



**הטכניון - מכון טכנולוגי לישראל**  
**הפקולטה למדעי המחשב**

אביב תשס"ו  
20 בספטמבר 2006

דר' אלדר פישר  
סעאב מנסור  
לינה זריבץ'

**מערכות מסדי נתונים – 236363**

**מועד א' - פתרון**

הזמן: 3 שעות  
במבחן זה 11 עמודים

שאלה	נקודות
שאלה 1 – ERD	17
שאלה 2 – שפות שאילתה	23
שאלה 3 – תלויות פונקציונאליות RA+	12
שאלה 4 – Design	19
שאלה 5 – XML	21
שאלה 6 – Datalog	8
סה"כ	100

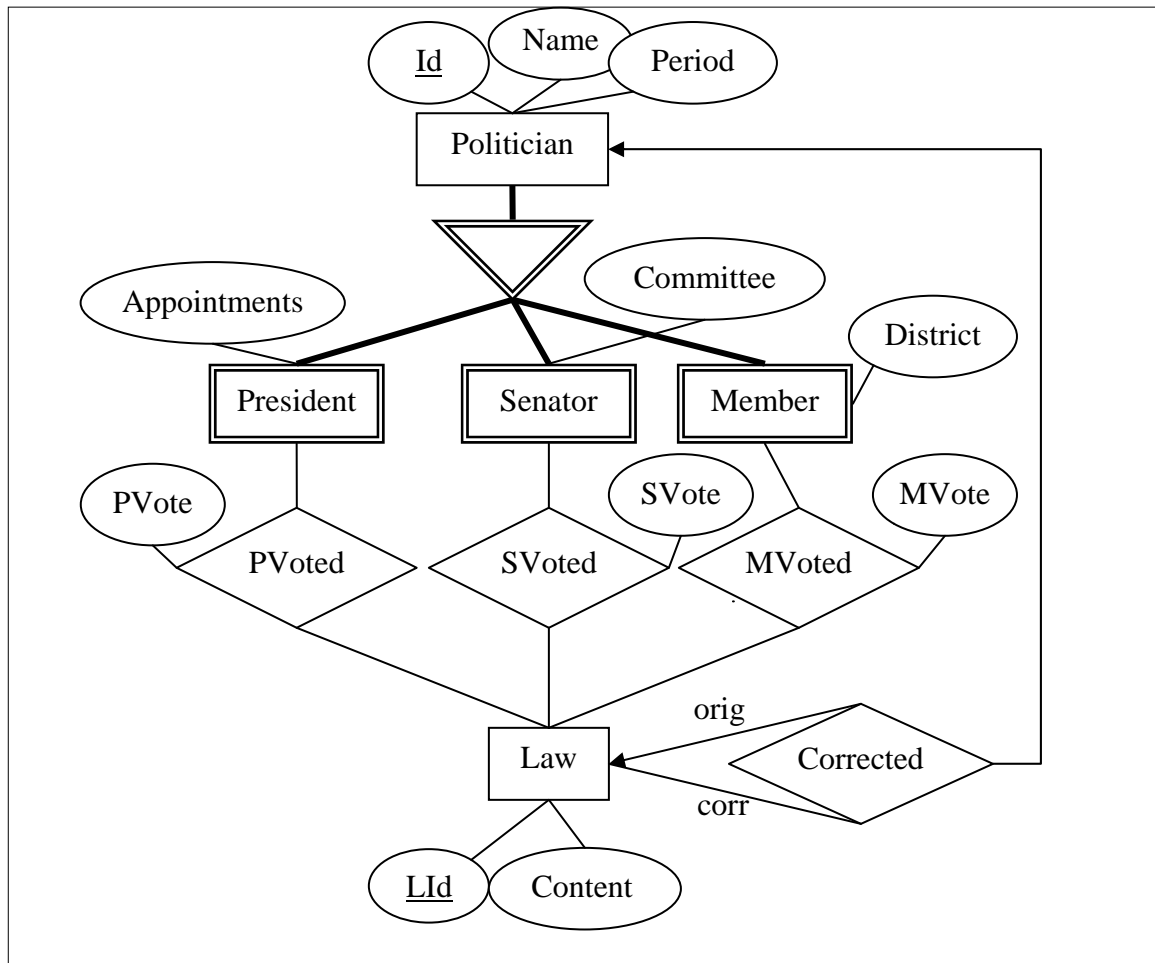
**הנחיות:**

1. יש לענות על כל השאלות **בטופס הבחינה**.
2. חומר עזר מותר: רק דברים שעשויים מנייר.
3. **אין להחזיק מכשיר אלקטרוני כלשהו לרבות מחשב כיס.**
4. קראו היטב את ההוראות שבתחילת כל שאלה ואת ההסברים לסכמות.
5. מומלץ שתכננו היטב את זמנכם, **לא תינתנה הארכות.**
6. ערעורים יש להגיש תוך שבועיים ממועד פרסום התוצאות.
7. לא יתקבלו ערעורים בנוסח "בדיקה מחמירה מדי".

**בהצלחה**

## שאלה 1 – ERD (17 נק')

נתונה סכימת ה-ERD הבאה שמתארת מסד נתונים של מערכת הממשל האמריקאית:



### הסברים לסכמה:

#### ישויות:

**Politician** – מייצג פוליטיקאי בממשל. לכל פוליטיקאי קיים מזהה ייחודי (Id), שם (Name) ותקופת כהונה (Period). קיימים מספר סוגים של פוליטיקאים שונים:

- **President** – נשיא – עבור הנשיא נשמר מספר הפגישות שהוא קיים (Appointments)
