

מרצה : פרופ' רועי פרידמן

סמסטר חורף התשע"ה

מתרגלים : מר רוני ליכר

גב' הדר פרנקל

מערכות מסדי נתונים

236363

מועד ב'

(ט"ו באדר התשע"ה, 6 במרץ 2015)

מס' סטודנט:

פירוט השאלות והניקוד:

| מס' | נושא | ניקוד |
|------|------------------------------|-------|
| 1 | ERD | 25 |
| 2 | שאליות מידע | 25 |
| 3 | פירוקים ותלויות פונקציונליות | 18 |
| 4 | XML | 20 |
| 5 | NoSQL | 12 |
| סה"כ | | 100 |

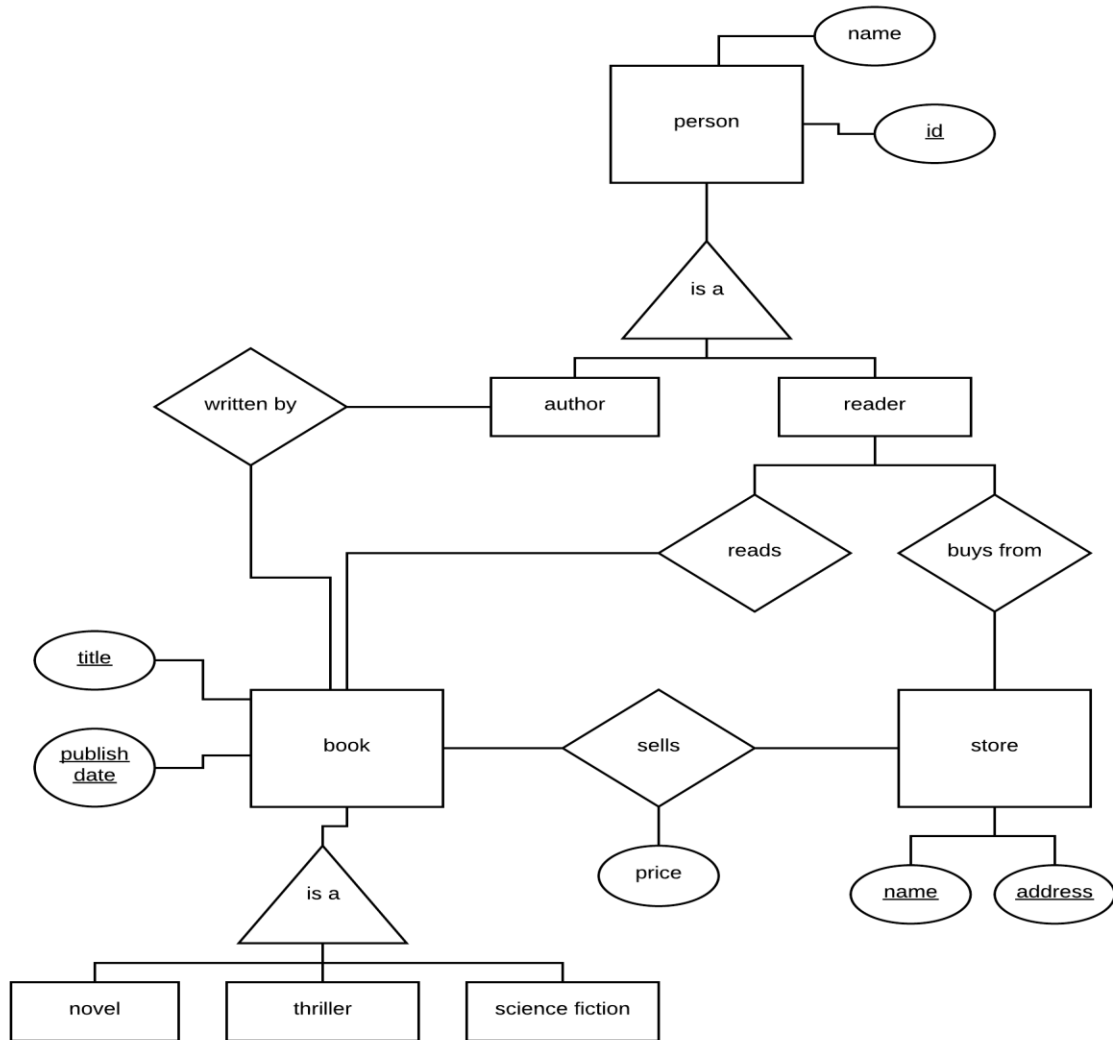
הנחיות לנבחנים

1. כתבו את התשובות אך ורק בטופס הבחינה, המחברת מיועדת לטייטה בלבד.
2. מותר ומומלץ לכתוב את התשובות בעפרון.
3. בדף האחרון יש מקום נוסף לתשובות. אם צריך מקום נוסף לתשובות, השתמשו במקום זה תוך ציון הדבר ליד השאלה המקורית.
4. כל חומר עזר כתוב על נייר בלבד מותר בשימוש.
5. אין לקבל או להעביר חומר עזר כלשהו בזמן הבחינה.
6. בבחינה חמש שאלות ללא בחירה. יש לענות עליהן במלואן.
7. בכל מקום שלא נאמר אחרת, יש לנמק את התשובות בקצרה. **תשובות לא מנומקות לא תתקבלנה, למעט במקומות שבהם אתם מתבקשים לכתוב שאלות.**
8. יש להשתמש רק בסימנים או פונקציות שנלמדו בתרגול או בהרצאה או שמופיעות בשקפים של הקורס. כל שימוש בסימון שאינו כזה מחייב הסבר מלא של משמעות הסימון.
