

מרצה : פרופ' רועי פרידמן

סמסטר חורף התשע"ה

מתרגלים :

מר רוני ליכר

גב' הדר פרנקל

מערכות מסדי נתונים

236363

מועד א'

(י"ט בשבט התשע"ה, 8 בפברואר 2015)

מס' סטודנט:

פירוט השאלות והניקוד:

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

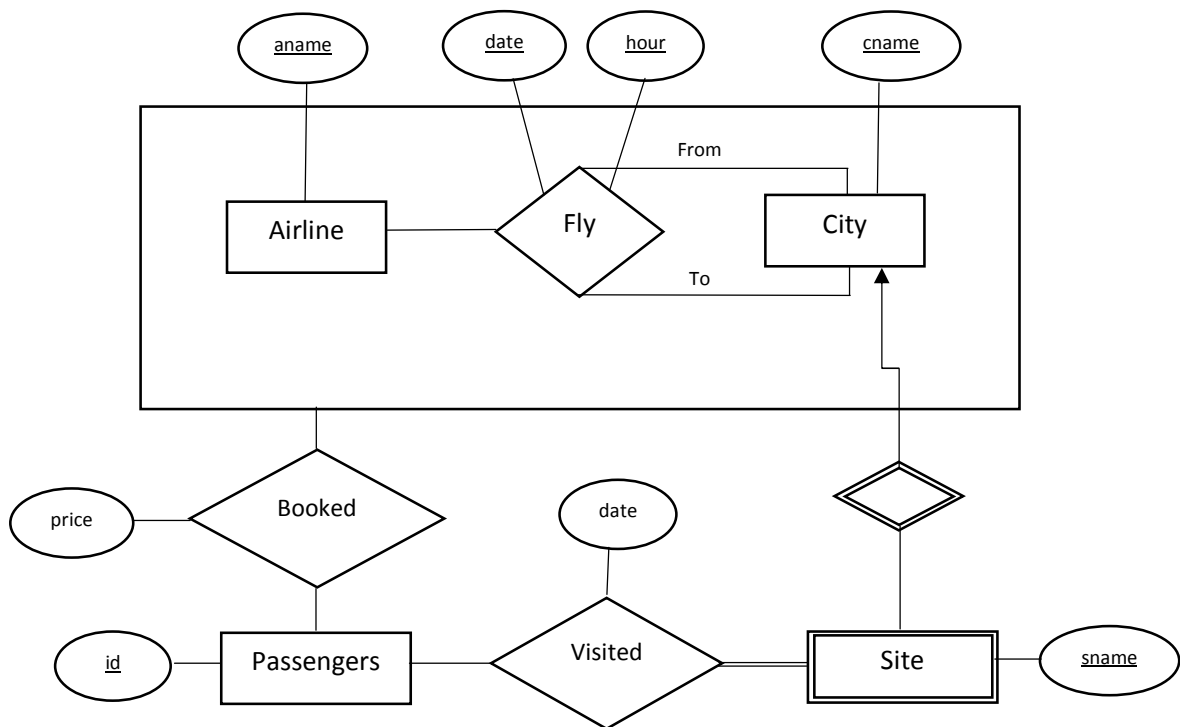
הנחיות לנבחנים

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 שאלתה.**
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):

ה-ERD הבא מייצג מסד נתונים של חברות תעופה (Airline) וערים (City) אליהן ומן הן טסות (Fly). בנוסף נוסעים (Passengers) שמזמינים כרטיסי טיסה (Booked). נוסעים גם מבקרים (Visited) באתרים (Site).



א. 6 נק' - השלימו את הטבלאות הבאות המתקבלות מתרגום ישיר של ה-ERD.

א.

שם הטבלה	Visited
שדות	
מפתחות	
תלויות פונקציונליות	

א.

