הטכניון - מכון טכנולוגי לישראל
הפקולטה למדעי המחשב

פרופ' עודד שמואלי
סעאב מנסור
לינה זריבץ'

חורף תשס"ו
6 בינואר 2006

מערכות מסדי נתונים

בוחן 1 - מועד א' – פתרון

הזמן: 100 דקות (1:40 שעות)
במבחן זה 7 עמודים

שם פרטי: ____________

שם משפחה: ____________

מס' סטודנט: ____________

פקולטה: ____________

שאלה	נקודות	מתוך
שאלה 1 – ERD		30
שאלה 2 – שפות שאילתא		40
שאלה 3 – תלויות פונקציונאליות		30
סה"כ		**100**

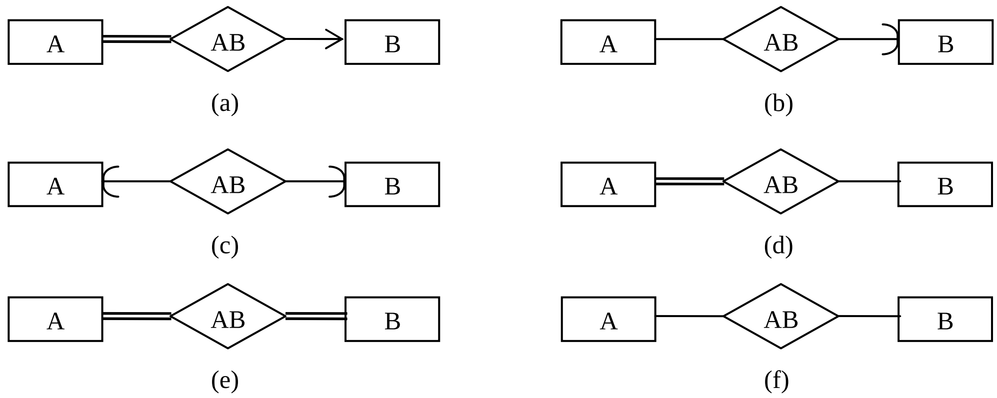
הערות:

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 שתתכננו היטב את זמנכם, **לא תינתנה הארכות**.
6. ערעורים יש להגיש תוך שבועיים ממועד פרסום התוצאות.
7. לא יתקבלו ערעורים בנוסח "בדיקה מחמירה מדי".
8. בשאלה של בחירה מרובה ("אמריקאית"), לכל סעיף יופעל כלל הניקוד הבא:
   - אם לא סומנה אף תשובה נכונה, הציון יהיה אפס, ואחרת:
     - על **אי** סימון תשובה נכונה יורדו 2 נקודות.
     - על סימון תשובה **לא** נכונה יורדו 2 נקודות.
   - בכל מיקרה, מספר הנקודות לסעיף כלשהו לא יקטן מאפס.

בהצלחה

שאלה 1 – ERD (30 נק')

נתונות סכמות ה-ERD הבאות:



נתבונן ברלציות המתאימות לטיפוס הישויות A, טיפוס הישויות B וליחס AB עבור כל אחת מהדוגמאות. נסתכל על מצב של המסד שבו ישנן $\{a_1,\ldots,a_m\}$ ישויות מטיפוס A, $\{b_1,\ldots,b_n\}$ ישויות מטיפוס B, ו-k רשומות בטיפוס הקשר AB.

לכל סעיף סמן רק את התשובות שבהכרח מתקיימות:

א. (5 נק') עבור הסכמה (a):

a. $m = n$
b. $m \le n$
c. $m \ge n$
d. $k = m$ ☑
e. $k = n$
f. $k \ge m+n$

ב. (5 נק') עבור הסכמה (b):

a. $m = n$
b. $m \le n$
c. $m \ge n$
d. $k = m$ ☑
e. $k = n$
f. $k \ge m+n$

ג. (5 נק') עבור הסכמה (c):

a. $m = n$ ☑
b. $m < n$
c. $m > n$
d. $k = m$ ☑
e. $k > m$
f. $k = n$ ☑
g. $k > n$
h. $k \ge m+n$

ד. (5 נק') עבור הסכמה (d):

a. $m = n$
b. $m \le n$
c. $m \ge n$
d. $k = m$
e. $k \ge m$ ☑
f. $k = n$
g. $k \ge n$
h. $k \ge m+n$

ה. (5 נק') עבור הסכמה (e):

a. $m = n$
b. $m \le n$
c. $m \ge n$
d. $k = m$
e. $k \ge m$ ☑
f. $k = n$
g. $k \ge n$ ☑
h. $k \ge m+n$

