



הטכניון - מכון טכנולוגי לישראל
הפקולטה למדעי המחשב

ד"ר אלדר פישר
סעאב מנסור
לינה זריבץ'

אביב תשס"ו
24 באוקטובר 2006

מערכות מסדי נתונים – 236363

פתרון – מועד ב'

הזמן: 3 שעות
במבחן זה 11 עמודים

שאלה	נקודות
שאלה 1 – ERD+ODL	24
שאלה 2 – שפות שאילתה	24
שאלה 3 – Design	30
שאלה 4 – XML	22
סה"כ	100

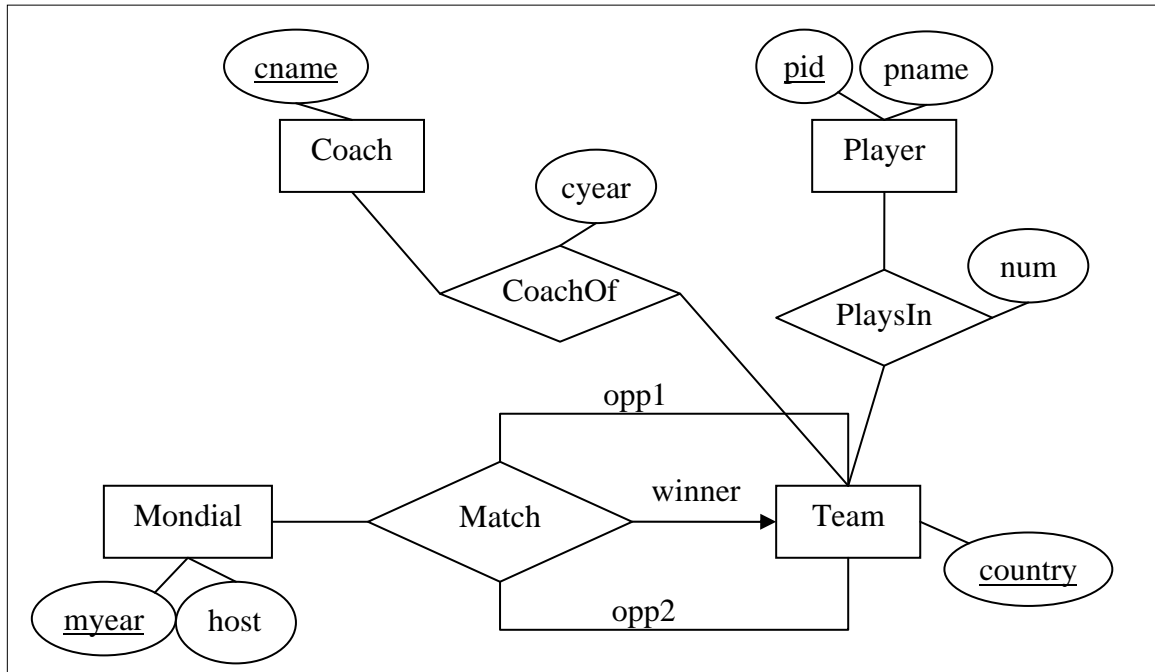
הנחיות:

1. יש לענות על כל השאלות **בטופס הבחינה**.
2. חומר עזר מותר: רק דברים שעשויים מנייר.
3. **אין להחזיק מכשיר אלקטרוני כלשהו לרבות מחשב כיס.**
4. קראו היטב את ההוראות שבתחילת כל שאלה ואת ההסברים לסכמות.
5. מומלץ שתכננו היטב את זמנכם, **לא תינתנה הארכות.**
6. ערעורים יש להגיש תוך שבועיים ממועד פרסום התוצאות.
7. לא יתקבלו ערעורים בנוסח "בדיקה מחמירה מדי".

בהצלחה

שאלה 1 – ERD+ODL (24 נק')

נתונה סכמת ה-ERD הבאה שמתארת מסד נתונים של מונדיאל הכדורגל:



הסברים לסכמה:

ישויות:

- Mondial** – מייצג מונדיאל שאירע בשנה myear. לכל מונדיאל אנו שומרים את המדינה המארחת (host).
- Team** – מייצג נבחרות שהשתתפו במונדיאלים. לכל נבחרת נשמר שם המדינה (country) אותה היא מייצגת.
- Player** – מייצג שחקנים המשתתפים בנבחרות במונדיאל. לכל שחקן קיים מזהה ייחודי (pid) ושם (pname).
- Coach** – מייצג מאמני נבחרות במונדיאל. לכל מאמן נשמר שם ייחודי (cname).

