

מבחן במערכות בסיסי נתונים
פרופ' טובה מילוא

סמסטר א', מועד א'
27 באפריל, 2008
חומר סגור, שלוש שעות
סה"כ 100 נקודות

נא לכתוב את כל התשובות על הדפים
(המבחן כולל 12 דפים, כולל דף זה)

מס מחברת: _____ מס סטודנט: _____

חלק	ציון
1	
2	
3	
סה"כ	

1. [40 נקודות]

חברת ביטוח מחזיקה מסד נתונים עם הסכמה הבאה:

Property(name, city)

Policy(pid, premium, propertyName, agent)

Claim(cid, pid, amount, year)

השדה premium מכיל את הפרמיה השנתית שהלקוח(ה) משלם(ת) כל שנה לחברת הביטוח, שערכה זהה בכל השנים. השדה amount מתאר תשלום חד פעמי, שחברת הביטוח משלמת ללקוח(ה), אם וכאשר הלקוח(ה) מגיש(ה) תביעה עבור פוליסת הביטוח.

א. נניח שסוכן ביטוח מטפל בתביעה שמספרה cid=03492 עבור לקוח מסוים. הסוכן היה רוצה לדעת אם היו תביעות אחרות עבור אותה הפוליסה שהוגשו באותה שנה. כתבו שאילתת SQL המחשבת זאת. השאילתה צריכה להחזיר את ה-cid-ים של אותן תביעות אחרות.

ב. כתבו את אותה שאילתה באלגברה רלציונית.

ג. כתבו שאילתת SQL שמחזירה לכל פוליסת ביטוח את השנים שבהן סכום התביעות היה יותר גבוה מהפרמיה השנתית של הפוליסה. התשובה צריכה להיות קבוצה של זוגות (pid, year)

ד. התבוננו בזוגות Q1,Q2 של השאילות הבאות. לכל זוג סמנו מי מהמשפטים הבאים מתקיים.

- (A) שתי השאילות מחזירות תמיד בדיוק אותה תשובה,
(B) התשובות המוחזרות ע"י Q1 מוחזרות גם ע"י Q2, אבל יתכן ש-Q2 תחזיר ערכים שאינם מוחזרים ע"י Q1,
(C) התשובות המוחזרות ע"י Q2 מוחזרות גם ע"י Q1, אבל יתכן ש-Q1 תחזיר ערכים שאינם מוחזרים ע"י Q2,
(C) אף אחד מהנ"ל.

לדוגמה אם

Q1=Select year from Claim where amount > 100
Q2=Select year from Claim where amount > 50

עליכם לענות (B)

i.

Q1 = Select Distinct P.pid
From Policy P, Claim C
Where P.pid = C.pid and C.amount = 300

Q2 = Select Distinct P.pid
From Policy P, Claim C1, Claim C2
Where P.pid = C1.pid and P.pid = C2.pid and
C1.amount > 100 and C2.amount < 2000

(A) (B) (C) (D)

ii.

Q1= **Select Distinct** A.name
From Property A, Policy P, Claim C
Where A.name=P.propertyName and
P.pid=C.pid and C.amoun > 5000

Q2= **Select Distinct** A.name
From Property A, Policy P, Claim C1, ClaimC2
Where A.name=P.propertyName and
P.pid=C1.pid and C1.amount=10000 and
P.pid=C2.pid and C2.amount < 300

(A) (B) (C) (D)

iii.

Q1= **Select Distinct** A.name
From Property A, Policy P, Claim C1, ClaimC2
Where A.name=P.propertyName and
P.pid=C1.pid and C1.year=1980 and
P.pid=C2.**pid** and C2.year=2002

Q2= **Select Distinct** A.name
From Property A, Policy P1,Policy P2, Claim
C1, Claim C2
Where A.name=P1.propertyName and
P1.pid=C1.pid and C1.year=1980 and
A.name=P2.propertyName and
P2.pid=C2.pid and C2.year=2002

(A) (B) (C) (D)

2. [30 נקודות]

א. נתונה הטבלה $R(A, B, C, D, E)$ עם התלויות הפונקציונאליות הבאות:

$AB \rightarrow C$

$BD \rightarrow A$

$E \rightarrow BD$

פרקו את הטבלה לטבלאות ב-BCNF ונמקו את תשובתכם. התשובה צריכה להכיל:

- רשימת טבלאות והשדות בכל טבלה
- המפתחות בכל אחת מהטבלאות

ב. עבור קבוצה X של שדות, נגיד ש X סגורה (ביחס לקבוצה נתונה של קבוצות פונקציונאליות) אם $X = X + X$. שימו לב שמידע שקבוצת שדות מסוימת הינה סגורה נותן לנו גם ידע מסוים על הצורה האפשרית של התלויות הפונקציונאליות הקיימות.

נניח שנתונה טבלה $R(A,B,C,D)$ שקבוצת התלויות הפונקציונאליות שלה אינה ידועה. לכל אחת מהטענות הבאות, תנו קבוצה אפשרית של תלויות פונקציונאליות שעבורה הטענה מתקיימת.

(1) כל הקבוצות האפשריות של שדות הן סגורות.

(2) הקבוצות היחידות שסגורות הן $\{A,B,C,D\}$ ו- $\{\}$.

(3) הקבוצות היחידות שסגורות הן $\{A,B,C,D\}$, $\{A,B\}$ ו- $\{\}$.

3. [30 נקודות]

נתונות שתי טבלאות $R(A,B)$ $S(\underline{C},D)$ כש C הוא המפתח ברלציה S ,

והשאלתה הבאה: $\sigma_{A=5}(R) \bowtie_{B=C} S$

נניח שגודל רשומה (tuple) בכל אחת מהרלציות הוא כגודל בלוק, ושכל אחת מהן מכילה 10000 רשומות (ולכן גודלה הוא 10000 בלוקים).

על השדה C הוגדר clustered index ברלציה S . גישה לכל ערך באינדקס לוקחת 4 פעולות I/O (כולל זמן הגישה להביא את הרשומה עצמה).

נתבונן בשלושת השיטות האפשריות הבאות לחישוב ה-join בשאלתא:

- nested loop join
- hash join
- index join

חשבו הערכה של מחיר ה-I/O לכל שיטת join בכל אחד מארבעת המקרים הבאים, וציינו בטבלה, בכל אחד מארבעת המקרים, את השיטה הצפויה להיות הטובה ביותר. נמקו את תשובתכם.

$Values(R.A)$ מציין כאן את מספר הערכים השונים בשדה A בטבלה R . M מציין את גודל הזיכרון.

	$Values(R.A) = 1$	$Values(R.A) = 10000$
$M = 101$		
$M = 5002$		

$M = 101$, $\text{Values(R.A)} = 1$

א.

$M = 5002$, $\text{Values(R.A)} = 1$

ב.

$M = 101$, $\text{Values(R.A)} = 10000$

.λ

$M = 5002$, $\text{Values(R.A)} = 10000$

.7