

סמסטר חורף התשע"ט

מרצה: פרופ' עודד שמואלי

מתרגלים: ענבר קסלסי

משה סבאג

אסף ישורון

מערכות מסד נתונים

236363

מועד א'

1 בפברואר 2019

פירוט החלקים והניקוד:

חלק	נושא	מס' שאלות בחלק	ניקוד
1	תכן מסדי נתונים	2	36
2	שאליות במודל היחסים	3	40
3	מודלים לא יחסיים	2	24

## הנחיות לנבחנים

1. כתבו את התשובות אך ורק בטופס הבחינה ובמקום המיועד להן, מחברת הטייטה לא תיבדק.
2. כל חומר עזר הכתוב על נייר בלבד מותר בשימוש.
3. אין לקבל או להעביר חומר כלשהו בזמן הבחינה.
4. יש להשתמש רק בסימנים או פונקציות שנלמדו בתרגול או בהרצאה והמופיעים בשקפים של הקורס. כל שימוש בסימון שאינו כזה מחייב הסבר מלא של משמעות הסימון.
5. משך הבחינה הינו שלוש שעות, תכננו את הזמן בהתאם.
6. בבחינה 7 שאלות בשלושה חלקים. נא וודאו שיש בידכם את כל הטופס.
7. אין לכתוב בעפרון.

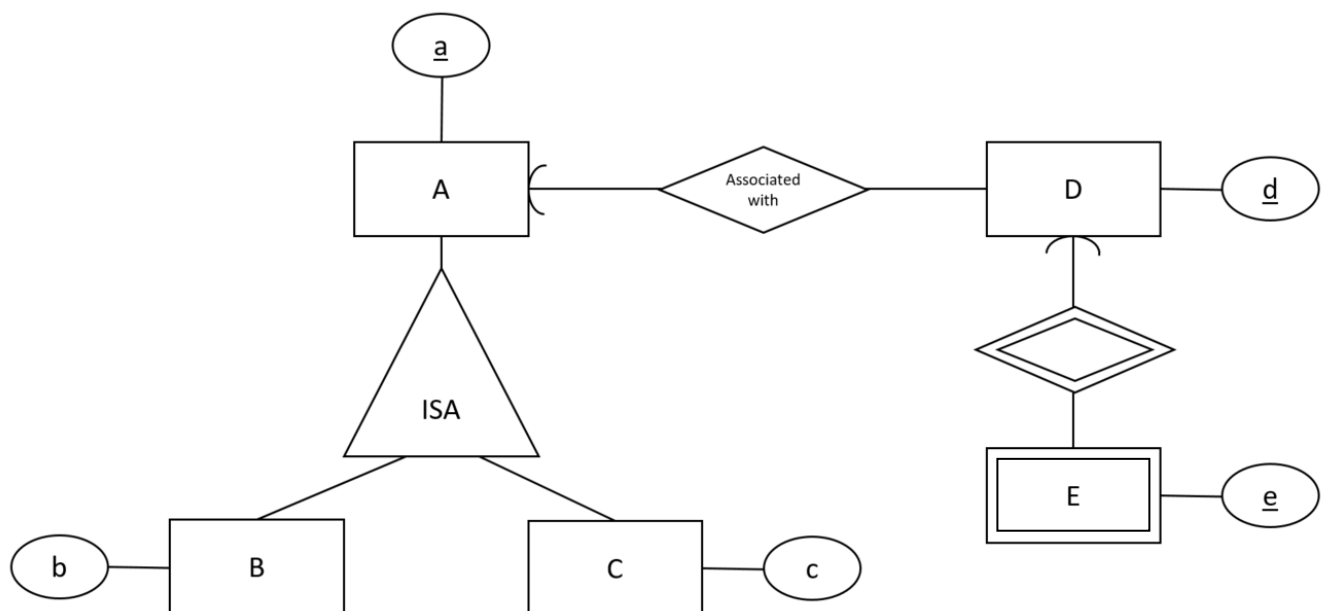
בהצלחה!

