

סמסטר חורף התשע"ב

פרופ' ח' אלדר פישר

מרצה :

מר יבגני אברמוביץ'  
מר דימה אלנבוגן

מתרגלים :

**מערכות מסדי נתונים  
236363**

מועד א' (ה' באדר התשע"ב, 28 בפברואר 2012)

<u>מס' ת.ז.</u>

מספר השאלה	נקודות
1	/27
2	/20
3	/24
4	/29
סה"כ	/100

**הנחיות לנבחנים**

1. כתבו את התשובות אך ורק בטופס הבחינה, המחברת מיועדת לטיוטה 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 מותר בשימוש.
5. אין לקבל או להעביר חומר עזר כלשהו בזמן הבחינה.
6. בבחינה ארבע 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 שאילתה.**
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 (ללא דף הסריקה) כוללת 12 דפים, כולל דף זה. נא לוודא שיש בידכם את כל הטופס.
11. כאשר הניקוד של 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**

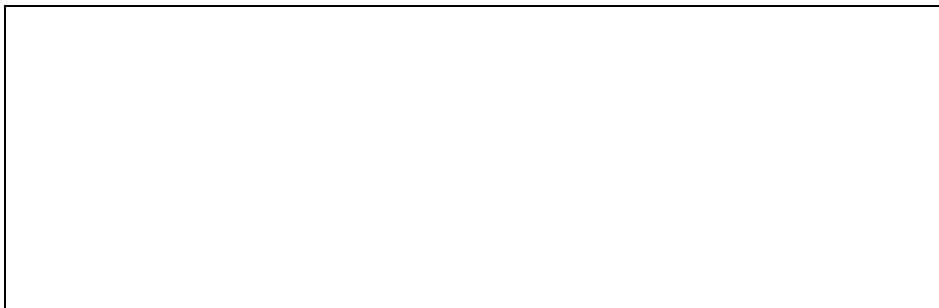
## שאלה 1 - ERD (27 נק')

נתונים מאפיינים של ממלכת גמדים:

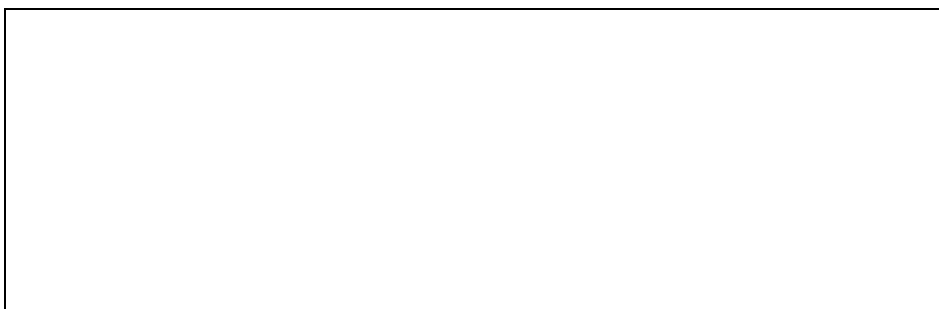
לכל גמד יש שם ייחודי, שבט וגובה. כל גמד-כורה עובד במכרה יחיד וכרה עד היום כמות מסוימת, הנמדדת בקילוגרמים, של משאבים. לכל מכרה יש מיקום ייחודי (נתון ע"י תיאורו), ומשאב אחד בדיוק שניתן לירות בו. לכל משאב יש שם ייחודי ומחיר לקילוגרם. מחיר המשאב אחיד בכל המכרות. גמד יכול להיות גמד-כורה ו/או גמד-מנהל. לכורה חייב להיות מנהל אחד לפחות, ולמנהל חייב להיות כורה אחד לפחות מנוהל על ידיו.

א. (6 נק') ענו ונמקו בקצרה על השאלות הבאות:

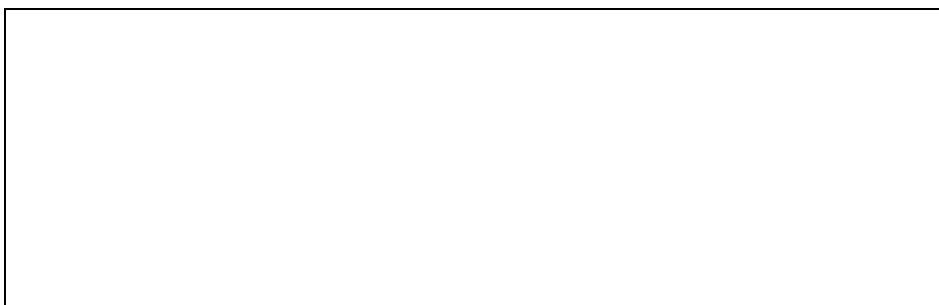
i. האם יכול להיות כורה המנוהל ע"י שני מנהלים בו זמנית?