9. משך הבחינה שלוש שעות. תכננו את הזמן בהתאם. **לא תינתנה הארכות זמן במהלך המבחן.**
10. הבחינה כוללת 8 דפים (כולל דף זה), **בהם 15 עמודים**. נא לוודא שיש בידכם את כל הטופס.
11. כאשר ניקוד תתי הסעיפים אינו מצוין, ניקוד הסעיף מתחלק שווה ביניהם.
12. הניקוד אינו נועד לשקף את קושי השאלה ולכן מומלץ לקרוא קודם את כל השאלות.

בהצלחה

שאלה 1 - ERD (25 נקודות):

נתונה הדיאגרמה הבאה, המתארת אנשים (person) שיכולים להיות סופרים (author) או קוראים (reader). הספרים (book) נקראים (reads) ע"י קוראים או נכתבים (written by) ע"י סופרים. בנוסף קיימות חנויות (store) המוכרות את הספרים. ישנם ספרים מסוג מדע בדיוני (science fiction), מותחנים (thriller) ורומנים (novel).



א. 8 נק' - ענו על השאלות הבאות על פי דיאגרמת ה-ERD, יש לנמק בקצרה:
1. האם יתכן שאותה חנות תמכור את אותו הספר במחירים שונים?

לא, המפתחות של היחס sells הם name, address, title, publish date ולכן הם קובעים גם את המחיר של הספר.

2. האם יתכנו שתי חנויות המוכרות את אותם ספרים באותו מחיר?

כן, מה שקובע את המחיר זה הצירוף של חנות וספר.

