

סמסטר חורף התש"פ

מרצה: פרופ' עודד שמואלי

עוזרי הוראה: אסף ישורון

שובל לגזיאל

משה סבאג

רואי קיסוס

מערכות מסד נתונים

236363

מועד ב'

1 במרץ 2020

פירוט החלקים והניקוד:

שאלה	נושא	ניקוד
1	ERD	10
2	RA	20
3	SQL	20
4	Design Theory	20
5	RDF	10
6	CC	20

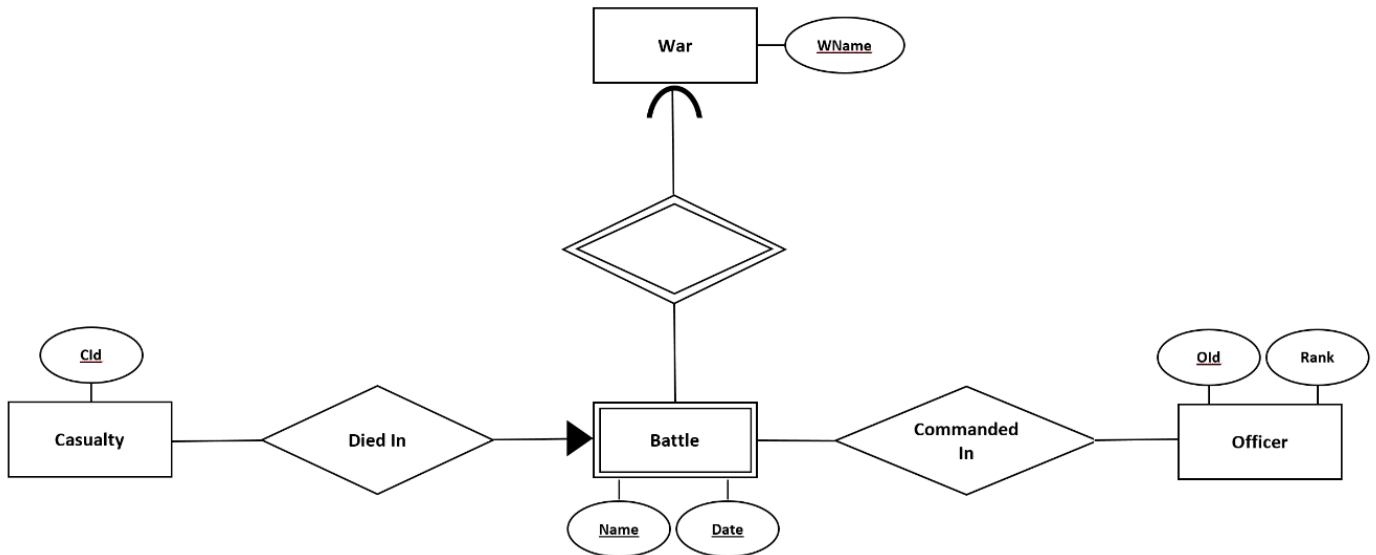
## הנחיות לנבחנים

1. כתבו את התשובות אך ורק בטופס הבחינה ובמקום המיועד להן, מחברת הטיוטה 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 בסמסטר זה 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. אין לכתוב בעפרון.

בהצלחה!

## ERD – 10 נקודות

המחלקה להיסטוריה יצרה את תרשים ה-ERD הבא:



א. (8 נק') תרגמו את התרשים (על כל מרכיביו) לסכמות של יחסים. סמנו מפתחות בקו תחתון וציינו מפתחות זרים במקומות המתאימים.

Officer(Oid, Rank)

Commanded In(Oid, Name, Date, WName), referencing Officer.Oid, Battle.Name, Battle.Date, War.WName, respectively.

Battle(Name, Date, WName), WName references War.WName

War(WName)

Died In(Cid, Name, Date, WName), referencing Casualty.Cid, Battle.Name, Battle.Date, Battle.WName, respectively.

Casualty(Cid).