יחסים:

- Match** – יחס שמייצג משחק במונדיאל. בכל משחק משתתפות שתי נבחרות (בתפקידים opp1 ו-opp2), ובנוסף אנו שומרים את הנבחרת המנצחת (winner) מבין אלו.
- PlaysIn** – יחס שמייצג השתתפות של שחקן בנבחרת. לכל יחס כזה נשמר מספר החולצה של השחקן (num).
- CoachOf** – יחס שמייצג אימון של נבחרת ע"י מאמן. לכל יחס כזה נשמור את שנת תחילת האימון (cyear) של המאמן לנבחרת אותה אימן.

א. (6 נק') מלאו את הטבלאות הבאות בהתייחס לסכמה הנתונה:

PlaysIn

שדות	pid, country, num
מפתחות	pid, country

Match

שדות	myear, opp1, opp2, winner
מפתחות	myear, opp1, opp2

ב. (8 נק') בתת הסעיפים הבאים מופיעה דרישה חדשה שה-ERD של המונדיאל צריך לקיים. עבור כל דרישה הסבירו במילים אילו שינויים יש לבצע ל-ERD המקורי.

שני תתי הסעיפים הבאים הם בלתי תלויים!

i. (4 נק') נרצה לשמור הבקעת גולים של שחקנים במשחקים. לכל שחקן נרצה לשמור כמה גולים הוא הבקיע בכל משחק.

נוסיף הקבצה ליחס Match ונקשר הקבצה זו ע"י יחס (נקרא לו Scored) לישות Player.

נוסיף ליחס Scored את התכונה (שאינה מפתח) numgoals אשר תשמור את מספר

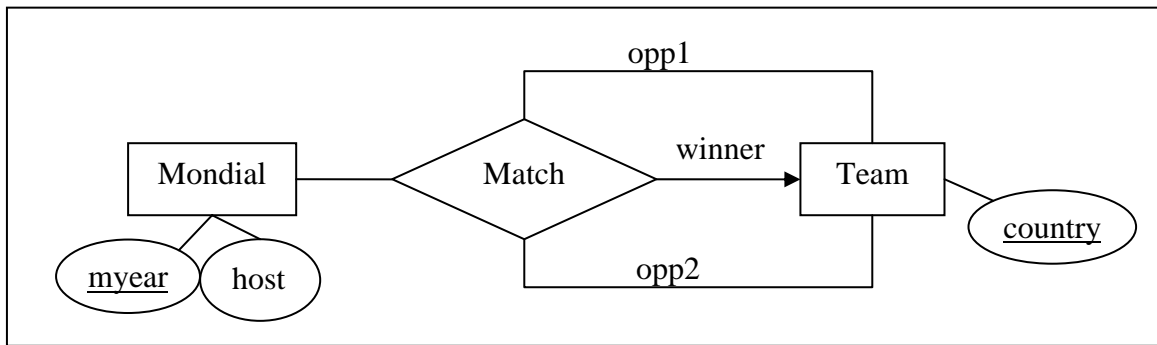
הגולים שהבקיע השחקן במשחק מסויים.

ii. (4 נק') נשים לב שב-ERD המקורי כל שתי נבחרות יכולות להיפגש לכל היותר פעמיים באותו מונדיאל (על ידי היפוך תפקידים של opp1, opp2). נרצה לאפשר לנבחרות להיפגש מספר פעמים לא מוגבל. לכל מפגש כזה נרצה לשמור רישום בנפרד.

נוסיף תכונת מפתח matchnum ליחס Match אשר תייצג מספר ייחודי לכל מפגש. במקרה

ושתי נבחרות ייפגשו פעמיים או יותר, לכל מפגש נשמור מספר שונה.

ג. (10 נק') תרגמו בצורה מדויקת ככל הניתן את ה-ERD הקטן הבא ל-ODL:



```

interface Mondial (extent Mondials key myear) {
    attribute integer myear;
    attribute string host;
    relationship Set<Match> myMatches inverse Match::inMondial;
}

interface Team (extent Teams key country) {
    attribute integer country;
    relationship Set<Match> asOpp1 inverse Match::opp1;
    relationship Set<Match> asOpp2 inverse Match::opp2;
    relationship Set<Match> asWinner inverse Match::winner;
}

interface Match (extent Matches key (inMondial, opp1, opp2)) {
    relationship Mondial inMondial inverse Mondial::myMatches;
    relationship Team opp1 inverse Team::asOpp1;
    relationship Team opp2 inverse Team::asOpp2;
    relationship Team winner inverse Team::asWinner;
}
    
```

שאלה 2 – שפות שאילתה (24 נק')

הערה: בסעיפים א', ב' ו-ג' הניחו תרגום סטנדרטי של סכמת ה-ERD משאלה 1 לסכמות רלציוניות: Mondial, Team, Match, Player, PlaysIn, Coach, CoachOf.