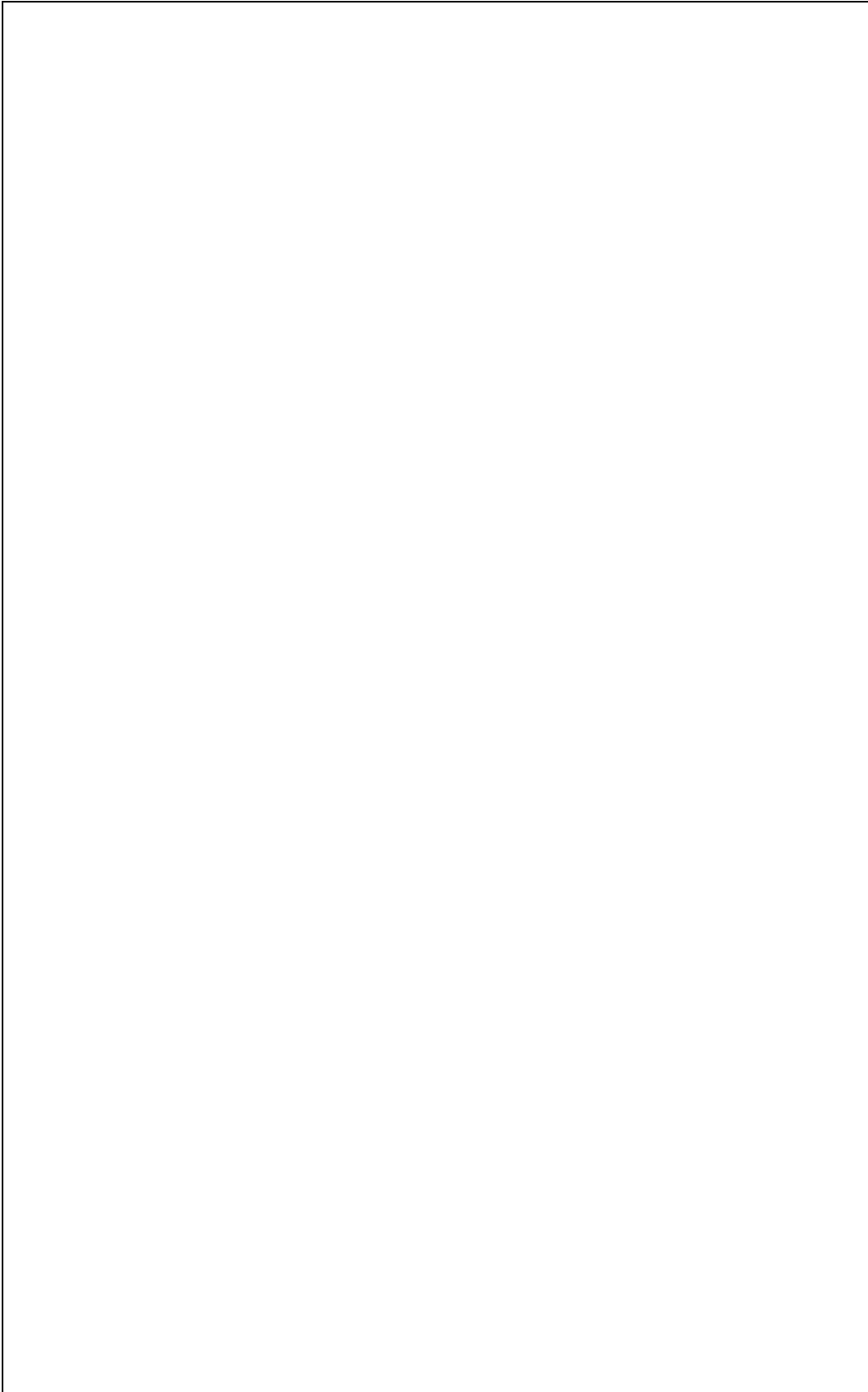
ii. האם יכול להיות כורה שכרה שני משאבים שונים?



iii. האם יכול להיות מנהל שכוריו כרו שני משאבים שונים?



