

הטכניון - מכון טכנולוגי לישראל
הפקולטה למדעי המחשב



חורף תשס"ה
10 במרץ 2005

פרופ' עודד שמואלי
מר גיא קורלנד
גברת לינה פבזנר

מערכות מסדי נתונים מועד ב'

הזמן: 3 שעות
במבחן זה 11 עמודים

שם פרטי: _____
שם משפחה: _____
מס' סטודנט: _____
פקולטה: _____

שאלה	נקודות	מתוך
שאלה 1 – ERD + ODL		21
שאלה 2 – SQL		20
שאלה 3 – RA + RC		16
שאלה 4 – Design		23
שאלה 5 – XML		20
סה"כ		100

הערות:

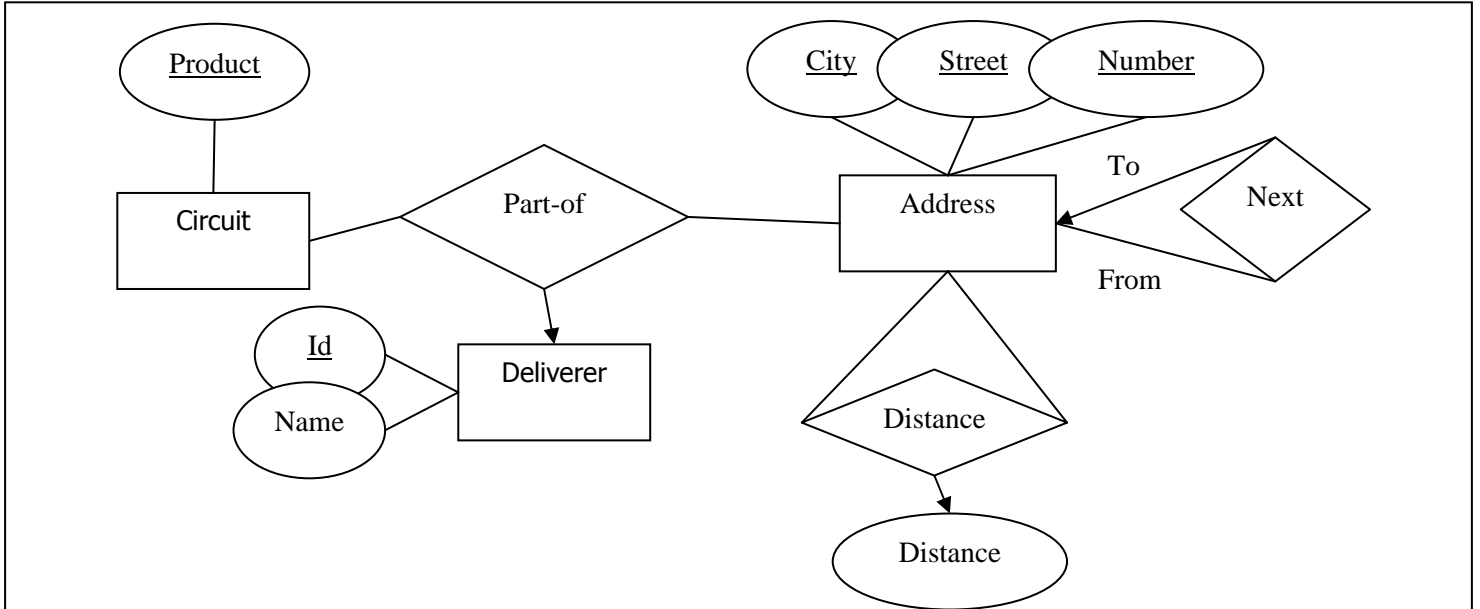
- יש לענות על כל השאלות **בטופס הבחינה**.
- חומר עזר מותר: רק דברים שעשויים מנייר.
- אין להחזיק מכשיר אלקטרוני כלשהו (לדוגמא, טלפון או מחשבון) במצב ON ואו מחוץ לילקוטב-פק.
- קראו היטב את ההוראות שבתחילת כל שאלה ואת ההסברים לדיאגרמות.
- מומלץ שתתכננו היטב את זמנכם, **לא תינתנה הארכות**.
- ערעורים יש להגיש תוך שבועיים ממועד פרסום התוצאות.
- לא יתקבלו ערעורים בנוסח "בדיקה מחמירה מדי".