ב. (2 נק') בתרשים הנוכחי, אם ברצוננו לתעד קצין שמת בקרב, עלינו לשמור את המזהה שלו פעמיים – פעם אחת תחת המפתח Oid של Officer ופעם שנייה תחת המפתח Cid של Casualty. הציעו שינוי בתרשים שימנע את הצורך בכפילות זו:

ניצור ישות Soldier בעלת מפתח Sid אליה יהיו מקושרות גם Officer וגם Casualty בקשר ISA.

## RA, RC – 20 נקודות

נתונות סכמות  $R(A,B,C)$  ו-  $S(C,D,E)$ .

בנוסף נתון היחס  $R_{>}(a,b)$  שמתקיים עבור כל המספרים  $a, b$  כך ש-  $a > b$ .

הניחו כי הרלציות מעל הסכמות  $R$  ו-  $S$  אינן ריקות.

א. בהינתן שאילתת ה- **SQL** הבאה:

```
(SELECT DISTINCT R.B, R.C
FROM R, S WHERE R.C = S.E AND B > 3)
INTERSECT
(SELECT DISTINCT R.B, R.C FROM R WHERE A=B)
```

a. (2 נק') נסחו שאילתת **RA** שקולה:

$$\pi_{B,R.C} \sigma_{A=B \wedge B>3} (R \bowtie_{R.C=S.E} S)$$

b. (3 נק') נסחו שאילתת **RC** שקולה:

$$\{(b, c1) \mid \exists c2, d (R(b, b, c1) \wedge S(c2, d, c1) \wedge R_{>}(b, 3))\}$$

כעת, הניחו כי נתון גם היחס  $R_{\neq}(a,b)$  שמתקיים עבור כל  $a \neq b$ .

ב. בהינתן שאילתת ה- **SQL** הבאה:

```
SELECT DISTINCT R.A, R.B
FROM R
WHERE R.C NOT IN (SELECT S.C FROM S)
```

a. (3 נק') נסחו שאילתת **RA** שקולה:

$$\pi_{A,B}(R \setminus (R \bowtie S))$$

b. (4 נק') נסחו שאילתת **RC** שקולה בה אין שימוש בשלילה (NOT, -):

$$\{(a, b) \mid \exists c R(a, b, c) \wedge \forall c2, d, e [S(c2, d, e) \rightarrow R_{\neq}(c, c2)]\}$$

ג. נתון כי:

i. השאילות:

```
SELECT COUNT(*) FROM R;
SELECT COUNT(*) FROM (SELECT DISTINCT * FROM R) AS RDistinct;
```

מחזירות שתיהן  $n > 0$ .

ii. השאילות:

```
SELECT COUNT(*) FROM S;
SELECT COUNT(*) FROM (SELECT DISTINCT * FROM S) AS SDistinct;
```

מחזירות שתיהן  $m > 0$ .

iii. השאילתא:

$$\pi_C S$$

מחזירה  $k > 0$  רשומות.

עבור כל אחת מהשאלות הבאות, ציינו אם הביטוי של השאלתה מוגדר היטב ב-RA, ואם כן כתבו מהו המספר המקסימלי של N'יות (tuples) שיכולות לחזור מהרצת השאלתה:

a.  $\pi_{A,B} R \setminus \pi_{A,B} ((\pi_{A,B} R \times \pi_C S) \setminus R)$  (2 נק')

מוגדר/ לא מוגדר (הקף),

אם בחרתם מוגדר, ציינו את מספר ה-N'יות (tuples) המקסימלי:  $\begin{bmatrix} n \\ k \end{bmatrix}$

b.  $R \setminus \pi_C S$  (1 נק')

מוגדר/ לא מוגדר (הקף),

אם בחרתם מוגדר, ציינו את מספר ה-N'יות (tuples) המקסימלי:

c.  $\sigma_{D>5} \pi_{D,E} S \bowtie \pi_{A,B} R$  (2 נק')