ו. (5 נק') עבור הסכמה (f) מצא יחס לא טריוויאלי המתקיים בין m, n ו-k. (יחס יקרא טריוויאלי אם הוא מתקיים לכל שלושה מספרים טבעיים. למשל: m+k+n≥0)

$$\boxed{k \leq m*n}$$

שאלה 2 – שפות שאילתא (40 נק')

שאלה זאת מתייחסת למסד נתונים ובו רלציה Location(Id, X, Y, Type) המתארת מס' מזהה ייחודי לאובייקטים, מיקומם במישור (כלומר, קואורדינאטות X ו-Y מטיפוס float) וסוגם. כשנתייחס ל- "אובייקט Id" נתכוון לאובייקט ש- Id הנו המזהה שלו. לדוגמא, הרשומה (1, 5.2, 3.4, 'Sch') מייצגת אובייקט מסוג בית ספר ('Sch'), בעל מספר מזהה 1 והנמצא בנקודה (5.2, 3.4).

א. (5 נק') הגדירו ב-**SQL** מבט בשם Distance אשר מכיל את המרחק בין כל שני אובייקטים <u>שונים</u> מהרלציה Location וכן את סוגי האובייקטים.
הסכמה של המבט תהא Distance(Id1, Type1, Id2, Type2, Dist).
מותר להשתמש בפונקציות הבאות:
- float sqrt(float) לחישוב שורש ריבועי.
- float pow(float number, float exponent) אשר מעלה את המספר number בחזקה exponent.

```
CREATE VIEW Distance As
SELECT l1.Id AS Id1, l1.Type AS Type1,
       l2.Id AS Id2, l2.Type AS Type2,
       sqrt(pow(l1.X-l2.X,2.0)+ pow(l1.Y-l2.Y,2.0)) AS Dist
FROM Location l1, Location l2
Where l1.Id <> l2.Id
```

<u>שימו לב</u>: בסעיפים הבאים ניתן להשתמש ברלציה Distance מסעיף א'.

ב. (10 נק') כתבו ב-**RA** שאילתא המחזירה את כל הזוגות (id_1, id_2) כך ש-id_1 הינו אובייקט מטיפוס בית ספר ('Sch') ו-id_2 הינו אובייקט מטיפוס תחנת אוטובוס ('BS'), כך שאין תחנת אוטובוס אחרת id_3 שמרחקה אל id_1 קטן ממש מהמרחק בין id_2 ל- id_1 (שימו לב, לכל בית ספר יתכנו יותר מתחנת אוטובוס אחת כזאת).

$\pi_{Id1,Id2}\ (\sigma_{Type1\ =\ 'Sch'\ \wedge Type2='BS'}\ Distance)\backslash$

$\pi_{Id1,Id2}\ (\sigma_{Id1=Id3\wedge T1='Sch'\wedge T2='BS'\wedge T4='BS'\wedge\ D1>D2}\ (\ (\rho_{Type1\rightarrow T1,\ Type2\rightarrow T2,Dist\rightarrow D1}\ Distance)\times$

$(\rho_{Id1\rightarrow Id3,Type1\rightarrow T3,Id2\rightarrow Id4,Type2\rightarrow T4,Dist\rightarrow D2}\ Distance)\ )\ )$

ג. (10 נק') זוג אובייקטים שונים id_1,id_2 יקרא *קרוב* אם לא קיים אובייקט שונה מהם כך שהמרחק, בינו ובין id_1 <u>או</u> בינו ובין id_2, קטן ממש מהמרחק בין id_1 ו-id_2.
כתבו שאילתא ב-**DRC** המחזירה את הרשומות (id_1,id_2) כך ש id_1 ו-id_2 הינם אובייקטים *קרובים*.
יהא **Close** השם של הרלציה המתקבלת.

$\{<id1,\ id2>|\ \exists\ (t1,\ t2,\ d)\ (Distance(id1,t1,id2,t2,d)\wedge$

$(\ \neg\exists\ (id',\ t',\ d')\ (Distance(id1,t1,id',t',d')\ \wedge\ (d'\ <d)\ )\ )\wedge$

$(\neg\exists\ (id',\ t',\ d')\ (Distance(id',t',id2,t2,d')\ \wedge\ (d'\ <d)\ )\ )\ )$