א. (6 נק') כתבו במילים פשוטות את משמעות השאילתה הבאה:
 $\{ \langle \text{myear}, \text{cname}, \text{country} \rangle : \exists \text{host Mondial}(\text{myear}, \text{host}) \wedge$
 $\exists \text{cyear} (\text{CoachOf}(\text{cname}, \text{country}, \text{cyear}) \wedge (\text{cyear} \leq \text{myear}) \wedge$
 $\neg \exists \text{cname1}, \text{cyear1} (\text{CoachOf}(\text{cname1}, \text{country}, \text{cyear1}) \wedge (\text{cyear1} \leq \text{myear}) \wedge$
 $(\text{cyear1} > \text{cyear}))) \}$

השאילתה מחזירה לכל מונדיאל ונבחרת את המאמן שאימן אותה לפני תחילת המונדיאל והכי קרוב לתחילת המונדיאל (מבין המאמנים שאימנו את הנבחרת לפני המונדיאל).

ב. (6 נק') כתבו שאילתת ב-RA אשר מחזירה לכל מונדיאל, את הנבחרות שלא הפסידו במונדיאל זה.

$(\text{Team} \times \pi_{\text{myear}} \text{Mondial}) \setminus (\rho_{\text{opp1} \rightarrow \text{country}} (\pi_{\text{opp1}, \text{myear}} (\sigma_{\text{opp1} \neq \text{winner}} \text{Match}))$
 $\cup \rho_{\text{opp2} \rightarrow \text{country}} (\pi_{\text{opp2}, \text{myear}} (\sigma_{\text{opp2} \neq \text{winner}} \text{Match})))$
 בנוסף התקבל הפיתרון הבא שדורש השתתפות של הנבחרת במונדיאל.
 $(\rho_{\text{opp1} \rightarrow \text{opp}} (\pi_{\text{opp1}, \text{myear}} \text{Match}) \cup \rho_{\text{opp2} \rightarrow \text{opp}} (\pi_{\text{opp2}, \text{myear}} \text{Match})) \setminus$
 $(\rho_{\text{opp1} \rightarrow \text{opp}} (\pi_{\text{opp1}, \text{myear}} (\sigma_{\text{opp1} \neq \text{winner}} \text{Match}))$
 $\cup \rho_{\text{opp2} \rightarrow \text{opp}} (\pi_{\text{opp2}, \text{myear}} (\sigma_{\text{opp2} \neq \text{winner}} \text{Match})))$

ג. (6 נק') כתבו שאילתה ב-SQL אשר מחשבת לכל שנת מונדיאל ולכל נבחרת שמופיעה ב-Team, את מספר הניצחונות של הנבחרת במונדיאל זה. סכימת הפלט תהיה [myear, country, numwins]

```
SELECT MT.myear, MT.country, COUNT(M.winner) as numwins
FROM (SELECT myear, country FROM Mondial, Team) MT LEFT OUTER
JOIN Match M
ON (M.myear=MT.myear AND M.winner=MT.country)
GROUP BY MT.myear, MT.country
```

ד. (6 נק') נתונות שתי רלציות $R(A,B)$ ו- $S(C)$. כמו כן, נתונה שאילתת ה-SQL הבאה:

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

כתוב שאילתת שקולה לשאילתת הנ"ל המשתמשת בהוראת SELECT אחת בלבד, כלומר המילה SELECT מופיעה בה רק פעם אחת.

```
SELECT DISTINCT A
FROM R LEFT OUTER JOIN S
ON (B= C)
WHERE C IS NULL
```

דרך שנייה:

```
SELECT DISTINCT R.A
FROM R, S AS S1, S AS S2
WHERE R.B != S1.C
GROUP BY R.A, R.B
HAVING COUNT(DISTINCT S1.C)=COUNT(DISTINCT S2.C)
```

שאלה 3 – Design (30 נק')

א. (16 נק') תהי R סכמה רלציונית, ותהי F קבוצת תלויות פונקציונליות מעל R . בסעיף זה נבחן הגדרות אחרות לצורות נורמליות של סכמות.

i. (4 נק') הגדרה של **FNF**:

סכמה R היא ב-FNF בהינתן F אם לכל תלות פונקציונלית לא טריוויאלית $X \rightarrow Y \in F^+$ מתקיים כי גם $X \rightarrow \{R \setminus Y\} \in F^+$.
האם FNF שקולה ל-BCNF?
אם כן, הוכיחו בקצרה. אם לא, תנו דוגמא נגדית.

