

הטכניון - מכון טכנולוגי לישראל
הפקולטה למדעי המחשב



חורף תשס"ה
6 בפברואר 2005

פרופ' עודד שמואלי
מר קורלנד גיא
גברת פבזנר לינה

מערכות מסדי נתונים מועד א'

הזמן: 3 שעות
במבחן זה 12 עמודים

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

שאלה	נקודות	מתוך
שאלה 1 – ERD + DATALOG + ODL		22
שאלה 2 – SQL		20
שאלה 3 – RA + RC		13
שאלה 4 – Design		25
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 לילקוטבק-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 + DATALOG + ODL (22 נק')

חברת רהיטים מתמחה בייצור מוצרים מודרניים, קשים להרכבה עצמית ויקרים. המוצרים מיועדים לשימוש ברחובות הערים ובשטחים ציבוריים. החברה הזמינה הצעות לתכנון מסד נתונים לניהול מלאי וצוות ההרכבה.

דרישות החברה ממסד הנתונים הינם כדלקמן (כל סעיף מציין טיפוס מידע ורכיביו):

ב. עיר: שם (ייחודי במסד), מדינה, גודל האוכלוסייה ושם ראש העיר.

ג. קיימים שני סוגי מוצרים:

a. פסל סביבתי: מספר (ייחודי במסד), שם הפסל, מחיר, שם האמן וסגנון הפסל.

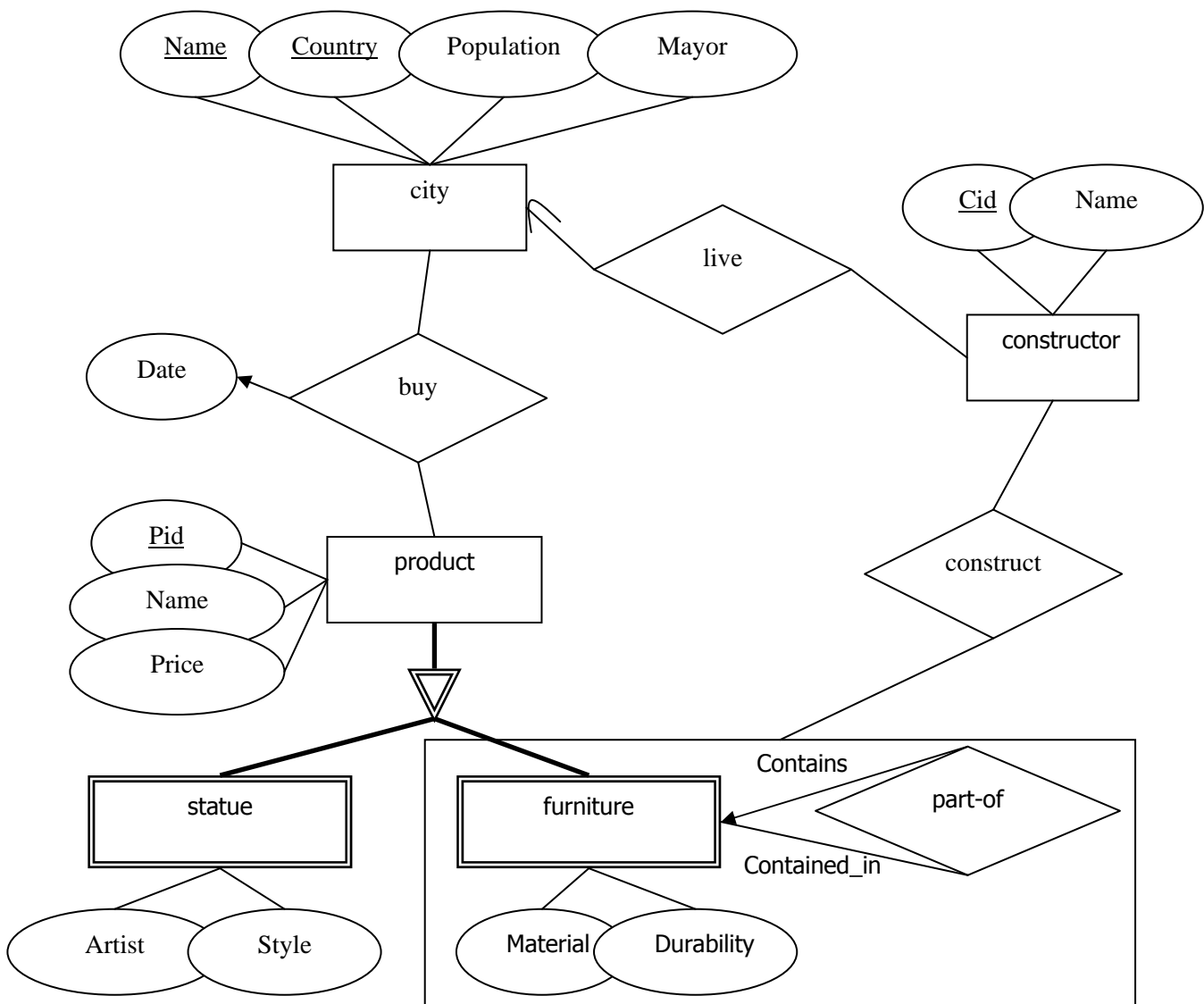
b. רהיט: מספר (ייחודי במסד), שם הרהיט, מחיר, החומר (העיקרי ממנו עשוי הרהיט) ועמידות (בשנים).