- **Senator** – סנטור – עבור הסנטור נשמרת הוועדה בה הוא השתתף (Committee)
- **Member** – חבר – עבור החברים של בית הנבחרים נשמר המחוז אותו הם מייצגים (District)

**Law** – מייצג חוקים המוצעים בממשל. לכל חוק נשמר מזהה ייחודי (LId) וטקסט תוכן (Content).

#### יחסים:

**PVoted, SVoted, MVoted** – יחס הצבעה שמכיל לכל פוליטיקאי את ההצבעה שלו על החוק. הערכים של PVote הם veto (ווטו) או sign (חותם). הערכים של SVote ו-MVote הם for (בעד), against (נגד), או abstain (נמנע).

**Corrected** – יחס תיקון חוק שמקשר בין תיקון החוק (corr) לחוק המקורי (orig) והפוליטיקאי שהציע את התיקון.

א. (9 נק') מלאו את הטבלאות הבאות בהתייחס לסכמה הנתונה:

Senator	
שדות	Id, Committee
מפתחות	Id
SVoted	
שדות	SVote, LId, Id
מפתחות	LId, Id
Corrected	
שדות	LIdOrig, LIdCorr, Id
מפתחות	(LIdCorr, LIdOrig) או (LIdCorr, Id)

ב. (8 נק') נתונים 4 תכנים של היחס Corrected:

טבלה 1:		
Orig	Corr	Politician
I1	I2	p1
I1	I2	p2
I1	I3	p1

טבלה 2:		
Orig	Corr	Politician
I1	I2	p1
I2	I2	p1

טבלה 3:		
Orig	Corr	Politician
I1	I2	p1
I1	I3	p1

טבלה 4:		
Orig	Corr	Politician
I1	I2	p1
I3	I2	p2

סמנו את כל הטבלאות אשר מתאימות לתכנון ה-ERD. אם הטבלה אינה מתאימה, רשמו אילו תלויות פונקציונליות ה-ERD דורש והטבלה מפרה.