שם הטבלה	Booked
שדות	
מפתחות	
תלויות פונקציונליות	

ב. 6 נק' - ענו על השאלות הבאות על פי דיאגרמת ה-ERD, יש לנמק בקצרה:
1. האם נוסעים יכולים לבקר באתרים הנמצאים בערים עבור הנוסעים לא רכשו כרטיס טיסה?

כן, כל נוסע יכול לבקר בכל אתר.

2. נתון אתר במסד הנתונים בעיר חיפה, האם יתכן שאף לקוח לא טס לחיפה?

כן, בהכרח מישו ביקר באתר אך אין הכרח שמישהו טס אליו.

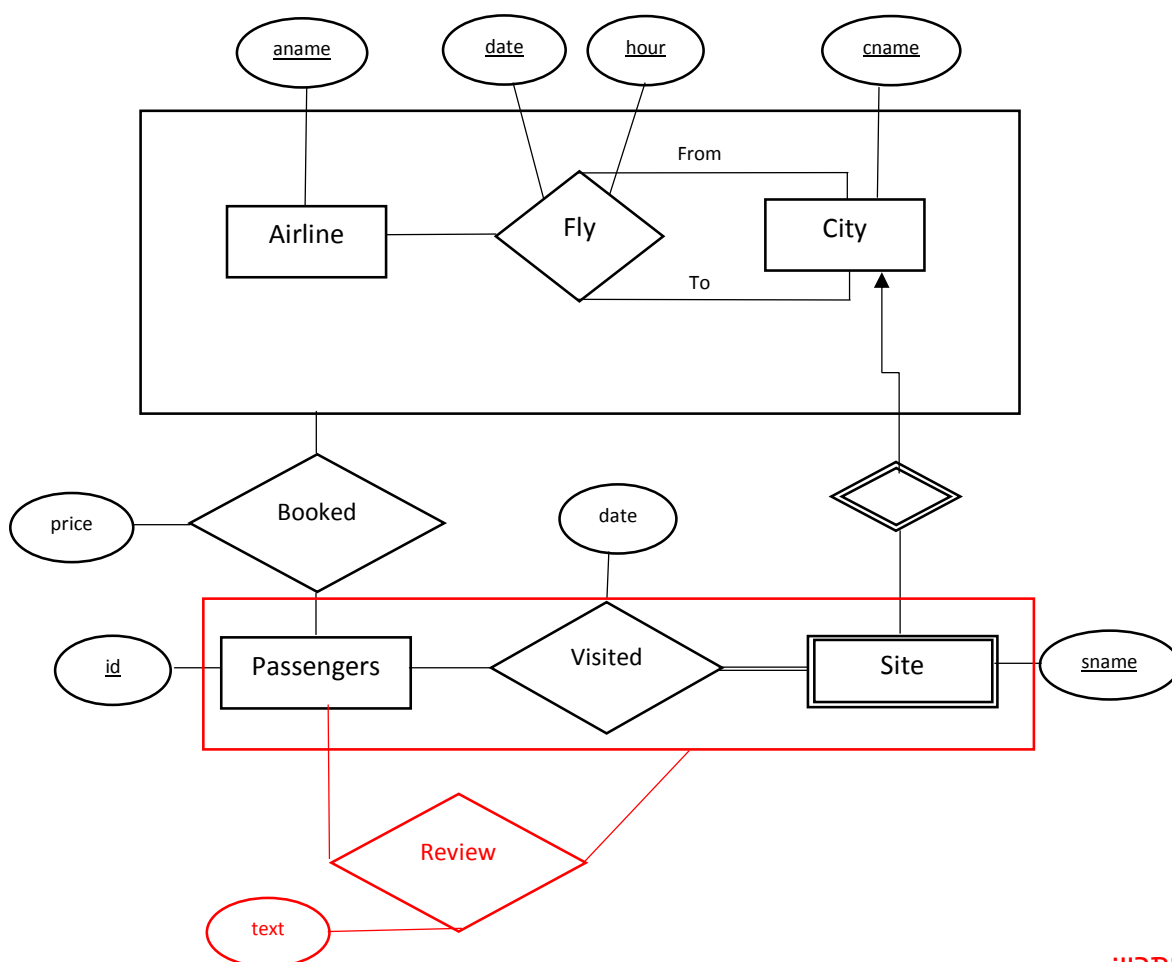
3. נתון שקיימת טיסה של חברת "אם-עם" מישראל לארה"ב, האם מובטח שקיימת טיסה מארה"ב לישראל עם אותה החברה?