ד. מרכיב רהיטים: מספר (ייחודי במסד), שם.

בנוסף:

- עבור כל מרכיב רהיטים יש לשמור נתונים על עיר מגוריו.
- כל מוצר הינו ייחודי ולכן יכול להיקנות פעם אחת לכל היותר. במידה ואכן המוצר ניקנה, יש לתעד את העיר שרכשה אותו ותאריך הרכישה.
- רהיט יכול להיות מורכב מרהיטים (ייחודיים אף הם) אחרים.
- עבור כל הרכבה שבוצעה, חובה לשמור מידע על המרכיב שביצע את ההרכבה.

חברת DBMS הציעה את סכמת ה-ERD הבאה:



א. (6 נק')
השוו את הסכמה המוצעת לדרישות שהתקבלו מהחברה. ציינו שלוש דרישות אשר אינן מתקיימות בסכמה המוצעת.

1. עיר: שם (ייחודי במסד)
2. כל מוצר הינו ייחודי ולכן יכול להיקנות פעם אחת לכל היותר.
3. עבור כל הרכבה שבוצעה, חובה לשמור מידע על המרכיב שביצע את ההרכבה.

ב. (6 נק')	
מלאו את הטבלאות הבאות בהתייחס לדיאגרמה 1.	
furniture	
שדות	Pid, Material, Durability
מפתחות	Pid
construct	
שדות	Cid, Pid_Contained_in
מפתחות	Cid, Pid_Contained_in
buy	
שדות	Name, Country, Date, Pid
מפתחות	Pid, Name, Coutry

ג. (6 נק')

הוחלט להוסיף לתכנון רלציה אשר תתאר עבור מרכיב את הרהיטים השלמים (אלו אשר אינם חלק של רהיט אחר), אשר נטל חלק בהרכבתם (הרכיב חלק אשר מהווה תת-חלק מהרהיט השלם).

כיתבו תוכנית ב- Datalog אשר תגדיר פרדיקט segment(Cid, Pid).

segment פרדיקט דו מקומי כאשר Cid – מספר המרכיב ו-Pid – מספר הרהיט.

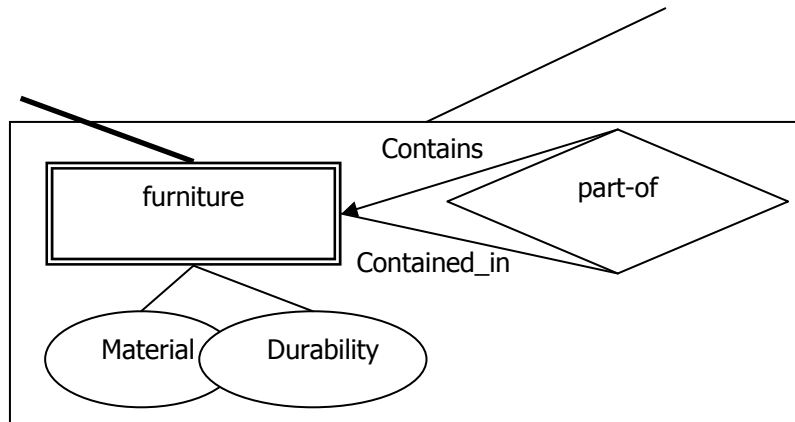
(הניחו תרגום סטנדרטי של סכמת ה ERD לסכמות רלציוניות וכי עבור כל סכמת רלציה הוגדר פרדיקט מתאים).

```

full(Pid) ← furniture(Pid, Name, Price, Material, Durability),
    ⌈ Part_of(Pid_Contained_in, Pid_Contains).
sub_construct(Cid, Pid_Contains) ←
    construct(Pid_Contained_in, Pid_Contains, Cid).
sub_construct(Cid, Pid_Contains) ←
    construct(Pid_Contained_in1, Pid_Contains, Cid1),
    sub_construct(Cid, Pid_Contained_in1).
segment(Cid, Pid) ← full(Pid), sub_construct(Cid, Pid).
    
```