טבלה	מתאים? (כן/לא)	תלויות מופרות
1	לא	מפר את התלות $\{LIdOrig, LIdCorr\} \rightarrow Id$
2	לא	מפר את התלות $\{LIdCorr, Id\} \rightarrow LIdOrig$
3	כן	לא מפר אף תלות
4	כן	לא מפר אף תלות

## שאלה 2 – שפות שאילתה (23 נק')

א. (5 נק') נניח שנתונה הטבלה NumVotesS(Id, Num) (שימו לב כי Id הינו מפתח ב-NumVotesS) שמכילה לכל סנטור בעל מזהה Id את מספר הפעמים שהוא הצביע. כתבו שאילתה ב-SQL שמחשבת את חציון רשימת מספרי ההצבעות. חציון של רשימת מספרים הוא איבר ברשימה כך שלפחות חצי מאיברי הרשימה קטנים או שווים לו, ופחות מחצי האיברים גדולים ממנו.

הגדרה מתוקנת שניתנה בזמן המבחן: החציון הוא האיבר המינימאלי מבין אלה שגדולים לפחות מ-1/2 האיברים האחרים.

```
CREATE VIEW MedNom AS (  
SELECT Id, Num FROM NumVotesS N  
WHERE (SELECT COUNT(*) FROM NumVotesS  
WHERE N.Num > Num) >= (SELECT COUNT(*)/2 FROM NumVotesS)  
)
```

פתרון:

```
SELECT Id, Num FROM MedNom  
WHERE Num = (SELECT MIN(Num) FROM MedNom)
```

ב. (6 נק') בנוסף לטבלה NumVotesS(Id,Num) מסעיף א', נתונה כעת הטבלה NumVotesM(Id,Num) אשר מוגדרת בצורה דומה עבור חברי בית הנבחרים. כמו כן נתונות שתי השאילתות הבאות ב-SQL:

Q1:

```
SELECT Id FROM NumVotesS  
WHERE Num > ALL (SELECT Num FROM NumVotesM)
```

Q2:

```
SELECT Id FROM NumVotesS  
WHERE Num > ANY (SELECT Num FROM NumVotesM)
```

האם:

1. Q1 תמיד מוכלת ב-Q2
2. Q1 מוכלת ב-Q2 רק אם מובטח ש-NumVotesM לא ריקה
3. Q1 מוכלת ב-Q2 רק אם מובטח ש-NumVotesS לא ריקה
4. אף אחד מהנ"ל אינו נכון

נמקו בקצרה.

התשובה הנכונה היא 2.

---

אם NumVotesM ריקה Q1 תחזיר את NumVotesS ואילו Q2 תחזיר תוצאה ריקה,

---

אחרת Q2 יכיל את Q1.

---

---

הערה: בסעיפים הבאים הניחו תרגום סטנדרטי של סכמת ה-ERD משאלה 1 לסכמות רלציוניות: Politician, President, Senator, Member, PVoted, SVoted, MVoted, Law, Corrected.

ג. (6 נק') כתבו שאילתא ב-DRC אשר מחזירה שלשות (Id, LId, Vote) כך שהסנטור בעל המזהה Id הצביע Vote על החוק LId. לכל חוק שעלה להצבעה יהיה רישום עבור כל הסנטורים. לכן, אם קיים סנטור אחד לפחות שהצביע על LId אז לכל הסנטורים שלא הצביעו כלל נחזיר "abstain" (ואם לא קיים סנטור כזה אז לא יהיה רישום הצבעה עבור אותו חוק). הניחו כי הסדר של התכונות ברלציות הינו אלפביתי.

פתרון 1:

$$\{ \langle Id, LId, Vote \rangle : \exists C \text{ Senator}(Id, C) \wedge \exists Id1, V1 \text{ SVoted}(Id1, LId, V1) \wedge ( \begin{aligned} & (\exists V1 \text{ SVoted}(Id, LId, V1)) \rightarrow \text{SVoted}(Id, LId, Vote) \wedge \\ & (\neg \exists V1 \text{ SVoted}(Id, LId, V1)) \rightarrow \text{Vote} = \text{"abstain"} \end{aligned} ) \}$$