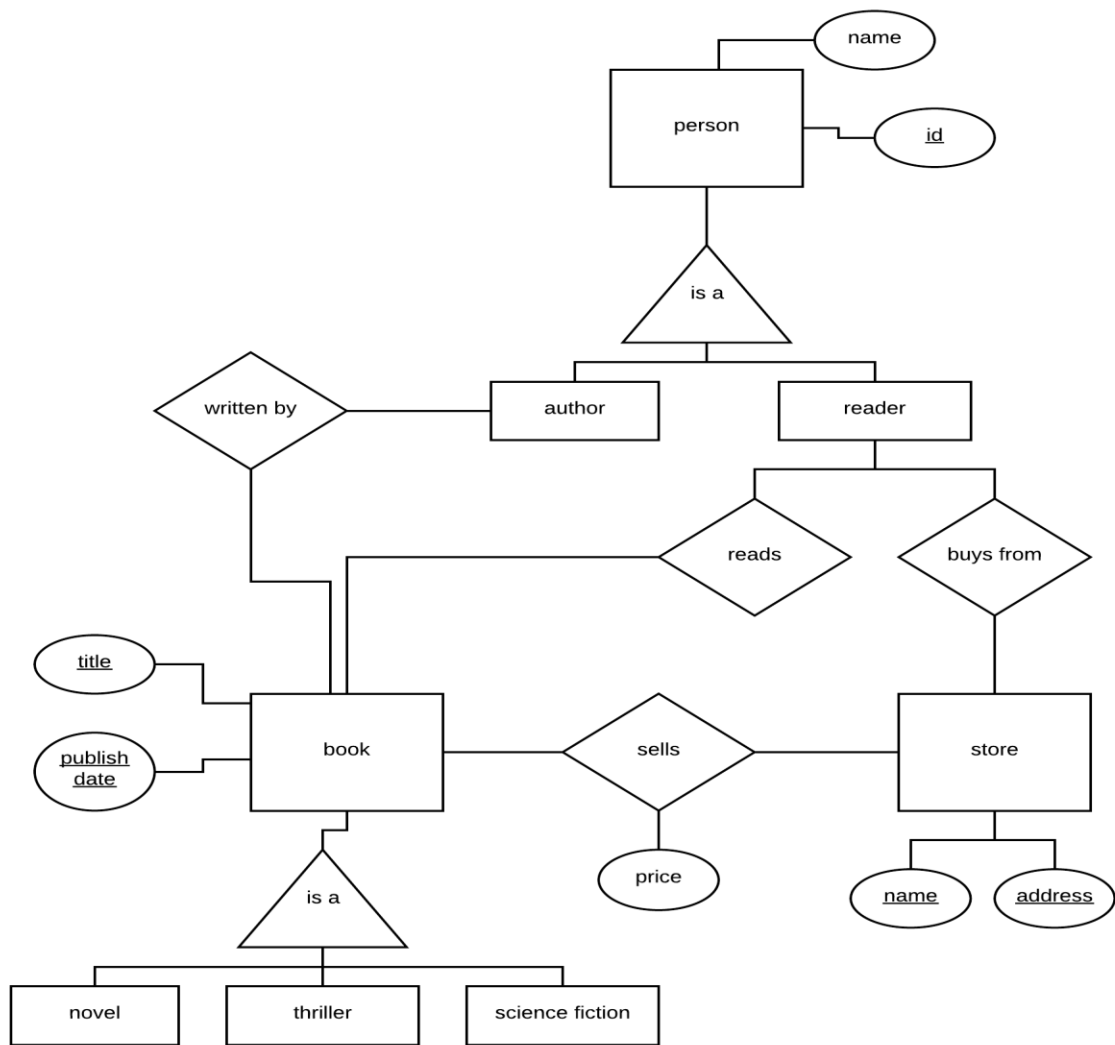
3. האם יתכן ספר שהוא גם מדע בדיוני וגם מותחן?

לא, לפי קשר הירושה היחיד כל ספר הוא לכל היותר אחד מהשלושה.

4. כמה ספרים יכולים להיכתב על ידי אותו סופר?

אין הגבלה על היחס בין מספר הספרים למספר הסופרים, קיים קישור רבים לרבים.

ב. 10 נק' - שנו את הדיאגרמה כך שכל ספר יכתב על ידי לפחות סופר אחד, סופר יכול להיות גם קורא, וספר חייב להיות מאחד משלושת הסוגים המצויינים.



פתרון:

נוסיף קו כפול בין **book** ל**written by**, נפריד את משולש היחשה של **person** לשני משולשים ונוסיף קו עבה בין **book** לשלושת היורשים שלו

ג. 7 נק' - כתבו את תרגום הדיאגרמה לטבלאות המכילות את התכונות והמפתחות של כל יחס וישות, עבור הגורמים הבאים: novel, reader, person, reads, sells.
יש לסמן בכל סכמה את המפתחות בקו תחתון.

novel (title, publish date)

reader (id)

person(id, name)

reads (title, publish date, id)

sells (name, address, title, publish date, price)

שאלה 2 - שאלות מידע (25 נקודות):

הסעיפים בשאלה זו מתייחסים לדיאגרמת ה-ERD שניתנה בשאלה 1.

