



הטכניון - מכון טכנולוגי לישראל
הפקולטה למדעי המחשב

דר' אמיר שפילקה
סעאב מנסור
ארינה לווה

אביב תשס"ז
24, אוקטובר 2007

מערכות מסדי נתונים – 236363

מועד ב'
הזמן: 3 שעות
במבחן זה 17 עמודים

| שאלה | נקודות |
|----------------------|--------|
| שאלה 1 – ERD | 12 |
| שאלה 2 – שפות שאילתא | 24 |
| שאלה 3 – תלויות | 33 |
| שאלה 4 – XML | 21 |
| שאלה 5 – DATALOG | 10 |
| סה"כ | 100 |

הנחיות:

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

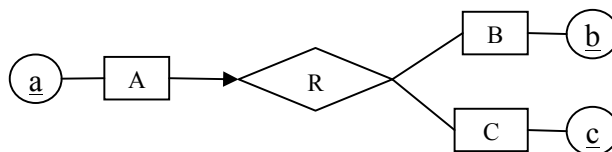
שאלה 1 – ERD (12 נק')

בהינתן הטבלאות הבאות (כל טבלה מופיעה בשמה העברי ובתרגום לאנגלית):

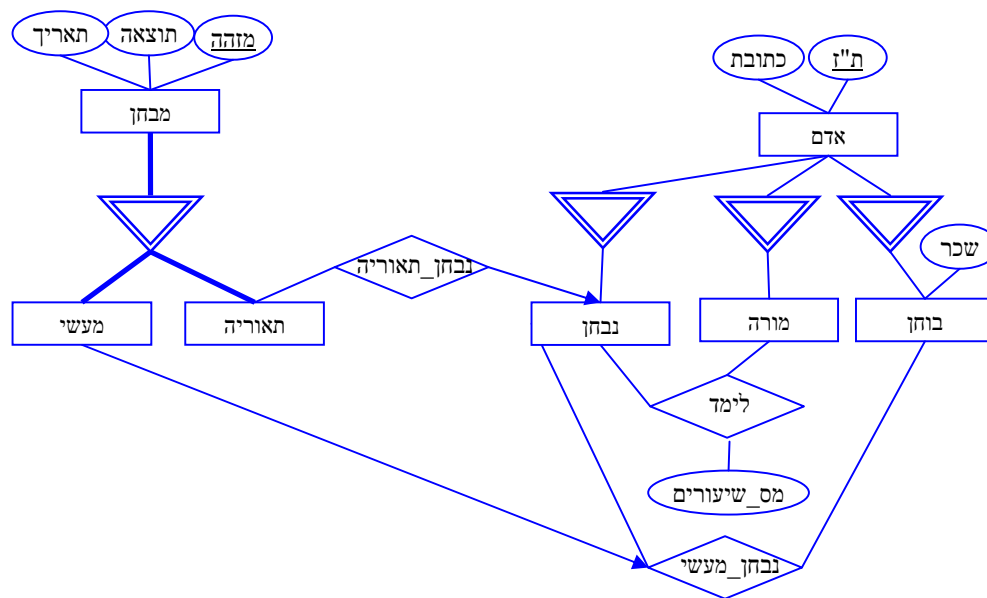
| | |
|---|--|
| Person(<u>pid</u> , address) | אדם(ת"ז, כתובת) |
| Examinee(<u>eid</u>) | נבחן(ת"ז) |
| Tester(<u>tstid</u> , salary) | בוחן(ת"ז, שכר) |
| Teacher(<u>tchid</u>) | מורה(ת"ז) |
| Test(<u>tid</u> , date, result) | מבחן(מזהה, תאריך, תוצאה) |
| Theoretical(<u>tid</u>) | תאוריה(מזהה) |
| Practical(<u>tid</u>) | מעשי(מזהה) |
| Did_theo(eid, <u>tid</u>) | נבחן_בתאוריה(ת"ז(נבחן), מזהה) |
| Did_prac(eid, <u>tstid</u> , <u>tid</u>) | נבחן_במעשי(ת"ז(נבחן), ת"ז(בוחן), מזהה) |
| Taught(<u>tchid</u> , <u>eid</u> , num) | לימד(ת"ז(מורה), ת"ז(נבחן), מס_שיעורים) |

כל מבחן הוא או תיאורטי או מעשי, אך לא שניהם.
אדם יכול לבצע מספר תפקידים. למשל, אדם יכול להיות גם בוחן וגם מורה.