פתרון 2:

$$\{ \langle Id, LId, Vote \rangle : \exists C \text{ Senator}(Id, C) \wedge \exists Id1, V1 \text{ SVoted}(Id1, LId, V1) \wedge ( \begin{aligned} & ( \exists V1 (\text{SVoted}(Id, LId, V1) \wedge \text{Vote} = V1) ) \vee \\ & (\neg \exists V1 \text{ SVoted}(Id, LId, V1) \wedge \text{Vote} = \text{"abstain"}) \end{aligned} ) \}$$

ד. (6 נק') כתבו שאילתא ב-RA אשר בודקת האם המזהה של החוק המתוקן קובע חד משמעית את המזהה של החוק המקורי ואת מזהה הפוליטיקאי שהציע את התיקון. אם כן אז יש להחזיר רלציה לא ריקה ואחרת יש להחזיר רלציה ריקה. ניתן להשתמש ברלציה חסרת התכונות TRUE אשר מכילה שורה אחת ריקה או ברלציה חסרת התכונות FALSE אשר אינה מכילה אף שורה.

נסמן

$$A = \pi_{\lambda} ( \sigma_{LIdC1=LIdC2 \wedge ( (LIdO1 \neq LIdO2) \vee (Id1 \neq Id2) ) } (Corrected \times Corrected) )$$

נסמן ב- TRUE את הרלציה עם השורה הריקה, התשובה הסופית היא:

TRUE<sub>A</sub>

### שאלה 3 – תלויות פונקציונאליות + RA (12 נק')

נתונות הרלציות  $r$  מעל  $R(A,B,C)$  ו- $s$  מעל  $S(B,C,D)$ . בסעיפים הבאים רשומים ביטויים ב-RA. כמו כן נתונות התלויות הפונקציונאליות המתקיימות ב- $r$  וב- $s$ . רשמו כיסוי מינימאלי של כל התלויות הפונקציונאליות אשר בהכרח מתקיימות ברלציה המחושבת ע"י הביטויים הנ"ל. יש ללוות את תשובתכם בנימוק קצר (בן 25 מילים לכל היותר). שימו לב שהסעיפים בלתי תלויים!

א. (4 נק')  $r$  מקיימת את התלויות הפונקציונאליות  $\{B \rightarrow C, C \rightarrow B\}$  ו- $s$  מקיימת את התלות הפונקציונאלית  $B \rightarrow C$ :

$$\pi_{B,C} r \cup \pi_{B,C} s$$

אף תלות.

לדוגמא:  $r = \{(0,0,0), (0,1,1)\}$  ו- $s = \{(0,1,0)\}$ . ניתן לראות כי הרלציה

$$\pi_{B,C} r \cup \pi_{B,C} s = \{(1,1), (1,0), (0,0)\}$$

ב. (4 נק')  $r$  מקיימת את התלות הפונקציונאלית  $A \rightarrow C$  ו- $s$  מקיימת את התלות הפונקציונאלית  $B \rightarrow C$ :

$$r \bowtie (r \div \pi_{B,C} s)$$

$A \rightarrow C$

הצירוף הוא תת קבוצה של  $r$  ולכן מקיים את התלויות של  $r$

התלות של  $s$  לא תתקיים אם לאותו ערך של  $A$  ש"נבחר" בצירוף יש ערכים

ל- $B$  ו- $C$  שאינם מקיימים התלות.

ג. (4 נק')  $r$  מקיימת את התלות הפונקציונאלית  $A \rightarrow C$  ו- $s$  מקיימת את התלות הפונקציונאלית  $B \rightarrow D$ :

$$r \bowtie \sigma_{B=C} s$$

$A \rightarrow C, B \rightarrow D, B \rightarrow C, C \rightarrow B$

הצירוף שומר על התלויות של  $r$  והבחירה מ- $s$ . הבחירה הנ"ל שומרת על תלויות  $s$

(כתת קבוצה) ומוסיפה את התלויות בין  $B$  ו- $C$  עקב תנאי הבחירה

## שאלה 4 – Design (19 נק')

א. (9 נק') תהא  $R(A, B, C, D, E, G)$  סכמה רלציונית ו-F קבוצת התלויות  $\{AC \rightarrow D, BD \rightarrow G, G \rightarrow ABCDE\}$ .