ב. (14 נק') שרטטו ERD המתאר את הממלכה.



ג. (7 נק') בכל אחד **מטיפוסי הקשר** ששרטטתם: כתבו כיסוי לקבוצת כל התלויות הפונקציונאלית המתקיימות בקשר.

## שאלה 2 – שאלות מידע (20 נק')

נתונות הרלציות הבאות, אשר אינן בהכרח תואמות את התיאור וה-ERD שבניתם בשאלה הקודמת:

gnome(name, gtribe) – גמד ששמו name (הניחו שאין שתי שורות ב-gnome עם אותו שם גמד) הוא חבר השבט gtribe.

mine(mname, mresource, mKgPrice) – מכרה ששמו mname (הניחו שאין שתי שורות ב-mine עם אותו שם מכרה), המשאב שלו הוא mresource, והמחיר לק"ג של המשאב הוא mKgPrice.

works(gname, mname, amountKg) – גמד ששמו gname (הניחו שאין שתי שורות ב-works עם אותו שם גמד) עובד במכרה ששמו mname, וכרה עד עכשיו amountKg ק"ג מהמשאב של המכרה.

managedBy(ename, mname) – גמד ששמו ename מנוהל ע"י גמד ששמו mname.

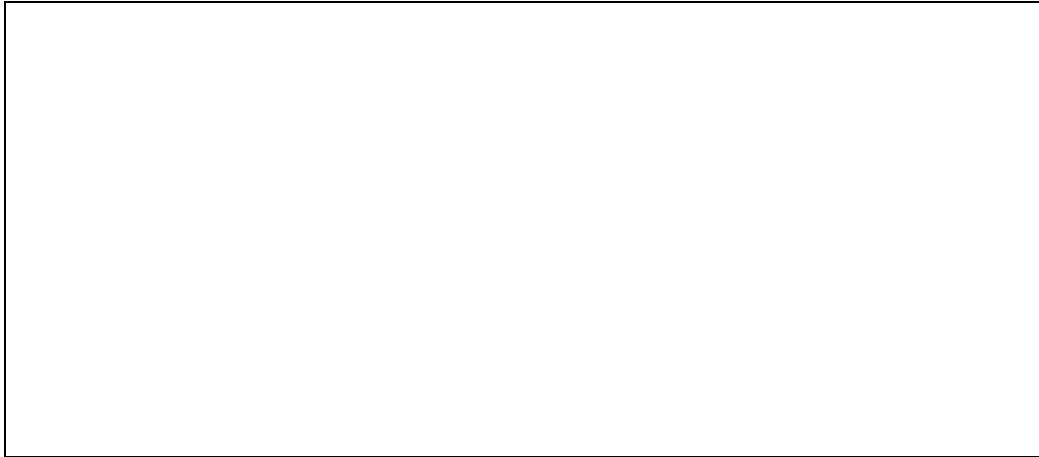
א. (12 נקודות) **גמד מנהל זוטרי מצטיין** הינו גמד שיש תחתיו 5 עובדים או פחות, כך שעובדיו כרו ביחד את המשקל המרבי של משאבים (מבין אלה שיש תחתם 5 גמדים או פחות). כתבו שאילתת SQL המוצאת את השבט אליו שייד המנהל המצטיין.

אם יש כמה מנהלים מצטיינים משבטים שונים, יש להחזיר את כל השבטים.

מותר ליצור ולהשתמש במבטים, אבל רק אם הם אינם מיותרים. מבט יחשב למיותר אם שאר הביטוי משתמש בו לא יותר מפעם אחת. אין ליצור טבלאות חדשות בשום מקרה.

ב. (8 נקודות) – כתבו שאילתת DRC המוצאת מנהלים המקיימים את התנאי הבא: לכל משאב (שנכרה באיזשהו מכרה) יש לפחות שני עובדים של המנהל שכורים אותו.

- מותר להגדיר קבוצות עזר באמצעות שאילתות DRC בלבד.
- אסור שהשאילתה תהיה תלוית תחום (domain-dependent). אין צורך להוכיח ששאילתה אינה תלוית תחום.



### שאלה 3 - נושאים מעורבים (24 נק')

שימו לב: אין תלות בין סעיפי השאלה (א', ב' ו-ג') וכן אין תלות ביניהם לשאלות קודמות.