ד. (4 נק')

הוחלט לממש את התכנון כפי שהוצע בסכמת ה ERD בעזרת ODL.
כתבו את הממשקים ב ODL המתארים רק את חלק הסכמה המתואר להלן.



```
Interface furniture : product{
    attribute string Material;
    attribute integer Durability;
    relationship part-of Contained_in inverse part-of::Contains;
    relationship Set<part-of> Contains inverse part-of::Contained_in;
}
```

```
Interface part-of : Auction{
    relationship furniture Contained_in inverse furniture::Contains;
    relationship furniture Contains inverse furniture::Contained_in;
    relationship constructor construct inverse Constructor::constructed;
}
```

הפתרון הבא גם התקבל

```
Interface furniture : product{
    attribute string Material;
    attribute integer Durability;
    relationship furniture Contained_in inverse furniture::Contains;
    relationship Set<furniture> Contains inverse furniture::Contained_in;
}
```

שאלה 2 – SQL (20 נק')

נתונה רלציה (קבוצת ניות ללא חזרות) $R(a, b)$ וארבע שאילתות ה-SQL הבאות a, b הם מטיפוס Integer):

Query1:

```
SELECT a, b
FROM R;
```

Query2:

```
SELECT r1.a, r2.b
FROM R r1, R r2
WHERE r1.a = r2.a AND r2.b > r1.b AND
      NOT EXISTS (SELECT *
                  FROM R
                  WHERE a = r1.a AND b < r2.b AND b > r1.b)
UNION
SELECT r3.*
FROM R r3
WHERE b = (SELECT MIN(b)
          FROM R
          WHERE a = r3.a);
```

Query3:

```
SELECT r1.a
FROM R r1
WHERE NOT EXISTS
(SELECT *
 FROM R r2
 WHERE a = r1.a AND
       NOT EXISTS (SELECT * FROM R WHERE a = r2.a AND b != r2.b)
);
```

Query4:

```
SELECT r1.a
FROM R r1, R r2
WHERE r1.a = r2.a AND r1.b != r2.b
GROUP BY r1.a, r1.b
HAVING COUNT(*) > 1;
```

א. (10 נק') ענו על השאלות הבאות :

a. (5 נק') Query1 שקולה ל-Query2 נכון / לא נכון
נימוק קצר לשקילות או דוגמת רלציה עליה אין שקילות :

נכון. "החלק הראשון" של השאילתא השניה מחזיר את הרשומות (a, b) כך ש-b הוא לא מינימלי מבין כל ה-b שמופיעים עם a מסויים. "החלק השני" מחזיר (a, b) כך ש-b הוא מינימלי.

b. (5 נק') Query3 שקולה ל-Query4 נכון / לא נכון
נימוק קצר לשקילות או דוגמת רלציה עליה אין שקילות :

לא. שאילתא 3 מחזירה a-ים שמופיעים יותר מפעם אחת בטבלה. שאילתא 4 מחזירה a-ים שמופיעים יותר מפעמיים. להלן רלציה עליה אין שקילות :

A	B
2	1
2	3

ב. (10 נק') בהינתן שרלציה R מספקת את התלות $A \rightarrow B$, ענו על השאלות הבאות :

a. (5 נק') Query1 שקולה ל-Query2 נכון / לא נכון
נימוק קצר :

כן. הפעם רק "החלק השני" של השאילתא השניה מחזיר רשומות.

b. (5 נק') Query3 שקולה ל-Query4 נכון / לא נכון
נימוק קצר :

כן. שתי השאילתות יחזירו תוצאה ריקה.

שאלה 3 – RA + RC (13 נק')

1. (3 נק') נתונה רלציה $R_1(A_1, A_2, A_3)$ המוגדרת מעל תחום סופי והמספקת את התלות $A_1 \rightarrow A_2A_3$. כתבו שאילתא ב-DRC אשר מחזירה תוצאה לא ריקה אם לא ניתן להוסיף חיה כלשהי ל- R_1 ועדיין לספק את התלות $A_1 \rightarrow A_2A_3$.

$$\forall a_1 \exists a_2, a_3 R_1(a_1, a_2, a_3)$$

2. (3 נק') נתונות הרלציות הבאות:

$$R_1(A_1, A_2, A_3), R_2(A_2, A_3, A_4)$$

המוגדרות מעל תחום סופי $D = \{a, b, c\}$ והמספקות את התלויות הבאות:
 $\{A_1 \rightarrow A_2A_3, A_2 \rightarrow A_3A_4\}$

