

מסדי נתונים – 89-281 תרגול 11 - נרמול מסדי נתונים

עמיעד רוזנברג

If all is lost and life are breaking down on you,
you better check for missing semi-colons.

הבהרה

- כל מה שנכתב בקובץ זה מבוסס על הבנתי והידע שלי.
בכל התנגשות בין מה שכתוב כאן לבין מה שהמרצה לימדה בכיתה – עליכם לפעול לפי הוראות המרצה.
כמובן שאשמח שתפנו אלי במקביל ע"מ שאני אוכל לברר את העניין.

הגדרה

• תלות פונקציונאלית

- בכל יחס R , אם יש לנו קב' שדות $A = \{A_1, A_2, \dots, A_n\}$ וקבוצת שדות $B = \{B_1, B_2, \dots, B_n\}$ כך שאם לכל 2 רשומות שבהם יש את אותם ערכים לשדות הקב' A אז יש את אותו הערך לשדות הקב' B , אזי ניתן להגיד שישנה תלות פונקציונאלית בין A לבין B .
- תלות זו תסומן כך: $\{A_1, A_2, \dots, A_n\} \rightarrow \{B_1, B_2, \dots, B_n\}$.
- התלויות הפונקציונאליות ייקבעו **לרוב** ע"פ ההיגיון. (ולעיתים ע"פ הגדרות הלקוח)
- נאמר שתלות פונקציונאלית היא מלאה כאשר B תלוי ב- A אבל אינו תלוי באף תת קבוצה (לא טריוויאלית) של A .

חברת BDMS משלוחים בע"מ

כותרת ההזמנה (Header)

מספר הזמנה
תאריך הזמנה
מספר הלקוח
שם הלקוח
כתובת הלקוח

פירוט
ההזמנה

מס' המוצר	שם המוצר	כמות	מחיר מוצר
סה"כ			

הערה: במהלך ההסבר לא נרשום בכל פעם את התלות הטריטוריאלי של שדות הטבלה בשדות המפתח. (במקרה שלנו: {מספר הזמנה} ← {כל שאר הפרטים}) פשוט נזכור שתלות זו תמיד קיימת.

- {מספר הזמנה} ← {כל שאר הפרטים}
- {מספר לקוח} ← {שם הלקוח, כתובת הלקוח}
- {מספר המוצר} ← {שם המוצר, מחיר מוצר}
- {מס' הזמנה, מס' מוצר} ← {כמות}

בעיות עם עבודה מול קובץ

- יתירות במידע.
- מופיע מידע כפול. למשל – מס' המוצר ושם המוצר.
- בעיה בהכנסת \ מחיקת מידע חדש .
- כיצד נעדכן מוצר שאף אחד לא הזמין?
- זמן רב בעדכון מידע.
- אם שם של פריט כלשהו משתנה – יש לעדכנו בהרבה מקומות.

שלבים עיקריים

- זיהוי מועמדים למפתחות
- בחירת מפתח ראשי
- ביצוע תהליך נרמול

זיהוי מועמדים למפתחות

- המפתח הוא קבוצת שדות (יכולה להיות בגודל 1) המכילה ערך ייחודי לכל רשומה. כלומר – לא ייתכנו 2 רשומות עם אותם הערכים עבור סט השדות שבקבוצה.
- **מפתח** הוא בעצם תת קבוצה של שדות המגדירה בצורה חד ערכית את שאר השדות בכל אחת מהרשומות בטבלה.
- במקרה שלנו – השדה "מס' הזמנה" הוא מועמד למפתח. (Candidate Key)
- **הגדרה:** מועמד למפתח K עבור היחס R היא קבוצת שדות בעלת 2 התכונות הבאות:
 - ייחודיות – בכל רשומה ביחס R הערכים המופיעים בשדות של K מזהים בצורה ייחודית את אותה העמודה (אין שתי רשומות שונות בעלות אותן ערכים בשדות K).
 - מינימאליות – לכל תת קבוצה של K אין את תכונת הייחודיות.
- המועמד למפתח שיבחר יהיה **המפתח הראשי**. המפסיד יהיה המשני \ חילופי.

