



אביב תשס"ט  
13 ביולי 2009

פרופ' ח' אלדר פישר  
גב' נעמה טפר  
מר נדב עמית

מרצה:  
מתרגלים:

## 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 + ODL		20
שאלה 2 – שפות שאילתה		36
שאלה 3 – תלויות ופירוקים		23
שאלה 4 – XML		21
סה"כ		100

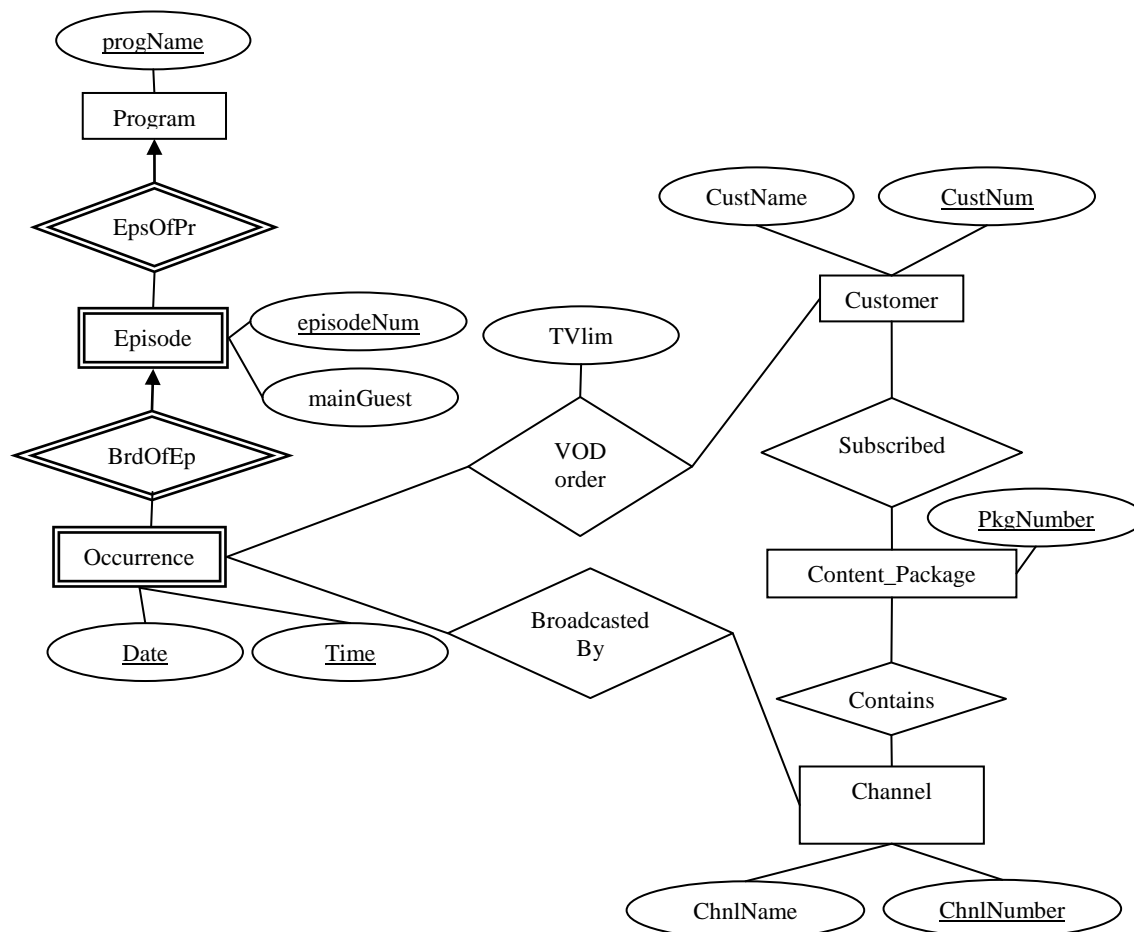
### הנחיות:

- 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 לתרשימים.
- 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 ו-ODL (20 נק')

נתון מסד נתונים של ספק כבלים:



### תיאור הישויות:

Customer – לקוחות. לכל לקוח יש מספר לקוח ושם

Content\_Package – חבילת תוכן. לכל חבילה יש מספר חבילה

Program – תוכנית, לכל תוכנית יש שם ייחודי

Episode – פרק בתוכנית. לכל פרק יש מספר ואורח ראשי שמופיע בפרק

Occurrence – מופע של פרק בתוכנית. לכל שידור יש את תאריך השידור ואת השעה בה שודר

### תיאור היחסים:

Subscribed – יחס המציין מינוי של לקוח לחבילת תוכן

Contains – יחס המציין שידור של ערוץ לחבילת תוכן

VOD\_Order – יחס המציין הזמנות של שידורים ע"י לקוח. עבור כל הזמנה מצוין גם בכמה מכשירי טלוויזיה רשאי הלקוח לצפות בתוכנית (TVlim).