לא, המידע הנשמר הינו עבור כיוון אחד.

ב-2 הסעיפים הבאים יש לתאר איך יש לשנות את ה-ERD כדי לתמוך בדרישות הסעיף. אין קשר בין הסעיפים ובשניהם יש להתייחס לדיאגרמה המקורית.

ג. 7 נק' - יש להוסיף תמיכה ביכולת הוספת ביקורות של נוסעים על ביקור נוסע באתר מסוים בתאריך מסוים.

- יתכן ונוסע יוסיף ביקורות על ביקור באתר שהתבצע על ידי נוסע אחר.
- כל נוסע יכול לשמור במסד הנתונים ביקורת אחת לכל היותר לכל ביקור.



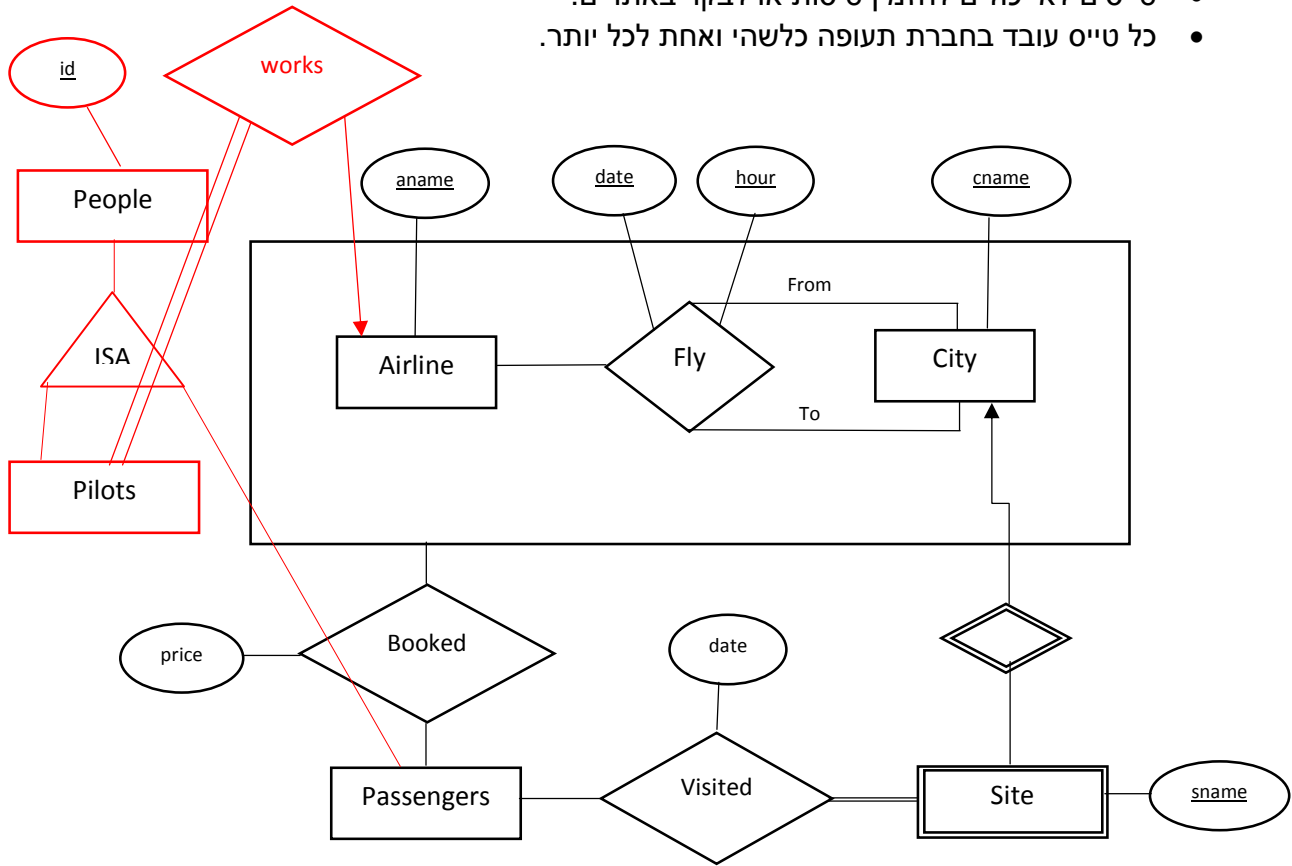
פתרון:

קיבוץ על הקשר Visited וחיבור של הקיבוץ והנוסעים ע"י קשר.

פתרון בעזרת חץ מ-Review אל ההקבצה שגוי שכן מאפשר לכל נוסע לכתוב ביקורת אחת בלבד.

ד. 6 נק' - יש להוסיף תמיכה ביכולת ניהול טייסים.

- אנשים יכולים להיות או טייסים או נוסעים אך לא גם וגם.
- טייסים לא יכולים להזמין טיסות או לבקר באתרים.
- כל טייס עובד בחברת תעופה כלשהי ואחת לכל יותר.



פתרון:

פיצול הנוסעים עם ISA וחיבור נכון עם קו כפול וחץ.

שאלה 2 - שאילות מידע (25 נקודות):

השאלות בסעיף זה מתייחסות לדיאגרמת ה-ERD שניתנה בשאלה 1.

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

א. 12 נק' - חברת התעופה "אם-עם" (באנגלית "if-with") מעוניינת להקים מועדון לקוחות "הנוסע המקפיד". כל מי שטס בחברה לפחות 10 פעמים יהיה חבר בו באופן אוטומטי. חברי המועדון שטסו מעל הממוצע מבין חברי המועדון וגם ביקרו באתר המבקרים של הטכניון יקבלו מעמד זהב.

כתבו שאילתת SQL שמחזירה:

לקוחות במעמד זהב במועדון "הנוסע המקפיד" ואת מספר הטיסות שלהם.

```
CREATE VIEW MAKPID as
SELECT id, count(*) as counter
FROM Booked
WHERE aname = "if-with"
GROUP BY id
HAVING count(*) > 90

SELECT M1.id, counter
FROM MAKPID M1, Visited V
WHERE counter > (SELECT AVG(counter)
                  FROM MAKPID)
AND M1.id = V.id
AND sname = "Technion"
```

ב. 5 נק' - כתבו במילים מה שאילתת ה-RA הבאה מחזירה:

$$\pi_{passenger_id1}(\sigma_{from='Israel'}Booked \bowtie \begin{array}{l} to1=from2 \\ passenger_id1=passenger_id2 \\ (date1<date2) \vee (date1=date2 \wedge hour1 \leq hour2) \end{array} \sigma_{to='Romania'}Booked \bowtie \begin{array}{l} date1 \leq date3 \wedge \\ date3 \leq date2 \wedge \\ to1=city \wedge \\ passenger_id1=passenger_id3 \end{array} Visited)$$

פתרון: השאילתה המחזירה נוסעים שטסו מישראל לרומניה דרך מדינה שלישית בה ביקרו באתר כלשהו. מתקיים שהטיסה השניה התקיימה לאחר הטיסה הראשונה.

ג. 8 נק' - כתבו שאילתת DRC המחזירה זוגות של חברות תעופה שמפעילות את אותן הטיסות (מקור ויעד וללא חשיבות לתאריכים ושעות) חוץ מטיסה אחת שהחברה הראשונה מפעילה והשנייה לא מפעילה.

