

## תרגיל 2 – DATABASES

כרמל גרוס, ליאור שפירא

שאלה א':

- 1)  $\pi_{\text{name}}(\sigma_{\text{character} = \text{"George"}}(\text{Actors} \bowtie \text{PlaysIn}))$
- 2)  $\pi_{\text{movieId}, \text{title}}(\sigma_{\text{dyear} = \text{year} \wedge (\text{genre} = \text{"Documentary"} \vee \text{genre} = \text{"Drama"})}(\text{Actors} \bowtie \text{PlaysIn} \bowtie \text{Movies}))$
- 3)  $(\pi_{\text{actorId}}((\sigma_{\text{name} = \text{"Charles Chaplin"}}(\text{Actors} \bowtie \text{PlaysIn} \bowtie \text{Movies}))) \bowtie \text{Actors} \bowtie \text{PlaysIn})) \wedge [\pi_{\text{actorId}}(\sigma_{\text{duration} > 90}(\text{Actors} \bowtie \text{PlaysIn} \bowtie \text{Movies}))]$
- 4)  $\pi_{\text{actorId}}(\text{PlaysIn}) - \pi_{\text{actorId}}(\sigma_{\text{rating} \leq 7}(\text{PlaysIn} \bowtie \text{Movies}))$
- 5)  $\pi_{\text{name}}(\sigma_{\text{actorId} \neq 100}((\text{PlaysIn} \div \pi_{\text{movieId}}(\sigma_{\text{actorId} = 100}(\text{PlaysIn}))) \bowtie \text{Actors}))$
- 6)  $\text{Over70} = \pi_{\text{title}, \text{movieId}, \text{ActorId}}(\sigma_{\text{year} - \text{byear} > 70}(\text{Actors} \bowtie \text{PlaysIn} \bowtie \text{Movies}))$   
 $\rho_{T1}(\text{title}, \text{movie}, \text{actor1}) \text{Over70}$   
 $\rho_{T2}(\text{title}, \text{movie}, \text{actor2}) \text{Over70}$   
 $\rho_{T3}(\text{title}, \text{movie}, \text{actor3}) \text{Over70}$   
 $\text{AtLeastTwo} = \pi_{\text{title}}(\sigma_{\text{actor1} \neq \text{actor2}}(T1 \bowtie T2))$   
 $\text{AtLeastThree} = \pi_{\text{title}}(\sigma_{\text{actor1} \neq \text{actor2} \wedge \text{actor1} \neq \text{actor3} \wedge \text{actor2} \neq \text{actor3}}(T1 \bowtie T2 \bowtie T3))$   
 $\text{AtLeastTwo} - \text{AtLeastThree}$

שאלה ב':

1) נוכיח הכלה דו כיוונית:

$\supseteq$

יהי  $(A1, D1)$  ביטוי המוכל בטבלה לאחר הסינון הימני. אזי בטבלה  $S$ , בשורה של  $D1$   $C$  היה שווה 8 (כי זה תנאי סינון סופי על היחס  $(A,D)$ ). אזי  $D1$  (יחד עם ערך אחדאו יותר של  $B$ ) עובר את הסינון:

$$\pi_{B,D}(\sigma_{C=8}(S))$$

כעת,  $A1$  ו- $D1$  היו בטבלאות  $R$  ו- $S$  בהתאמה מול לפחות ערך  $B$  משותף אחד. זאת, משום שהם עוברים בביטוי הימני סינון של natural join. על כן ב-natural join המופיע בביטוי השמאלי תיווצר לפחות שורה אחת הכוללת את  $A1$  ואת  $D1$ , ולכן הביטוי  $(A1, D1)$  יהיה מוכל בטבלה לאחר הסינון השמאלי כנדרש.

$\subseteq$

יהי  $(A2, D2)$  ביטוי המוכל בטבלה לאחר הסינון השמאלי. ראשית, לאחר ה-natural join של הביטוי הימני תהיה לפחות שורה אחת עם זוג ערכי  $A2$  ו- $D2$ , משום שהם עברו natural join בביטוי השמאלי (רק עם תנאים). שנית, נודא ש- $C$  שהיה בשורה של  $D2$  הוא 8, ובכך נסיים להראות שהביטוי מוכל בסינון הימני. זה נכון כי את ה-natural join בביטוי השמאלי ביצענו מראש רק על ערכי  $D$  שמול ערך  $C$  שווה ל-8, ולכן אם  $D2$  עבר את הסינון הזה, השורה שמכילה את  $A2$  ו- $D2$  תעבור את הסינון הימני, כנדרש.

2) נוכיח כי תוצרי הביטוי השמאלי מוכלים בתוצרי הביטוי הימני, אך לא להפך:

$\subseteq$

יהי  $(A1, D1)$  ביטוי המוכל בטבלה לאחר הסינון השמאלי. אזי  $A1$  הוא ערך בעמודה  $A$  של  $R$  ו- $D1$  הוא ערך בעמודה  $D$  של  $S$ , וזה למעשה תנאי מספיק לכך ש- $(A1, D1)$  מוכל בטבלה לאחר הסינון הימני, משום שהמשמעות של natural join של שתי עמודות שונות משתי טבלאות הוא למעשה מכפלה קרטזית רגילה, כלומר כל צירוף של ערכי  $A$  וערכי  $D$  מהטבלאות המתאימות.

הכיוון השני לא נכון - נביא דוגמא נגדית :

$R$ :

A	B
1	2

$S$ :

B	C	D
3	4	4

תוצאת הפעלת הביטוי השמאלי על טבלאות אלה היא  $\emptyset$ , כי אין אף שורה ב- $R$  עם ערך  $B$  ששווה לערך  $B$  של שורה ב- $S$ . תוצאת הפעלת הביטוי הימני על הטבלאות הנ"ל תהיה :

A	D
1	4

כלומר הביטויים לא שווים.