בהצלחה!

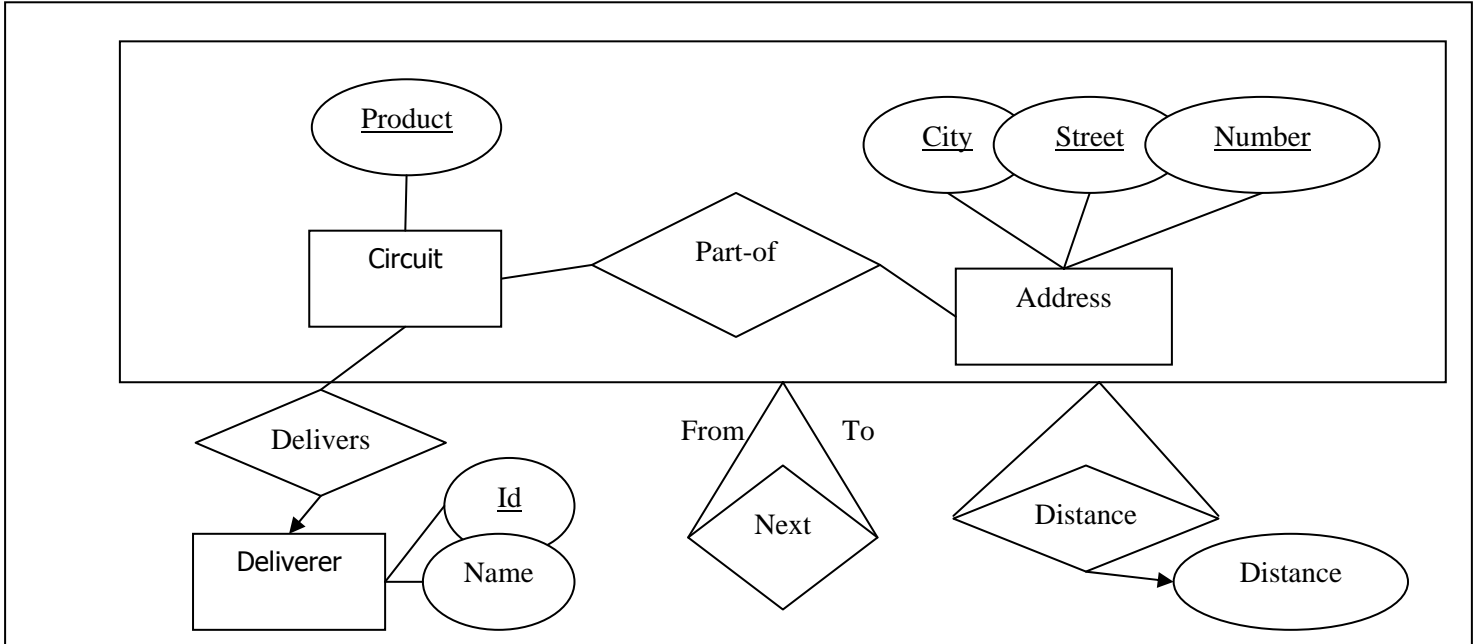
שאלה 1 – ERD + ODL (21 נק')

חברת "משלוח לכל" מספקת שירות הפצה של מוצרים למינויים עד בית הלקוח (עיתונים, לחמניות, שוקו וכו'). החברה הזמינה הצעה לניהול מערך ההפצה והמוצרים. חברת DBMS שוב נרתמה למשימה והציעה את סכמות ה ERD הבאות:

דיאגרמה 1.