הציעו תוכן עבור R_1 ו- R_2 כך שהביטוי הבא מחזיר רלציה ובה מספר חיות גדול ככל האפשר:

$$\pi_{A_2, A_3, A_4}(R_1 \bowtie R_2)$$

$$R_1 = \{ (a, a, a), (b, b, b), (c, c, c) \}$$

$$R_2 = \{ (a, a, a), (b, b, b), (c, c, c) \}$$

3. (7 נק') נתונות הרלציות הבאות:

$$R_1(A_1, A_2, A_3), R_2(A_2, A_3, A_4)$$

המוגדרות מעל תחום סופי כלשהו והמספקות את התלויות הבאות:
 $\{A_1 \rightarrow A_2A_3, A_2 \rightarrow A_3A_4\}$

כתבו שאילתא ב-DRC אשר מחזירה תוצאה לא ריקה אם ורק אם תוכן הרלציות R_1, R_2 הוא כזה שהביטוי $\pi_{A_2, A_3, A_4}(R_1 \bowtie R_2)$ מחזיר מספר חיות גדול ככל האפשר. אין להשתמש ב- $=$ או ב- \neq בשאלה זו.

$$\forall a_1 (\exists a_2, a_3 R_1(a_1, a_2, a_3) \wedge (\forall a_1, a_2, a_3) (R_1(a_1, a_2, a_3) \rightarrow \exists a_4 R_2(a_2, a_3, a_4)))$$

שאלה 4 – Design (25 נק')

א. (7 נק')

קבוצת אטריביוטים X בסכמה R המקיימת תלויות פונקציונאליות F תקרא **סגורה** אם

$$X_F^+ = X \quad \text{הוכח או הראה דוגמא נגדית:}$$

אם X היא קבוצה סגורה ו- Y היא קבוצה סגורה, אזי גם $X \cap Y$ היא קבוצה סגורה.

כנ"ל

נימוק:

הוכחה:

נניח בשלילה ש- $X \cap Y$ אינה קבוצה סגורה $\Leftrightarrow X \cap Y \subsetneq (X \cap Y)_F^+$. נתבונן

באטריביוט A כך ש- $A \in ((X \cap Y)_F^+) \setminus (X \cap Y)$.

$$\Leftrightarrow Y \rightarrow A \in F^+ \text{ וגם } X \rightarrow A \in F^+ \Leftrightarrow A \in (X \cap Y)_F^+$$

כיוון ש X ו- Y סגורות נובע, $A \in X$ וגם $A \in Y$

$$A \in X \cap Y$$

סתירה.

ב. נתונה הסכמה R, F כאשר R הנם האטריביוטים M, L, G, E

ו- F היא: $\{GL \rightarrow E, G \rightarrow L, E \rightarrow G\}$

(1) (2 נק') מי הם כל המפתחות הקבילים בסכמה?

המפתחות הם GM ו- EM .

(2) (8 נק') מצא לסכמה הנתונה פרוק משמר מידע ותלויות לסכמות ב- $3NF$ הכולל

מספר מינימאלי של תתי-סכמות.

(4 נק') הפירוק: $[GL] [EGM]$

נימוק ל- $3NF$:

המפתחות בסכמה $[EGM]$ הם GM ו- EM .

(4 נק') נימוק למינימאליות של מספר תתי-הסכמות:

הסכמה המקורית איננה ב- $3NF$ בשל $G \rightarrow L$.

3) (3 נק') האם יש לסכמה הנתונה פרוק משמר מידע לסכמות ב- BCNF שבו יש בדיוק שתי תתי-סכמות.

כנולא לא.

נימוק:

יש 4 אטריביוטים. אם יש חלוקה לשתי תתי-סכמות אז חייב להיות להן אטריביוט משותף. אם לאחת יש שני אטריביוטים אז לשנייה יש שלושה!

האפשרויות: $(G \rightarrow L)$ MLG, $(G \rightarrow L)$ LGE, $(E \rightarrow G)$ GEM, $(E \rightarrow L)$ EML.

4) (5 נק') מצא לסכמה הנתונה פרוק משמר מידע לסכמות ב- BCNF שבו יש מספר מינימאלי של תתי-סכמות.

[EG] [EL] [EM]

הפירוק:

בינארי.

נימוק ל- BCNF:

פרוק לשתיים בלתי אפשרי – סעיף קודם.