א. (8 נק') נתונה טבלה  $Values(Name, Val)$ . הטיפוס של  $Name$  הוא מחרוזת שאורכה מוגבל ל-64 תווים. הטיפוס של  $Val$  הוא  $REAL$ . אף אחד מהערכים בטבלה לא יכול להכיל ערך  $NULL$ .

i. (3 נק') כתבו ביטוי  $SQL$  היוצר את הטבלה  $Values$  על סמך כל העובדות הנ"ל.

ii. (5 נק') בסעיף זה הנכם מתבקשים לכתוב שאילתת  $SQL$  שמחשבת את הערך השני בגודלו מכל הערכים השונים של העמודה  $Val$  (כלומר, תוך השמטת כפילויות). במידה ויש בטבלה פחות משני ערכים שונים של  $Val$ , על השאילתא להחזיר  $NULL$ . אין ליצור טבלאות זמניות או מבטים.

ב. (8 נק') נתונה רלציה  $r$  מעל סכמה  $R(A,B,C)$ . לאחר פירוק הסכמה לתתי-סכמות  $S(A,B)$  ו- $T(B,C)$  התקבלו רלציות  $s$  ו- $t$  בהתאמה.

i. (2 נק') נמקו בקצרה מדוע על סמך הנתונים הנ"ל אי-אפשר באופן כללי לדעת האם הפירוק של  $R$  ל- $S$  ו- $T$  משמר מידע.

ii. (6 נק') נסמן ב- $r'$  את הרלציה שעבורה מתקיימים כל אחד משני התנאים הבאים :

- לאחר הפירוק מ- $r'$  היו מתקבלות בדיוק אותן הרלציות  $s$  ו- $t$  בהתאמה.
- הפירוק היה משמר את התוכן של  $r'$ .

הביעו את  $r'$  כתלות ב- $r$  בלבד (ללא שימוש ב- $s$  או  $t$ ) כביטוי באלגברה רלציונית (RA).

ג. (8 נק') בשאלה זאת תכתבו תוכנית Datalog (ללא שלילות) בטוחה.

הניחו כי בהינתן סכמה  $R$  וקבוצת תלויות פונקציונליות לא טריביאליות  $F$  מעל הסכמה נתונים מראש הפרדיקטים המפורשים הבאים :

- $dep1(A, C)$  שמתקיים אם ורק אם  $A \rightarrow C \in F$
- $dep2(A, B, C)$  שמתקיים אם ורק אם  $AB \rightarrow C \in F$
- $inR(A)$  שמתקיים אם ורק אם  $A \in R$

הניחו גם שאין ב- $F$  תלויות למעט אלו המתוארות ע"י הפרדיקטים  $dep1$  ו- $dep2$ . בהסתמך על הפרדיקטים המפורשים עליכם לכתוב את הכללים לפרדיקט סתום  $aClose(A, B)$  שמתקיים אם  $B \in A_F^+$ .



## שאלה 4 - XML (29 נק')

נתון ה- DTD הבא, המתאר דו"חות הניתנים ע"י המנהלים על העובדים. אלמנט המסמך הוא file.

**הבהרה בזמן המבחן: הניחו כי התאריך הוא מספר שלם.**

```
<!ELEMENT file (review)*>
<!ATTLIST file updated CDATA #REQUIRED>
<!ELEMENT review (manager, worker*, rtext)>
<!ATTLIST review date CDATA #REQUIRED
            id ID #REQUIRED>
<!ELEMENT manager (#PCDATA)>
<!ELEMENT worker (#PCDATA)>
<!ELEMENT rtext (#PCDATA|emph|point)*>
<!ELEMENT emph (#PCDATA)>
<!ELEMENT point EMPTY>
<!ATTLIST point to IDREF #REQUIRED>
```

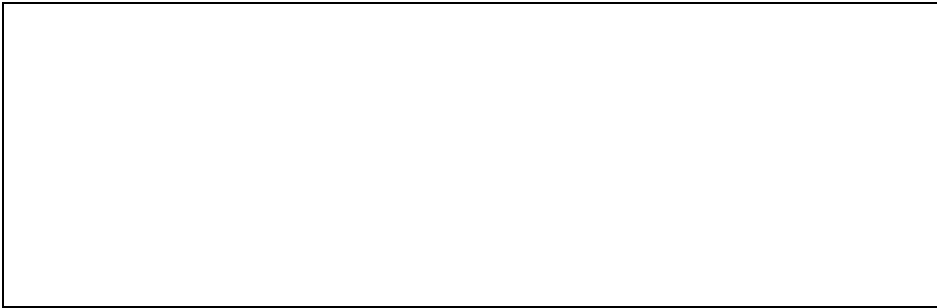
הסבר ל- DTD:

הקובץ מכיל דו"חות (review) שנכתבו ע"י מנהלים באתר. לכל דו"ח מצוין המנהל (manager), העובדים שאליהם הוא מתייחס (צומת worker לכל אחד), וטקסט הדו"ח. הטקסט יכול להכיל הדגשות (emph) ומצביעים לדו"חות אחרים (point). לכל דו"ח מזהה ייחודי שיכול לשמש להצבעות עליו. כמו כן שומרים תאריך עדכון אחרון של הקובץ (תכונה updated של הצומת file) ותאריך כתיבה של כל דו"ח (תכונה date של צומת הדו"ח). אין רישום נפרד של פרטי העובדים והמנהלים פרט למה שמופיע בדו"חות.

א. (8 נק') ענו ונמקו בקצרה על השאלות הבאות באשר ל- DTD.

i. האם יתכן שמנהל ייתן דו"ח בלי לציין עובדים?

ii. האם יתכן שאותו מנהל יתן דו"ח על אותו עובד יותר מפעם אחת?



iii. האם יתכן שדו"ח יכלול הצבעה על עצמו?



iv. האם תתכן הצבעה באמצע קטע מודגש?



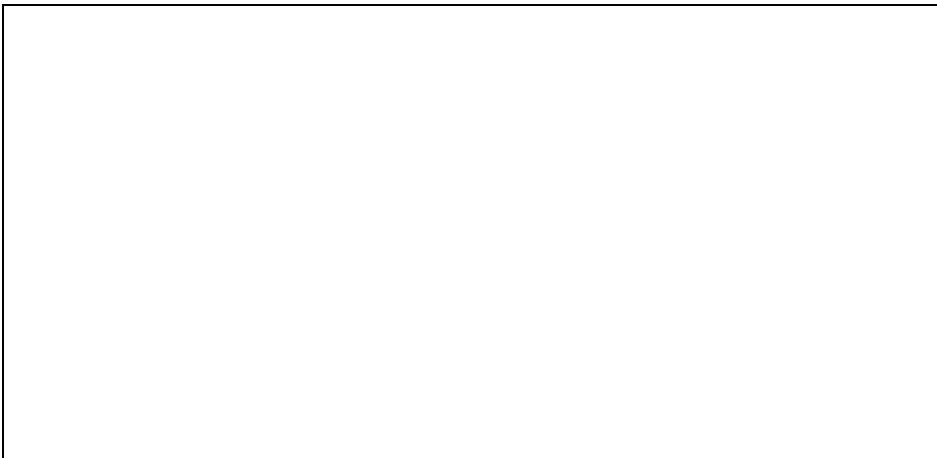
ב. (9 נק') דו"ח יקרא "תמוה" אם התאריך שלו הוא אחרי תאריך עדכון הקובץ (התכונה של file), או לפני תאריך של דו"ח כל שהוא שהוא מצביע עליו. כתבו שאילתת **XPath1.0** המחזירה את קבוצת צמתי ה-review של כל הדוחות התמוהים.



ג. 12 נק') נתונה שאילתת ה-XQuery1.0 הבאה (file.xml הוא הקובץ מהשאלות הקודמות):

```
document {<list updated="{ doc("file.xml")/file/@updated }">
for $w in distinct-values(doc("file.xml")//worker)
order by $w
return (<worker><name>$w/text()</name>
      for $r in doc("file.xml")/file/review
      where $r/worker=$w
      order by $r/@date
      return ($r/manager,<rev>($r/rtext/text() union $r/rtext/emph)</rev>)
</worker>)
</list>}
```

i. כתבו במילים מה השאילתה מחזירה. יש לציין את כל הפרטים הרלוונטיים.



ii. כתבו DTD המתאים למסמך שהשאילתה מחזירה. יש להשתמש במרב האילוצים אשר תמיד מתקיימים בפלט השאילתה עבור מסמך file.xml תקין. עבור צמתים המכילים בן #PCDATA בלבד אפשר לציין זאת במרוכז במקום לכתוב שורה לכל אחד מהם.



### מקום נוסף לתשובות

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

שאלה: \_\_\_\_\_ סעיף: \_\_\_\_\_

שאלה: \_\_\_\_\_ סעיף: \_\_\_\_\_