Broadcasted\_By – יחס המציין שידור של פרק בערוץ

EpsOfProg – יחס המציין שייכות של פרק לתוכנית

BrdOfEp – יחס המציין שייכות של שידור לפרק

שאלות:

1. (4 נק') הציגו את הטבלאות הנדרשות עבור הישויות והיחסים הבאים:

טבלה: VOD Order
שדות: CustNum, Date, Time, episodeNum, progName, TVlim
מפתחות: CustNum, Date, Time, episodeNum, progName

טבלה: Episode
שדות: ProgName, episodeNum, MainGuest
מפתחות: ProgName, episodeNum

טבלה: BroadcastedBy
שדות: ChnlNumber, Date, Time, episodeNum, progName
מפתחות: ChnlNumber, Date, Time, episodeNum, progName

2. (5 נק') האם ניתן להזמין מופע של תוכנית שאינה משודרת במסגרת של אף ערוץ? נמקו.

כן, תתכן תוכנית שאף מופע של פרק שלה אינו מופיע ב-Broadcasted_By, ומופעים שלה כן יכולים להשתתף ב-VOD_Order.
--

3. (5 נק') ספק הכבלים רוצה לעקוב מתי מופע של תוכנית הוא שידור חוזר של פרק תוכנית ששודר במסגרת אחד הערוצים (וביחס לאותו שידור). תארו איך ניתן להוסיף מעקב כזה לדיאגרמה

באמצעות קשר חדש בין Occurance לבין הקבצה על Broadcasted\_By.

4. (6 נק') כתבו ב-ODL interface עבור VOD\_Order. על כל הקשרים להיות דו-כיווניים.

```
interface VOD_Order (key (Customer, Occurance)) {  
    attribute integer tvsNum;  
    relationship Customer customer inverse Customer::vod_Order;  
    relationship Occurance occurance inverse Occurance::vod_Order;  
};
```

## שאלה 2 – שפות שאילתה (36 נק')

- הנחיות: אין להשתמש באף פונקציה SQL שלא נלמדה בהרצאות או בתרגול בקורס. בפרט אין להשתמש בפונקציות ייחודיות ל-SQL 3.0.
- אין ליצור מבטים אלא רק במקרים שבהם הדבר הותר במפורש

1. (8 נקודות) נתונה הטבלה T במסד נתונים רלציוני SQL המוגדרת כך:

```
CREATE TABLE t (  
    val integer NOT NULL  
);
```

אנו מעוניינים למצוא את החציון של הערכים בטבלה. הוגדר מפתח על הערך val, כך שמובטח לנו מראש שכל הערכים שונים זה מזה.

תזכורת – החציון של קבוצת מספרים בת  $2n$  איברים הוא הממוצע של האיבר ה- $n$  והאיבר ה- $n+1$  בגודלם בקבוצה. החציון של קבוצת מספרים בת  $2n-1$  איברים הוא פשוט האיבר ה- $n$  בגודלו. כתבו שאילתת SQL המוצאת את החציון.

```
SELECT AVG(mid.val)  
FROM  
(SELECT t.val FROM t, t AS t2, (SELECT COUNT(*) AS cnt FROM t) AS t3  
WHERE t.val < t2.val GROUP BY t3.cnt, t.val HAVING COUNT(t2.val) BETWEEN  
t3.cnt/2 AND (t3.cnt+1)/2) AS mid
```

עבור השאלות הבאות, להלן סכמה אפשרית לחלק ממסד הנתונים (כולל סדר המשתנים עבור DRC). סכמה זו אינה בהכרח תואמת לזו המתקבלת בתרגום ישיר מה-ERD בשאלה 1.