פ.

אם R, F ב-BCNF אזי לכל תלות לא טריוויאלית $X \rightarrow Y \in F^+$ מתקיים כי

$X \rightarrow R \in F^+$ ובפרט גם מתקיים ש- $X \rightarrow \{R \setminus Y\} \in F^+$.

אם R, F ב-FNF אזי לפי ההגדרה לכל תלות לא טריוויאלית $X \rightarrow Y \in F^+$

מתקיים ש- $X \rightarrow \{R \setminus Y\} \in F^+$, כלומר $X_F^+ = R$. מכאן נובי כי X הוא על-מפתח.

ii. (4 נק') הגדרה של **ZNF**:

סכמה R היא ב-ZNF בהינתן F אם לכל תלות פונקציונלית לא טריוויאלית $X \rightarrow Y \in F$ מתקיים כי X הוא על מפתח של R .
האם ZNF שקולה ל-BCNF?
אם כן, הוכיחו בקצרה. אם לא, תנו דוגמא נגדית.

פ. ברור שאם R, F ב-BCNF, אזי גם R, F ב-ZNF.

תהי R, F ב-ZNF. ניקח את F וניצור ממנה F' בצורה הבאה:

$$F' = \{X \rightarrow A \mid X \rightarrow Y \in F \wedge A \in Y\}$$

ברור כי F' שקולה ל- F ולכן R, F ב-BCNF אם ורק אם R, F' ב-BCNF.

לפי המשפט שהיה בתרגיל בית מספיק לבדוק רק את התלויות $X \rightarrow A \in F'$.

כיוון ש- R, F ב-ZNF אזי X הוא על מפתח גם ב- R, F' .

iii. (4 נק') הגדרה של **MNF**:
 סכמה R היא ב-MNF בהינתן F אם לכל תלות פונקציונלית לא טריוויאלית
 כך ש- $X \rightarrow A \in F^+$, X הוא על מפתח של R או A שייך לעל מפתח של R
 המכיל מספר מינימאלי של אטריביוטים.
 האם MNF שקולה ל-3NF?
 אם כן, הוכיחו בקצרה. אם לא, תנו דוגמא נגדית.

לא.

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

ברור כי R, F נמצאת ב-3NF אך היא אינה נמצאת ב-MNF שכן התלות

$C \rightarrow B$ מפרה את MNF כי BA הוא מפתח קביל אבל יש גם מפתח קביל

קטן יותר, D .

iv. (4 נק') הגדרה של **ENF**:
 סכמה R היא ב-ENF בהינתן F אם לכל תלות פונקציונלית לא טריוויאלית
 כך ש- $X \rightarrow Y \in F^+$, X הוא על מפתח של R או קיים ב- R מפתח קביל Z כך
 ש- $Y \cap Z \neq \emptyset$.
 הוכח או הפרך:
 האם ENF שקולה ל-3NF?
 אם כן, הוכיחו בקצרה. אם לא, תנו דוגמא נגדית.

כן. נניח כי R, F נמצאת ב-3NF. אזי לפי ההגדרה של 3NF, לכל תלות

$X \rightarrow A \in F^+$, X הוא על מפתח של R או A שייך למפתח קביל של R .

תהי $X \rightarrow Y \in F^+$. בהכרח קיימת התלות $X \rightarrow B \in F^+$ כך ש- $B \in Y$. לכן או ש- X

הוא על מפתח או ש- B שייך למפתח קביל K של R, F , כלומר קיים מפתח

קביל $Z=K$ על פי הדרישה כך ש- $Y \cap Z \neq \emptyset$.

נניח כי R, F נמצאת ב-ENF. תהי $X \rightarrow A \in F^+$. לפי ENF נובע שאם ש- X הוא

על מפתח או שקיים מפתח קביל Z כך ש- $\{A\} \cap Z \neq \emptyset$. מכאן ש- Z הוא מפתח

קביל של R, F וגם מתקיים ש- $A \in Z$.

ב. (14 נק') נתונה סכמה $R[A,B,C,D,E,H]$ וקבוצת תלויות $F = \{AB \rightarrow C, H \rightarrow D, BD \rightarrow E, HE \rightarrow C, DE \rightarrow H, C \rightarrow D\}$

i. (5 נק') הוכיחו כי $F \vdash AB \rightarrow H$ באמצעות אקסיומות ארמסטרונג.

- | | |
|------------------------|-----------|
| 1. $AB \rightarrow C$ | |
| 2. $C \rightarrow D$ | |
| 3. $AB \rightarrow D$ | $A3\ 1,2$ |
| 4. $AB \rightarrow BD$ | $A2\ 3$ |
| 5. $BD \rightarrow E$ | |
| 6. $BD \rightarrow DE$ | $A2\ 5$ |
| 7. $AB \rightarrow DE$ | $A3\ 4,6$ |
| 8. $DE \rightarrow H$ | |
| 9. $AB \rightarrow H$ | $A3\ 7,8$ |

ii. (4 נק') ציינו את כל המפתחות הקבילים עבור F, R .

AB

iii. (5 נק') מצא עבור R, F פירוק ב-3NF, משמר מידע ותלויות המכיל שתי סכמות בדיוק.

$[A,B,C], [D,E,H,B,C]$

שאלה 4 – XML (22 נק')

א. (7 נק') כתבו DTD עבור מאגר פסקי דין. על המאגר לקיים את התנאים הבאים:

- צומת המסמך הוא court. הוא מכיל את שם אב בית הדין (צומת president) ולאחריו מספר כלשהוא של פסקי דין (צמתי decision).
- כל צומת פסק דין מכיל מספר אי-זוגי של מחברים (צמתי author) ולאחריהם תוכן (צומת content).
- בנוסף, לפעמים נרצה לציין מהם התקדימים שפסק הדין מתייחס אליהם. לשם כך אפשר באופן אופציונאלי לתת לפסק דין תכונה מזהה, וכן יש לאפשר לכתוב מצביעים למספר לא קבוע של פסקי דין המשמשים תקדים אם היו כאלו. עליכם לדאוג לקיום התכונות המתאימות בצמתיים (מותר לכם לבחור את שמות התכונות כרצונכם).

הערה: אם הנכם רוצים לחסוך בכתיבה, מותר לכם לצורך המבחן לכתוב בסוף הפיתרון מהם שמות האלמנטים המוגדרים כצמתי טקסט ללא תכונות, במקום לכתוב שורה נפרדת ב-DTD לכל אחד מהם.

```
<!ELEMENT court (president, decision*)>
<!ELEMENT decision (((author, author)*, author), content)>
<!ATTLIST decision
    did ID #IMPLIED
    precedents IDREFS #IMPLIED
>
<!ELEMENT president (#PCDATA)>
<!ELEMENT author (#PCDATA)>
<!ELEMENT content (#PCDATA)>
```

ב. (7 נק') השאלתה הבאה מתייחסת לקובץ file.xml אשר לא נתון עבורו DTD, אולם ידוע שלחלק מצמתי האלמנט שלו יתכנו תכונות בעלות השם label ו/או תכונות בעלות השם reference. עליכם לבנות שאלתת XPath 1.0 בוליאנית אשר תבדוק שמתקיימים עבור הקובץ התנאים הבאים:

- כל הערכים של תכונות label בקובץ שונים זה מזה.
 - לכל תכונת reference בקובץ יש ערך המופיע באחת מתכונות ה-label בקובץ.
- אם התנאים מתקיימים, יש להחזיר true, אחרת החזירו false.

```
not(//@reference[not(=//@label)] or //*[@label = (preceding::*/@label |
following::*/@label | descendant::*/@label | ancestor::*/@label)])
```

ג. (8 נק') נתון מסמך exam.xml שמכיל צמתי טקסט וצמתי אלמנט בלבד. כתבו שאילתא ב-XQuery אשר מחזירה את כל אלמנטי ה-comment במסמך אשר אינם מצייתים לדרישות המתוארות ע"י מקטע ה-dtd הבא.

```
<!ELEMENT comment (content, comment*)>
<!ATTLIST comment
    title CDATA #IMPLIED
    author CDATA #REQUIRED>
<!ELEMENT content (#PCDATA)>
```

אין חובה לבדוק צמתיים אחרים (כולל צמתי content).

```
for $c in doc("exam.xml")//comment
where not($c/@author) or not($c/*) or
    $c/*[1][name()!="content"] or
    $c/*[position()=1][name()!="comment"] or
    $c/@*[name()!="title" and name()!="author"]
return
$c
```