יש לענות על השאלות בהתאם לתרגום הישיר של הטבלאות כפי שנלמד בכיתה.

א. 12 נק' - כתבו שאלת SQL המוצאת את כל הקוראים (שם ותעודת זהות) שקראו לפחות ספר אחד שנכתב ע"י סופר אחד בלבד.

```
select name, reads.id  
from reads, person  
where person.id = reader.id and (title, publishDate) in  
  (select title, publishDate  
   from writtenBy  
   group by title, publishDate  
   having count(*) = 1)
```

ב. 8 נק' - כתבו שאילתת DRC המוצאת את כל הספרים כך שלפחות אחד מהסופרים שכתב אותם כתב עוד ספר אחד לפחות.

$$\{ \langle title, pDate \rangle \mid book(title, pDate) \wedge \exists aid (writtenBy(title, pDate, aid) \wedge \exists t_2, pd_2 (book(t_2, pd_2) \wedge (t_2 \neq title \vee pd_2 \neq pDate) \wedge writtenBy(t_2, pd_2, aid))) \}$$

ג. 5 נק' - בסעיף זה עליכם להתייחס לדיאגרמה המתוקנת, משאלה 1ב. כתבו שאילתת RA המחזירה את שמות כל הספרים שנכתבו ע"י לפחות סופר אחד שמקיים את התנאים הבאים:

- הסופר לא קרא אף ספר.
- המחיר המקסימלי בו נמכר ספר שלו הוא מעל 50 ₪.

$$\pi_{title} ((writtenBy - (writtenBy \bowtie reads)) \bowtie \pi_{id} (writtenBy \bowtie \sigma_{price > 50} sells))$$

שאלה 3 – פירוקים ותלויות פונקציונליות (18 נקודות):

א. 6 נק' - נתונה סכמה R וקבוצת תלויות פונקציונליות F. תהי $X \rightarrow A$ תלות ב F. הוכיחו או תנו דוגמא נגדית:

אם X מפתח ב R אזי תת הסכמה XA היא ב 3NF

דוגמא נגדית:

$$R = R_1 = ABC$$

$$F = \{BC \rightarrow A, C \rightarrow A\}$$

BC הינו מפתח ולכן R אכן מהצורה הרצויה, אך התלות $C \rightarrow A$ סותרת את תנאי ה 3NF שכן A אינו חלק ממפתח קביל ו C אינו מפתח.

ב. 6 נק' - נתונה סכמה R וקבוצת תלויות F. נתון פירוק R_1, R_2 של R. הוכיחו או תנו דוגמא נגדית:

נניח ש R הוא ב BCNF. אזי שתי תת הסכמות R_1, R_2 גם הן ב BCNF

הטענה נכונה. הוכחה: נניח ש F אינה ריקה (שכן אם F ריקה הטענה נכונה תמיד). תהי תלות פונקציונלית $f = X \rightarrow Y$ בהיטל של F על אחת מתת הסכמות, בה"כ על R_1 . אזי $f \in F^+$ ולכן לפי הנחת ה BCNF Y מוכל ב X או X הוא מפתח על של R. אם Y מוכל ב X, זה מתקיים גם ב R_1 . לכן נניח כי Y לא מוכל ב X. מכאן, X הוא מפתח על של R. בפרט, $X \rightarrow R_1$ כנדרש.

ג. 6 נק' - נתונה סכמה $R = \{ABCD\}$ ופירוק שלה לשתי תת סכמות: $R_1 = \{AB\}$, $R_2 = \{BCD\}$. מצאו סט תלויות פונקציונליות כך שהפירוק הנ"ל משמר מידע אך לא משמר תלויות (והוכיחו את תשובתכם), או הוכיחו שלא יתכן סט כזה.

$$F = \{B \rightarrow CD, A \rightarrow C\}$$

הפירוק משמר מידע שכן החיתוך בין שתי תת הסכמות הוא B, וזהו מפתח ל R_2 . מנגד, הפירוק אינו משמר תלויות שכן: התלות $B \rightarrow CD$ נשמרת ע"י R_2 אך מהרצת האלגוריתם לבדיקת שימור תלויות נגלה שהתלות השנייה אינה נשמרת.

