

סמסטר אביב 2011

פרופ' יוהן מקובסקי

מרצה :

גב' נעמה טפר
מר יבגני אברמוביץ'

מתרגלים:

**מערכות מסדי נתונים
236363**

מועד ב' (19 בספטמבר 2011)

| <u>מס' ת.ז.</u> |
|-----------------|
| |

| מספר השאלה | נקודות |
|------------|--------|
| 1 | /26 |
| 2 | /36 |
| 3 | /25 |
| 4 | /16 |
| סה"כ | /103 |

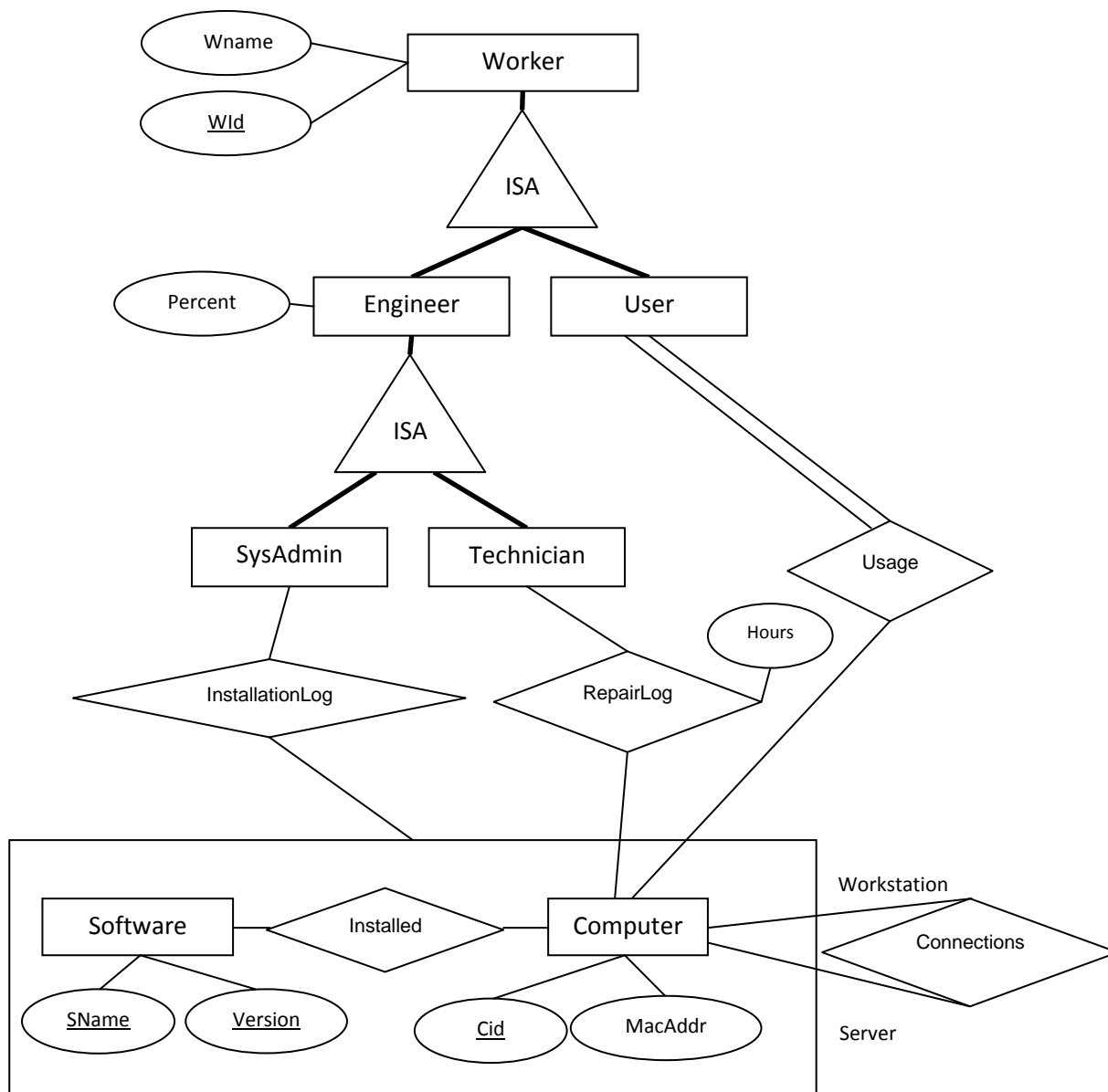
הנחיות לנבחן

1. כתבו את התשובות אך ורק בטופס הבחינה, המחברת מיועדת לטיוטה בלבד.
2. כל חומר עזר כתוב על נייר מותר בשימוש.
3. אין לקבל או להעביר חומר עזר כלשהו בזמן הבחינה.
4. בבחינה ארבע שאלות. יש לענות עליהן במלואן.
5. יש לנמק את התשובות בקצרה. תשובות לא מנומקות לא תתקבלנה.
6. אין להשתמש בפונקציות שלא נלמדו בתרגול ולא מופיעות בשקפים של הקורס.
7. משך הבחינה שלוש שעות. תכננו את הזמן בהתאם.
8. הבחינה כוללת 13 דפים כולל דף זה. נא לוודא שיש בידכם את כל הטופס.
9. הניקוד אינו משקף את קושי השאלה.

בהצלחה

שאלה 1: ERD (26 נקודות)

נתונה דיאגרמת ה-ERD הבאה המתארת מידע על מערכת של חוות מחשבים.



פירוט הישויות:

Worker – מייצג עובד בחוות המחשבים. לכל עובד יש שם (WName) ומספר תעודת זהות (WId) ייחודי. קיימים כמה סוגים של עובדים שונים:

User – משתמש

Engineer – מהנדס – עבור כל מהנדס יש מידע לגבי אחוז המשרה בה מועסק (Percent)

Technician – טכנאי – אחראי על תיקוני חומרה במחשבים

