



הטכניון - מכון טכנולוגי לישראל
הפקולטה למדעי המחשב

אביב תשס"ח
8, ספטמבר 2008

ד"ר אמיר שפילקה
נדב שרגאי

מערכות מסדי נתונים – 236363

מועד א'
הזמן: 3 שעות
במבחן זה 10 עמודים

עדכונים שנכתבו על הלוח במהלך הבחינה כתובים באדום

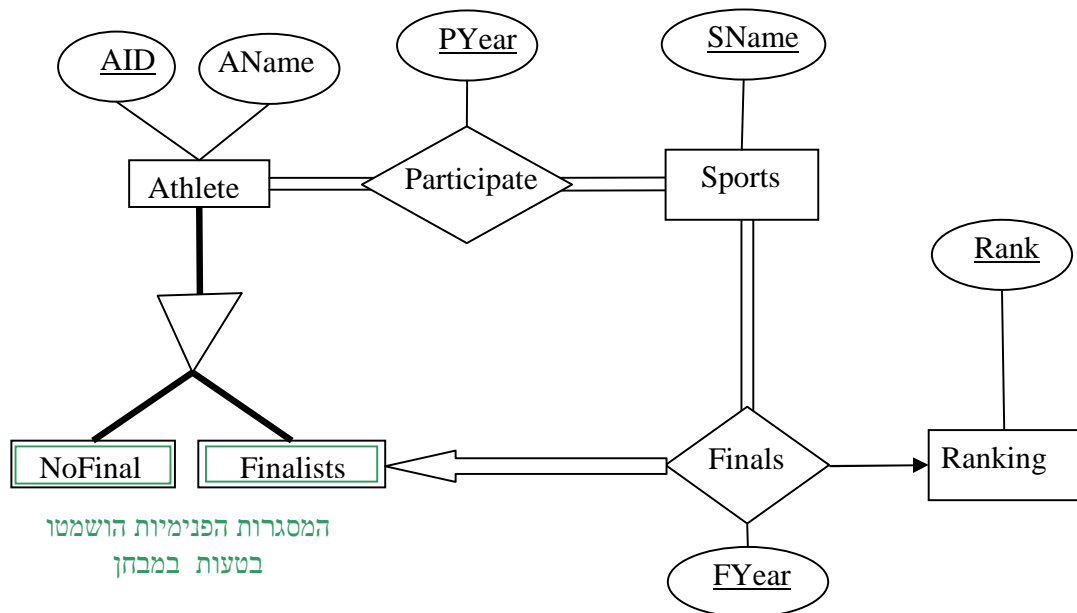
| שאלה | נקודות |
|----------------------------------|--------|
| שאלה 1 – ERD | 20 |
| שאלה 2 – שפות שאילתא | 36 |
| שאלה 3 – פירוקים וצורות נורמליות | 24 |
| שאלה 4 – XML | 20 |
| סה"כ | 100 |

הנחיות:

- יש לענות על כל השאלות **בטופס הבחינה** מחברות הטיוטה לא תיאספנה.
 - חומר עזר מותר: דף נוסחאות A3 יחיד מודפס משני הצדדים.
 - אין להחזיק מכשיר אלקטרוני כלשהו לרבות מחשב כיס.**
 - קראו היטב את ההוראות שבתחילת כל שאלה ואת ההסברים לסכמות.
 - מומלץ שתכננו היטב את זמנכם, **לא תינתנה הארכות.**
 - ערעורים יש להגיש תוך שבועיים ממועד פרסום התוצאות.
 - לא יתקבלו ערעורים בנוסח "בדיקה מחמירה מדי".
- בהצלחה

שאלה 1 – ERD (20 נק')

נתונה סכמת ה-ERD הבאה המתארת ספורטאים שהגיעו לגמרים באולימפיאדות ומיקומם הסופי.



הסבר לסכמה:

הסימון מתאר קו כפול (השתתפות מלאה) וחץ (חד ערכיות).

ישויות:

Athlete - ישות המתארת אתלטים. אתלט מאופיין ע"י מזהה ייחודי (**AID**) ושם (**AName**).
 האתלטים יכולים להיות אתלטים שהגיעו לגמר במקצוע כלשהוא (**Finalists**) או אתלטים שלא הגיעו לאף גמר (**NoFinal**).
Ranking - ישות המתארת את המיקום הסופי של האתלט בגמר. המיקום מאופיין ע"י הדירוג (**Rank**) שהוא שדה מספרי.
Sports - ישות המכילה ענפי ספורט. ענף ספורט מאופיין ע"י שמו (**SName**).