1. (3 נק') מהם כל המפתחות הקבילים של R?

---

G, BD, ACB

---

2. (3 נק') האם הסכמה נמצאת ב-BCNF? אם כן, נמקו. אחרת, נמקו וכתבו פירוק ל-BCNF המשמר מידע ותלויות.

הסכמה אינה נמצאת ב-BCNF כיוון שהתלות  $AC \rightarrow D$  מפרה את

תנאי ה-BCNF. להלן הפירוק הנדרש ל-BCNF:

---

ACD, BDG, GEAC

---

3. (3 נק') האם הסכמה נמצאת ב-3NF? אם כן, נמקו. אחרת, נמקו וכתבו פירוק ל-3NF המשמר מידע ותלויות.

הסכמה נמצאת ב-3NF, כי אין שום תלות המפרה את תנאי ה-3NF.

---

---

---

ב. (10 נק') תהא  $R, F$  סכמה אשר אינה נמצאת ב-3NF.  
ידוע כי לכל תלות  $X \rightarrow Y \in F$  מתקיים  $|X| = |Y| = 1$ .  
יהי  $\rho$  פירוק 3NF של  $R, F$  המתקבל אחרי הרצת האלגוריתם לפירוק 3NF שהוצג בתרגול.  
הוכח או הפרך:  
 $\rho$  הינו גם פירוק ל-BCNF.  
להזכירכם, להלן האלגוריתם שהוצג בתרגול:

בהינתן כסוי מינימאלי  $F_C$  של תלויות פונקציונאליות:

1. אם קיימת ב- $F_C$  תלות פונקציונאלית שכוללת את כל התכונות ב- $R$ , התשובה היא  $\{R\}$  - עצור.
2. לכל קבוצת תלויות פונקציונאליות  $X \rightarrow A_1, X \rightarrow A_2, \dots, X \rightarrow A_n$ , צור סכמה  $X \cup \{A_1 A_2 \dots A_n\}$ .
3. אם אין אף סכמה המכילה מפתח קביל של  $R$ , הוסף סכמה שהיא מפתח קביל כלשהו של  $R$ .

הוכחה: יהא  $F_C$  הכיסוי המינימאלי של  $F$ , המועבר כקלט לאלגוריתם. ברור שכל תלות  $f = X \rightarrow Y \in F_C$  מקיימת כי  $|X| = |Y| = 1$ . יתרה מכך, ניתן להוכיח כי כל תלות השייכת לכיסוי המינימאלי של  $F_C$   $\pi_S F_C$  היא גם מהצורה הנ"ל עבור  $S \subseteq R$ .

נתבונן בתת-הסכמה  $R_i$  השייכת לפירוק המוחזר מהאלגוריתם. נסמן ב- $F_i$  את הכיסוי המינימאלי של  $\pi_{R_i} F_C$ . אם  $R_i$  נוצרה בצעד ה-3, אזי ברור כי היא נמצאת ב-BCNF.

אחרת, תת-הסכמה  $R_i$  נוצרה בצעד ה-2, כלומר קיימות התלויות  $L \rightarrow M_1, \dots, L \rightarrow M_n$  בתוך  $F_C$  כך ש- $\{L, M_1, \dots, M_n\} = R_i$  (זכרו כי  $L, M_1, \dots, M_n$  מסמנות אטריביוטים בודדים).

נניח בשלילה כי  $R_i$  אינה ב-BCNF. לכן בהכרח קיימת תלות בתוך  $F_i$  אשר מפירה את תנאי ה-BCNF. (ראו שאלה 3 בתרגיל בית האחרון).

התלות הנ"ל היא בהכרח מהצורה  $M_j \rightarrow M_k$  בה  $M_j$  אינו על מפתח של  $R_i$ .