מוגדר/ לא מוגדר (הקף),

אם בחרתם מוגדר, ציינו את מספר ה-N'יות (tuples) המקסימלי:  $n * m$

ד. (3 נק') עבור שאלתת ה-RA הבאה:

$$\pi_A R \setminus \pi_A (\sigma_{D>7} (\pi_{A,B} R \bowtie_{A=E} \pi_{D,E} S))$$

בחרו מבין האפשרויות את שאלתת ה-RC השקולה לשאלתת הנתונה (בחרו אפשרות אחת בלבד):

a.  $\{a \mid \exists b, c, c2, d R(a, b, c) \wedge \neg S(c2, d, a) \wedge \neg R_{>}(d, 7)\}$

b.  $\{a \mid \exists b, c R(a, b, c) \wedge \forall a2, c2, d [\neg S(c2, d, a) \wedge \neg R_{>}(a2, 7) \wedge R_{\neq}(a, a2)]\}$

c.  $\{a \mid \exists b, c R(a, b, c) \wedge \forall c2, d, e [S(c2, d, e) \rightarrow R_{\neq}(a, e)]\}$

d.  $\{a \mid \exists b, c R(a, b, c) \wedge \forall c2, d, e [(S(c2, d, e) \wedge R_{>}(d, 7)) \rightarrow R_{\neq}(a, e)]\}$

e. שאלתת ה-RA הנתונה לא ניתנת לביטוי על ידי שאלתת RC כלשהיא.

f. אף אחת מהתשובות הנ"ל אינה נכונה.

## SQL – 20 נקודות

בהמשך לתרגיל הבית הרביעי, הוחלט כי ייצוג סרטים ואנשים בתור גרף Neo4j לא מתאים למערכת שנבחרה ולכן נבחרה מערכת מבוססת SQL.

נתונות הטבלאות (הלא ריקות) הבאות כאשר קו מודגש מסמל מפתח:

Persons Table – Person

<u>PersonID</u> (INTEGER)	PersonName (TEXT)
---------------------------	-------------------

טבלת האנשים המכילה עבור כל אדם את המזהה שלו ושמו.

Movies Table – Movie

<u>MovieID</u> (INTEGER)	MovieName (TEXT)	ReviewersRating (INTEGER)	Prequel (INTEGER)
--------------------------	------------------	---------------------------	-------------------

טבלת הסרטים המכילה עבור כל סרט את המזהה שלו, שמו, הציון שקיבל ממבקרי ו הסרט הקודם לו בסדרת הסרטים (כלומר, ה-MovieID של הסרט הקודם לו). הניחו כי קיימת רשומה (שמציינת סרט דמה) בעלת מזהה 0 וכאשר ברשומה של סרט m כלשהו  $Prequel = 0$  אזי m הוא סרט ראשון בסדרה.

Watch Table – Watch

<u>PersonID</u> (INTEGER)	<u>MovieID</u> (INTEGER)	Rank (INTEGER)
---------------------------	--------------------------	----------------

בטבלת הצפיות, לכל צפייה של אדם מסוים בסרט מסוים ישנה רשומה המכילה את המספר המזהה של האדם (PersonID) הצופה בסרט בעל מזהה MovieID, כאשר אותו אדם דירג את הסרט בדירוג Rank.

בעת יצירת מסד הנתונים הוגדרו התכונות הבאות:

Watch.PersonID, Watch.MovieID הינם מפתחות זרים ל-Movie.MovieID, Person.PersonID, בהתאמה.

בנוסף השדה "Prequel" ב-Movie הינו מפתח זר ל-Movie.MovieID.