זיהוי מועמדים למפתחות

- דוגמא:

- יחס: $R(A,B,C,D,E)$

- תלויות: $A, B, D \rightarrow C$; $B, C \rightarrow D$; $C, D \rightarrow E$

L	M	R
A,B	C,D	E

- נייצר טבלת L, M, R:

- A,B תמיד בשמאל

- C, D בצדדים שונים

- E תמיד בימין

זיהוי מועמדים למפתחות

L	M	R
A,B	C,D	E

• דוגמא:

• יחס: $R(A,B,C,D,E)$

• תלויות: $A, B, D \rightarrow C$; $B, C \rightarrow D$; $C, D \rightarrow E$

• נבדוק האם L יכול להיות מפתח.

• A,B מגדירים רק את עצמם ולכן הם אינם יכולים להיות המפתח.

• כעת ננסה להוסיף את איברי M לפי הסדר, כאשר קודם נבדוק את האפשרות של הוספת שדה אחד (יש לבדוק את כל האפשרויות), לאחר מכן הוספת 2 שדות (יש לבדוק את כל הזוגות) וכן הלאה.

זיהוי מועמדים למפתחות

L	M	R
A,B	C,D	E

• דוגמא:

• יחס: $R(A,B,C,D,E)$

• תלויות: $A, B, D \rightarrow C$; $B, C \rightarrow D$; $C, D \rightarrow E$

• נעבור על איברי M וננסה להוסיף אחד מהם לאיברי L.

• נבדוק את האפשרות A, B, C לפי התלויות:

• A, B, C קובעים את D (כי B, C קובעים את D)

• C, D קובעים את E.

• לכן A, B, C קובעים את כל R. לכן A, B, C הוא מועמד למפתח.

זיהוי מועמדים למפתחות

L	M	R
A,B	C,D	E

• דוגמא:

• יחס: $R(A,B,C,D,E)$

• תלויות: $A, B, D \rightarrow C$; $B, C \rightarrow D$; $C, D \rightarrow E$

• נעבור על איברי M וננסה להוסיף אחד מהם לאיברי L.

• נבדוק את האפשרות A, B, D לפי התלויות:

• A, B, D קובעים את C.

• C, D קובעים את E.

• לכן A, B, D קובעים את כל R ולכן הוא מועמד למפתח.

זיהוי מועמדים למפתחות

L	M	R
A,B	C,D	E

- דוגמא :

- יחס : $R(A,B,C,D,E)$

- תלויות : $A, B, D \rightarrow C$; $B, C \rightarrow D$; $C, D \rightarrow E$

- קיבלנו 2 אפשרויות למפתח ראשי :

- A, B, C

- A, B, D

- המפתח הראשי ייקבע מבין האפשרויות האלה לפי השימוש השכיח יותר.

זיהוי מועמדים למפתחות

L	M	R
B	A,C,D,E	

- הערה חשובה:

- נשנה טיפה את הדוגמא:

- יחס: $R(A,B,C,D,E)$

- תלויות:

$A, B \rightarrow C, D, E$; $D, B, C \rightarrow A$; $E, B, C \rightarrow D, A$; $D, B \rightarrow E$

- בדוגמא זו B אינו המפתח וכאשר נבדוק את B וערך אחד מתוך M נגלה בקלות ש-A, B הוא מועמד למפתח.

- חשוב לשים לב שאנחנו לא עוצרים בשלב הזה! כעת עלינו לבדוק את B עם זוג ערכים אפשרי של M.

- עם זאת, נשים לב שבגלל דרישת המינימאליות של מועמד למפתח – אנחנו יכולים לוותר על הערך A (כלומר, אין צורך לבדוק את B עם זוג ערכים של M הכולל את A משום שברור שאפשרות זו לא תעמוד בדרישת המינימאליות).

- כאשר נמשיך את התהליך נגלה שגם D, B, C וגם E, B, C הם מועמדים למפתח.