אבל אז נקבל סתירה למינימליות של  $F_C$  שכן בהינתן  $M_j \rightarrow M_k \in F_C$  התלות  $L \rightarrow M_k$  מיותרת.



## אפשרות שניה:

הוכחה: יהא  $F_c$  הכיסוי המינימאלי של  $F$ , המועבר כקלט לאלגוריתם. ברור שכל תלות

$$f=X \rightarrow Y \in F_c \text{ מקיימת כי } |X| = |Y| = 1.$$

נתבונן בתת-הסכמה  $R_i$  השייכת לפירוק המוחזר מהאלגוריתם. אם  $R_i$  נוצרה בצעד ה-3,

אזי ברור כי היא נמצאת ב-BCNF. אחרת, תת-הסכמה  $R_i$  נוצרה בצעד ה-2, כלומר

$$\{L, M_1, \dots, M_n\} = R_i \text{ כך ש-} F_c \text{ בתוך } L \rightarrow M_1, \dots, L \rightarrow M_n$$

(זכרו כי  $L, M_1, \dots, M_n$  מסמנות אטריביוטים בודדים).

נניח בשלילה כי  $R_i$  איננה ב-BCNF. נתבונן בתלויות שיכולות להיות בתוך  $F_i$ ,  $F_i = \pi_{R_i} F$ .

תלות מהצורה  $X \rightarrow L$  אינה מהווה סתירה ל-BCNF שכן  $X$  הוא על מפתח.

תלות מהצורה  $X \rightarrow M_k$  כאשר  $X \subseteq \{M_1, \dots, M_n\}$  ו- $1 \leq k \leq n$  יכולה להוות סתירה ל-BCNF.

אך קיומה של תלות כזאת סותרת את המינימליות של  $F_c$  שכן

$$\text{בהינתן } X \rightarrow M_k \in F^+ \text{ התלות } L \rightarrow M_k \text{ מיותרת.}$$

## שאלה 5 – XML (21 נק')

שאלה זאת מתייחסת לקבצי xml המתארים חוקה (constitution) ותיקוני החוקה (amendments). להלן מסמך ה-DTD המתאר את מבנה המסמכים הנ"ל:

```
<!ELEMENT constitution (section+, amendment*)>
<!ELEMENT section (content)>
<!ATTLIST section
    title CDATA #REQUIRED>
<!ELEMENT content (#PCDATA)>
<!ELEMENT amendment (content, amendment*)>
<!ATTLIST amendment
    title CDATA #REQUIRED
    year CDATA #REQUIRED>
```

מותר להניח כי התוכן של כל אלמנטי ה-content הוא ייחודי ואינו ריק.

א. (6 נק') כתבו שאילתא ב-XPath 1.0 אשר מחזירה תוצאה ריקה אם תיקוני החוקה מסודרים לפי סדר כרונולוגי לא יורד, ומחזירה תוצאה לא ריקה אחרת. נגיד שתיקוני החוקה מסודרים לפי סדר כרונולוגי לא יורד אם מתקיימים כל התנאים הבאים:

- לא יתכן תיקון ב' (לאו דווקא ישיר) לתיקון א' כך ששנת החקיקה של א' הינה מאוחרת מזאת של תיקון ב'.
- לא יתכן תיקון א' שתוקן ישירות ע"י תיקון ב' וישירות ע"י תיקון ג', כאשר תיקון ג' הינו תיקון מאוחר יותר מתיקון ב' לפי סדר המסמך, ושנת החקיקה של ב' הינה מאוחרת מזאת של תיקון ג'.

```
//amendment [ @year > amendment/@year or amendment[ @year >
following-sibling::amendment/@year ] ]
```

ב. (8 נק') נתונה השאילתא הבאה:

```
//amendment[ amendment != amendment ]
```

1. (4 נק') כתבו בקצרה מה המשמעות של השאילתא הנתונה.

מחזיר את כל התיקונים שלהם עצמם היו לפחות שני תיקונים לא כולל תיקון לתיקון.