ציירו ERD מתאים ביותר שיתאר את היחסים ואת הקשרים בין הטבלאות הנ"ל (יש להשתמש בסעיף זה בשם העברי של הטבלה). לשם כך היעזרו בסימון החדש הבא:



הסימון (חץ נכנס ליחס) אומר ש-a מפתח של R.



שאלה 2 – שפות שאילתא (24 נק')

עבור סעיפים א', ב', השתמשו בתרגום לאנגלית של הטבלאות משאלה.

- א. (8 נק') כתבו שאילתא ב-SQL שבודקת אם התנאים הבאים מתקיימים:
- האיחוד של המזהים של תאוריה ומעשי שווה לקבוצת המזהים של מבחן.
 - החיתוך בין המזהים של תאוריה ומעשי ריק.

אם התנאים מתקיימים, יש להחזיר תוצאה לא ריקה, אחרת החזירו תוצאה ריקה. הניחו ש-Test לא ריקה. מותר להשתמש במבט עזר אחד לכל היותר.

```
CREATE VIEW un as
SELECT tid FROM Theoretical Union SELECT tid FROM Practical

SELECT *
FROM Test
WHERE
NOT EXISTS (SELECT tid FROM Test EXCEPT SELECT * FROM un)
AND NOT EXISTS (SELECT * FROM un EXCEPT SELECT tid FROM Test)
AND NOT EXISTS (SELECT * FROM Practical INTERSECT SELECT * FROM
Theoretical)
```

- ב. (8 נק') כתבו שאילתא ב-DRC שבודקת אם התנאים הבאים מתקיימים:
- האיחוד של המזהים של תאוריה ומעשי שווה לקבוצת המזהים של מבחן.
 - החיתוך בין המזהים של תאוריה ומעשי ריק.
- אם התנאים מתקיימים, יש להחזיר תוצאה לא ריקה, אחרת החזירו תוצאה ריקה.
אין להשתמש בחישובי עזר.

$$\{ \langle \rangle \mid \forall t ((\exists d, r \text{Test}(t, d, r)) \rightarrow (\text{Theoretical}(t) \vee \text{Practical}(t))) \wedge$$

$$\forall t ((\text{Theoretical}(t) \vee \text{Practical}(t)) \rightarrow (\exists d, r \text{Test}(t, d, r))) \wedge$$

$$\neg \exists t (\text{Theoretical}(t) \wedge \text{Practical}(t)) \}$$

הסעיף הבא בלתי תלוי בסעיפים הקודמים.

- ג. (8 נק') נתונה הרלציה $R(\text{doc}, \text{word})$ המציינת שמילה word מופיעה במסמך doc .
הסבירו במלים פשוטות מה מחושב בשאילתא הבאה (אין להתייחס לרלציות העזר שחושבו במהלך השאילתא).

$$R_1(d, w) = \rho_{\text{doc} \rightarrow d, \text{word} \rightarrow w}((\pi_{\text{doc}}(R) \times \pi_{\text{word}}(R)) \setminus R)$$

$$R_2(d_1, d_2) = \rho_{\text{doc} \rightarrow d_1, d \rightarrow d_2}(\pi_{\text{doc}, d}(R \bowtie_{w=\text{word}} R_1))$$

$$R_3(d_1, d_2) = \pi_{d_1, d_2}(\rho_{\text{doc} \rightarrow d_1} R \times \rho_{\text{doc} \rightarrow d_2} R)$$

$$\text{Answer}(d_1, d_2) = R_3 \setminus R_2$$

התשובה היא כל זוגות המסמכים $(d1, d2)$ כך שכל מלה המופיעה ב- $d1$ מופיעה גם ב- $d2$

שאלה 3 – תלויות פונקציונליות (33 נק')

בשאלה הזו נשתמש בסימונים הבאים: תכונות בודדות נסמן באותיות לועזיות מתחילת הא"ב: A, B, \dots . קבוצות של תכונות נסמן באותיות מסוף הא"ב: X, Y, \dots .

