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

מועד א'

25 ביולי 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

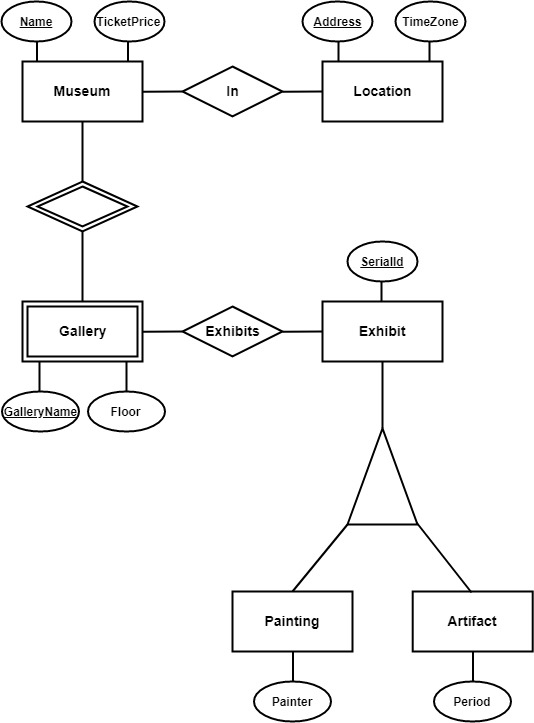
הקורס. כל שימוש בסימון שאינו כזה מחייב הסבר מלא של משמעות הסימון.

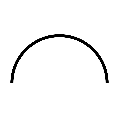
1. משך הבחינה הינו שלוש שעות, תכננו את הזמן בהתאם.
2. אין לכתוב בעפרון.

בהצלחה!

**שאלה 1: ERD (12 נק')**

התבוננו בתרשים ה-ERD שלפניכם:





**ISA**

1. תרגמו את התרשים למסד נתונים על פי הכללים שנלמדו. עבור כל ישות, עליכם לרשום את סכמת הטבלה שתתקבל בתרגום, **כולל סימון מפתחות בקו תחתון וציון מפתחות זרים**:
2. ענו על השאלות הבאות. **נמקו בקצרה את תשובתכם**:
3. האם תתכנה במסד הנתונים שתי גלריות שונות עם GalleryName זהה?
4. נתון כי קיימים, בסך הכול, 2 גלריות (Gallery) ו-3 אתרים (Location) במסד הנתונים. כמה מוזיאונים לכל הפחות קיימים במסד הנתונים?
5. האם ייתכן מוצג (Exhibit) בגלרייה (Gallery) שהוא אינו Painting ואינו Artifact?
6. הקיפו בעיגול את הטענה הנכונה:
7. כל מיקום (Location) יכול לעמוד ביחס In עם 0 או 1 מוזיאונים.
8. כל מוזיאון (Museum) יכול לעמוד ביחס In עם 0 או 1 מיקומים.
9. חייב להיות לפחות מוזיאון אחד שעומד ביחס In עם מיקום כלשהו.
10. חייב להיות לפחות מיקום אחד שעומד ביחס In עם מוזיאון כלשהו.
11. תשובות b,c נכונות.

**שאלה 2:RA/RC (20 נק')**

תהי WorksIn(person, company) סכמה, ו-Q שאילתה מעל הסכמה המחזירה את כל האנשים (person) העובדים **בשתי חברות או יותר**.

1. נסחו את Q ב-Relational Calculus.
2. הקיפו בעיגול את האפשרות הנכונה:
3. **ניתן** לנסח את Q ב-RA. **בנוסף,** **ניתן** לנסח את Q ב-RC כך ש-Q תהיה תלויה בתחום (כלומר לא תהיה DI).
4. **לא ניתן** לנסח את Q ב-RA. **בנוסף, ניתן** לנסח את Q ב-RC כך ש-Q תהיה תלויה בתחום (כלומר לא תהיה DI).
5. **ניתן** לנסח את Q ב-RA. **בנוסף, לא ניתן** לנסח את Q ב-RC כך ש-Q תהיה תלויה בתחום (כלומר לא תהיה DI).
6. **לא ניתן** לנסח את Q ב-RA. **בנוסף, ניתן** לנסח את Q ב-RC כך ש-Q תהיה תלויה בתחום (כלומר לא תהיה DI).
7. **הוכיחו** כי Q לא יכולה להיות מנוסחת ב-RA מעל הסכמה WorksIn על ידי שימוש אך ורק בקבוצת האופרטורים { {𝜋כלומר: איחוד, מכפלה קרטזית, שינוי שם, והטלה. קחו בחשבון כי סינגלטונים הינם חוקיים בשפה.

**הדרכה**: התבוננו במסד הנתונים: .

1. תהי OwnsHouseIn(person, city) סכמה נוספת במסד הנתונים. כתבו שאילתת RA מעל

WorksIn, OwnsHouseIn המחזירה את כל הזוגות {person, company} בהם person הוא אדם אשר בבעלותו דירה בכל אחת מן הערים הקיימות במסד הנתונים, ו- company היא חברה בה אותו אדם עובד.

**שאלה 3: SQL (20 נק')**

התבוננו במסד הנתונים הבא המכיל מידע על תקציבים, הזמנות ומחירי פריטים אותם מזמינות מחלקות בחברה מסוימת:

* **Budgets (deptId, year ,budget)**

משמעות התכונות: מס' מזהה של המחלקה, שנה, תקציב המחלקה לאותה שנה.

* **Acquisitions (aId, deptId, item, year)**

משמעות התכונות: מס' מזהה של ההזמנה, מס' מזהה של המחלקה המזמינה, שם הפריט, שנת ההזמנה.

* **Items (item, price)**

משמעות התכונות: שם הפריט, מחיר הפריט. (אנחנו מניחים כי מחירו של פריט אינו משתנה לאורך השנים.)

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

Budgets:

|  |  |  |
| --- | --- | --- |
| **deptId** | **year** | **budget** |
| 1 | 2019 | 1,000,000 |
| 2 | 2019 | 800,000 |
| 1 | 2018 | 900,000 |

Acquisitions:

|  |  |  |  |
| --- | --- | --- | --- |
| **aId** | **deptId** | **item** | **year** |
| 100 | 3 | Iron bar | 2019 |
| 101 | 3 | Iron bar | 2019 |
| 102 | 2 | Cement sack | 2018 |
| 103 | 1 | Steel plate | 2017 |

Items:

|  |  |
| --- | --- |
| **item** | **price** |
| Iron bar | 500 |
| Cement sack | 200 |
| Steel plate | 1,000 |

**שימו לב:** אף אחת מן הטבלאות אינה ריקה.

1. כתבו שאילתה המחזירה את המספרים המזהים של המחלקות (deptId) אשר ביצעו לפחות 50 הזמנות שונות בשנת 2017.
2. כתבו שאילתה המחזירה את המס' המזהים של המחלקות אשר הזמינו בעבר שק חול (sand bag) או פריט אחר שמחירו זהה למחירו של שק חול.
3. כתבו שאילתה המחזירה את המספרים המזהים של המחלקות אשר הזמינו בשנת 2018 את כל הפריטים הקיימים במסד הנתונים. ניתן להניח שכל הפריטים מופיעים בטבלה Items.
4. מנכ"ל החברה מעוניין לדעת אילו מחלקות חרגו בעבר מן התקציב השנתי שלהם.

כתבו שאילתה המחזירה את המס' המזהים של המחלקות אשר חרגו מן התקציב שלהן בשנה מסוימת (ורק אותן), יחד עם סכום החריגה עבור כל מחלקה. על תוצאות השאילתה להיות ממוינות מן החריגה החמורה ביותר לקלה ביותר. הניחו כי עבור כל מחלקה ועבור כל שנה קיימת רשומה מתאימה בטבלה Budgets.

**הבהרה:** חריגה של מחלקה מתקציבה מתרחשת כאשר סך מחירי ההזמנות השנתיות של המחלקה גדול ממש מתקציבה השנתי. במידה וקיימת חריגה, ערכה יהיה ההפרש בין שני גדלים אלה.

**שאלה 4: Design (28 נק')**

נתונה הסכמה :

1. מצאו כיסוי מינימלי עבור . אין צורך להוכיח נכונות.
2. **הוכיחו** כי לסכמה מפתח קביל יחיד.
3. לפניכם טבלה בעלת שתי עמודות. עבור כל תלות בעמודה הימנית, סמנו "+" בעמודה השמאלית אם או "-" אם .

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |

1. *התבוננו בפירוק הבא של לשלוש תת סכמות: .*
2. *הוכיחו כי הפירוק אינו משמר מידע, באמצעות האלגוריתם הנלמד בכיתה.*

1. הראו דוגמה ליחס עבורו אין שימור מידע ע״י הפירוק (כלומר, דוגמא נגדית לשימור מידע):

.

1. *עבור כל אחד מבין הפירוקים הבאים של , הקיפו בעיגול את הטענה הנכונה.*

1. *הפירוק* ***אינו*** *פירוק BCNF* ***ואינו*** *פירוק NF3.*
2. *הפירוק* ***הינו*** *פירוק BCNF* ***אך אינו*** *פירוק NF3.*
3. *הפירוק* ***הינו*** *פירוק NF3* ***אך אינו*** *פירוק BCNF.*
4. *הפירוק* ***הינו*** *פירוק NF3* ***וגם הינו*** *פירוק BCNF.*

*2.*

1. *הפירוק* ***אינו*** *פירוק BCNF* ***ואינו*** *פירוק NF3.*
2. *הפירוק* ***הינו*** *פירוק BCNF* ***אך אינו*** *פירוק NF3.*
3. *הפירוק* ***הינו*** *פירוק NF3* ***אך אינו*** *פירוק BCNF.*
4. *הפירוק* ***הינו*** *פירוק NF3* ***וגם הינו*** *פירוק BCNF.*
5. *תהיינה קבוצות תלויות פונקציונליות מעל סכמה המקיימות , ונתון כי קיים אטריביוט המופיע באגף שמאל של כל אחת מן התלויות ב-.*

*תהי תלות לא טריוויאלית ב-. הוכיחו כי .* (6 נק')

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

עקב הביקוש הגבוה, הנהלת הפקולטה למדעי המחשב מימשה מפה של הטכניון באמצעות גרף Neo4j. הגרף מומש לפי הכללים הבאים:

1. כל צומת בגרף מחזיק בתווית (label) אחת בדיוק מבין האפשרויות הבאות: **Faculty**, **Dorms**, **Other**.
2. כל אחד מן הצמתים מחזיק בתכונה (Attribute) בשם **Name**.
3. צמתים בעלי תווית **Faculty** מחזיקים בתכונה בשם **NumStudents**.
4. צמתים בעלי תווית **Dorms** מחזיקים בתכונה בשם **Price**.
5. צמתים בגרף מחוברים על ידי קשרים בעלי label אחד בדיוק מבין האפשרויות הבאות: **Road**, **Trail**.

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

:Road

:Trail

:Trail

1. הוחלט להוסיף לצומת המייצג את הפקולטה למדעי המחשב תכונה בשם **MainBuilding** אשר תחזיק את הערך **‘Taub’**. כתבו קטע קוד שמבצע פעולת הוספה זו.
2. כתבו שאילתה המחזירה את כל המעונות (Dorms) מהם ניתן להגיע **גם** לפקולטה לרפואה **וגם** לפקולטה להנדסה אזרחית במעבר על פני כביש (Road) בודד (**שימו לב:** לאו דווקא מדובר על אותו קטע כביש בשני התנאים).
3. כתבו שאילתת Cypher המחזירה את כל המסלולים המורכבים מכבישים (Road) בלבד ומחברים בין הפקולטה למדעי המחשב לכל המעונות שמחירם הוא לפחות 1,000 שקלים בחודש.
4. כתבו שאילתת Cypher המחזירה את שמות המקומות בטכניון מהם ניתן להגיע לכל פקולטה ע״י מסלול בעל לכל היותר ארבעה צמתים ובו כל הקשתות הן מסוג שביל (Trail).

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

לרשותכם מסד נתונים ובו אוסף (Collection) בודד הנקרא Films אשר מכיל מידע על סרטים. כל מסמך באוסף הוא מהצורה הבאה:

{

“\_id”: <ObjectId>

“name”: <string>

“year”: <int>

“country”: <string>

“genre”: <string>

“revenue”: <int>

“actors”: [

{

“actorName”: <string>

“role”: <string>

}

…

{

“actorName”: <string>

“role”: <string>

}

}

דוגמא למסמך אפשרי:

{

“\_id”: ObjectId(056ab83901a09b)

“name”: “Titanic”

“year”: 1997

“country”: “USA”

“genre”: “Drama”

“revenue”: 2,187,000,000

“actors”: [

{

“actorName”: Leonardo DiCaprio

“role”: “Jack”

}

{

“actorName”: “Kate Winslet”

“role”: “Rose”

}

}

1. כתבו שאילתה המחזירה עבור כל מדינה (country) את מספר הסרטים באוסף אשר שייכים לה. החזירו את הרשימה ממוינת כך שהמדינה בעלת מספר הסרטים הרב ביותר תופיע ראשונה והמדינה בעלת מספר הסרטים הנמוך ביותר תופיע אחרונה.
2. השתמשו ב-Map Reduce על מנת להחזיר **זוגות של ז'אנר** (genre) **והכנסות** (revenue), כך שעבור כל ז'אנר מוחזרות סך ההכנסות של סרטים איטלקיים השייכים לז'אנר. השלימו את תבנית השאילתה:

DB.Films.mapReduce(

)

1. התבוננו בשאילתה הבאה:

DB.Films.aggregate({

{$unwind: “$actors”},

{$project: {“genre”: 1, “actors.role”: 1}}

})

נניח כי באוסף Films קיים רק מסמך הדוגמא מן העמוד הקודם. **כמה זוגות סך הכול** של key-value תחזיר השאילתה? (לדוגמא, אם השאילתה תחזיר ארבע רשומות ובכל אחת מהן שני זוגות של key-value, התשובה הנכונה תהיה 8). הקיפו בעיגול את האפשרות הנכונה:

1. 0
2. 2
3. 4
4. 6
5. 8

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

בשאלה זו הניחו את קיום ה-namespaces הבאים:

* 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/>

התבוננו בגרף ה-RDF הבא:

|  |  |  |
| --- | --- | --- |
| dbr:Roosevelt | rdf:type | dbo:person |
| dbr:Roosevelt | dbp:birthPlace | dbr:NY |
| dbr:Roosevelt | dbp:deathPlace | dbr:GA |
| dbr:Truman | rdf:type | dbo:person |
| dbr:Truman | dbp:birthPlace | dbr:Missouri |
| dbr:Truman | dbp:deathPlace | dbr:Missouri |
| dbr:Obama | rdf:type | dbo:person |
| dbr:Obama | dbp:birthPlace | dbr:Hawaii |
| dbr:Hawaii | rdf:type | dbo:location |

1. רשמו את תוצאת ההפעלה של כל אחת מן השאילתות הבאות על הגרף הנתון תחת סמנטיקת RDF רגילה (ללא סמנטיקת RDFS):

|  |
| --- |
| 1.  SELECT ?person ?loc {  ?person rdf:type dbo:person.  OPTIONAL { ?person dbp:deathPlace ?loc. }  ?person dbp:birthPlace ?loc.  } |
|  |
| 2.  SELECT ?person ?loc {  ?person rdf:type dbo:person.  OPTIONAL { ?person dbp:deathPlace ?loc. }  ?**someone** dbp:birthPlace ?loc.  } |
|  |
| 3.  SELECT ?person ?loc {  { ?person rdf:type dbo:person.  OPTIONAL { ?person dbp:deathPlace ?loc. }  }  MINUS  { ?person dbp:birthPlace ?loc. }  } |
|  |

1. התבוננו בשתי השאילתות הבאות:

|  |
| --- |
| 1.  SELECT ?person ?loc {  ?person rdf:type dbo:person.  ?person dbp:birthPlace ?loc.  ?loc rdf:type dbo:location.  } |
| 2.  SELECT ?person ?loc {  ?person rdf:type dbo:person.  ?person dbp:birthPlace ?loc.  OPTIONAL { ?loc rdf:type dbo:location. }  } |

מצאו שלשה בודדת אשר הנחת קיומה תהפוך את שתי השאילתות לעיל **לשקולות** **תחת סמנטיקת RDFS**. הסבירו בקצרה את תשובתכם.

בהצלחה!