• לשם פשטות, הניחו שהטבלה fly מוגדרת כך: $fly(airline, route)$

```
{airline1, airline2 | airline(airline1) ∧ airline(airline2) ∧
(¬∃routeX((¬fly(airline1, routeX) ∧ fly(airline2, routeX)) ∧
∃routeY((fly(airline1, routeY) ∧ ¬fly(airline2, routeY) ∧
∀routeZ((fly(airline1, routeZ) ∧ ¬fly(airline2, routeZ) → routeY = routeZ)))
}
```

שאלה 3 – פירוקים ותלויות פונקציונליות (18 נקודות):

א. 9 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 $R_1 = XA$ היא ב BCNF

דוגמא נגדית:

$$R = R_1 = ABC$$

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

המפתחות הם BC, AB ולכן R אכן מהצורה הרצויה, אך התלות $A \rightarrow C$ סותרת את תנאי ה BCNF

ב. 9 נק' - נתונה הסכמה $R(A,B,C,D,E,F)$ וקבוצת תלויות פונקציונליות:

$$F = \{A \rightarrow BC, D \rightarrow E, B \rightarrow D, A \rightarrow E, F \rightarrow BC\}$$

ונתונים שני פירוקים:

$$\{R_1(A,B,C), R_2(D,E), R_3(B,D), R_4(F,B,C), R_5(F,A)\}$$

$$\{R_1(A,B,C), R_2(A,D,F), R_3(D,E)\}$$

ידוע שאחד מהפירוקים הנ"ל נוצר ע"י הרצת אלגוריתם לפירוק ל-BCNF שנלמד בכיתה והאחר נוצר ע"י אלגוריתם לפירוק ל-3NF הנלמד בכיתה. באיזה מהפירוקים הייתם מעדיפים להשתמש לסכמה בה יש עדכונים רבים בשדה D? נמקו את תשובתכם.

הפירוק הראשון הוא פירוק ל-3NF ואי לכך הפירוק השני הוא ל-BCNF. בפירוק השני התלות $B \rightarrow D$ לא נשמרת. היות ואנחנו מעדכנים הרבה תכונה בתלות שאינה נשמרת נעדיף את הפירוק הראשון שישמר תלויות.

שאלה 4 – XML (20 נקודות):

נתון ה- DTD הבא המתאר לוח מודעות עם צומת מסמך db:

```
<!ELEMENT db (user*, ad*)>
```

```
<!ELEMENT user (name, tel)>
```

```
<!ATTLIST user Id ID #REQUIRED>
```

```
<!ELEMENT ad (type, photo, date, review)>
```

```
<!ATTLIST ad publisher IDREF #REQUIRED>
```

```
<!ELEMENT name #PCDATA>
```

```
<!ELEMENT tel #PCDATA>
```

```
<!ELEMENT type #PCDATA>
```

```
<!ELEMENT photo #PCDATA>
```

```
<!ELEMENT date #PCDATA>
```

```
<!ELEMENT review #PCDATA>
```

```
<!ATTLIST review reviewer IDREF #REQUIRED>
```

הסבר: במאגר (db) יש משתמשים (user) ומודעות (ad). לכל מודעה יש סוג, תמונה, תאריך וביקורת.

א. 3 נק' - כתבו XML חוקי וקצר הנאכף ע"י ה DTD, על כל אלמנט מוצהר ב DTD להופיע ב- XML.

```
<?xml version="1.0">
<db>
  <user ID = "a10">
    <name> Yossi </name>
    <tel> *5555 </tel>
  </user>
  <ad publisher = "a10">
    <type> A </type>
    <photo> ☺ </photo>
    <date> 08/02/2015 <date>
    <review reviewer = "a10"> Nice! <review>
  </ad>
</db>
```

ב. 5 נק' - הוחלט לשנות את ה- DTD כך שכל מודעה תכיל 1 או יותר מחזורים של האלמנטים

הבאים ובסדר הבא:

1. סוג.
2. 2 תמונות.
3. 0 ביקורת או יותר.
4. תאריך אופציונלי.