א. (5 נק') כזכור, בהינתן קבוצת תלויות פונקציונליות F , קבוצת תכונות X היא סגורה אם $X^+ = X$. נתונות שתי קבוצות תכונות X ו- Y הסגורות ביחס ל- F . האם גם $X \cap Y$ סגורה? הוכיחו או תנו דוגמא נגדית.

כן. נניח כי $Z \rightarrow W$ נובע מ- F עבור קבוצת תכונות Z המוכלת ב- $X \cap Y$ וקבוצת תכונות W . אז מסגירות X נקבל כי W מוכלת ב- X ומסגירות Y נקבל כי W מוכלת ב- Y (כי Z מוכלת הן ב- X והן ב- Y) ולכן W מוכלת ב- $X \cap Y$, כנדרש.

ב. (28 נק') תהי $R(A_1, \dots, A_n)$ סכמה עם קבוצת תלויות F . ידוע כי כל התלויות ב- F הן מהצורה $A_k \rightarrow A_m$.

i. (5 נק') הראו כי אם $X \rightarrow Y$ נובע מ- F אז לכל תכונה B ב- Y קיימת תכונה A ב- X כך ש- $A \rightarrow B$ נובע מ- F . רמז: אינדוקציה על אורך ההוכחה.

כל הוכחה היא אוסף של תלויות שחלקן ב- F וחלקן נובע מהפעלת אקסיומות ארמסטרונג על

תלויות קודמות. לפיכך הוכחה באורך 1 היא תלות ב- F או אחת מאקסיומות ארמסטרונג

(רפלקסיביות). לפי הגדרת התלויות שלנו ומהגדרת אקסיומות ארמסטרונג הטענה

נובעת למקרה זה.

נניח נכונות להוכחות מאורך $K \geq$ ונביט בהוכחה מאורך $K+1$. השורה האחרונה בהוכחה היא

תלות מ- F (ואז סיימנו) או הפעלה של אחת מהאקסיומות. במקרה של רפלקסיביות הטענה

ברורה. במקרה של הכללה, נובע שבהוכחה היתה שורה $X' \rightarrow Y$ וע"י הכללה קיבלנו את $X \rightarrow Y$.

בפרט יש Z כך ש- $X = X' \cup Z$ ו- $Y = Y' \cup Z$. מנכונות הטענה לגבי $X' \rightarrow Y'$ (שבהכרח יש לה

הוכחה באורך $K \geq$) הטענה נובעת. במקרה שהשורה האחרונה בהוכחה נבעה מטרנזיטיביות

נובע שיש Z כך שבמהלך ההוכחה הראנו $X \rightarrow Z$ ו- $Z \rightarrow Y$ (ולכל אחת מהן יש הוכחה באורך

$K \geq$). יהי B ב- Y . לפי האינדוקציה יש C ב- Z כך ש- $C \rightarrow B$ ויש A ב- X כך ש- $A \rightarrow C$.

מטרנזיטיביות נובע כי $A \rightarrow B$ כנדרש.

[illegible]

ii. (5 נק') נגדיר את גרף התלויות המכון של F להיות גרף בן n קדקודים עם צלע מכוונת $k \rightarrow m$ אם ורק אם ב- F יש תלות מהצורה $A_k \rightarrow A_m$. לשם הנוחות נזהה תכונה עם הקודקוד המתאים לה. לכל תכונה A_k נגדיר את הרכיב היוצא ממנה להיות כל התכונות A_m כך שיש מסלול מכוון מ- k ל- m . הראו כי $F \models A \rightarrow B$ אם ורק אם B ברכיב שיוצא מ- A . רמז לאחד הכיוונים: תנו דוגמא של רלציה המראה שאם B אינה ברכיב היוצא מ- A אז B לא נובעת מ- A . (שימו לב, יש עוד כיוון לטענה).

נשים לב כי מטרנזיטיביות נובע שאם B ברכיב של A אז $A \rightarrow B$. להוכחת הכיוון השני נחלק את

התכונות לשתי קבוצות, אלו שברכיב של A ואלו שלא. נשים לב שמהרכיב של A לא יוצאת אף

צלע (אחרת הרכיב היה גדל). נביט בהשמה הנותנת ערך 0 לכל התכונות ובהשמה אחרת

הנותנת ערך 0 לתכונות ברכיב של A וערך 1 לתכונות מחוץ לרכיב. נשים לב ששתי ההשמות

קונסיסטנטיות עם F (אכן, כל תלות היא או בתוך אחד הרכיבים או לתוך הרכיב של A). בפרט

נובע שאין B מחוץ לרכיב של A עבורה $A \rightarrow B$, כנדרש.