שאלה 4 – XML (20 נקודות):

נתון ה- DTD הבא המתאר מאגר סרטים עם צומת מסמך mdb.

```
<!ELEMENT mdb (movie*)>
<!ELEMENT movie (title,part*)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT part (title,(part+|cut))>
<!ELEMENT cut (shot,time,duration)>
<!ELEMENT shot EMPTY>
<!ATTLIST shot number ID #REQUIRED>
<!ELEMENT time (#PCDATA)>
<!ELEMENT duration (#PCDATA)>
```

א. 4 נק' כתבו מסמך XML הנאכף ע"י חוקי ה DTD הנ"ל. על כל אלמנט מוצהר להופיע ב DTD

```
<?xml version="1.0">
<mdb>
  <movie>
    <title>
      Lord of the Rings
    </title>
    <part>
      <title> The Fellowship of the Ring </title>
      <cut>
        <shot number = "a17" />
        <time> 01:30 </time>
        <duration> 12:01 </duration>
      </cut>
    </part>
  </movie>
</mdb>
```

ב. 6 נק' – ענו על השאלות הבאות לפי מסמך ה DTD
1. כמה חלקים (part) יכולים להיות לכל סרט (movie)?

לכל סרט יכולים להיות כל מספר של חלקים, או 0

2. מהם הבנים של part? הסבר

ל part יש כותרת, וכמו כן יש לו או: part אחד או יותר, או: קטע אחד

3. כמה צילומים (shot) בעלי אותו מספר (number) יכולים להיות? נמקו

יכול להיות לכל היותר צילום אחד עם אותו מספר שכן זהו שדה ID

ג. 10 נק' – כתבו שאילתת XPath2.0 המחזירה את כל שמות (title) הסרטים שיש להם לפחות שני חלקים (part) שהם בנים ישירים של הסרט שלכל אחד מהם יש בן מסוג קטע (cut). בנוסף על הסרטים המוחזרים לקיים שהאורך (duration) של הקטע בחלק הראשון הוא המקסימלי מבין כל הקטעים של כל החלקים הנ"ל.

```
/mdb/movie[count(/part[cut]) >1 and every $d in ./part/cut/duration satisfies  
number($d)<= number(/part[1]/cut/duration)]/title
```

שאלה 5 – NoSQL (12 נקודות):

נתון מסד הנתונים הגרפי עבור עץ תורשה המכיל את הצמתים הבאים:

| |
|---------------|
| :PERSON |
| Name |
| Date-of-Birth |
| Gender |

צמתים אלו יכולים להיות מקושרים ע"י שלושה סוגי קשרים (relationship):

ל-MARRIED: יש תכונה שנקראת duration שערכה הוא משך הזמן שהנ"ל נשואים (או היו נשואים אם כבר נפטרו). הניחו שכל זוג נשוי מחובר בקשת אחת בכיוון שרירותי כלשהו בין בני הזוג.

לקשרים FATHER: ו-MOTHER: אין תכונות. הקשרים מכוונים מהילדים אל ההורים.

נניח שמסד הנתונים תקין ושכל צומת יכול להיות מקושר לאב אחד לכל היותר ולאם אחת לכל היותר.

שימו לב: בשאלה זו עלולות לרדת נקודות על תשובות מסורבלות.

א. 6 נק' - כתבו שאילתא ב-Cypher המוצאת את כל הזוגות (ללא כפילויות) של אנשים שיש להם אב קדמון משותף או אם קדמונית משותפת והם נמצאים באותה רמה של תת העץ המתחיל באב או באם הקדמונים.

```
MATCH (a:Person)-[:FATHER|MOTHER*]->(f:Person)<-[:FATHER|MOTHER*]-(b:Person)
WHERE id(a) < id(b)

And length(shortestPath((a)-[:FATHER|MOTHER*]->(f))) =
length(shortestPath((b)-[:FATHER|MOTHER*]->(f)))

RETURN a, b
```

ב. 6נק' - כתבו שאילתא המחזירה את כל הצמתים כך שכל הוריהם הקדמונים הרשומים במסד הנתונים נשואים (או היו נשואים) מעל 15 שנים.