## חלק 1 - תכן מסדי נתונים 36 נק'

### שאלה 1 - ERD, 8 נק'

לפניכם תרשים ERD ורשימת טענות. הניחו כי תרשים ה-ERD תורגם לטבלאות לפי הכללים אשר נלמדו בקורס, באופן כזה שכל סוג ישות ב-ERD תורגם לטבלה בעלת אותו שם (לדוגמא סוג הישות A תורגם לטבלה בשם A).

עבור כל טענה מתוך רשימת הטענות קבעו האם היא בטוח נכונה, אולי נכונה בנסיבות מסוימות, או בטוח שגויה. הסבירו בקצרה את תשובותיכם. (2 נק' לכל טענה)



$$1. \pi_a(E) \subseteq \pi_a(A)$$

$$2. \pi_d(D) = \pi_d(E)$$

$$3. \pi_a(A) \setminus \pi_a(B) = \pi_a(C)$$

$$4. |\pi_a(A)| < |\pi_a(B)| + |\pi_a(C)| \text{ (כאשר } |S| \text{ מציין את מספר האלמנטים בקבוצה } S)$$

**שאלה 2 - פירוקים ותלויות פונקציונליות, 28 נק'**

א. נתונה סכמת הרלציה  $R = (A, B, C, D, E, F)$  וקבוצת התלויות פונקציונליות הבאה:

$$AB \rightarrow C, A \rightarrow B, C \rightarrow F, BC \rightarrow A, A \rightarrow E, ABE \rightarrow F, F \rightarrow A$$

1. (4 נק') מצאו כיסוי מינימלי לקבוצת התלויות. פרטו את שלבי ריצת האלגוריתם:

כיסוי מינימלי:
פירוט אלגוריתם:

2. (2 נק') ציין את כל המפתחות עבור  $R$  תחת קבוצת התלויות הנתונה.

--

ב. לסעיף זה אין קשר ישיר לסעיף הקודם. הוכח או הפרך את הטענות הבאות:

1. (3 נק') קיימת סכמת רלציה  $R$  וקבוצת תלויות פונקציונאליות  $F$  כך שקיים עבורן פירוק BCNF אשר משמר תלויות אך אינו משמר מידע.

2. (3 נק') עבור כל סכמת רלציה  $R$  וקבוצת תלויות פונקציונאליות מינימלית  $F$  (כלומר,  $F$  מהווה כיסוי מינימלי של עצמה) שעבורן  $R, F$  היא ב-BCNF, מתקיים כי לכל תלות  $X \rightarrow A$  ב- $F$ ,  $X$  הוא מפתח.

3. (4 נק') עבור כל סכמה רלציונית המכילה  $n \geq 4$  אטריביוטים  $R = (A_1, \dots, A_n)$  קיימת קבוצת תלויות פונקציונאליות  $F$  כך ש:

- i.  $F$  מהווה כיסוי מינימלי של עצמה (כלומר, מינימלית).
- ii. כל אטריביוט  $A_i$  מופיע בלפחות תלות אחת של  $F$ .
- iii.  $R, F$  היא ב-3NF.
- iv. בכל תלות  $X \rightarrow A$  ב- $F$ ,  $X$  אינו מפתח על.

ג. נגדיר INTEGRITY CONSTRAINT מסוג חדש בשם תלות שוללת.

בהינתן סכמת רלציה R עם אטריביוטים  $A_1, \dots, A_n$  תלות שלילה היא מהצורה  $X \rightarrow \neg A$  כאשר X היא קבוצה של אטריביוטים מתוך  $A_1, \dots, A_n$  ו-A הוא אטריביוט מתוך  $A_1, \dots, A_n$ . רלציה r עבור סכמה R מספקת את תלות השלילה  $X \rightarrow \neg A$  אם לכל שתי ניות t1, t2 ב-r מתקיים שאם  $t1[X] = t2[X]$  אז  $t1.A \neq t2.A$ . נסמן זאת ב-  $X \rightarrow \neg A$  r. עבור תלות  $X \rightarrow \neg A_1 \dots A_n$  אם  $t1[X] = t2[X]$  אז  $t1[A_1 \dots A_n] \neq t2[A_1 \dots A_n]$ .