הערה: שימו לב כי ההוכחה משתמשת בעובדה (שהוכחנו בכיתה) כי $F \models X \rightarrow Y$ אם ורק אם

$F \models X \rightarrow Y$.

iii. (5 נק') הוכיחו כי קבוצת תכונות $Z \subseteq R$ היא מפתח אם ורק אם לכל תכונה $B \in R$ קיים $A \in Z$ כך ש- B נמצאת ברכיב היוצא מ- A .

מפתח הוא קבוצת תכונות שכל תכונה אחרת נובעת ממנו. לפי סעיף א' כל תכונה ביחס נובעת

ישירות מאחת מתכונות המפתח. לפי סעיף ב' נקבל כי כל תכונה נמצאת ברכיב היוצא מאחת

(או יותר) מתכונות המפתח. לכן כל מפתח מקיים את התנאי. מאידך ברור שכל קבוצת תכונות

המקיימת את התנאי היא מפתח (כי מהסעיף הקודם נקבל שכל תכונה נובעת מאברי המפתח).

iv. (5 נק') הוכיחו כי קבוצת תכונות $Z \subseteq R$ היא מפתח קביל אם ורק אם לכל תכונה $B \in R$ קיים $A \in Z$ כך ש- B נמצאת ברכיב היוצא מ- A , ואין מסלול מכוון בין אף זוג תכונות ב- Z .
 לפי הסעיף הקודם התנאי בודאי מגדיר מפתח. נשים לב כי המפתח קביל, מאחר ואם ניתן היה להוריד ממנו תכונה אחת ועדין לקבל מפתח, הרי שלפי הסעיף הקודם היה מסלול מכוון אל התכונה שהורדנו מאחת (או יותר) מהתכונות שנשארו במפתח, בסתירה לתנאי.
 מאידך כל מפתח קביל הוא מפתח ולכן מקיים את התנאי מהסעיף הקודם. בנוסף אם היה מסלול מכוון בין שתי תכונות במפתח $A \rightarrow B$ הרי שניתן היה להוריד את B ועדיין להישאר עם מפתח.
 ובאופן מדויק יותר. אם $A \rightarrow B$ אז מהכללה נקבל כי $Z \setminus B \rightarrow Z$ וע"י טרנזיטיביות נקבל כי $Z \setminus B$ מפתח, בסתירה לכך ש- Z קביל.

v. (5 נק') הראו שלכל המפתחות הקבילים של R אותו הגודל.

יהיו $S = \{A_1, \dots, A_k\}$ ו- $T = \{B_1, \dots, B_m\}$ שני מפתחות קבילים. נניח כי $k \leq m$. מאחר ו- S מפתח הרי

שלכל B_i יש A_{k_i} ב- S כך ש- B_i ברכיב של A_{k_i} . בפרט יש תת קבוצה של S בגודל $m \geq$ ממנה

יוצאים מסלולים מכוונים לכל אברי T . לפי הסעיפים הקודמים נובע כי תת קבוצה זו היא מפתח.

מכך ש- S קביל נובע כי תת הקבוצה היא S כולה ולכן בפרט $k \leq m$. מכאן נובע ש- $m = k$ כנדרש.

vi. (3 נק') תנו דוגמא לסכמה וקבוצת תלויות (ללא הגבלה על צורת התלויות) עם מפתחות קבילים בגדלים שונים.

$\{R=\{A,B,C\}$ ו- $F=\{A \rightarrow BC, BC \rightarrow A\}$. ברור כי A הוא מפתח קביל וכי $\{B,C\}$ גם מפתח קביל.