הפתרון הבא אינו נכון שכן אי אפשר לסנן תכונה ע"פ ערך שאינו קבוע.

```
MATCH (a:Person)-[:FATHER|:MOTHER*]->(f:Person)
WITH a, collect(f) as ancestors
WHERE ALL (f IN ancestors WHERE (f)-[:MARRIED]{duration  $\geq$  15} )-( )
RETURN a
```

פתרון אפשרי א:

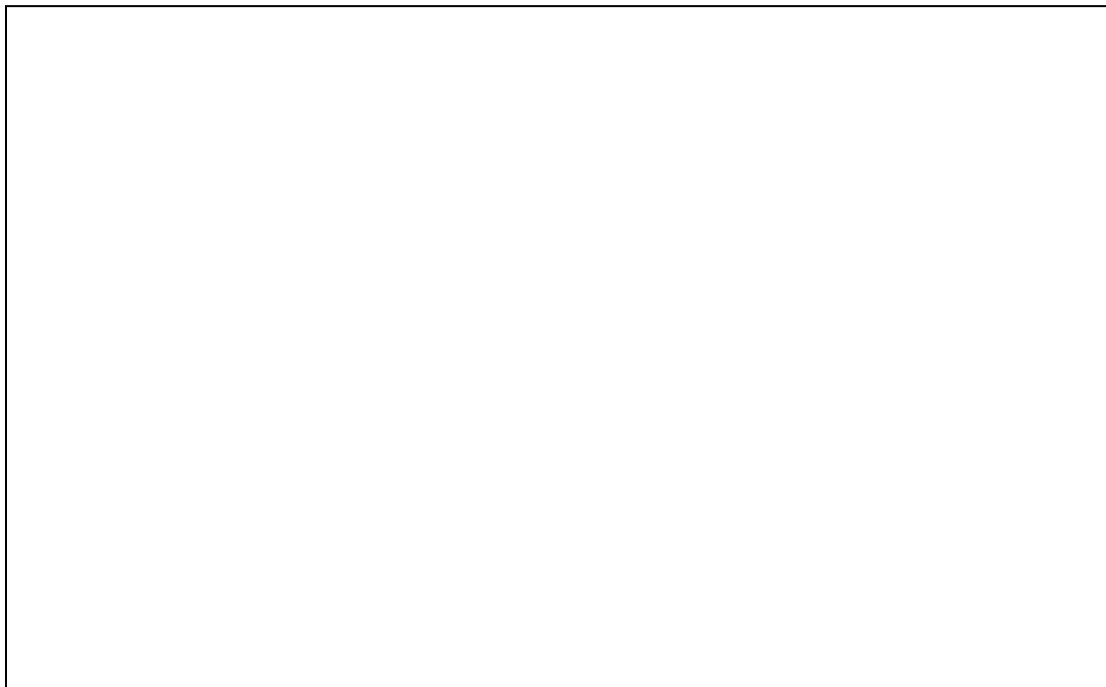
```
MATCH      (a:Person)-[:FATHER|:MOTHER*]->(f1:Person),
            (a:Person)-[:FATHER|:MOTHER*]->(f2:Person)-[r:MARRIED]-()
WITH        a ,
            count(f1) as num_ancestors,
            count(f2) as num_marriedAncestors,
            collect(r) as merriedRs
WHERE       num_ancestors = num_marriedAncestors
            AND
            ALL (r IN merriedRs WHERE r.duration >= 15 )
RETURN      a
```

פתרון אפשרי ב:

```
MATCH      (a:Person)-[:FATHER|:MOTHER*]->(f:Person)-[p:MARRIED*0..1]-(:Person)
WITH        a,
            collect(p) AS ancestors
WHERE       ALL (p IN ancestors
                WHERE length(p) = 1
                AND
                ALL (r IN p
                    WHERE r.duration >= 15
                    )
                )
RETURN      a
```

אם אתם משתמשים בדף זה, ציינו זאת ליד השאלה/השאלות המקוריות, וציינו כאן את מספר/י השאלה/השאלות.

שאלה: _____ סעיף: _____



שאלה: _____ סעיף: _____



שאלה: _____ סעיף: _____