SysAdmin – מנהלן – תפקידו בעיקר לבצע התקנות על מחשבים

Computer – מייצג מחשב השייך לחווה. לכל מחשב יש מספר משהה ייחודי (Cid) וכתובת ה-Mac (MacAddr)

Software – יישות שמתארת תוכנה. לכל תוכנה יש שם שמתאר אותה (SName) ומספר הגרסה (Version)

פירוט הקשרים:

Installed – קשר שמתאר אילו תכונות מוקנות באילו מחשבים.
Connections – קשר שמייצג רשת מחשבים בחווה, כלומר תחנות עבודה (Workstations) המחוברות לשרתים (Servers).
Usage – קשר המייצג מול אילו מחשבים עובדים המשתמשים.
InstallationLog – קשר שמייצג אצ כל ההתקנות שבוצעו בחווה.
RepairLog – קשר שמתאר את כל העבודות אשר בוצעו ע"י הטכנאים על מחשבי החווה והזמן שגולה כל עבודה (Hours).

תזכורת: תכונות של קשרים שאינן מסומנות בקו תחתון אינן חלק מהמפתח של הקשר.

1. (8 נק') בהנתן טבלה, רשמו כיסוי לכל התלויות הפונקציונאליות המתקיימות בין תכונותיה. הניחו כי בתרגום קשרי ה-IS-A לטבלאות, יוצרו טבלאות רק עבור טיפוס היישות המוכללים.
א. (4 נק') RepairLog ⋈ (TechnicianUSysAdmin)

Wid,cid -> hours

Wid -> percent, wname

ב. (4 נק') Usage ⋈ User

Wid -> wname

2. (6 נק') ענו על השאלות הבאות ונמקו בקצרה. תשובות לא מנומקות לא תתקבלנה.
א. (3 נק') האם אותו מנהלן (SysAdmin) יכול להתקין את אותה התוכנה (בעלת אותו שם ואותה גרסה) פעמיים על אותו מחשב?