לדוגמא, עבור הסכמה R עם אטריביוטים A, B, C הרלציה r הבאה מספקת את תלות השלילה  $AB \rightarrow \neg C$  אך איננה מספקת את תלות השלילה  $A \rightarrow \neg C$ .

r:

A	B	C
5	8	7
5	8	9
5	7	9
1	1	1

בשאלה זו, קבוצות של אטריביוטים ואטריביוטים בודדים מתייחסים לאטריביוטים מ-  $A_1, \dots, A_n$ .

הוכח/הפוך את הטענות המתמטיות הבאות:

1. (3 נק')  $Y \subseteq X$  גורר לוגית את  $X \rightarrow \neg Y$ :

2. (3 נק') לכול קבוצה של אטריביוטים Z, התלות  $X \rightarrow \neg Y$  גוררת לוגית את  $XZ \rightarrow \neg YZ$ :

3. (3 נק')  $X \rightarrow \neg Y$  וגם  $Y \rightarrow \neg Z$  גורר לוגית את  $X \rightarrow \neg Z$

4. (3 נק')  $X \rightarrow \neg Y$  גורר לוגית את  $Y \rightarrow \neg X$



## חלק 2- שאלות מידע רלציוניות – 40 נק'

### שאלה 3 - SQL, RA 20 נק'

בעקבות כמות הווידויים הגדלה ב- Technion Confessions (דף פייסבוק שאליו משתמשי פייסבוק יכולים להעלות וידויים אנונימיים) החליט הקורס לפתוח אתר וידויים מתחרה בשם Maman Confessions. לצורך הקמת מסד הנתונים לאתר הוחלט להיעזר בסטודנטים בקורס.

האתר יעבוד באופן הבא:

כל אדם יכול להירשם כמשתמש לאתר ולכתוב וידוי (שלא יהיה אנונימי למי שיש גישה למסד הנתונים). כל משתמש יכול לסמן בלייק (כלומר לסמן שאהב) וידויים שונים (ויכול גם לסמן שאהב וידויים של עצמו). שימו לב, כי כל משתמש יכול לסמן כל וידוי בלייק פעם אחת בלבד (כלומר לא ייתכן שמשתמש יעשה שני לייקים לאותו וידוי).

במסד הנתונים קיימות הטבלאות הבאות:

Users:

<u>uid</u>	Name	Faculty
------------	------	---------

Confessions:

<u>cid</u>	uid	Content
------------	-----	---------

Likes:

<u>uid</u>	<u>cid</u>
------------	------------

תוכן הטבלאות:

הטבלה Users מחזיקה רשומות עם מספר זיהוי של סטודנט (uid), שם הסטודנט (Name), והפקולטה אליה שייך הסטודנט (Faculty).

הטבלה Confessions מחזיקה רשומות עם מספר הזיהוי של הווידוי (cid), מספר הזיהוי של הסטודנט שכתב את הווידוי (uid) ותוכן הווידוי (Content). בנוסף, uid הוא מפתח זר לטבלה Users.

הטבלה Likes מסמנת שסטודנט בעל מספר זהות uid עשה לייק לוידוי (כלומר סימן שאהב את הווידוי) עם מספר זיהוי cid. בנוסף, uid הוא מפתח זר לטבלה Users ו-cid הוא מפתח זר לטבלה Confessions.

שימו לב כי מפתחות מסומנים בקו תחתון בטבלאות.

א. (4 נק') הגדירו View שמציג לכל וידוי את מספר הווידוי (cid), כותב הווידוי (uid) ומספר הלייקים (likesNum) שהווידוי קיבל (כלומר, מספר האנשים שאהבו את הווידוי).  
תזכורת: COUNT על סט ערכים יספור רק ערכים שאינם NULL ואם כולם NULL יחזיר אפס.  
עליכם להשתמש בשלד השאילתה הבא:

```
CREATE VIEW LikesPerConf AS  
SELECT
```

ב. (4 נק') כתבו שאילתה שמחזירה את כל המזהים של הווידויים (cid) שקיבלו מעל 10 לייקים ולא קיבלו לייק מכותב הווידוי עצמו (כלומר מעל 10 אנשים אהבו את הווידוי וגם כותב הווידוי לא סימן שאהב אותו בעצמו). מותר (וכדאי) להשתמש ב-View שהוגדר בסעיף 1 כטבלה גם אם לא עניתם על סעיף 1.

