



הטכניון - מכון טכנולוגי לישראל
הפקולטה למדעי המחשב

פרופ' ח' אלדר פיושר
נדב שרגאי
סעאב מנסור

חורף תשס"ח
27, אפריל 2008

מערכות מסדי נתונים – 236363

מועד א'
הזמן: 3 שעות
במבחן זה 10 עמודים

עדכונים שנכתבו על הלוח במהלך הבחינה כתובים באדום

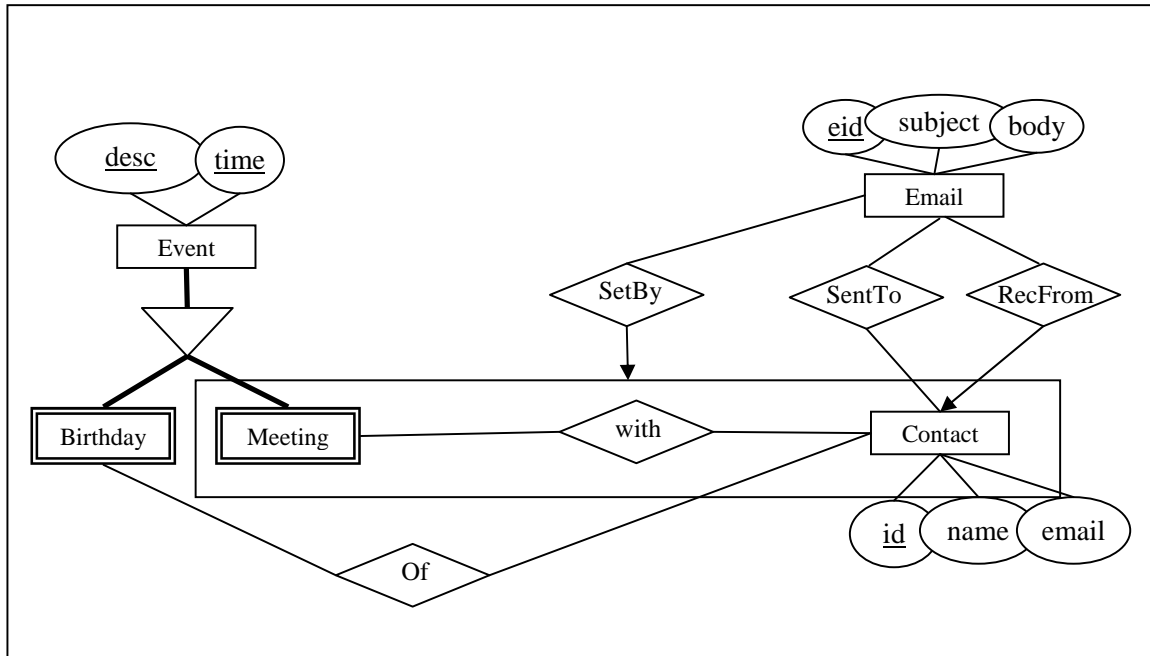
שאלה	נקודות
שאלה 1 – ERD	18
שאלה 2 – שפות שאליתא	32
שאלה 3 – תלויות	28
שאלה 4 – XML	22
סה"כ	100

הנחיות:

- יש לענות על כל השאלות **בטופס הבחינה** מחברות הטייטה לא תיאספנה.
 - 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 (18 נק')

נתונה סכמת ה-ERD הבאה שמייצגת מערכת לניהול מידע אישי עבור **משתמש** יחיד.



הסבר לסכמה:

ישויות:

Contact – ישות המתארת אנשי קשר במערכת. לכל איש קשר יש מזהה ייחודי (id), שם וכתובת דוא"ל (אימייל).

Event – מתאר אירועים שמזוהים ע"י תאור (desc) וזמן (time). קיימים שני סוגים של אירועים:

▪ **Meeting** – מתאר פגישות.

▪ **Birthday** – מתאר ימי הולדת.

Email – מתאר פריטי דוא"ל בעלי מזהה ייחודי eid, נושא (subject) וגוף הדוא"ל (body).

יחסים:

SentTo – יחס המתאר שליחה של דוא"ל לאיש קשר.

RecFrom – יחס המתאר קבלה של דוא"ל מאיש קשר. שימו לב שיכול להיות פריט דוא"ל שגם נשלח לאנשי קשר וגם התקבל מאנשי קשר (למשל דוא"ל עם ממוענים מרובים).

With – יחס המתאר קביעת פגישה עם איש קשר.

Of – יחס המתאר יום הולדת של איש קשר.

SetBy – יחס המתאר שפגישה מתקיימת בנוגע לפריט דוא"ל מסוים.

א. (6 נק') מלאו את הטבלה הבאה שמתייחסת לשדות ולמפתחות (הקבילים) של יחסים בסכמת ה-ERD.