לא מכיוון והמנהלן, המחשב, שם וגרסת התוכנה מהווים את המפתח של הקשר installationLog

ב. (3 נק') האם טכנאי יכול להיות גם משתמש?

לא, מכיוון ויש משולש ISA יחיד בין עובד, מהנדס ומשתמש.

3. (12 נק') בסעיפים הבאים, יש לשנות את ה-ERD המקורי (ללא תלות בסעיפים האחרים), כך שהתנאי יתקיים. יש להסביר את השינויים. **שימו לב:**

- למעט השינוי הנדרש, אסור להוסיף אילוצים נוספים שלא היו קיימים ב-ERD המקורי.
- מותר לתאר את השינוי במילים אם התיאור ברור וחד משמעי.

א. (6 נק') לא ייתכנו שני מחשבים (וגם לא מחשב אחד) שבהם מותקנות גרסאות שונות לאותה התוכנה.

יש להוריד את version מהמפתח

ב. (6 נק') לא ייתכן מצב שבו באותו מחשב מותקנות גרסאות שונות של אותה התוכנה.

יש להפוך את הקשר installed לחד ערכי ל-software

שאלה 2 : שפות שאילתא (36 נקודות)

עבור השאלות הבאות, להלן סכמה אפשרית לחלק ממסד הנתונים (כולל סדר המשתנים עבור ש.DRC).

user(uid, wname)
technician(tid, wname, percent)
connections(server, workstation)
repairLog(tid, cid, hours)
usage(uid,cid)

הסבר :

User – טבלת המשתמשים. Uid – מזהים של עובדים שהם משתמשים.

Technician – טבלת טכנאים. tid – מזהים של עובדים שהם טכנאים.

Connections – server ו- workstation הם מזהים של מחשבים.

1. (8 נק') כתבו שאילתה ב-SQL המחזירה לכל טכנאי את שמו, מספר המחשבים עליהם עבד ואת מספר השעות הממוצע שעבד על כל מחשב (מבין המחשבים בהם טיפל). לדוגמא, אם טכנאי טיפל ב-3 מחשבים ועבד בסה"כ 30 שעות, אזי מספר השעות הממוצע שעבד על כל מחשב הוא 10 שעות.

```
SELECT user.wname,COUNT(DISTINCT  
repairLog.cid),SUM(repairLog.hours)/COUNT(DISTINCT repairLog.cid)  
  
FROM user,technician,repairLog  
  
WHERE user.uid = technician.tid AND technician.tid = repairLog.tid  
  
GROUP BY user.uid,user.wname
```

2. (20 נק') נתונה השאילתה הבאה :

$$Q = \{u \mid \text{usage}(u) \wedge \forall c, s (\text{connection}(s, c) \rightarrow \text{usage}(c, u))\}$$

א. (8 נק') כתבו במילים מה מבצעת השאילתה Q

השאילתה מחזירה את מזהי המשתמשים שהשתמשו בכל המחשבים מסוג
.workstation

ב. (8 נק') תרגמו את השאילתה Q הכתובה ב-DRC ל-RA.

$\text{usage} \div \rho_{\text{workstation} \rightarrow \text{cld}}(\text{connections})$

ג. (4 נק') האם השאילתה Q בטוחה? נמקו.

3. (8 נק') כתבו שאילתה ב-Datalog המגדירה פרדיקט $busy(s)$ שיקבל ערך true עבור שרתים עמוסים. שרת מוגדר עמוס אם יותר מ- n מחשבים מחוברים אליו. אסור להשתמש בשלילה.