ג. (4 נק') כתבו שאילתה שתחזיר את המזהים (cid) של כל הווידויים שקיבלו לייק מלפחות 5 סטודנטים ממדמ"ח (Faculty='CS') אן שהם בעלי מספר הלייקים הגדול ביותר בקרב כל הווידויים. גם בסעיף זה מותר (וכדאי) לכם להשתמש ב- View שהוגדר בסעיף 1 כטבלה (גם אם לא עניתם על סעיף 1). השתמשו בשלד השאילתה הבא:

SELECT

HAVING

UNION

SELECT

ד. (4 נק') כתבו שאילתת RA המחזירה את המזהים של כל משתמש המספק את התנאי הבא:  
המשתמש (uid) עשה לייק לכל הווידויים שמשתמשים אחרים כתבו (כלומר, בין היתר המשתמש סימן שאהב את כל הווידויים שמשתמשים אחרים כתבו).

ה. (4 נק') מה מחזירה שאילתת ה-RA הבאה? הקף בעיגול את התשובה הנכונה.

$$(\pi_{uid}(Users)) \setminus (\pi_{uid}((\rho_{uid \rightarrow uid1} Likes) \bowtie Confessions))$$

- א. את כל מזהי המשתמשים שעשו לייק לווידי כלשהו.
- ב. את כל מזהי המשתמשים שכתבו ווידוי כלשהו.
- ג. את כל מזהי המשתמשים שקיבלו לייק על ווידוי כלשהו שלהם.
- ד. את כל מזהי המשתמשים שקיבלו לייק על כל הווידויים שלהם.
- ה. את כל מזהי המשתמשים שלא קיבלו לייק על שום ווידוי שלהם.
- ו. את כל מזהי המשתמשים שלא עשו לייק לווידי כלשהו.
- ז. אף תשובה מהנ"ל אינה נכונה.

#### שאלה 4 - DRC , 10 נק'

תזכורת: מטריצה היא מערך דו-מימדי של אלמנטים.  
נתונה הסכמה הבאה:

Matrix:

SerialId	NumElements	NumRows	Rank
----------	-------------	---------	------

המספר הסידורי SerialId של מטריצה הינו מפתח לצורך זיהויה במסד הנתונים, מספר זה ייחודי לכל מטריצה ואינו קשור לערכיה או לדרגתה.  
NumElements מציין את מספר האלמנטים במטריצה.  
NumRows מציין את מספר השורות במטריצה.  
Rank מציין את דרגת המטריצה שהינה תכונה מתמטית שלה.

בנוסף נתון היחס הדו מקומי  $R_=(a, b) \leftrightarrow a = b$ , ונתון יחס נוסף, תלת מקומי,  $R_d$ , המקיים:

$$R_d(a, b, c) \leftrightarrow \frac{a}{b} = c$$

מותר להשתמש בשני יחסים אלו בתשובותיכם.

א. (4 נק') תזכורת: מטריצה הפיכה היא מטריצה ריבועית  $N \times N$  בעלת דרגה  $N$ .  
כתבו שאילתה ב- RELATIONAL CALCULUS (RC) המחזירה את המספרים הסידוריים (SerialId) של כל המטריצות ההפיכות.

ב. (6 נק') התבוננו בשאילתה הבאה:

$$\{s | \exists nr, ne, r [Matrix(s, ne, nr, r) \wedge [[R_d(ne, ne, ne) \vee \neg R_d(ne, nr, nr)] \rightarrow R_=(r, 5)]]\}$$

הקיפו בעיגול את כל הטענות הנכונות:

1. השאילתה תחזיר, בין היתר, את המספרים הסידוריים של כל המטריצות הריבועיות אשר בהן יותר מאיבר אחד.
2. השאילתה תחזיר, בין היתר, את המספרים הסידוריים של כל המטריצות בעלות איבר בודד.
3. השאילתה תחזיר, בין היתר, את המספרים הסידוריים של כל המטריצות הריבועיות שדרגתן אינה 5.
4. השאילתה תחזיר, בין היתר, את המספרים הסידוריים של כל המטריצות אשר אינן ריבועיות ואשר דרגתן 5.