פרטו את השינויים שיש לבצע:

יש לבצע שינוי של שורה אחת:

<!ELEMENT ad (type, photo, photo, review*,date?)+>

ג. 3 נק' - האם user יכול לפרסם את אותה המודעה פעמיים, הסבירו?

כן, במודעה אין תכונה מסוג ID ולכן אין מניעה מפרסום בדיוק אותה המודעה פעמיים.

ד. 9 נק' - כתוב שאילתת XPath2.0 אשר מחזירה זוגות של מודעות מסוג רכב שבינהן (מבחינת תאריך פרסום המודעה) התפרסמו רק מודעות מסוג נדל"ן.

* הניחו כי date היא מחרוזת המציינת מספר וניתן להשוות אותו עם הכמתים הרגילים ($\dots, =, <, >$)
* אין להניח דבר על סדר הופעת האלמנטים במסמך.

```
for $c1 in //ad[type = "cars"], $c2 in //ad[type = "cars"]
return
  if (number($c1/date) < number($c2/date)
    and every $ad in //ad[number($c1/date) < number(date)
      and number(date) < number($c2/date)]
      satisfies $ad/type = "Real Estate")
  then ($c1,$c2)
  else()
```

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):

FATHER, MOTHER, ו-MARRIED - לקשרים אלו אין תכונות.

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 של בני דודים שניים בעץ (כאלו שלהם סבתא רבה או סבא רבה משותפים, אך הוריהם אינם אחים). יש להימנע מהדפסה כפולה של שמות. שימו לב: הדפסת זוג דודים כ- (a,b) וגם כ- (b,a) מהווה הדפסה כפולה. (אין חשיבות באיזה סדר תבחרו להדפיס את הזוג כל עוד הוא מודפס פעם אחת בלבד).

```
MATCH (a:Person)-[:FATHER|:MOTHER*3]->(:Person)<-[:FATHER|:MOTHER*3]-(b:Person)
WHERE id(a) < id(b)
      And NOT ((a:Person)-[:FATHER|:MOTHER*..2]->(:Person)<-[:FATHER|:MOTHER*..2]-(b:Person))
RETURN DISTINCT a, b
```

ב. 6 נק' - כתבו שאילתא ב-Cypher המוצאת את כל הדודים והדודות של יוסי (דרגה ראשונה בלבד). שימו לב שדודה יכול להיות גם דרך חתונה עם דוד/דודה בקשרי דם (אח/אחות של אחד ההורים). לא מובטח שכל ההורים נשואים.

