הטכניון – מכון טכנולוגי לישראל הפקולטה למדעי המחשב

מרצה: פרופ' בני קימלפלד סמסטר אביב התשע"ט

מתרגלים: אסף ישורון

משה סבאג

אליאס טנוס

ניר אליאס

מערכות מסד נתונים

236363

פתרון מועד ב'

15 בספטמבר 2019

פירוט החלקים והניקוד:

|  |  |  |  |
| --- | --- | --- | --- |
| שאלה | נושא | ניקוד | הערות |
| 1 | ERD | 12 |  |
| 2 | RA | 20 |  |
| 3 | SQL | 20 |  |
| 4 | Design Theory | 28 |  |
| 5 | Neo4j | 10 | **יש לבחור 2 שאלות מתוך 5,6,7** |
| 6 | MongoDB | 10 | **יש לבחור 2 שאלות מתוך 5,6,7** |
| 7 | RDF | 10 | **יש לבחור 2 שאלות מתוך 5,6,7** |

הנחיות לנבחנים

1. כתבו את התשובות אך ורק בטופס הבחינה ובמקום המיועד להן, מחברת הטיוטה לא תיבדק.
2. כל חומר עזר הכתוב על נייר בלבד מותר בשימוש.
3. אין לקבל או להעביר חומר כלשהו בזמן הבחינה.
4. יש להשתמש רק בסימנים או פונקציות שנלמדו בתרגול או בהרצאה והמופיעים בשקפים של

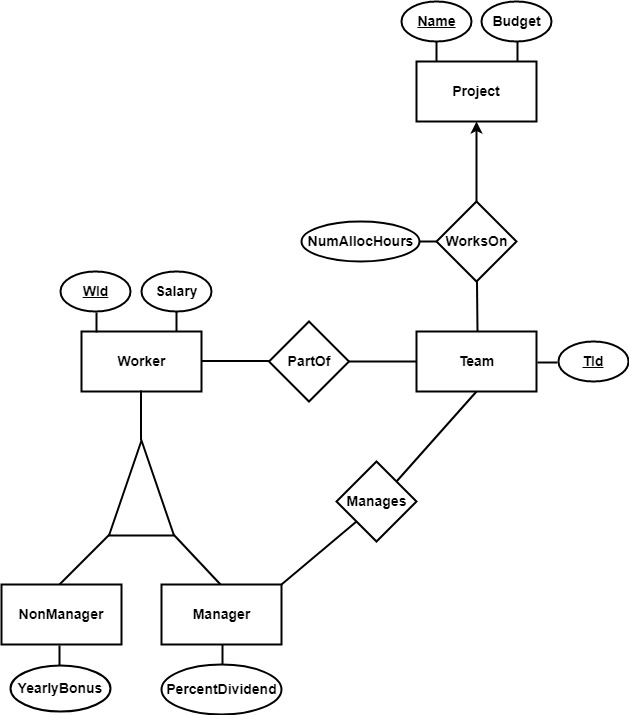
הקורס. כל שימוש בסימון שאינו כזה מחייב הסבר מלא של משמעות הסימון.

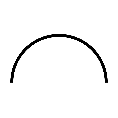
1. משך הבחינה הינו שלוש שעות, תכננו את הזמן בהתאם.
2. אין לכתוב בעפרון.

בהצלחה!

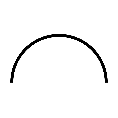
**שאלה 1: ERD (12 נק')**

לפניכם תרשים ERD המגדיר מסד נתונים של חברה:





**ISA**



1. רשמו את הסכמות המתקבלות מתרגום הישויות והיחסים המופיעים בתרשים על פי הכללים שנלמדו. הקפידו על סימון מפתחות בקו תחתון וציון מפתחות זרים: (7 נק')

Project(Name, Budget)

Team(TId, WId), WId references Manager

Worker(WId, TId, Salary), TId references Team.

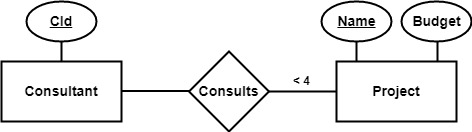
NonManager(WId, YearlyBonus), WId references Worker.

Manager(WId, PercentDividend), WId references Worker.

WorksOn(TId, NumAllocHours, Name), TId references Team, Name references Project.

1. נתון כי החברה מעסיקה בדיוק 50 עובדים (Workers). השלימו מספרים במקומות החסרים כך שתתקבלנה טענות נכונות: (2 נק')
2. החברה עובדת לכל הפחות על \_0\_ פרויקטים (Projects).
3. החברה עובדת (ע״י צוותים לא ריקים) על לכל היותר \_50\_ פרויקטים (Projects).
4. מנהלי החברה מעוניינים לשכור יועצים חיצוניים (שאינם בקבוצת העובדים בחברה) שיחוו את דעתם על העבודה ויעזרו בייעול הפרויקטים. הציעו תוספות ו/או שינויים לתרשים ה-ERD כך שיתקיימו התנאים הבאים: (3 נק')

* יועצים יוכלו לחוות את דעתם על פרויקטים.
* אין חובה כי לכל פרויקט יותאם יועץ.
* לא יהיה אף יועץ שיחווה את דעתו על יותר מ-3 פרויקטים במקביל.



נחבר ליישות Project את המבנה הבא:

את שאר התרשים נשאיר כפי שהיה.

**שאלה 2: RA & RC (20 נק')**

נתונים היחסים הבאים:

Posting(user,post)

Likes(user, post)

בנוסף נתון כי Likes.post הינו מפתח זר לPosting.post -.

1. כתבו שאילתה באלגברת היחסים המחזירה את המשתמשים (user) שפרסמו לפחות 2 פוסטים שונים. (3 נק')
2. כתבו שאילתה באלגברת היחסים המחזירה את המשתמשים שפרסמו פוסט לו הם עצמם **לא** עשו לייק (וייתכן כי הם פרסמו פוסטים נוספים להם הם כן עשו לייק). (4 נק')
3. נתונה השאילתה הבאה באלגברה הרלציונית:

נסחו שאילתה שקולה אך ורק בעזרת אופרטורים מתוך הטלה, איחוד, מכפלה קרטזית, בחירה ושינוי שם. (3 נק')

1. נתונה השאילתה הבאה באלגברה הרלציונית:

הסבר במילים אילו משתמשים מחשבת השאילתה, ונסח את השאילתה בתחשיב היחסים. (4 נק')

כל המשתמשים שפרסמו פוסט אחד לפחות, ובנוסף עשו לייק לכל הפוסטים.

1. הוכח כי את השאילתה שבסעיף b לא ניתן לנסח אך ורק בעזרת הטלה, איחוד, מכפלה קרטזית, בחירה ושינוי שם. (6 נק')

נגדיר אופרטור אונארי כמונוטוני אם:

ונגדיר אופרטור בינארי כמונוטוני אם:

השאילתה אינה מונוטונית. הוכחה על ידי דוגמא נגדית:

מתקיים אך תוצאת הפעלת השאילתה על (u) אינה מוכלת בתוצאת הפעלת השאילתה על (קבוצה ריקה).

חמשת האופרטורים הנתונים הינם מונוטונים ולכן לא ניתן לבנות בעזרתם את השאילתה הנדרשת, שהיא אינה מונוטונית.

**שאלה 3: SQL (20 נק')**

לפניכם מסד נתונים של ספרייה:

* Books (Name ,Genre, Year, NumPages)

משמעות התכונות: שם הספר, הז'אנר אליו משתייך הספר, שנת הוצאה, מספר עמודים.

* Readers (RId, FavouriteGenre, FavouriteBook, YearsOfMembership)

משמעות התכונות: מס' מזהה של הקורא, הז'אנר האהוב על הקורא, הספר האהוב על הקורא, מספר שנות החברות של הקורא בספרייה.

להלן, **בתור דוגמא**, השמה אפשרית לתוכן הטבלאות:

Books:

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Genre** | **Year** | **NumPages** |
| Of Mice And Men | Novel | 1937 | 187 |
| A Streetcar Named Desire | Play | 1947 | 224 |
| Paradise Lost | Epic | 1667 | 112 |
| War And Peace | Epic | 1869 | 1225 |

Readers:

|  |  |  |  |
| --- | --- | --- | --- |
| **RId** | **FavouriteGenre** | **FavouriteBook** | **YearsOfMembership** |
| 1 | Novel | Of Mice And Men | 0 |
| 2 | Novel | A Streetcar Named Desire | 3 |
| 3 | Epic | War And Peace | 2 |
| 4 | Epic | War And Peace | 8 |

**שימו לב:**

1. הניחו כי אף אחת מן הטבלאות אינה ריקה.
2. הז'אנר האהוב על כל קורא מופיע גם בעמודה Genre בטבלה Books.
3. הספר האהוב על כל קורא מופיע גם בעמודה Name בטבלה Books.
4. כתבו שאילתה המחזירה את המספרים המזהים של הקוראים (RId) שהספר האהוב עליהם שייך לז'אנר האהוב עליהם. (4 נק')

SELECT Rid FROM Readers JOIN Books

ON FavouriteBook = Name

WHERE FavouriteGenre = Genre

1. קורא ותיק הוא קורא בעל חברות בספרייה של 5 שנים לפחות. כתבו שאילתה המחזירה את שמות הקוראים שהספר האהוב עליהם הוא גם הספר האהוב על קורא אחר (אחד או יותר) מלבדם שהינו קורא ותיק. (5 נק')

SELECT R1.Rid FROM Readers R1, Readers R2

WHERE R1.Rid <> R2.Rid AND R1.FavouriteBook = R2.FavouriteBook

AND R2.YearsOfMembership >= 5

1. כתבו שאילתה המחזירה את שם הז'אנר (או שמות הז'אנרים במקרה של שוויון) בעל מספר עמודים ממוצע לספר גבוה ביותר. (5 נק')

SELECT DISTINCT Genre, AVG(NumPages) AS AvgPages FROM Books

GROUP BY Genre

HAVING AvgPages >= ALL(SELECT AVG(NumPages) FROM Books GROUP BY Genre)

1. עליכם לכתוב שאילתה המחזירה שמות של ספרים, ועליה להחזיר את שמו של ספר X אם ורק אם הוא מקיים את התכונה הבאה:

כל הקוראים שהספר האהוב עליהם הוא X, כבר היו חברים בספרייה כאשר X יצא לאור.

הניחו שהשאילתה נכתבת היום, ב-2019. (6 נק')

SELECT Name FROM Books B

WHERE (2019 – Year) < ALL (SELECT YearsOfMembership FROM Readers

WHERE FavouriteBook = B.Name)

**שאלה 4: Design Theory (28 נק')**

1. נתונה הסכמה כך ש וקבוצת התלויות היא
2. ציינו 2 מפתחות קבילים עבור הסכמה: (2 נק')
3. מצאו כיסוי מינימלי עבור , היעזרו באלגוריתם הנלמד בכיתה. (אין צורך להראות את צעדי האלגוריתם) (4 נק')
4. עבור התלויות הבאות, קבעו האם התלות נובעת מ-או לא. (4 נק')

|  |  |
| --- | --- |
|  | כן/לא |
|  | כן/לא |
|  | כן/לא |
|  | כן/לא |

1. קבעו האם הטענות הבאות נכונות או שגויות וסמנו בהתאם: (2 נק')

|  |  |
| --- | --- |
| קיימת קבוצה בת **4** אטריביוטים שהסגור שלה ביחס לקבוצת התלויות F **לא** כולל את האטריביוט D. | נכון/לא נכון |
| קיימת קבוצה בת **2** אטריביוטים שהסגור שלה ביחס לקבוצת התלויות F **לא** כולל את האטריביוט A. | נכון/לא נכון |

1. האם הפירוק משמר מידע? הוכיחו את תשובתכם. (6 נק')

הפירוק אכן משמר מידע. הוכחה בעזרת האלגוריתם:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | H |
| R1 | a | b1 | c1 | d1 | e | h1 |
| R2 | a2 | b | c | d2 | e2 | h |
| R3 | a3 | b3 | c3 | d | e | h |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | H |
| R1 | a | b1 | c1 | d1 | e | h1 |
| R2 | a2 | b | c | d2 | e2 | h |
| R3 | a | b | c | d | e | h |

1. מצאו פירוק BCNF עבור הסכמה. אין צורך להוכיח את נכונות תשובתכם. (4 נק')

R1(A,B,C), R2(C,H), R3(H,D,E)

1. תהי קבוצת תלויות פונקציונליות מעל סכמה כך שבאגף שמאל של כל תלות ב-מופיע אטריביוט בודד.

הוכיחו כי אם ו-הם אטריביוטים ב-, אזי . (6 נק')

מתקיים תמיד. כעת, נריץ את האלגוריתם לחישוב סגור של קבוצת האטריביוטים ונניח בשלילה כי בשלב מסוים בריצת האלגוריתם מתווסף לסגור אטריביוט שאינו נמצא ב-. נסמן ב- את הצעד הראשון בריצת האלגוריתם שבו זה קורה. נשים לב כי מהגדרת , עד הצעד ה- הסגור שחישבנו מכיל אך ורק אטריביוטים אשר נמצאים גם ב-. בצעד ה- אנו משתמשים בתלות ב- על מנת להרחיב את הסגור. התלויות ב- הן מהצורה כאשר אטריביוט יחיד, ו- מכילה אטריביוט אשר לא נמצא ב-.

נמצא בהכרח גם ב-, ולכן בפרט . בה"כ נניח שהוא נמצא ב-. מכאן שגם נמצא ב- ולכן גם ב- בסתירה להנחת השלילה.

**שאלה 5: Neo4j (10 נק')** שאלת בחירה מבין שאלות 5,6,7. יש לבחור 2 מתוך 3!

להלן תיאור של גרף Neo4j המחזיק צמתים של קורסים וסטודנטים. הגרף מומש לפי הכללים הבאים:

1. כל צומת בגרף מחזיק בתווית (label) אחת בדיוק מבין האפשרויות הבאות: **Student, Course**.
2. כל אחד מן הצמתים מחזיק בתכונה (Attribute) בשם **Name** המזהה באופן חד-חד-ערכי כל קורס או סטודנט.
3. צמתי קורסים בגרף מחוברים על ידי קשרים בעלי label מסוג **Prerequisite** המסמנים כי קורס מסוים הוא קורס קדם לקורס אחר.
4. צומת של סטודנט יכול להיות מקושר לצומת של קורס על ידי קשר בעל label מסוג **Took** שמשמעותו היא שהסטודנט לקח את הקורס אליו הוא מקושר.

להלן **דוגמא** לגרף המקיים את ארבעת הכללים:

:Prerequisite

:Took

1. כתבו שאילתה בשפת Cypher המחזירה את כמות הקורסים שאינם דורשים אף קדם. (3 נק')

MATCH (a :Course)  
WHERE NOT () – [:Prerequisite]-> (a)  
RETURN COUNT(a)

2. כתבו שאילתה בשפת Cypher המחזירה את צמתי הסטודנטים שלקחו קורס מסוים אך לא את הקדם שלו. (3 נק')

MATCH (a :Student) – [:Took] - > (b : Course) <- [:Prerequisite] – (c : Course)   
WHERE NOT (a) – [ :Took] - > (c)  
RETURN a

חשוב לבדוק שלקורס b יש קורס קדם c.

3. כתבו שאילתה בשפת Cypher המחזירה את צמתי הקורסים שיש להם לפחות 3 קדמים ישירים (כלומר 3 קורסים עם קשת מסוג Prerequisite אליהם). (4 נק')

MATCH (a :Course) – [:Prerequisite] -> (b: Course)  
WITH b, count(\*) as num\_ prerequisites  
WHERE num\_prerequisites >= 3  
RETURN b

**שאלה 6: MongoDB (10 נק')** שאלת בחירה מבין שאלות 5,6,7. יש לבחור 2 מתוך 3!

לרשותכם מסד נתונים המשמש חנות עתיקות ובו אוסף (Collection) בודד הנקרא Antiques אשר מכיל מידע על פריטים המוצעים למכירה בחנות, כמו גם על פריטים שנמכרו לאחרונה. כל מסמך באוסף הוא מהצורה הבאה:

{

“\_id”: <ObjectId>

“name”: <string>

“estimated\_year”: <int>

“condition”: <string>

“estimated\_worth”: <int>

“worth\_error\_margin”: <int>

“selling\_price”: <int>

“sold\_on”: <date> **This field only exists for items no longer available in the shop**

}

דוגמא למסמך אפשרי:

{

“\_id”: ObjectId(849a9cc012abcd)

“name”: “A letter signed by Abraham Lincoln”

“estimated\_year”: 1856

“condition”: “Poor”

“estimated\_worth”: 150

“worth\_error\_margin”: 30

“selling\_price”: “200”

}

1. מנהלי החנות החליטו שאין להם צורך לשמור במסד הנתונים פריטים שכבר נמכרו. כתבו פקודת MongoDB המסירה מהאוסף Antiques את כל הפריטים שכבר אינם מוצעים למכירה בחנות.(3 נק')

DB.Antiques.remove({

‘sold\_on’: { $exists: true }

})

1. כתבו שאילתה המחזירה עבור כל מצב אפשרי (condition) את מחירו של הפריט ה**יקר** ביותר (בעל selling\_price גבוה ביותר) הנמצא במצב זה, **ובנוסף,** את מחירו של הפריט ה**זול** ביותר הנמצא במצב זה. עליכם למיין את הערכים המוחזרים לפי מחירו של הפריט הזול ביותר עבור כל מצב, בסדר עולה. (3 נק')

DB.Films.aggregate({

{$group: {‘\_id’: ‘$condition’, ‘max\_selling\_price’: {$max: ‘$selling\_price’},

‘min\_selling\_price’: {$min: ‘$selling\_price’}},

{$sort: {‘min\_selling\_price’: 1}}

})

1. פריט רווחי הוא פריט שמחירו (selling\_price) גדול או שווה להערכה האופטימית ביותר של שוויו. השתמשו ב-mapReduce על מנת להחזיר לכל מצב אפשרי (condition) את סך מחירי הפריטים הרווחיים: (4 נק')

DB.Antiques.mapReduce(

function() { emit(this.condition, this.price); },

function(key, values) {return Array.sum(values) },

{

query: { “price” : {$gte : $estimated\_worth + $worth\_error\_margin}},

out: “total\_prices\_per\_condition”

}

)

# **שאלה 7: RDF/SPARQL (10 נק')** שאלת בחירה מבין שאלות 5,6,7. יש לבחור 2 מתוך 3!

In this question, we assume the following namespaces:

* rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
* rdfs: <http://www.w3.org/2000/01/rdf-schema#>
* dbp: <http://dbpedia.org/property/>
* dbo: <http://dbpedia.org/ontology/>
* dbr: <http://dbpedia.org/resource/>

1. לכל אחת משתי השאילתות הבאות, הוכיחו כי היא מוכלת באחרת, או לחילופין, הראו דוגמא לגרף המוכיח **אי**-הכלה. (3 נק' כל כיוון)

ראשית, .

נסמן את חלק השאילתה אשר נמצא תחת האופרטור MINUS ב-, ואת חלק השאילתה אשר נמצא תחת האופרטור MINUS ב-. מכאן ששאילתה היא מהצורה , בעוד ששאילתה 2 היא מהצורה

(עבור אותו בשני המקרים). מאופן פעולת האופרטור OPTIONAL מתקיים , ולכן

כלומר .

ההכלה ההפוכה לא מתקיימת. דוגמא נגדית:

|  |  |  |
| --- | --- | --- |
| dbr:p1 | rdf:type | dbo:person |
| dbr:b1 | dbp:member | dbr:p1 |
| dbr:a1 | dbp:artist | dbr:b1 |

על מסד נתונים זה תחזיר , בעוד ש- תחזיר קבוצה ריקה

|  |
| --- |
| 1.  SELECT ?p ?band {  ?p rdf:type dbo:person.  ?band dbp:member ?p.  MINUS {  ?album dbp:artist ?band.  ?album dbp:producer ?p. }  } |
| 2.  SELECT ?p ?band {  ?p rdf:type dbo:person.  ?band dbp:member ?p.  MINUS {  ?album dbp:artist ?band.  OPTIONAL {?album dbp:producer ?p. } } } |

1. נתון המאגר הבא:

|  |  |  |
| --- | --- | --- |
| Default graph: | | |
| dbr:p1 | rdf:type | dbo:person |
| dbr:b1 | dbp:member | dbr:p1 |
| dbr:p2 | rdf:type | dbo:person |
| dbr:b2 | dbp:member | dbr:p2 |

|  |  |  |
| --- | --- | --- |
| Graph: http://ex1.edu | | |
| dbr:a1 | dbp:artist | dbr:b1 |
| dbr:a1 | dbp:producer | dbr:p1 |
| dbr:a2 | dbp:artist | dbr:b2 |

|  |
| --- |
| Graph: http://ex2.edu |

|  |  |  |
| --- | --- | --- |
| dbr:a3 | dbp:artist | dbr:b2 |
| dbr:a3 | dbp:producer | dbr:p3 |

מהי תוצאת הפעלת השאילתה הבאה על המאגר? (4 נק')

|  |
| --- |
| SELECT ?person ?g ?album {  ?person rdf:type dbo:person.  ?band dbp:member ?person.  GRAPH ?g {  ?album dbp:artist ?band.  OPTIONAL { ?album dbp:producer ?person. }  }  } |