יחסים:

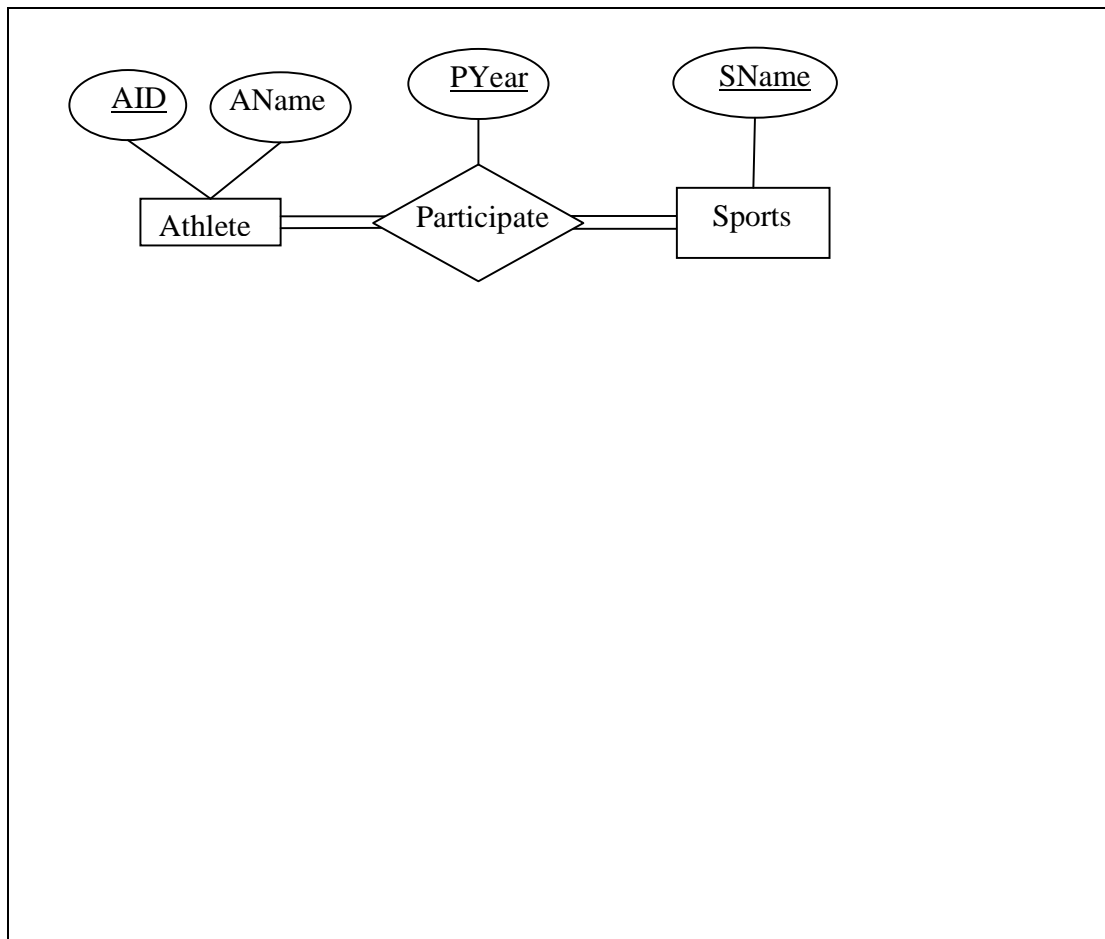
Participate - יחס המתאר התחרות של אתלט במקצוע מסוים בשנה מסוימת (**PYear**).
Finals - יחס המתאר השתתפות של אתלט במקצוע מסוים בשנה מסוימת (**FYear**) ומיקומו הסופי.

א. (8 נק') מלאו את הטבלה הבאה שמתייחסת לשדות ולמפתחות (הקבילים) של יחסים בסכמת ה-ERD.

| Participate | |
|-------------|--------|
| | שדות |
| | מפתחות |

| Finals | |
|--------|--------|
| | שדות |
| | מפתחות |

ב. (6 נק') שרטטו ERD כך שנתונים על גמר ודירוג סופי יינתנו רק לספורטאים שהתחרו באותה שנה (לפי Participate).



מתייחס לדיאגרמה המקורית

ג. (6 נק') נתון כי לאחר המרת הדיאגרמה לטבלאות בטבלה Finalists יש f שורות, בטבלת Ranking יש r שורות, בטבלה Sports יש s שורות. בנוסף נתון שב Finals יש נתונים רק משתי אולימפיאדות. מהו המספר המקסימאלי של שורות בטבלה Finals כתלות ב-f, r, s?

שאלה 2 – שפות שאילתא (36 נק')

$$R=R(A,B,C)$$

א. (6 נק') נתונים הביטויים הבאים:

1. $(\pi_A(R) \times \pi_B(R) \times \pi_C(R)) \div \pi_A(R)$

2. $[(\pi_A(R) \times \pi_{B,C}(R)) \cup (\pi_{A,B}(R) \times \pi_C(R)) \cup (\pi_{A,C}(R) \times \pi_B(R))] \div \pi_A(R)$

3. בחרו את האפשרות הנכונה ונמקו את תשובתכם:

i. הביטויים שקולים

ii. התוצאה של ביטוי 1 תמיד מוכלת בתוצאה של ביטוי 2

iii. התוצאה של ביטוי 2 תמיד מוכלת בתוצאה של ביטוי 1

iv. אף אחד משני הביטויים אינו מוכלל בביטוי השני (התוצאה של 1 לא תמיד מוכללת בתוצאה של 2 ולהפך).