Customers(CustNum, CustName)  
VOD\_order(CustNum, ProgNum, Date, TVlim)  
Broadcast\_by(ChNum, Date, ProgNum)  
Subscribed (CustNum, PkgNum)  
Contains(PkgNum, ChNum)  
Content\_Package(PkgNum)

2. (8 נקודות) כתבו שאילתת SQL המחזירה עבור כל חבילה רלוונטית את מספרה וכמה לקוחות רשומים לחבילה זו בלבד. חבילה רלוונטית הינה חבילה אשר קיימים לקוחות הרשומים לחבילה זו בלבד.

```
SELECT T1.PkgNumber, COUNT(T1.cnt, 0) FROM  
Subscribed AS S1  
WHERE NOT EXISTS (SELECT 1 FROM Subscribed AS S2 WHERE  
S1.CustNumber=S2.CustNumber AND S1.PkgNumber<>S2.PkgNumber)  
GROUP BY T1.PkgNumber
```

3. (13 נק') (a)

מה מחזירה השאילתה הבאה:

$$\left\langle PkgNumber \mid \left( \neg \exists chnlNumber (Contains(PkgNumber, chnlNumber)) \right. \right. \\ \left. \rightarrow \exists CustNumber (Subscribed(CustNumber, PkgNumber)) \right) \\ \wedge Content\_Package(PkgNumber) \left. \right\rangle$$

השאילתא מחזירה את כל חבילות התוכן שיש בהן ערוצים או שיש להן מנויים

(b) (3 נק') האם השאילתה ב- safe DRC? נמקו.

השאילתה אינה ב-safe DRC מכיוון וגרירה אינה מקיימת את חוקי safe DRC

(c) (5 נק') האם השאילתה בלתי תלויה בתחום? הוכיחו.

$$\left\langle PkgNumber \mid (\neg \exists chnlNumber (Contains(PkgNumber, chnlNumber))) \rightarrow \exists CustNumber (Subscribed(CustNumber, PkgNumber))) \wedge Content\_Package(PkgNumber) \right\rangle$$

שקול ל:

$$\left\langle PkgNumber \mid (\exists chnlNumber (Contains(PkgNumber, chnlNumber))) \vee \exists CustNumber (Subscribed(CustNumber, PkgNumber))) \wedge Content\_Package(PkgNumber) \right\rangle$$

שהוא ביטוי בטוח:

יש קשר "וגם" בין ביטוי "או" בטוח עם אותם משתנים חופשיים ובין ביטוי בטוח

4. (7 נק') כתבו שאילתא ב-RA המחזירה את כל הלקוחות שהזמינו רק פרקים שמשודרים בערוצים אליהם הם מנויים.

$$\pi_{custnum}(\text{customers}) / \pi_{custnum}(\pi_{custnum, prognum}(\text{vodorder}) / \pi_{custnum, prognum}(\text{subscribed} \bowtie \text{contains} \bowtie \text{broadcastby}))$$

### שאלה 3 – תלויות ופירוקים (23 נק')

1. (15 נק') נתונה הסכמה הבאה  $R(A, B, C, D, E)$  וקבוצת התלויות  $F = \{A \rightarrow BE, AC \rightarrow D, D \rightarrow BA, B \rightarrow C\}$

a. (7 נק') בהנתן הפירוק  $R_1(A, B), R_2(A, C, D, E)$ , האם הפירוק משמר מידע? האם משמר תלויות? נמקו.

הפירוק משמר מידע: החיתוך A הוא מפתח עבור  $R_1$ .

הפירוק לא משמר תלויות:  $B \rightarrow C$  אינו נשמר (הרצה של האלגוריתם מהתרגול על תלות זו לא תעבור אפילו איטרציה אחת).



b. (8 נק') האם הסכמה המקורית (ללא הפירוק מהסעיף הקודם) ב-BCNF? אם כן הוכיחו, אם לא פרקו את הסכמה המקורית פירוק משמר מידע לסכמות ב-BCNF.

הסכימה לא ב-BCNF בגלל התלות  $B \rightarrow C$ .

נפרק לפיה ל-BC ול-ABDE. הראשונה ב-BCNF כי יש לה רק שתי תכונות (ואין ב-F תלויות מהקבוצה הריקה), והשניה ב-BCNF בגלל ש-A ו-D הן מפתחות (אפילו עבור כל R), ואין תלות לא טריביאלית מאף קבוצה המוכלת ב-BE (מספיק לבדוק ל-BE עצמה).

2. (8 נק') נניח ש-X,Y,Z הוא פירוק של רלציה R[U] לשלוש רלציות שמשמר מידע. האם נובע מכך בהכרח ש-X,Y הוא פירוק משמר מידע של איחוד שתי הקבוצות הנ"ל? הוכח או תן דוגמא נגדית.

לא. לדוגמא עבור הפירוק  $\rho = \{R_1(A,C), R_2(B,D), R_3(A,B)\}$  וקבוצת התלויות  $F = \{A \rightarrow C, B \rightarrow D\}$  האיחוד של  $R_1$  ו- $R_2$  לא משמר מידע (חיתוכן ריק). לעומת זאת הפירוק המלא ל-3 רלציות משמר מידע כי הוא משמר תלויות (קל לראות) וכולל את המפתח AB.