נימוק למינימאליות:

שאלה 5 – XML (20 נק')

בשאלה זו נעסוק במסמכים המצייתים ל- DTD המוצג ב- DOCTYPE של המסמך המוצג להלן (אותו נכנה "מסמך הדוגמא"). במסמך מוצגים נתוני מכירות של חברה לייצור מקררים תעשייתיים. אלמנט **sale** מכיל נתונים עבור שנה מסוימת (ניתן להניח ששנה מופיעה פעם אחת לכל היותר). אלמנט **theyear** מכיל את השנה. אלמנט **units** מכיל את מספר המקררים שנמכרו ואלמנט **amount** את סכום המכירות בדולרים. הנח שאלמנט **sale** לא יכול להכיל שני אלמנטים מסוג **region** עם אותו ערך באטריביוט **name**.

```
<?xml version="1.0"?>
<!DOCTYPE salesdata [
  <!ELEMENT salesdata (sale*)>
  <!ELEMENT sale (theyear, region*)>
  <!ELEMENT theyear (#PCDATA)>
  <!ELEMENT region (units, amount)>
  <!ATTLIST region name (Southeast | Southwest | Northeast | Northwest) #IMPLIED
>
  <!ELEMENT units (#PCDATA) >
  <!ELEMENT amount (#PCDATA) >
]>
<salesdata >
  <sale>
    <theyear>2000</theyear>
    <region name="Northeast">
      <units>374</units>
      <amount>12500.26</amount>
    </region>
    <region name="Southeast">
      <units>512</units>
      <amount>17692</amount>
    </region>
    <region name="Southwest">
      <units>161</units>
      <amount>8349.72</amount>
    </region>
    <region name="Northwest">
      <units>465</units>
      <amount>15239.6</amount>
    </region>
  </sale>
</salesdata>
```

בסעיפים הבאים עליך לכתוב או לפרש שאילות ב- XPATH 1.0 :
א. (3 נק') מה מחשבת השאילתא הבאה?

$$\frac{\text{sum}(\text{//region}[\text{@name}=\text{"Northwest"}]/\text{amount})}{\text{sum}(\text{//region}[\text{@name}=\text{"Northwest"}]/\text{units})}$$

תשובה (עד 14 מילים):

את המחיר הממוצע למקרר על פני כל השנים באזור Northwest.

ב. (3 נק') מה מחשבת השאילתא הבאה?

```
//units [ not ( . < preceding::units) and not ( . < following::units)] [last()]
```

תשובה (עד 14 מילים):

מספר מקררים מקסימלי שנמכר באזור כלשהו בשנה כלשהי.

ג. (4 נק') כתוב שאילתא המחזירה את שמו של כל אזור כך שלא קיימות שתי שנים בהן נמכר אותו מספר מקררים באזור זה.

```
//region[@name="Northeast"][ not(./units =  
./following::region[@name="Northeast"]/units) and  
not(./units = ./preceding::region[@name="Northeast"]/units )]/@name | ...
```

ד. (5 נק') כתוב שאילתא המחזירה את השנים בהן מספר היחידות שנמכר באזור Northwest היה הגדול ביותר בין כל השנים באזור Northwest, במידה ובשנה מסוימת אין נתוני מכירה עבור אזור Northwest הנח שנמכרו אפס יחידות בשנה המסוימת. הנח שישנה שנה בה נמכרו מקררים באזור Northwest.

```
//sale[./region[@name="Northwest"]/amount and  
not( ./region[@name="Northwest"]/amount <  
.. //sale/region[@name="Northwest"]/amount)] /theyear/text()
```

ה. (5 נק')

הסבירו בקצרה מה מחשבת השאילתא הבאה ב- Xquery – (עד 30 מילים):

הערה: הפונקציה fn:number מתרגמת מחרוזות למספרים.
הפונקציה fn:string מחזירה את ה- string value של הארגומנט שלה.

```
for $a in doc("sales.xml")//region,  
    $b in doc("sales.xml")//region  
where (fn:number($b/parent::*/theyear/text()) - 1 =  
    fn:number($a/parent::*/theyear/text()) ) and  
    ($b/@name = $a/@name) and  
    (fn:number($b/units) < fn:number($a/units))  
order by $b/@name ascending,$b/parent::*/theyear/text() ascending  
return  
<region>  
    { fn:string($b/@name) },  
    { fn:number($a/units) },  
    { fn:number($b/units) },  
    { $b /../theyear/text() }  
</region>
```

**השאלתא מוצאת אזורים בהם חלה ירידה במספר היחידות הנמכרות בשנים עוקבות. האזורים
ממוינים בסדר עולה. בתוך כל אזור השנים ממוינות בסדר עולה.**