א. (5 נק') כתבו שאילתת SQL המחזירה את כל מזהי הסרטים אשר קיימים לפחות 2 סרטים קודמים להם בסדרה (Prequels). הניחו כי ברשומות של Movie,  $Prequel = 0$  מציין כי זהו הסרט הראשון, או הבודד, בסדרה.

```
SELECT DISTINCT m.MovieID
FROM Movie m JOIN Movie m1,
ON m.Prequel = m1.MovieID AND m.MovieID <> 0 AND m.Prequel <> 0
AND m1.Prequel <> 0
```

ב. (5 נק') כתבו שאילתת SQL המחזירה את המזהים של כל סרט שממוצע דירוגו ע"י האנשים שצפו בו (Rank) גבוה ממש מהדירוג שניתן לאותו הסרט על ידי המבקרים (ReviewersRating).

```
SELECT m.MovieID
FROM Movie m
WHERE m. ReviewersRating < (SELECT AVG(Rank)
FROM Watch w
GROUP BY w.MovieID
HAVING w.MovieID =m. MovieID)
```

ג. (5 נק') כתבו שאילתת SQL המחזירה את מזהי האנשים שצפו בסרט כלשהו שיש לו לפחות 2 סרטים קודמים בסדרה (Prequels) וגם מתקיים שמספר הסרטים שבהם צפה כל אחד מהם גדול או שווה ל-10. הניחו כי הפתרון לסעיף א' נתון לכם כ-VIEW בשם GotPrequels גם אם לא פתרתם אותו.

```
SELECT DISTINCT w. PersonID
FROM Watch w
GROUP BY w. PersonID
HAVING COUNT(*) >= 10 AND
EXISTS (SELECT * FROM
(SELECT w1.MovieID FROM Watch w1 WHERE w1. PersonID = w. PersonID) f1 JOIN
(SELECT * FROM GotPrequels) pr ON f1. MovieID = pr. MovieID)
```

ד. (2.5 נק') הקיפו בעיגול את תיאור קבוצת הרשומות המוחזרות על ידי השאילתה הבאה:

```
1 SELECT DISTINCT PersonID
2 FROM Watch
3 WHERE MovieID
4 IN (SELECT MovieID
5     FROM Movie
6     WHERE MovieID
7     NOT IN (SELECT Prequel FROM Movie))
```

- a. כל מזהי האנשים שצפו בסרט בעל סרט המשך כלשהו.
- b. כל מזהי האנשים שצפו בכל הסרטים שאין להם סרט המשך.
- c. כל מזהי האנשים שצפו בכל הסרטים בעלי סרט המשך כלשהו.
- d. כל מזהי האנשים שצפו בסרט שאין לו סרט המשך.
- e. אף אחת מהתשובות הנ"ל אינה נכונה.

ה. (2.5 נק') הסבירו בקצרה, מה היו צריכים מתכנני מסד הנתונים לשנות בהגדרת הטבלאות על מנת שלא יהיו שני סרטים (Movie) או שני אנשים (Person) בעלי אותו השם? הניחו כי כל מה שלא ציינתם כשינוי – נשאר זהה.

לשנות בהגדרת הטבלאות Person, Movie את הגדרת PersonName, MovieName בהתאמה ל-UNIQUE.