### שאלה 5 - Datalog , 10 נק'

נתון גרף מכוון  $G = (V, E)$ , כאשר  $V$  היא קבוצה של צמתים ו- $E$  היא קבוצה של קשתות מכוונות מהצורה  $(v, u)$ , המציינת קשת מ- $v$  אל  $u$ . הגרף  $G$  מיוצג על ידי הרלציה  $edge$  ב-EDB כך שמתקיים:

$$edge(v, u) \leftrightarrow (v, u) \in E$$

הגרף מוגדר רק ע"י הרלציה  $edge$  ולכן לא קיימים צמתים נוספים שאינם מוזכרים ברלציה  $edge$ .

א. (1 נק') כתבו תכנית Datalog המגדירה את היחס החד מקומי  $node(u)$  המתקיים עבור כל הצמתים בגרף.

ב. (4 נק') כתבו תכנית Datalog המגדירה את היחס הדו-מקומי  $notSame(u, v)$  המתקיים עבור כל שני צמתים  $u, v$  השונים זה מזה. אין להשתמש ב- $\neq$  או ב- $=$  בשאלה זו. ניתן להשתמש בפרדיקט  $node(u)$  מהסעיף הקודם גם אם לא פתרת אותו.

ג. (5 נק') כתבו תכנית Datalog המגדירה את היחס הדו-מקומי  $notSimplePath(u, v)$  שמשמעותו היא שבין  $u$  ל-  $v$  קיים מסלול שאינו פשוט.

תזכורת:

**מסלול** הוא סדרת צמתים  $v_1, \dots, v_n$  כאשר  $(v_i, v_{i+1})$  היא קשת ב- $E$  כך ש  $n > 1, 1 \leq i < n$ .  
**מסלול פשוט** הוא מסלול בו אין צומת המופיע יותר מפעם אחת (כלומר, המסלול לא מכיל מעגל).

### חלק 3- מודלים לא רלציוניים – 24 נק'

#### שאלה 6 XML, 16 נק'

נתון ה DTD הבא (כאשר a הוא אלמנט השורש) בתוך מסמך (שיכונה מעתה מסמך הדוגמה).  
הניחו שההפרדה לשורות היא לצרכי תצוגה בלבד.

```
<?xml version="1.0" standalone="yes"?>
<!DOCTYPE a [
  <!ELEMENT a (b+)>
  <!ELEMENT b (c+, d*, e)>
  <!ELEMENT c (#PCDATA)>
  <!ELEMENT d (#PCDATA)>
  <!ELEMENT e (#PCDATA)>
  <!ATTLIST d
    ida ID #IMPLIED
  >
  <!ATTLIST c
    ref IDREF #REQUIRED
  >
]>
```

```
<a>
  <b>
    <c ref="d1"> 1000 </c>
    <d> a1 </d>
    <d> a2 </d>
    <d ida="d2"> a3 </d>
    <d> a4 </d>
    <e> e1 </e>
  </b>
  <b>
    <c ref="d2"> 3000 </c>
    <d ida="d3"> b1 </d>
    <d ida="d1"> b2 </d>
    <e> e2 </e>
  </b>
  <b>
    <c ref="d2"> 4000 </c>
  </b>
</a>
```



א. (2 נקודות) האם מסמך הדוגמה מספק את ה- DTD שלו? הקף בעיגול **כן** **לא** אם תשובתך "לא" הסבר:

בסעיפים הבאים יש לציין מה השאילתות יחזירו על מסמך הדוגמה. שים לב שניתן לבצע שאילות Xpath על מסמך גם כאשר המסמך אינו מספק את ה- DTD שלו.