ד. (15 נק') נסמן ב-**InClose** את קבוצת כל האובייקטים המופיעים ברשומה כלשהי ברלציה **Close**. זוג אובייקטים שונים id_1,id_2 יקרא *קרוב מסדר שני* אם
(a) $(id_1,id_2) \notin$ **Close** (כלומר הזוג id_1,id_2 אינו *קרוב*).
(b) אם נזרוק ממסד הנתונים שלנו את כל האובייקטים הנמצאים ב-**InClose** אזי id_1,id_2 יהוו זוג *קרוב* במסד החדש.
כתבו שאילתא ב-**SQL** המחזירה את כל הזוגות *הקרובים מסדר שני*.
מותר להשתמש ברלציה **Close** מהסעיף הקודם וכמו כן להגדיר מבטי עזר.

```
CREATE VIEW InClose AS
SELECT DISTINCT Id1 FROM Close

CREATE VIEW NotInClose AS
```

```sql
SELECT Id1, Id2, Dist
FROM Distance
WHERE Id1 NOT IN (SELECT * FROM InClose)
AND Id2 NOT IN (SELECT * FROM InClose)
```

```sql
SELECT Id1, Id2
FROM NotInClose d
WHERE Dist < (SELECT MIN(Dist) FROM NotInClose
          WHERE d.Id1 = Id1 AND Id2<>d.Id2) AND
     Dist < (SELECT MIN(Dist) FROM NotInClose
                  WHERE d.Id2 = Id1 AND Id2<>d.Id1)
```

פתרון נוסף:

```sql
SELECT Id1, Id2
FROM NotInClose d
WHERE Dist < (SELECT MIN(Dist) FROM NotInClose
          WHERE (d.Id1 = Id1 OR d.Id2=Id1) AND Id2<>d.Id2 AND Id2<>d.Id1)
```

שאלה 3 – תלויות פונקציונאליות (30 נק')

נתונה סכמת יחס $R(A,B,C,D)$ עם קבוצת תלויות פונקציונאליות F הנכונות בה. נגדיר שקבוצת אטריבוטים X היא *סגורה* אם מתקיים ש- $X^+=X$ (בשאלה זו כל הסגורים מחושבים יחסית ל-F).

נתון שרק קבוצות האטריבוטים Ø ו- {A,B,C,D} הן *סגורות*, וכל קבוצה אחרת $S\subseteq\{A,B,C,D\}$ אינה *סגורה*, כלומר $S\neq S^+$.
לדוגמא: $\{A\}\neq\{A\}^+$.

א. (10 נק') מהו $\{B\}^+$?

נראה כי $\{B\}^+=\{A,B,C,D\}$

לפי הגדרת סגור של קב' אטריביוטים $\{B\}\subseteq\{B\}^+$ ולכן $\{B\}^+\neq Ø$.

מכיוון ש-{B} אינה קב' סגורה לפי נתוני השאלה, מתקיים כי $\{B\}\neq\{B\}^+$.

נניח בשלילה כי למשל $\{B\}^+=\{A,B\}$. אך לפי הגדרת הסגור, נובע כי $\{A,B\}^+=\{A,B\}$, וזה

בסתירה לנתוני השאלה. ניתן לחזור על הטיעון הזה לכל תת קבוצה ממש של { A,B,C,D},

ולכן האפשרות היחידה שנותרה היא $\{B\}^+=\{A,B,C,D\}$.

ב. (7 נק') מצא כיסוי מינימאלי (minimal cover) של F.

ניתן לחזור על הטיעון מסעיף ב' לכל אטריבוט ב-R,ולכן $\{A\}^+=\{B\}^+=\{C\}^+=\{D\}^+=\{A,B,C,D\}$

כיסוי מינימאלי שמקיים זאת הוא למשל $\{A\to B, B\to C, C\to D, D\to A\}$

ג. (5 נק') האם הכיסוי מסעיף ב' הוא יחיד או שיש יותר מאחד כזה? נמק.

לא, הכיסוי הנ"ל אינו יחיד. למשל גם הכיסוי $\{B\to A, A\to C, C\to D, D\to B\}$ הוא כיסוי

מינימאלי שמקיים את דרישות השאלה.

ד.   (8 נק') הוצע להוסיף את חוק ההיסק הבא לאקסיומות של ארמסטרונג:
אם XZ→YZ, אזי X→Y כאשר $X,Y,Z \subseteq U$.
האם החוק שהוצע הינו נאות (sound)? אם כן, נמק. אם לא, תן דוגמא נגדית.

החוק אינו נאות. נדגים זאת ע"י כך שנראה תלות f וקבוצת תלויות F, כך שמ- F ניתן להסיק את f במערכת החדשה, אך F אינה גוררת לוגית את f.

לדוגמה, F={AC→BC}, f=A→B. ברור כי ע"י מערכת האקסיומות החדשה F מוכיח את f. אבל, F אינה גוררת לוגית את f, כיוון שלמשל עבור הרלציה r={(0,1,2),(0,0,3)}, r מספקת את F, אך r אינה מספקת את f.