1. (2 נק') תהי סכמה  $R(A,B,C,D)$  עם התלויות הפונקציונליות הבאות:  
 $F = \{A \rightarrow B, B \rightarrow C, C \rightarrow A\}$   
מי מהקבוצות הבאות שקולה ל- $F$ ? (כלומר, בעלת סגור השווה  $F^+$ ) סמנו בעיגול את כל התשובות הנכונות.
  - a.  $\{A \rightarrow BC, B \rightarrow AC, C \rightarrow AB\}$
  - b.  $\{A \rightarrow B, B \rightarrow A, B \rightarrow C\}$
  - c.  $\{A \rightarrow BC, B \rightarrow AC\}$
  - d.  $\{A \rightarrow BC, C \rightarrow AB\}$
  - e. אף אחת מהקבוצות הנ"ל איננה שקולה ל- $F$ .
2. (2 נק') תהי סכמה  $R(A,B,C)$  בעלת התלויות הפונקציונליות  $F = \{A \rightarrow B, B \rightarrow C\}$ . יהי  $r$  מופע (רציה) של  $R$  בעל רשומה אחת  $(0,0,0)$  בלבד. את מי מהרשומות הבאות ניתן להכניס לטבלה  $r$  באופן חוקי (מבלי להפר את  $F$ )? (את הרשומות יש לקרוא משמאל לימין. כלומר האטריביוט השמאלי הינו  $A$ , מימינו  $B$  ולבסוף, בצד ימין,  $C$ )  
סמנו בעיגול את כל התשובות הנכונות.
  - a.  $(1,0,1)$
  - b.  $(1,1,0)$
  - c.  $(0,2,0)$
  - d.  $(0,0,1)$
  - e. לא ניתן להכניס אף אחת מהרשומות הנ"ל לטבלה  $r$  באופן חוקי.
3. (2 נק') תהי סכמה  $R(A,B,C,D,E)$  בעלת התלויות הפונקציונליות  
 $F = \{D \rightarrow CE, CE \rightarrow A, D \rightarrow A, AE \rightarrow D\}$   
מי מהבאים הוא מפתח של  $R$ ? סמנו בעיגול את כל התשובות הנכונות.
  - a.  $AB$
  - b.  $CDE$
  - c.  $BD$
  - d.  $BCE$
  - e. אף אחת מהתשובות הנ"ל אינה נכונה.
4. (2 נק') עבור פירוק שאינו משמר מידע מתקיים כי (סמנו את התשובה הנכונה):
  - a. ב- JOIN של תוצרי הפירוק עלולים לאבד רשומות שהיו בטבלה המקורית.
  - b. ב- JOIN של תוצרי הפירוק עלולים ליצר רשומות שלא היו קיימות בטבלה המקורית.
  - c. ב- JOIN יתכן שנקבל אטריביוטים שלא היו בסכמה המקורית לפני הפירוק.
  - d. אף אחת מהתשובות הנ"ל אינה נכונה.
5. (5 נק') תהי סכמה  $R(A,B,C,D,E,H)$  בעלת התלויות  
 $F = \{ABC \rightarrow CDEH, C \rightarrow E, A \rightarrow B, D \rightarrow H\}$   
מצאו כיסוי מינימלי ל- $F$ .  
 $G = \{AC \rightarrow D, C \rightarrow E, A \rightarrow B, D \rightarrow H\}$

6. (7 נק') ניזכר במשפט הבא לגבי פירוק בינארי משמר מידע מהרצאה:

יהי  $\{X_1, X_2\}$  פירוק של הסכמה  $\{U, F\}$ . הטענות הבאות הינן שקולות:

1.  $X_1 \cap X_2 \rightarrow X_1$  או  $X_1 \cap X_2 \rightarrow X_2$ .

2. הפירוק  $\{X_1, X_2\}$  משמר מידע.

ננסה להכליל את המשפט עבור פירוק ל-3 תתי סכמות:

יהי  $\{X_1, X_2, X_3\}$  פירוק של הסכמה  $\{U, F\}$ . הטענות הבאות הינן שקולות:

1.  $X_1 \cap X_2 \cap X_3 \rightarrow X_1$  או  $X_1 \cap X_2 \cap X_3 \rightarrow X_2$  או  $X_1 \cap X_2 \cap X_3 \rightarrow X_3$  או  $X_1 \cap X_2 \cap X_3 \rightarrow X_3$

2. הפירוק  $\{X_1, X_2, X_3\}$  משמר מידע.

a. האם כאשר 1 מתקיים אז בהכרח 2 מתקיים? הוכיחו את תשובתכם..

### 1 לא גורר בהכרח את 2 :

$R(ABCDEF)$   $\{B \rightarrow F\}$   $X_1=ABC, X_2=BDE, X_3=BF$  נתקעים אחרי איטרציה אחת באלגוריתם.

b. האם כאשר 2 מתקיים אז בהכרח 1 מתקיים? הוכיחו את תשובתכם.

### 2 לא בהכרח גורר 1:

$R(A,B,C,D), F = \{B \rightarrow C, C \rightarrow D\}$

פירוק  $X_1(A, B), X_2(B, C), X_3(C, D)$ . מראים לפי האלגוריתם שמשמר מידע, אך החיתוך ריק ובוודאי לא גורר את אחת הקבוצות.