## שאלה 4 – XML (21 נק')

נתונה הסכימה הבאה עבור מפעל, כאשר יש להניח שצומת המסמך factory הוא בן יחיד של השורש. הסכימה מתארת את החלקים part המורכבים במפעל. לכל חלק מתוארים או תתי החלקים המרכיבים אותו, או הספק supplier ממנו הוא נקנה (במקרה שחלק זה עצמו כבר לא מורכב במפעל).

<!ELEMENT factory (part)+>

<!ELEMENT part (part+ |supplier)>

<!ATTLIST part

|    |    |           |
|----|----|-----------|
| id | ID | #REQUIRED |
|----|----|-----------|

|         |       |           |
|---------|-------|-----------|
| quality | CDATA | #REQUIRED |
|---------|-------|-----------|

>

<!ELEMENT supplier EMPTY>

<!ATTLIST supplier

|      |    |           |
|------|----|-----------|
| name | ID | #REQUIRED |
|------|----|-----------|

|       |       |           |
|-------|-------|-----------|
| phone | CDATA | #REQUIRED |
|-------|-------|-----------|

>

1. (6 נק') כתבו סכימה חדשה שבה לא תהיה כפילות בשמירת פרטי הספקים, כאשר אלו יכולים לספק יותר מחלק אחד למפעל. עדיין יש לשמור על ההיררכיה של חלקים המרכיבים חלקים. לנוחותכם, ה-DTD המקורי כתוב בחלון התשובה, ועליכם למחוק ממנו ולהוסיף אליו בהתאם.

(חלון התשובה נמצא בעמוד הבא)

~~<!ELEMENT factory (part)+>~~  
<!ELEMENT factory (part+,supplier+)>

~~<!ELEMENT part (part+|supplier)>~~  
<!ELEMENT part (part|suppliedpart)+>

<!ATTLIST part

|    |    |           |
|----|----|-----------|
| id | ID | #REQUIRED |
|----|----|-----------|

|         |       |           |
|---------|-------|-----------|
| quality | CDATA | #REQUIRED |
|---------|-------|-----------|

>

<!ELEMENT supplier EMPTY>

<!ATTLIST supplier

|      |    |           |
|------|----|-----------|
| name | ID | #REQUIRED |
|------|----|-----------|

|       |       |           |
|-------|-------|-----------|
| phone | CDATA | #REQUIRED |
|-------|-------|-----------|

>

<!ELEMENT suppliedpart EMPTY>

<!ATTLIST suppliedpart

|    |    |           |
|----|----|-----------|
| id | ID | #REQUIRED |
|----|----|-----------|

|          |       |           |
|----------|-------|-----------|
| supplier | IDREF | #REQUIRED |
|----------|-------|-----------|

|         |       |           |
|---------|-------|-----------|
| quality | CDATA | #REQUIRED |
|---------|-------|-----------|

>

2. (7 נק') נקבע כי איכות של כל חלק מוצגת כערך מספרי. כתבו שאילתת XPath 1.0 המחזירה את צומת הספק אצלו מוזמן החלק שאיכותו היא הגדולה ביותר (quality) מבין החלקים המסופקים על-ידי הספקים. במידה ויש מספר חלקים בעלי אותה איכות, יש להחזיר את הראשון ביניהם. עשו זאת עבור ה-DTD המקורי:

```
/factory//part/supplier[not (../@quality < //part/supplier/../@quality)][1]
```

3. (8 נק') חלק המורכב במפעל יקרא מורכב היטב אם איכותו (לפי quality, נניח שערכי תכונות שלו הן תמיד מספרים אי-שליליים) גדולה או שווה לסכום האיכויות של המוצרים המרכיבים אותו, וכן כל אלו גם מורכבים היטב בעצמם (חלק שאינו מורכב במפעל הוא תמיד מורכב היטב). חלק יקרא ראשי אם הוא לא תת-חלק של חלק אחר (ז"א אם הוא בן של הצומת factory). כתבו שאילתת XPath 1.0 או שאילתת XQuery 1.0 המחזירה את כל החלקים הראשיים המורכבים היטב. עשו זאת עבור ה-DTD המקורי. ציינו באיזו שפת שאילתא השתמשתם.

תשובה קצרה יחסית ניתנת להיכתב ב-XPath 1.0:

```
/factory/part[count(../descendant-or-self::part)=count(../descendant-or-self::part[@quality>=sum(part/@quality)])]
```

יש גם שיטות אחרות (ב-XPath או XQuery) להשגת התוצאה המבוקשת.