2. (4 נק') כתבו שאילתא ב-XPath 1.0 אשר אינה משתמשת בסימני השוואה כל שהם והשקולה לשאילתא הנ"ל.

```
//amendment[amendment[2] ]
```

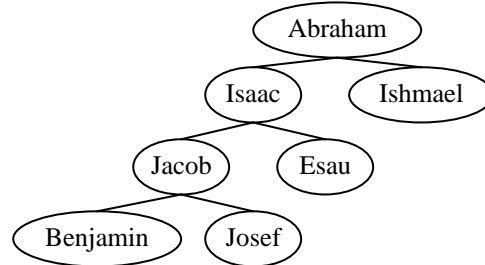
ג. (7 נק') כתבו שאילתא ב-XQuery הפועלת על הקובץ cons.xml המציית לקובץ ה-DTD הנתון עם צומת מסמך constitution (שהוא הבן היחידי של השורש) והמחזירה את כל התיקונים (amendments) אשר הוצאו בשנים בהן בוצע מספר מקסימאלי של תיקוני חוקה. יש להשתמש באופרטורים והפונקציות אשר נלמדו בהרצאה ובתרגולים בלבד! בפרט אין להשתמש בפונקצית max().

```
xquery version "1.0";
<result>
{
  for $a in doc("cons.xml")//amendment
  let $num:=count(doc("cons.xml")//amendment[@year = $a/@year] )
  where not (
    for $a1 in doc("cons.xml")//amendment
    where count( doc("cons.xml") //amendment[@year = $a1/@year]) >$num
    return $a1
  )
  return $a
}
</result>
```

## שאלה – Datalog<sup>-</sup> (8 נק')

נתון הפרדיקט  $\text{father}(F,S)$  שמייצג כי  $F$  הוא אב של  $S$ . האב הקדמון המשותף המאוחר ביותר (Latest Common Ancestor –  $\text{lca}$ ) מוצג בדוגמה הבאה (מימין שושלת היוחסין, משמאל הייצוג ב-DATALOG) שאחריה מופיעה ההגדרה המדויקת:

$\text{father}('Abraham', 'Isaac').$   
 $\text{father}('Abraham', 'Ishmael').$   
 $\text{father}('Isaac', 'Jacob').$   
 $\text{father}('Isaac', 'Esau').$   
 $\text{father}('Jacob', 'Benjamin').$   
 $\text{father}('Jacob', 'Josef').$



חלק מהיחסים שיתקיימו:

$\text{lca}('Benjamin', 'Esau', 'Isaac').$   
 $\text{lca}('Benjamin', 'Jacob', 'Jacob').$   
 $\text{lca}('Josef', 'Ishmael', 'Abraham').$

הגדרה: האב הקדמון המשותף המאוחר ביותר של  $p_1, p_2$  הוא  $p_3$  אם  $p_3$  הוא אב קדמון של  $p_1, p_2$  וגם לא קיים  $p_4$  כך ש- $p_4$  הוא צאצא של  $p_3$  (ושונה ממנו) וגם  $p_4$  הוא אב קדמון של  $p_1, p_2$ .

שימו לב כי האבות הקדמונים של צומת כוללים את הצומת עצמו.

כתוב תוכנית ב- Datalog<sup>-</sup> אשר מגדירה את הפרדיקט  $\text{lca}(P_1, P_2, P_3)$  כך ש-  $P_3$  הוא האב הקדמון המשותף המאוחר ביותר של  $P_1, P_2$ .

```

ancestor(A,A) ← father(A,_).
ancestor(A,A) ← father(_,A).
ancestor(A, D) ← father(A,D).
ancestor(A, D) ← father(A,M), ancestor(M, D).

ca(P1,P2,P3) ← ancestor(P1,P3), ancestor(P2,P3).
nlca(P1,P2,P3) ← ca(P1,P2,P3), father(P3,P4), ca(P1,P2,P4).
lca(P1,P2,P3) ← ca(P1,P2,P3), ¬nlca(P1,P2,P3)
  
```