דיאגרמה 2.



הסברים:

מסלול חלוקה (Circuit) עבור מוצר (Product) מורכב ממספר כתובות (Address). כתובות מורכבות מעיר (City) רחוב (Street) ומספר בית (Number). עבור שתי כתובות רצופות במסלול יש סדר ביניהן (Next) כמו כן ניתן לשמור מרחק (Distance). מפיץ (Deliverer) במסלול חלוקה מורכב ממספר אישי (Id) ושמו.

א. 9 נק') מלאו את הטבלאות הבאות, בהתייחס לדיאגרמת ה-ERD.

Delivers		
דיאגרמה 2.	דיאגרמה 1.	
		שדות:
		מפתחות:

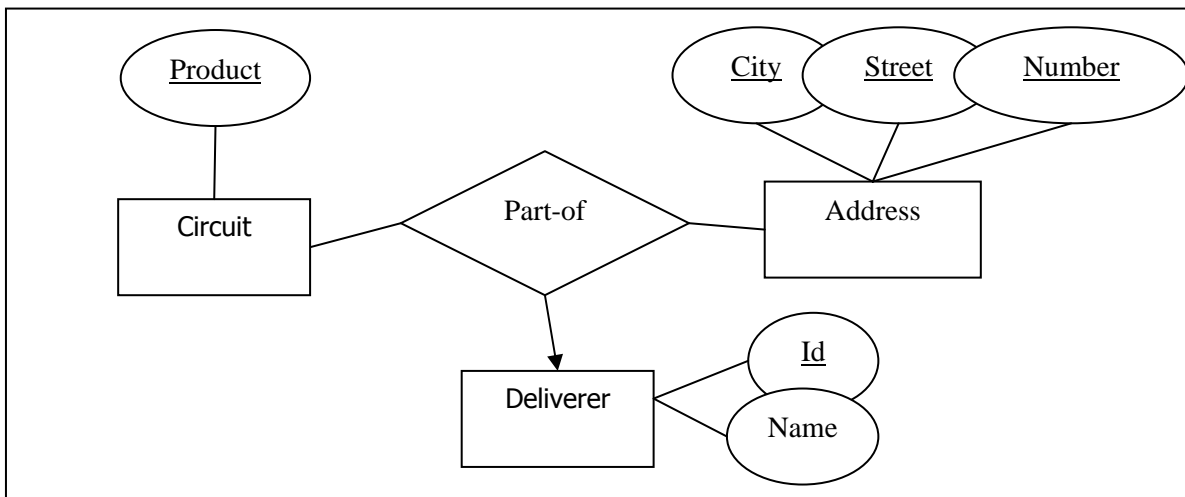
Part-of		
דיאגרמה 2.	דיאגרמה 1.	
		שדות:
		מפתחות:

Next		
דיאגרמה 2.	דיאגרמה 1.	
		שדות:
		מפתחות:

ב. 6 נק') השוו בין שני התכנונים השונים, סמנו נכון/לא נכון עבור כל משפט:

דיאגרמה 2.	דיאגרמה 1.	
		לכל מסלול מפיץ אחד בדיוק
		לכל מסלול יתכנו מספר מפיצים
		האם יתכן זוג כתובות סדור אשר במסגרת שני מסלולים שונים המרחק בין הכתובות הנו שונה
		שתי כתובות הנמצאות באותו מסלול יופיעו תמיד באותו סדר (NEXT) בכל מסלול

ג. 6 נק') השלימו את התרגום (המופיע בהמשך) עבור החלק הנ"ל של הסכמה ל ODL:



```
Interface Circuit (key Product){  
    attribute string Product;
```

```
}
```

```
Interface Address {  
    attribute string City;  
    attribute string Street;  
    attribute integer Number;
```

```
}
```

```
Interface Deliverer {  
    attribute string Name;  
    attribute string Id;
```

```
}
```


ג. (5 נק')

כיתבו שאילתא המחזירה את שני הערכים השכיחים ביותר ב-R. בסעיף זה ניתן להניח כי מספר המופעים של כל ערך הינו ייחודי.
למשל, עבור רלצית הדוגמא מהסעיף הקודם, השאילתא אמורה להחזיר 1 ו-2.