Meeting	
שדות	desc, time
מפתחות	desc, time
RecFrom	
שדות	eid, id
מפתחות	eid
SetBy	
שדות	eid, desc, time, id
מפתחות	eid

ב. (12 נק') תארו במילים פשוטות אילו שינויים יש לעשות בסכמה המקורית, כך שה-ERD החדש יענה על הדרישות הבאות. שימו לב שהסעיפים הינם בלתי תלויים.

i. (6 נק') המערכת תאפשר למשתמש מספר חשבונות דוא"ל, שמאופיינים ע"י כתובת הדוא"ל. כל פריט דוא"ל שייך לחשבון אחד בדיוק שאינו בהכרח חשבון שלפי כתובתו הדוא"ל התקבל או נשלח.

מוסיפים טיפוס ישות Account עם תכונה יחידה address שהיא מפתח.

לאחר זאת מוסיפים טיפוס קשר Belongs בין Email ו-Account עם חץ לאחור

וקו כפול בין Belongs ל-Email.

ii. (6 נק') המערכת תכפה שכל פריט דוא"ל יהיה פריט שנשלח או פריט שהתקבל (אך לא שניהם).

יוצרים יחס הכללה מהישות Email לשתי ישויות חלשות SentEmail ו-RecEmail.

מקשרים את SentTo ל-SentEmail ואת RecFrom ל-RecEmail, במקום הקישור ל-Email.

ראוי גם להשתמש בקווים כפולים בחיבורים ל-SentEmail ו-RecEmail,

אבל לא הורדו נקודות בבדיקה על אי-שימוש כזה.

שאלה 2 – שפות שאילתא (32 נק')

הערה: בסעיפים הבאים הניחו תרגום סטנדרטי של סכמת ה-ERD המקורית משאלה 1 לסכמות רלציוניות: Contact, Email, Meeting, Birthday, SentTo, RecFrom, SetBy, Of, With. בשאלות העוסקות ב-DRC הניחו סדר אלפביתי על התכונות של הרלציות.

א. (8 נק') כתבו שאילתא ב-SQL, אשר לכל איש קשר תחזיר את מספר ההתכתבויות איתו (שליחה או קבלה של דוא"ל). יש למיין את התוצאה בסדר עולה לפי מספר ההתכתבויות. עבור אנשי קשר ללא התכתבויות יש להחזיר אפס.

אין להשתמש במבטים

```
SELECT C.id, COUNT(eid)
FROM Contact C LEFT OUTER JOIN
      (SELECT * FROM SentTo UNION
       SELECT * FROM RecFrom) H
ON C.id=H.id
GROUP BY C.id
ORDER BY COUNT(eid)
```

ב. (8 נק') כתבו שאילתא ב-DRC שמחזירה אנשי קשר, כך שאם נשלח להם דוא"ל, אז גם נתקבל מהם דוא"ל.

$$\{ \langle id \rangle \mid \exists name, email \text{ Contact}(email, id, name) \wedge ((\exists eid \text{ SentTo}(eid, id)) \rightarrow (\exists eid \text{ RecFrom}(eid, id))) \}$$

ג. (8 נק') להלן נתונה שאילתא ב-DRC אשר אמורה לפלוט זוגות של שמות אנשים וימי ההולדת שלהם. השאילתא הזו אינה נכונה. כתבו מהי הבעיה איתה ורשמו מהו התיקון הנדרש.

$$\{ \langle name, time \rangle : \exists id, email (\text{Contact}(email, id, name) \wedge \exists desc \text{ Birthday}(desc, time)) \}$$

בשאילתא אין התאמה בין שם איש הקשר לזמן יום ההולדת ולכן יהיו גם זוגות של שם של איש.

עם תאריך יום הולדת של איש אחר.

יש להחליף את Birthday ביחס $\text{Of}(desc, id, time)$ (ניתן להשאיר את הכמתים כמו שהם).

ד. (8 נק') כתבו שאילתא ב-RA שמחזירה משתמשים שכל הפגישות שנקבעו איתם היו קשורות בפריטי דוא"ל.

ב-"משתמשים" כאן הכוונה היא בעצם לאנשי קשר.

כוונתנו הייתה שפגישה של איש קשר נחשבת לקשורה בפריטי דוא"ל אם קיים לפחות קשר אחד בין דוא"ל כלשהו לבין הקבצה שמכילה את הפגישה. ההקבצה הזאת אינה דווקא עם איש הקשר המדובר.

$\pi_{id}Contact \setminus \pi_{id}((\pi_{desc,time}Meeting \setminus \pi_{desc,time}SetBy) \bowtie With)$

קיבלנו גם את התשובה לפיה פגישה של איש קשר נחשבת לקשורה בפריטי דוא"ל אם קיים קשר בין דוא"ל כלשהו לבין ההקבצה שמכילה אותה ואת הקשר.