ב. (8 נק') תהי $R(A)$ רלציה כך שב-A רשומים מספרים. הראו כי לכל קבוע k ניתן לרשום ב- RA שאלתא המחזירה את k האיברים הגדולים ב- R .

הערה: בסעיפים הבאים הניחו תרגום סטנדרטי של סכמת ה-ERD המקורית משאלה 1 לסכמות רלציוניות: **S**=Sports, **R**=Rank, **A**=Athlete, **Fsts**=Finalists, **NF**= NoFinal, **F**=Finals, **P**=Participate. בשאלות העוסקות ב-DRC הניחו סדר אלפביתי על התכונות של הרלציות.

ג. (8 נק') נאמר כי שני ספורטאים הם יריבים ספורטיביים אם יש ענף ספורט בו שניהם התחרו יחדיו לפחות פעם אחת ובכל שנה בה שניהם התחרו בענף הספורט הנ"ל הם סיימו במקומות 1 ו-2 בגמר. כתבו שאלתת DRC המחזירה את כל השלשות $(SName, AID1, AID2)$ כאשר $AID1$ ו- $AID2$ הם מזהים של יריבים בענף הספורט $SName$ ו- $AID1 < AID2$.

ד. (8 נק') נאמר כי ספורטאי הצטיין בשנה מסוימת אם הוא הגיע לגמר ברוב המקצועות בהם הוא התחרה באותה השנה. כתבו מבט `excelled` ב-SQL המחזיר את כל הזוגות (AID,Year) עבור כל מזהה של אתלט AID ושנה בה הצטיין.

ה. (6 נק') הסבירו מה עושה שאילתת ה-SQL הבאה. הניחו כי `excelled(AID,Year)` הוא המבט שחושב בסעיף הקודם.

```
SELECT AVG(t.times)
FROM (SELECT COUNT(*) as times
      FROM excelled e
      WHERE year BETWEEN 1940 AND 2008
      GROUP BY AID
      HAVING COUNT(*) > 1
     ) t
```

שאלה 3 – פירוקים וצורות נורמליות (24 נק')

א. (6 נק') לכל אחד מהסעיפים הבאים אמרו האם הוא נכון או לא. נמקו!

i. (3 נק') כל פירוק BCNF הוא משמר מידע

ii. (3 נק') כל פירוק 3NF משמר תלויות

ב. (6 נק') יהי $R(A,B,C,D)$ יחס עם קבוצת תלויות F . נאמר שקבוצת תכונות X היא סגורה אם $X_F^+ = X$. נתון כי רק הקבוצות \emptyset , $\{A,B\}$ ו- $\{A,B,C,D\}$ הן סגורות ב- R . תנו כיסוי ל- F .
נמקו!

ג. (6 נק') נתונה סכמה R הנמצאת ב- 3NF אך לא ב- BCNF. הוכיחו כי ל-R שני מפתחות קבילים שונים.

ד. (6 נק') הראו כי הרלציה R(A,B,C,D,E,H) נמצאת ב- BCNF עבור קבוצת התלויות $F = \{A \rightarrow BC, CD \rightarrow EH, AB \rightarrow D, C \rightarrow AH\}$

שאלה 4 – XML (20 נק')

נתון ה- DTD הבא המתאר קובץ document (שהוא צומת המסמך) המכיל מאמרים (articles) שונים. כל מאמר מכיל טקסט ו/או ציטוטים ממאמרים אחרים.

```
<!ELEMENT document (article)* >
<!ELEMENT article (#PCDATA|citation)* >
<!ATTLIST article aid ID #REQUIRED >
<!ELEMENT citation (#PCDATA) >
<!ATTLIST citation ref IDREF #REQUIRED>
```

מאמר אחד לפחות ציטט לפחות פעמיים.

א. (7 נק') כתבו שאילתת xpath המחזירה את כל המאמרים שמאמר אחר ציטט אותם לפחות פעמיים.

ב. (7 נק') כתבו שאילתת xpath המחזירה את כל המאמרים המצטטים מאמר המצטט אותם (ז.א כל מאמר A כך שיש B המצוטט ב- A והמצטט את A).

ג. (6 נק') שנו את ה-DTD כך שכל מאמר יוכל להכיל גם הערות שוליים (footnotes). הערת footnote יכולה להכיל טקסט והפניה למאמר (לכל היותר למאמר אחד). לא ניתן לצטט footnote.

לנוחותכם, מצורף להלן ה-DTD המקורי. מותר להשאיר או לתקן ו/או למחוק שורות ישנות, וכן מותר להוסיף שורות חדשות.

```
<!ELEMENT document (article)* >
```

```
<!ELEMENT article (#PCDATA|citation)* >
```

```
<!ATTLIST article aid ID #REQUIRED >
```

```
<!ELEMENT citation (#PCDATA) >
```

```
<!ATTLIST citation ref IDREF #REQUIRED>
```