ד. (5 נק')

נסמן ב-R(a) את מספר המופעים של ערך a ברלציה R. נסתכל על כל הערכים השונים המופיעים ברלציה R, נניח a_1, a_2, \dots, a_k ונסדר אותם באופן הבא (יתכנו מספר סדרים כאלו עבור R):

$$R(a_{i1}) \leq R(a_{i2}) \leq \dots \leq R(a_{ik})$$

כיתבו שאילתא המחזירה את $a_{i1}, a_{i2}, \dots, a_{i100}$.
שימו לב שבסעיף זה אין הנחה שמספר המופעים של כל ערך הינו ייחודי.
אין לכתוב שאילתא ארוכה או המשתמשת ב- וכו' או ב-"...".

שאלה 3 – RA+DRC (16 נק')

נתונה רלציה R בעלת אטריביוטים A, B, C . התחום מתוכו נלקחים הערכים עבור A, B, C הנו המספרים הטבעיים. בכל סעיף כתוב ביטוי קצר ככל האפשר. זכור כי \setminus מציין חיסור.

א. (3 נק') כתוב ביטוי ב-DRC השקול לביטוי הבא: $(\rho_{S(A,B)}\pi_{B,C}R) \setminus \pi_{A,B}R$

כתוב ביטוי שקול ב-RA לביטויי ה-DRC הבאים (סעיפים ב'-ד'):
ב. (3 נק') $\{ \langle x, y \rangle : \exists z R(x, y, z) \wedge \forall z \neg R(z, x, y) \}$

ג. (5 נק') $\{ \langle x, y \rangle : \exists z R(x, y, z) \wedge \forall z', x', y' (R(x', y', z') \rightarrow R(x, y, z')) \}$

ד. (5 נק') $\{ \langle x, y \rangle : \exists z R(x, y, z) \wedge \forall z' (\exists x', y' R(x', y', z') \rightarrow R(x, y, z')) \}$

שאלה 4 – Design (23 נק')

א. (2 נק') תהא R, H סכמה מעל אטריביוטים R וקבוצה של תלויות פונקציונאליות H . נתון שכל אטריביוט ב- R הנו מפתח קביל עבור R, H . האם R, H בהכרח ב-BCNF? אם כן נמק. אם לא, תן דוגמא נגדית קצרה כן/לא

ב. (8 נק') יהיו R, F ו- R', F' סכמות הנמצאות ב-3NF. יהיו K ו- K' מפתחות קבילים כלשהם של שתי הסכמות בהתאמה. נגדיר תלויות פונקציונאליות $f = K \rightarrow K'$, $g = K' \rightarrow K$. נגדיר סכמה חדשה: $R'', F'' = R \cup R', F \cup F' \cup \{f, g\}$

טענה: R'', F'' נמצאת בהכרח ב-3NF.

"הוכחה": נתבונן בתלות כלשהי ב- F . נאמר ללא"כ, $X \rightarrow E$. כיוון ש- R, F ב-3NF, (1) X מפתח על עבור R, F , או ש- (2) E מוכל במפתח קביל כלשהו של R, F . במקרה (1), $X \rightarrow K \rightarrow K'$, ולכן X מפתח על בסכמה החדשה. במקרה (2), E שייך למפתח קביל U ב- R, F . אז $U \rightarrow K \rightarrow K'$ ולכן E שייך למפתח קביל בסכמה החדשה. הסיטואציה דומה לגבי תלות כלשהי ב- F' . התלויות החדשות, f, g , אינן סותרות 3NF כיוון שהאטריביוטים המופיעים בהן בצד שמאל מהווים מפתח על בסכמה החדשה R'', F'' .