שאלה 4 – XML (21 נק')

נתון מסמך ה- DTD "DrivingSchool.dtd" הבא :

```
<!ELEMENT DrivingSchool (Teacher*, Student*, Lesson+)>
<!ELEMENT Teacher (Name)>
<!ATTLIST Teacher id ID #REQUIRED>
<!ELEMENT Name (#PCDATA)>
<!ELEMENT Student (Name)>
<!ATTLIST Student id ID #REQUIRED>
<!ELEMENT Lesson EMPTY>
<!ATTLIST Lesson
    teacher IDREF #REQUIRED
    student IDREF #REQUIRED >
```

א. (5 נק') כתבו מסמך XML קטן שבנוי לפי ה-DrivingSchool.dtd DTD. השלימו את המסמך להלן.

```
<!DOCTYPE DrivingSchool SYSTEM "1.dtd">
<DrivingSchool>
  <Teacher id="t1">
    <Name/>
  </Teacher>
  <Student id="s1">
    <Name/>
  </Student>
  <Lessons teacher="t1" student="s1"/>
</DrivingSchool>
```

ב. (8 נק') נתונה השאילתא הבאה ב-XPATH. כתבו מה מחשבת השאילתא על מסמך XML הבנוי לפי ה- DrivingSchool.dtd DTD. הסבירו מה בודקת כל אחת מארבעת השורות המסומנות.

//Lesson[

1. @student="saib" and (
2. @teacher!=following-sibling::*[@student="saib"]/@teacher or
3. @teacher!=preceding-sibling::*[@student="saib"]/@teacher) and
4. not(@teacher=preceding-sibling::*[@student="saib"]/@teacher)
]/@teacher

השאילתא בודקת אם התלמיד "saib" למד אצל יותר ממורה נהיגה אחד. אם כן,

השאילתא מחזירה את מזהי המורים ללא חזרות, אחרת התוצאה ריקה.

1. השאילתא עוברת על כל השיעורים בהם למד סאיב.

2. בודקת האם קיים שיעור נוסף, באחים הבאים, שסאיב למד בו אך שם המורה שונה

3. כמו 2 אך עבור אחים קודמים, שורה זו נחוצה כדי להחזיר את כל המורים העונים על הדרישה

4. שורה זו מונעת כפילויות על ידי החזרת המופע הראשון של המורה.

ג. (8 נק') בהינתן מסמך XML d.xml הבנוי לפי ה- DTD DrivingSchool.dtd, כתבו שאילתה ב XQuery המחזירה את רשימת הסטודנטים שלקחו שיעורים אצל מורים שונים. יש לדאוג שאף תלמיד לא יופיע בפלט יותר מפעם אחת. הפלט יהיה מהצורה:

```
<res>
  <Student>..</Student>
  <Student>..</Student>
  ...
</res>
```

```
<res>{
for $s in doc("d.xml")//Student
let $teachers := doc("d.xml")//Lessons[@student = $s/@id]/@teacher
where $teachers != $teachers
return $s
}</res>
```


שאלה 5 – DATALOG (10 נק')

כזכור, בכיתה הגדרנו נקודת שבת מינימאלית להיות השמה ל-IDB המקיימת את כללי התוכנית אך שכל תת השמה מפרה את אחד הכללים.

נתונה תוכנית ה-DATALOG הבאה:

$$p(X) \leftarrow a(X), \neg b(X)$$

$$q(X) \leftarrow b(X), \neg p(X)$$

$$r(X) \leftarrow p(X), \neg q(X)$$

כאשר a ו- b הם ה-EDB ו- p, q, r הם ה-IDB. נתון כי $a = \{1, 2\}$ ו- $b = \{2, 3\}$. מצאו שתי נקודות שבת מינימאליות שונות לתוכנית. הוכיחו את טענתכם.

נקודת שבת מינימאלית אחת היא $p = \{1\}, q = \{2, 3\}, r = \{1\}$. נקודת שבת מינימאלית שניה היא

$p = \{1, 2, 3\}, q = \{\}, r = \{1, 2, 3\}$. הנקודה הראשונה מתקבלת ע"י ריבוד התוכנית וחישוב נקודת

השבת המתקבלת לפי האלגוריתם שנלמד בכיתה.

הבניה השניה בברור נקודת שבת, והיא מינימלית כי אם נוריד נקודה מ- p אז או שנפר את

$p(x) \leftarrow a(x), \neg b(x)$ או שנאלץ להוסיף ערכים ל- q . אם נקטין את r אז נפר את הכלל

$r(x) \leftarrow p(x), \neg q(x)$.