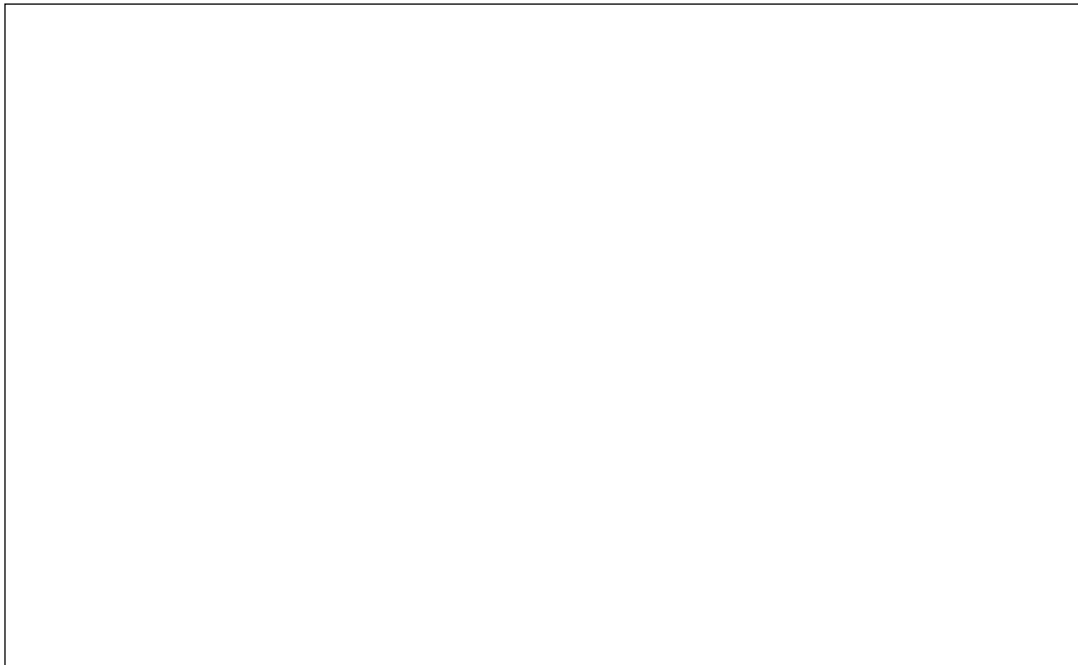
```
MATCH (:Person { name:"Yossi" })-[:FATHER|MOTHER]-> (parent:Person)-
[:FATHER|MOTHER]->(:Person)<-[:FATHER|MOTHER]-(Uncle1:Person)-[:MARRIED*0..1]-
(Uncle:Person)

WHERE Uncle1 <> parent

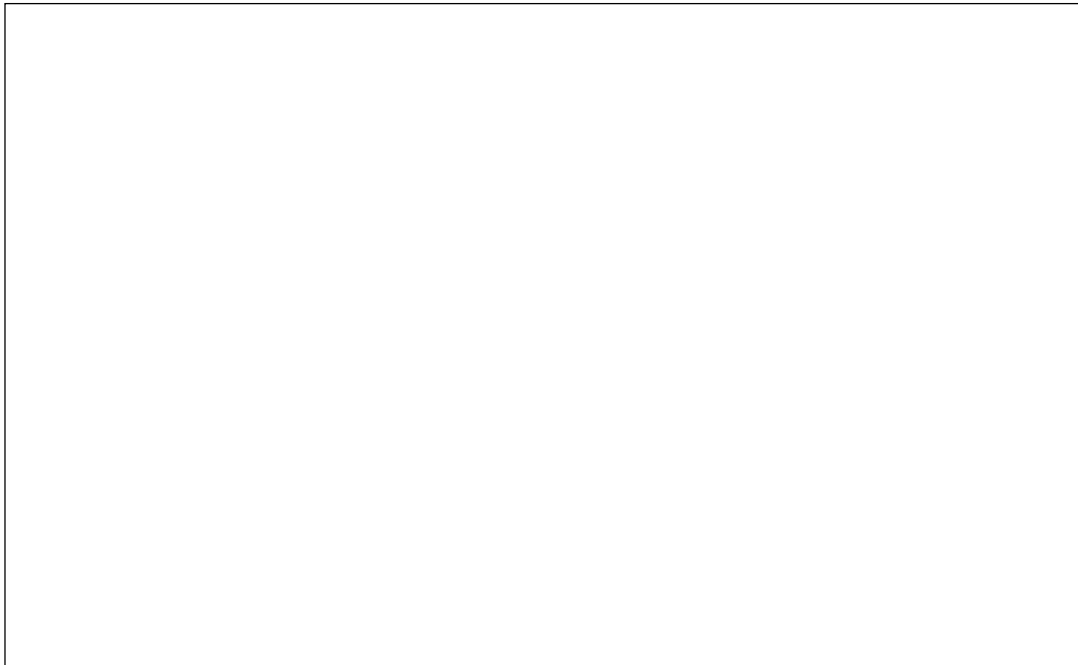
RETURN Uncle
```

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

שאלה: _____ סעיף: _____



שאלה: _____ סעיף: _____



שאלה: _____ סעיף: _____