ב. (2 נקודות) אילו אלמנטים יחזיר הביטוי  
`/a/b[child::d[@ida]] [position() < 3] [position() = 2]/d[2]`

ג. (2 נקודות) אילו אלמנטים יחזיר הביטוי  
`/a/b[child::c[id(@ref) [@ida]] /d[2]`

ד. (10 נקודות) בטבלה הבאה שני טורים ובהם ביטויים ב-XPATH. בטור השלישי עליך לציין האם הביטויים שקולים על מסמך הדוגמה ובטור הרביעי עליך לציין אם הם שקולים על כל מסמך המספק את ה-DTD הנתון (תזכורת - בשורה החמישית בטבלה למטה, | מציין OR).

האם זהים על כל מסמך המספק את ה-DTD של מסמך הדוגמה?	האם זהים על מסמך הדוגמה?	ביטוי שני	ביטוי ראשון
		//b[d]/d	//b[c]/d
		/a[//b[c]/d]	/a[//b[d]/c]
		//b[//a]/c	//b[a]/c
		//b[c/@ref]/c/id(@ref)	//b[c/d[@ida]
		//b[a   e]/c	//b[ /a   b]/c

## שאלה 7, RDF, 8 נק'

לפניכם גרף RDF בפורמט של TRIPLES (כל שורה בטבלה מייצגת שלשה בגרף):

ent:Technion	prop:department	ent:CS
ent:Technion	prop:department	ent:EE
ent:CS	prop:lecturer	ent:Ido
ent:CS	prop:lecturer	ent:Avi
ent:CS	prop:lecturer	ent:Oded
ent:EE	prop:lecturer	ent:Pnina
ent:EE	prop:lecturer	ent:Ido
ent:Ido	prop:employer	ent:Technion
ent:Ido	prop:employer	ent:Google
ent:Oded	prop:employer	ent:Technion
ent:Pnina	prop:employer	ent:FB
ent:Avi	prop:employer	ent:FB

א. (3 נק') כתבו את ההשמות (mappings) למשתנים שמחזירה השאילתה הבאה (אין צורך לרשום את אותה ההשמה יותר מפעם אחת).

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

PREFIX ent: <http://dbpedia.org/property/ent/>

PREFIX prop: <http://dbpedia.org/property/>

```
SELECT ?lecturer1 {
  ?d prop:lecturer ?lecturer1 .
  ?lecturer1 prop:employer ent:Technion .
  ?lecturer1 prop:employer ent:Google .
  OPTIONAL{
    ?lecturer1 prop:employer ent:FB .
  }
}
```

א. (4 נק') כתבו את ההשמות (mappings) שמחזירה השאילתה הבאה (אין צורך לרשום את אותה ההשמה יותר מפעם אחת).

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

PREFIX ent: <http://dbpedia.org/property/ent/>

PREFIX prop: <http://dbpedia.org/property/>

```
SELECT ?u ?lecturer {  
    ?u prop:department ?d1 .  
    ?u prop:department ?d2 .  
    ?d1 prop:lecturer ent:Avi .  
    ?d1 prop:lecturer ?lecturer .  
    ?d2 prop:lecturer ?lecturer .  
    ?lecturer prop:employer ?u .  
    FILTER ( ! (?d1 = ?d2) ). # different variable bindings to ?d1 and ?d2  
}
```



ב. (1 נק') כתבו את ההשמות (mappings) שמחזירה השאילתה הקודמת אם ent:Avi מוחלף בשאילתה הזו ב- ent:Pnina (אין צורך לרשום את אותה ההשמה יותר מפעם אחת).



**מקום לתשובות נוספות:**

<b><u>חלק:</u></b>	<b><u>שאלה:</u></b>	<b><u>סעיף:</u></b>
<b><u>חלק:</u></b>	<b><u>שאלה:</u></b>	<b><u>סעיף:</u></b>
<b><u>חלק:</u></b>	<b><u>שאלה:</u></b>	<b><u>סעיף:</u></b>
<b><u>חלק:</u></b>	<b><u>שאלה:</u></b>	<b><u>סעיף:</u></b>