## RDF – 10 נקודות

התבוננו בגרף ה-RDF הבא:

PREFIX exp:< http://www.ontotext.com/explicit>

exp: A	exp: Cooks	exp: Pasta
exp: B	exp: Cooks	exp: Lasagna
exp: C	exp: Cooks	exp: Gnocchi
exp: A	exp: Likes	exp: Pasta
exp: B	exp: Likes	exp: Gnocchi
exp: C	exp: Likes	exp: Lasagna
exp: C	exp: Hosts	exp: CEvent
exp: D	exp: Hosts	exp: Devent

א. (4 נק') לכל אחת מ-4 השאילות שלפניכם ציינו את מספר הרשומות (Mappings) שהן מחזירות. לדוגמא, אם לדעתכם שאילתה a מחזירה חמש רשומות, שאילתה b שתי רשומות, שאילתה c שלוש רשומות ושאילתה d ארבע רשומות, כתבו את התשובה "b-2, c-3, d-4, a-5".  
זכרו: ללא DISTINCT ייתכן ויהיו תשובות מרובות עם אותה רשומה (DUPLICATES).

a. PREFIX exp:<http://www.ontotext.com/explicit/>

```
select ?x {
  ?x exp:Cooks ?y.
  ?x exp:Likes ?z
}
```

b. PREFIX exp:<http://www.ontotext.com/explicit/>

```
select ?x {
  ?x exp:Hosts ?z
}
```

c. PREFIX exp:<http://www.ontotext.com/explicit/>

```
select ?x {
  ?x exp:Hosts ?z
  MINUS {
    ?x exp:Cooks ?y
  }
}
```

d. PREFIX exp:<http://www.ontotext.com/explicit/>

```
select ?x {
  ?x exp:Likes ?y.
  ?z exp:Hosts ?e.
}
```

c-1, b-2, a-3, d-6

ב. (4 נק') כתבו מה מחזירה השאילתה הבאה בהפעלה על הגרף הנתון:

```
PREFIX exp:<http://www.ontotext.com/explicit/>
select distinct ?x ?z {
  ?x exp:Cooks ?y.
  MINUS {
    ?x exp:Hosts ?b
  }
  OPTIONAL {
    ?z exp:Likes ?a
  }
}
```

$\{?x \leftarrow \text{exp:A}, ?z \leftarrow \text{exp:A}\}, \{?x \leftarrow \text{exp:A}, ?z \leftarrow \text{exp:B}\}, \{?x \leftarrow \text{exp:A}, ?z \leftarrow \text{exp:C}\}$   
 $\{?x \leftarrow \text{exp:B}, ?z \leftarrow \text{exp:A}\}, \{?x \leftarrow \text{exp:B}, ?z \leftarrow \text{exp:B}\}, \{?x \leftarrow \text{exp:C}, ?z \leftarrow \text{exp:C}\}$

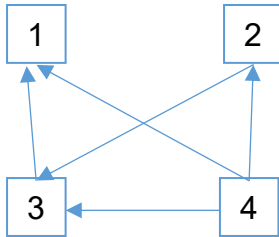
ג. (2 נק') כמה רשומות (Mappings) מחזירה השאילתה הבאה? נמקו.

```
PREFIX exp:<http://www.ontotext.com/explicit/>
select distinct ?a ?b ?c ?d ?e ?f {
  ?a exp:Cooks ?b.
  ?c exp:Likes ?d.
  ?e exp:Hosts ?f
}
```

18 מיפויים כיוון שיש 3 שלשות בעלות פרדיקט Cooks, 3 בעלות פרדיקט Likes, ו-2 בעלות פרדיקט Hosts, אז  $3 \times 3 \times 2 = 18$ .