$last(S,W,1) \leftarrow connection(S,W).$

$last(S,W,N) \leftarrow connection(S,W), connection(S,W'), load(S,W',N-1), W' < W.$

$busy(S) \leftarrow last(W,S,N).$

שאלה 3 : תלויות ופירוקים (25 נקודות)

נתונות הסכמות וקבוצות התלויות הבאות :

$$R_1[C, S, Z], F_1 = \{ CS \rightarrow Z \}$$

$$R_2[C, Z_c], F_2 = \{ C \rightarrow Z_c, Z_c \rightarrow C \}$$

$$R_3[S, Z, Z_c], F_3 = \{ Z \rightarrow Z_c S, Z_c S \rightarrow Z \}$$

$$R[C, S, Z, Z_c], F = F_1 \cup F_2 \cup F_3$$

1. (5 נק') מי משלושת (R_i, F_i) הן ב-BCNF?

שלושתן.

לכל i , כל התלויות ב- F_i הן תלויות במפתח של R_i .

2. (5 נק') בהנתן $R[C, S, Z, Z_c]$ ו- $F = F_1 \cup F_2 \cup F_3$ האם (R, F) ב-BCNF?

לא. $C \rightarrow Z_c$ אינה תלות בפתח של R ($C_F^+ = C$).

3. (5 נק') האם מתקיים השוויון $\pi_{R1}F = (F_1)^+$?

לא.

$$Z \rightarrow S \in \pi_{R1}(F) \setminus (F_1)^+$$

4. (5 נק') בהנתן $r \models F$ רלציה תחת הסכמה R , האם מתקיים השוויון

$$r = \pi_{CSZ}r \bowtie \pi_{CZc}r \bowtie \pi_{SZZc}r$$

$\{R_1, R_2\}$ הוא פירוק משמר מידע, כי $F = (R_1 \cap R_2) \rightarrow R_2$ (לפי התלות $C \rightarrow Z_c$).

$$r = \pi_{CSZ}r \bowtie \pi_{CZc}r, \text{ לכן}$$

$$r = r \bowtie \pi_{SZZc}r \text{ לכל } r.$$

$$\pi_{CSZ}r \bowtie \pi_{CZc}r \bowtie \pi_{SZZc}r = r \bowtie \pi_{SZZc}r = r, \text{ מכאן,}$$

5. (5 נק') האם $\rho = \{R_1, R_2, R_3\}$ הוא פירוק BCNF של (R, F) ?

כן.

לכל i , כל תלות לא טריביאלית ב- $\pi_{Ri}(F)$ היא תלות במפתח של R_i .

שאלה 4 : XML (16 נקודות)

עבור הסעיפים הבאים, להלן מסמך DTD של חוות מחשבים :

```
<?xml version="1.0" encoding="UTF-8"?>
<!ELEMENT farm (user*,computer*)>
<!ELEMENT user (#PCDATA)>
<!-- מזהה המשתמש והמחשבים שיש לו גישה אליהם -->
<!ATTLIST user
    uid ID #REQUIRED
    computers IDREFS #REQUIRED
>
<!ELEMENT computer (#PCDATA)>
<!-- מזהה המחשב והמחשבים אליהם הוא מחובר -->
<!ATTLIST computer
    cid ID #REQUIRED
    macAddr CDATA #REQUIRED
    connectedTo IDREFS #REQUIRED
>
```

1. (4 נק') האם מחשב יכול להיות מחובר לעצמו? לא להיות מחובר לאף מחשב?

מחשב יכול להיות מחובר לעצמו. מחשב חייב להיות מחובר לפחות למחשב אחד.

2. (4 נק') שנו את ה-DTD כך שלכל מחשב יוכל לגשת משתמש אחד בדיוק.

יש להוסיף תכונה למחשב :

User IDREF #REQUIRED

3. (8 נק') מתייחס ל-DTD המקורי: רשמו שאילתת XPath 1.0 שמוודאת שלמשתמש הראשון שמופיע במסמך יש גישה לכל המחשבים – מחזירה את המזהה שלו אם יש לו גישה, וקבוצה ריקה אם אין. השתמשו בפונקציית id().

/farm/user[position()=1 and
count(id(@computers))=count(./ancestor::farm/computer)]/@uid