$\pi_{id}Contact \setminus \pi_{id}((with \setminus \pi_{desc,time,id}SetBy))$

שאלה 3 – Design (28 נק')

נתונה הסכמה $R(A,B,C,D,E,H)$ וקבוצת התלויות הפונקציונאליות
 $F=\{AB \rightarrow H, E \rightarrow BC, D \rightarrow H, A \rightarrow DE, C \rightarrow E, D \rightarrow BH\}$

א. (7 נק') מצאו כיסוי מינימאלי ל-F

$F_m=\{E \rightarrow B, E \rightarrow C, D \rightarrow H, A \rightarrow D, A \rightarrow E, C \rightarrow E, D \rightarrow B\}$

ב. (7 נק') האם הפירוק $\{R_1(A,B,D), R_2(A,C), R_3(C,D,E,H)\}$ הוא ב-BCNF? הוכיחו את טענתכם.

לא, $D \rightarrow H$ קיים ב- R_3 אבל $D \rightarrow E$ אינו נובע מ-F (ולכן D אינו על-מפתח של R_3)

ג. (7 נק') האם הפירוק בסעיף ב' משמר מידע? הוכיחו את טענתכם.

כן.

נפרק לפי $A \rightarrow C$, נקבל $\{A, C\}$, $\{A, B, D, E, H\}$. נפרק את $\{A, B, D, E, H\}$ לפי $CD \rightarrow EH$, נקבל $\{A, B, D\}$, $\{C, D, E, H\}$. קיבלנו את הפירוק המבוקש באמצעות פירוק לפי תלויות ולכן זהו פירוק משמר מידע.

לחליפין היה אפשר להשתמש באלגוריתם לבדיקת שימור מידע על מנת לענות על השאלה.

ד. (7 נק') האם הפירוק הנ"ל משמר תלויות? הוכיחו את טענתכם.

לא. התלות $E \rightarrow B$ לא נשמרת. לפי האלגוריתם לבדיקת שימור, נתחיל עם $\{E\}$.

החיתוך עם $R1$ ו- $R2$ ריק ולכן לא יתווספו אטריבוטים. עבור $R3$, נוסיף C ע"י $E \rightarrow C$.

משלב זה לא יהיו שינויים. לכן התלות לא נשמרת.

שאלה 4 – XML (22 נק')

נתון קובץ meets.xml המציית ל-DTD הבא:

```
<!ELEMENT db (meet)*>
<!ELEMENT meet (date, name*, agenda)>
<!ELEMENT date (#PCDATA)>
<!ELEMENT name (#PCDATA)>
<!ELEMENT agenda (#PCDATA|time)*>
<!ELEMENT time (#PCDATA)>
```

הסבר: לכל פגישה meet יש תאריך, date, רשימת אנשים משתתפים, שמו של כל אחד name, וסדר יום agenda אשר כולל תיאור טקסטואלי של פעילויות יחד עם שעות שלהם, time.

א. (10 נק') כתבו שאילתת XPath 1.0 אשר תחזיר את האגדות הממוינות. אגדה תיקרא ממוינת אם כל הזמנים המופיעים בה הם בסדר לא יורד – ניתן להניח כאן שהזמנים נתונים בצורה מספרית כך שניתן להשוות ביניהם.
הכוונה היא לסדר מונטוני לא-יורד,

```
//agenda[not(time[.>following-sibling::time])]
```

ב. (12 נק') כתבו עתה שאילתת XQuery 1.0 אשר הפלט שלה הוא קובץ המתאר את תאריכון הפגישות לפי ימים ואנשים מעורבים. המיון צריך להיות בסדר עולה, קודם לפי ימים, ואח"כ לפי שמות האנשים (ללא כפילויות). על הפלט לציית ל-DTD הבא. שימו לב שזהות כל תאריך עוברת עתה לתכונות מתאימות. מותר להניח שכל התאריכים כתובים לפי אותו פורמט (ז"א שלתאריכים זהים יהיו מחרוזות זהות והתאריכים ניתנים להשוואה).

התשובה צריכה להיות ללא כפילויות אנשים, וללא כפילויות ימים.

```
<!ELEMENT db (date)*>
<!ELEMENT date (name)*>
<ATTLIST date d CDATA #REQUIRED>
<!ELEMENT name (#PCDATA)>
```

```
xquery version "1.0";
let $r:=doc("meets.xml")
return
<db>
{
for $d in distinct-values($r//date)
order by $d
return
<date d="{ $d }">
{
for $n in distinct-values($r//name[../date=$d])
order by $n
return <name>{ $n }</name>
}
</date>
}
</db>
```