## בקרת מקביליות – 20 נקודות

א. (3 נק') במסד הנתונים יש פריט נתונים בודד, Z. הראה תהליך ביצוע S של ארבע תנועות כך שה- precedence graph של S הוא:

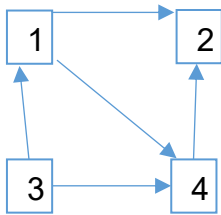


או, לחלופין, הסבר מדוע תהליך ביצוע כזה אינו קיים.

תהליך ביצוע:

$$s = W4[z] R2[z] W3[z] R1[z]$$

ב. (8 נק') נתון פורמט של תהליך ביצוע S' ובו ארבע תנועות, הסימון  $O_i$  מציינ פעולה R או W של תנועה מספר i:



$$s' = R_3(x) O_1(x) O_4(x) O_4(z) O_1(y) W_2(y) R_2(z)$$

ידוע גם כי ה- precedence graph של s הוא כדלהלן:

השלם את הטבלה הבאה:

הפעולה	האם הפעולה חייבת להיות W?	הסבר במשפט קצר
$O_1(x)$	Y	כדי שיהיה קונפליקט עם $R_3(x)$
$O_4(x)$	Y	כדי שיהיה קונפליקט עם $R_3(x)$
$O_4(z)$	Y	כדי שיהיה קונפליקט עם $R_2(z)$
$O_1(y)$	N	בשני המקרים קונפליקט עם $W_2(y)$

$$s' = R_3(x) W_1(x) W_4(x) W_4(z) R_1(y) W_2(y) R_2(z)$$

ג. (5 נק') התבונן בתהליך הביצוע הבא ( .... מציין חלק של התהליך שהושמט מהתצוגה,  $C_j$  מציין את ביצוע ה-COMMIT של תנועה j):

$s'' = W_3(x) \ R_2(x) \ W_2(y) \ c_2 \ R_4(y) \ W_4(z) \ R_3(z) \ W_3(z) \ c_3 \dots$

(1) האם התהליך  $s''$  הוא recoverable? הקף בעיגול כן לא  
הסבר קצר: לא.

תנועה 2 ביצעה COMMIT למרות שקראה את X מתנועה 3 שטרם ביצעה COMMIT.

תנועה 3 ביצעה COMMIT למרות שקראה את Z מתנועה 4 שטרם ביצעה COMMIT.

(2) האם התהליך  $s''$  יוצר ע"י מתזמן העובד על פי אלגוריתם two phase locking? הקף בעיגול כן לא

הסבר קצר: לא, תנועה 3 שחררה מנעול על x עבור תנועה 2. תנועה 4, מאוחר יותר, שחררה מנעול על z עבור תנועה 3 שרכשה אותו. לכן, תנועה 3 איננה דו-פאזית.

ד. (4 נק') התבונן בתהליך הביצוע החלקי הבא המתוזמן ע"י אלגוריתם 2-Phase Locking (הנח שמספר התנועה הוא תאריך התחלתה, קטן יותר אומר שהתאריך מוקדם יותר):

$s = RL_2(x)R_2(x)WL_2(y)W_2(y)WL_3(z)W_3(z)RU_2(x)WL_1(x)W_1(x) (*)$

בנקודה (\*), תנועות 1, 2 ו-3 עדיין פעילות ותנועה מספר 3 מתעתדת לכתוב ל- x ולכן מבקשת  $WL_3(x)$  מהמתזמן. מה יקרה תחת אלגוריתם Wait-die? בחר בתשובה אחת בלבד.

a. תנועה מספר 3 תחכה ותנועה מספר 1 תופסק (restarted).

b. תנועה מספר 3 תופסק (restarted).

c. תנועה מספר 3 תחכה ותנועה מספר 2 תופסק (restarted).

d. תנועה מספר 3 תקבל מידית את המנעול על x ותמשיך בפעילותה.

e. התשובות א-ד אינן נכונות. התשובה הנכונה היא:

**מקום לתשובות נוספות:**

שאלה:	סעיף:

שאלה:	סעיף:

**מקום לתשובות נוספות:**

שאלה:	סעיף:
שאלה:	סעיף:

שאלה:

סעיף: