

שאלה 1

- א. הטענה נכונה. ניקח לדוגמה $R(A), F = \{ \}$, הוא ב-BCNF שכן מתקיים באופן ריק, זאת גם כאשר מוסיפים אטריבוטים נוספים.
- ב. הטענה נכונה. ניקח לדוגמה $R(A, B), F = \{ A \rightarrow B \}$, הוא ב-3NF שכן A מפתח ובפרט מפתח על. בהוספת C מתקיים שהמפתח היחיד הינו AC, מה שאומר שקודם כל התלות אינה טריוויאלית וכן A כבר לא מפתח על שכן $C \notin A^+$ ו-B לא A-ולא במפתח. כלומר R החדשה לא ב-3NF ולא ב-BCNF.
- ג. הטענה אינה נכונה. יהי R שב-3NF ולא ב-BCNF. זה אומר שקיים תלות $X \rightarrow Y$ כך שאינה טריוויאלית, X אינו מפתח על וכן קיים $A \in Y$ לכך שאינה חלק מX אבל הוא כן חלק ממפתח. לאחר הוספת אטריבוט חדש M, אז המפתחות נשארים זרים מלבד זה שנוסף לכולם גם M*, שכן התלויות לא השתנו. אם בשלילה לאחר ההוספה R תהיה ב-BCNF, אז X יהיה מפתח על (שכן לא טריוויאלי), וזו סתירה לכך שבכל המפתחות יש M ובפרט גם לכל מפתחות העל, שכן $M \notin X$.

שאלה 2

- א. $\{iau_id1, year\}, \{eng_name, year\}, \{orig_name, year\}$.
- ב. בתלות הראשונה, $country \rightarrow countrycode, incomegroup$, התלות אינה טריוויאלית, country אינו מפתח על ולכן אינה בצורה BCNF ובנוסף countrycode לא חלק מהמפתחות שבסעיף א, לכן גם אינה ב-3NF.
- ג. ראשית למען הנוחות נשנה את השמות ל-ABC בהתאמה ללא F, $country=A$, $countrycode=B$ וכ.

$$F = \{A \rightarrow BD, B \rightarrow ACD, E \rightarrow GHI, G \rightarrow HEA, H \rightarrow GJK, EL \rightarrow MI\}$$

$$R_1 = EGHA, R_2 = ELMI, R_3 = ABCD, R_4 = GJK$$

קיבלנו את הטבלה הבאה במהלך הרצת האלגוריתם (מכאן אין טעם להמשיך לבדוק את המשך הריצה): בכחול זה ההתחלתי. היכן שריק אז זה $b_{i,j}$

	A	B	C	D	E	G	H	I	K	L	M
EGHA	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	
ELMI	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	a_{11}
ABCD	a_1	a_2	a_3	a_4							
GJK	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}	

יש שורות של a בלבד ולכן הפירוק הינו ללא אובדן.

- ד. $country \rightarrow countrycode$
 $countrycode \rightarrow country$
 $countrycode \rightarrow region$
 $countrycode \rightarrow incomegroup$
 $iau_id1 \rightarrow orig_name$
 $iau_id1 \rightarrow foundedyr$
 $eng_name \rightarrow iau_id1$
 $eng_name \rightarrow country$
 $orig_name \rightarrow eng_name$
 $orig_name \rightarrow latitude$

orig_name → longitude
 iau_id1, year → students5_estimated

ה. ניצור את את הסכימות ונמחק מיותרות (אחת שמוכלת באחרת):

R1=country, countrycode
~~countrycode, country~~
 R2=countrycode, region
 R3=countrycode, incomegroup
 R4=iau_id1, orig_name
 R5=iau_id1, foundedyr
 R6=eng_name, iau_id1
 R7=eng_name, country
 R8=orig_name, eng_name
 R9=orig_name, latitude
 R10=orig_name, longitude
 R11=iau_id1, year, students5_estimated

נשים לב שהאחרון מכיל את המפתח iau_id1, year ולכן תתי סכמות בפירוק זה הן לפחות בNF3. נשים לב שתתי סכמות R1 – R10 הן יחסים בעלי 2 עמודות בלבד, לכן בפרט הן ב-BCNF. R11 גם בBCNF שכן היחס היחיד שקיים הינו:
 iau_id1, year → students5_estimated
 וצד שמאל כמובן.

ו. המודגשים הם הסופיים כמובן.

נשים לב: country+ = {country, countrycode, incomegroup, region} מפר את הכלל, לכן ניצור: **R1 = country, countrycode, Incomegroup, region**

R2

= country, iau_id1, eng_name, orig_name, foundedyr, latitude, longitude, year, students5_estimated
עבור R1:

FR1= country → countrycode, incomegroup
 countrycode → country, region, incomegroup
 בשני היחסים צד שמאל הוא מפתח על ולכן תת סכימה זאת בBCNF.

עבור R2:

FR2= iau_id1 → eng_name, orig_name, foundedyr
 eng_name → orig_name, iau_id1, country
 orig_name → eng_name, latitude, longitude
 iau_id1, year → students5_estimated, foundedyr
 יש הפרה של הכלל:

(year, students5_estimated) ≠ iau_id1+ = {iau_id1, eng_name, orig_name, foundedyr, country, latitude, longitude}

לכן נגדיר: **R21=iau_id1, eng_name, orig_name, foundedyr, country, latitude, longitude**

R22= iau_id1, year, students5_estimated

עבור R21:

FR21= iau_id1 → eng_name, orig_name, foundedyr
 eng_name → orig_name, iau_id1, country
 orig_name → eng_name, latitude, longitude
 כל השמאליים מפתחות על ולכן תת סכימה זו בBCNF.

עבור R22:

FR22= iau_id1, year \rightarrow students5_estimated
 וצד שמאל כמובן מפתח על ולכן תת סכימה זאת ב-BCNF.
 ז. ראשית נשים לב שהתלויים הבאות נמצאות כולן בתת סכימה כלשהי ($XY \in R_i$) ולכן נשמרות:
 R1-ב country \rightarrow countrycode, incomegroup
 R1-ב countrycode \rightarrow country, region, incomegroup
 R21-ב iau_id1 \rightarrow eng_name, orig_name, foundedyr
 R21-ב eng_name \rightarrow orig_name, iau_id1, country
 R21-ב orig_name \rightarrow eng_name, latitude, longitude
 אבל לא מתקיים עבור: iau_id1, year \rightarrow students5_estimated, foundedyr
 נבדוק תלות זאת ע"י האלגוריתם שראינו בהרצאה, נסמן שוב ב-ABC, $EL \rightarrow MI$:

$Z = EL$
 $Z = Z \cup ((Z \cap R_{22})^+ \cap R_{22}) = EL \cup ((ELM \cap EL)^+ \cap ELM) = EL \cup (EL^+ \cap ELM)$
 $= EL \cup (ELMI \cap ELM) = ELM$
 $Z = Z \cup ((Z \cap R_{21})^+ \cap R_{21}) = ELM \cup ((AEGHIJK \cap ELM)^+ \cap AEGHIJK)$
 $= ELM \cup (E^+ \cap AEGHIJK)$
 $= ELM \cup (EGHIABCDJK \cap AEGHIJK) = ELM \cup AEGHIJK$
 $= AEGHIJKLM$
 מתקיים כי $MI \in Z = AEGHIJKLM$ ולכן תלות זאת גם כן נשמרת.
 כלומר הפירוק אכן משמר תלויות.

שאלה 3

- א. $F = \{ \text{Name} \rightarrow \text{Author}$
 $\text{Name} \rightarrow \text{Genre}$
 $\text{Name} \rightarrow \text{User_Rating}$
 $\text{Name} \rightarrow \text{Reviews}$
 $\text{Name} \rightarrow \text{Price} \}$
 ב. המפתח היחיד הינו $\{ \text{Name}, \text{Year} \}$.
 ג. בכל התלויים $X = \text{Name}$ והינו מפתח ובפרט מפתח על, לכן הצורה הנורמלית היא BCNF.
 ד. קבצים
 ה. המופרות:
 $\text{Name} \rightarrow \text{User_Rating}$
 $\text{Name} \rightarrow \text{Reviews}$
 $\text{Name} \rightarrow \text{Price} \}$
 הנשמרות:
 $\text{Name} \rightarrow \text{Author}$
 $\text{Name} \rightarrow \text{Genre}$
 ו. $R_1 = \{ \text{Name}, \text{Year} \}$
 $R_2 = \{ \text{Name}, \text{Author}, \text{Genre}, \text{User_Rating}, \text{Reviews}, \text{Price} \}$
 נשים לב שמצד אחד התלויים נשמרות שכן כל אחת מהן מוכלת בתת סכימה כלשהי כולה.
 בנוסף, הפירוק הוא ב-BCNF שכן הראשונה היא בעלת 2 עמודות והשניה, ההטלה של F עליה היא כל **F המקורית**, ונשים לב שמתקבל שהמפתח הוא $\{ \text{Name} \}$ כלומר בכולן השמאלי הוא מפתח על ולכן גם השניה ב-BCNF.
 $R_2 \subseteq (R_1 \cap R_2)^+ = \text{Name}^+ = \{ \text{Name}, \text{Author}, \text{Genre}, \text{User_Rating}, \text{Reviews}, \text{Price} \}$
 לכן הפירוק הוא גם ללא אובדן.
 שמירת הנתונים בפירוק היתה מונעת הכנסת שורות לא קונסיסטנטיות שכן כאן אפשר ש- r_2

תיווצר כך ש-Name יהיה מפתח ואז זה אומר שלא ניתן שיהיה ספר עם שם זהה ואחד השדות האחרים (ללא year) יהיה עם יותר ערך אחד – סותר את היותו מפתח. מצד שני, בטבלה של R1 נקבל שכל ספר יוכל להיות עם שנים שונות שבהן היה רב מכך כדרוש (והגיוני לשים אותו כמפתח זר).

הפעם כן דיברתי במונחים של אילוצים על הטבלאות שכן כפי שהובהר בפורום – אם היינו יוצרים רק טבלה אחת לכל הדאטה אז לא משנה כמה אילוצים היינו שמים בה, לא היינו מצליחים לגרום שהתלויית התקיימו. בפירוק שהצעתי כן הצלחנו לגרום להם להתקיים. מקווה שהבנתי נכון את השאלה.