמה

צירוף עם שמירת מידע – Outer Join –

דוגמא: ●

R_1	
name	Grade
Avi	89
Beni	95
Gadi	97

R_2		
name	Grade1	Grade2
Avi	94	91
Beni	87	98

$R_1 \bowtie R_2$			
name	Grade	Grade1	Grade2
Avi	89	94	91
Beni	95	87	98
Gadi	97	NULL	NULL

חילוק (מנה) – Division – \div

- פעולת החילוק היא פעולה בינארית הבוחרת מתוך היחס R_1 את הרשומות ה"מתאימות" **לכל הרשומות ב- R_2** ומחזירה את התכונות הקיימות ב- R_1 ולא ב- R_2 .
- בצורה פורמאלית:
 - $R_1(S_1)$ ו $R_2(S_2)$ הן יחסים כך שמתקיים $S_2 \subset S_1$.
 - היחס $R_1 \div R_2$ הוא בעל סכמה $S_1 \setminus S_2$.
- $$R_1 \div R_2 = \{t[S_1 \setminus S_2] : t \in R_1, \forall t_2 \in R_2 ((t[S_1 \setminus S_2] \cup t_2) \in R_1)\}$$

חילוק (מנה) – Division – \div

• דוגמא:

R_2		
B	C	D
b	c	d
b	cc	d
b	ccc	d

R_1			
A	B	C	D
1	b	c	d
1	b	cc	d
1	b	ccc	d
3	b	c	d

- נשים לב שמתקיים $S_2 \subset S_1$ ולכן הסכמה של $R_1 \div R_2$ תהיה $\{A, B, C, D\} \setminus \{B, C, D\} = \{A\}$

חילוק (מנה) – Division – \div

• דוגמא:

R_2		
B	C	D
b	c	d
b	cc	d
b	ccc	d

R_1			
A	B	C	D
1	b	c	d
1	b	cc	d
1	b	ccc	d
3	b	c	d

• נבדוק מהן הרשומות שנקבל:

$$R_1 \div R_2 = \{t[A] : t \in R_1, \forall t_2 \in R_2 ((t[A] \cup t_2) \in R_1)\}$$

חילוק (מנה) – Division – \div

• דוגמא:

R_2		
B	C	D
b	c	d
b	cc	d
b	ccc	d

R_1			
A	B	C	D
1	b	c	d
1	b	cc	d
1	b	ccc	d
3	b	c	d

• נסתכל על התנאי הראשון

$$R_1 \div R_2 = \{t[A] : t \in R_1, \forall t_2 \in R_2 ((t[A] \cup t_2) \in R_1)\}$$

• הערכים העונים עליו הם 1 ו-3

חילוק (מנה) – Division – \div

• דוגמא:

R_2		
B	C	D
b	c	d
b	cc	d
b	ccc	d

R_1			
A	B	C	D
1	b	c	d
1	b	cc	d
1	b	ccc	d
3	b	c	d

• נסתכל על התנאי השני

$$R_1 \div R_2 = \{t[A] : t \in R_1, \forall t_2 \in R_2 ((t[A] \cup t_2) \in R_1)\}$$

• מתוך הערכים 1 ו-3 רק הערך "1" עונה על התנאי השני.

$R_1 \div R_2$
A
1

חילוק (מנה) – Division – \div

• דוגמא:

R_2	
B	C
b	c
b	cc
b	ccc

R_1		
A	B	C
1	b	c
1	b	cc
1	b	ccc
3	b	c

$R_1 \div R_2$
A
1

• בד"כ משתמשים בחילוק בשאלות "כל".

• נחיה טיפה את הדוגמא:

- R_1 = טבלת הלוואות ; R_2 = טבלת סוגי הלוואות (חלקית).
- A = לוקח ההלוואה ; B = שם בנק ; C = סוג הלוואה ;
- החילוק ישמש אותנו כדי לתת מענה לשאלה איזה לקוח לקח את כל סוגי ההלוואות של בנק b.

דוגמא נוספת לחילוק (Division)

הסבר נוסף:

בהינתן שני יחסים $S[A_1, \dots, A_n, B_1, \dots, B_m]$ ו- $T[B_1, \dots, B_m]$,
(ז"א כאשר תכונות T מוכלות בתכונות S), נסמן ב- $S \div T$ את היחס
 $R[A_1, \dots, A_n]$ המכיל כל רשומה שעבורה יש רשומות ב- S לכל
הקומבינציות האפשריות עם רשומות מ- T .

S

צבע	חיה
לבן	סוס
שחור	סוס
שחור	חתול

T

צבע
לבן
שחור

$R = S \div T$

חיה
סוס

A

דוגמא נוספת לחילוק

נרצה לקבל את רשימת הספקים מ-A המספקים את כל החלקים המופיעים ברשימה B.

B

÷

=

<u>sno</u>	<u>pno</u>
S1	P1
S1	P2
S1	P3
S1	P4
S2	P1
S2	P2
S3	P2
S4	P2
S4	P4

<u>pno</u>
P2

<u>sno</u>
S1
S2
S3
S4

A

דוגמא נוספת לחילוק

נרצה לקבל את רשימת הספקים מ-A המספקים את כל החלקים המופיעים ברשימה B.

B

<u>sno</u>	<u>pno</u>	
S1	P1	
S1	P2	
S1	P3	
S1	P4	
S2	P1	
S2	P2	
S3	P2	
S4	P2	
S4	P4	

÷

<u>pno</u>
P2
P4

=

<u>sno</u>
S1
S4

A

דוגמא נוספת לחילוק

נרצה לקבל את רשימת הספקים מ-A המספקים את כל החלקים המופיעים ברשימה B.

<u>sno</u>	<u>pno</u>
S1	P1
S1	P2
S1	P3
S1	P4
S2	P1
S2	P2
S3	P2
S4	P2
S4	P4

B

<u>pno</u>
P1
P2
P4

÷

=

<u>sno</u>
S1