1. (4 נק') מה השגיאה העיקרית בהוכחה זו?

2. (4 נק') הראה דוגמא נגדית לטענה ובה ל- R בדיוק 3 אטריביוטים.

ג. (13 נק') תהא R, F סכמה ב-3NF אך לא ב-BCNF. תהא R, G סכמה ב-3NF אך לא ב-BCNF, כאשר R זהה בשני המקרים.

1. (6 נק') האם $R, F \setminus G$ בהכרח נמצאת ב-3NF? אם כן נמק בקצרה. אם לא תן דוגמא נגדית קצרה (לכל היותר 4 אטריביוטים ל- R).
כן/לא

2. (7 נק') האם $R, (F \cup G)$ יכולה להיות ב-BCNF? אם לא נמק בקצרה. אם כן תן דוגמא קצרה (לכל היותר 4 אטריביוטים ל- R).
כן/לא

שאלה 5 – XML (20 נק')

נניח מסמך XML כאוסף עובדות ל-DATALOG. לצורך שאלה זו, נניח שלמסמכים אין אטריביוטים אלא רק אלמנטים.

נשתמש בפרדיקטים הבאים ב-EDB:

- Root(ID) – ID is the unique id of the root element of the document.
ID הנו מזהה ייחודי של שורש המסמך (קיימת עובדה אחת בלבד לפרדיקט Root).
 - Tag(ID, TAG) – TAG is the tag of the XML element with unique id ID. Tag is **not** defined for the ID of the root.
TAG הנו שם האלמנט של אלמנט ה-XML בעל מזהה ייחודי ID. TAG אינו מוגדר עבור המזהה הייחודי של השורש.
 - Edge(ID1, ID2, POS) – The element with ID2 is the POSth child of the element with ID1.
האלמנט בעל ID2 הנו הבן ה-POS במספר (מספר הבן השמאלי ביותר הוא 1) של האלמנט בעל ID1.
 - Text(ID, TEXT) – Element ID has a text child element containing the text TEXT.
לאלמנט בעל מזהה ייחודי ID יש בן מסוג text הכולל את המחרוזת TEXT.
- לדוגמא:

```
<a>
  <b>55</b>
  <b>6</b>
  <d> <e>909</e></d>
</a>
```

ה-EDB ייוצג ע"י:

Root	Tag	Edge	Text
Root(1).	Tag(2,a).	Edge(1,2,1).	Text(3, 55).
	Tag(3,b).	Edge(2,3,1).	Text(4, 6).
	Tag(4,b).	Edge(2,4,2).	Text(6, 909).
	Tag(5,d).	Edge(2,5,3).	
	Tag(6,e).	Edge(5,6,1).	

א. (3 נק') כתוב תוכנית DATALOG ובה פרדיקט $r(X)$ המחשב את ה-ID של הבן הראשון של השורש של המסמך.

ב. (4 נק') כתוב תוכנית DATALOG ובה פרדיקט $p(X)$ המחשב את ה-IDs של כל

האלמנטים המוחזרים ע"י ביטוי ה-Xpath: $/a \mid /child::*[2]/child::b$

זכור כי \mid מציין UNION.

ג. (6 נק') כתוב תוכנית DATALOG ובה פרדיקט $q(X)$ המחשב את ה-IDs של כל האלמנטים המוחזרים ע"י ביטוי ה- $Xpath : //d$

ד. (7 נק') בטא ב- XPath את שאילתת ה- DATALOG הבאה כאשר S הינו פרדיקט השאילתה, כלומר $?s(X)$.
הערה : אות גדולה מציינת משתנה.

$s(X) \leftarrow r(R), \text{Edge}(R,X,P), \text{Tag}(X, a), t(X).$
 $t(X) \leftarrow \text{Edge}(X,Y,P), \text{Tag}(Y, b), u(Y).$
 $u(X) \leftarrow \text{Edge}(X,Y,P), \text{Tag}(Y, c).$