- בנקודה זו אנחנו יכולים לעצור משום שברור שלא קיימת קבוצה המכילה את B ו-3 שדות מתוך M שתעמוד בדרישת המינימאליות.

נרמול - הבהרה

- מטרת הנרמול היא להוריד יתירות הקיימת במידע.
- אין צורך ב-3 השדות: כמות, מחיר לתשלום, סכום לתשלום. יותר מכך – שימוש בשלושתם מייצר לנו סיכוי גדול לטעות.
- תהליך הנרמול מצמצם לנו את הסיכוי לטעויות שנובעות מהיתירות במידע. עקרונית, תמיד ניתן לבטל את הסיכוי הזה ע"י התוכנה שנשתמש איתה מעל בסיס הנתונים, אך מטרת הנרמול היא לייצר מצב שלא נצטרך להתעסק כלל בדברים אלו.
- מצד שני תהליך הנרמול מייצר לתוכנה עבודה (יותר חישובים, יותר פעולות JOIN וכו').
- במציאות, בגלל שהנרמול אכן מייצר לנו יותר עבודה – לעיתים נוותר על נקודות מסוימות בנרמול ע"מ לחסוך בפעולות (בד"כ בודקים קבצי לוג ומבצעים חישובים כדי לאתר מקומות בהם וויתור על נרמול והעברת האחריות לתוכנה יחסוך לנו חישובים).

חברת BDMS משלוחים בע"מ

מספר הזמנה
 תאריך הזמנה
 מספר הלקוח
 שם הלקוח
 כתובת הלקוח

כותרת ההזמנה (Header)

פירוט
 ההזמנה

מס' המוצר	שם המוצר	כמות	מחיר מוצר
סה"כ			

מס' הזמנה	תאריך הזמנה	מספר לקוח	שם לקוח	כתובת לקוח	מס' מוצר	שם מוצר	כמות	מחיר מוצר

- {מספר לקוח} ← {שם הלקוח, כתובת הלקוח}
- {מספר המוצר} ← {שם המוצר, מחיר מוצר}
- {מס' הזמנה, מס' מוצר} ← {כמות}

נרמול - 1st Normal Form

מחיר מוצר	כמות	שם מוצר	מס' מוצר	כתובת לקוח	שם לקוח	מספר לקוח	תאריך הזמנה	מס' הזמנה

- {מספר לקוח} ← {שם הלקוח, כתובת הלקוח}
- {מספר המוצר} ← {שם המוצר, מחיר מוצר}
- {מס' הזמנה, מס' מוצר} ← {כמות}

- **1NF – הגדרה:** כל תא בטבלה מכיל ערך יחיד.
- ניתן לראות כי – "מס' המוצר", "שם המוצר", "כמות" ו-"מחיר מוצר" יכילו 2 ערכים בכל תא. (זה מס' הפעמים שאלמנטים אלה חוזרים על עצמם בכל אחד מהטפסים האמיתיים שהיו לנו!)
- **בעיה:** במקרה זה החלטנו שאנחנו חוזרים על השדות שלנו פעמיים (בהתאם לטפסים האמיתיים).
אם מישהו רוצה להזמין 3 מוצרים – נצטרך לפתוח הזמנה חדשה למרות שמדובר באותה הזמנה!
אם למשל נרצה לדעת כמה הזמנות היו לנו מאותו איש (למשל, כדי לדעת כמה משלוחים נעשו אליו), נקבל שהיו לנו 2 הזמנות במקום הזמנה אחת שהיתה בפועל.

נרמול - 1st Normal Form

מחיר מוצר	כמות	שם מוצר	מס' מוצר	כתובת לקוח	שם לקוח	מספר לקוח	תאריך הזמנה	מס' הזמנה

- {מספר לקוח} ← {שם הלקוח, כתובת הלקוח}
- {מספר המוצר} ← {שם המוצר, מחיר מוצר}
- {מס' הזמנה, מס' מוצר} ← {כמות}

- כאשר אנחנו נכשלים להגיע לאיזשהו NF, עלינו ליצור טבלה חדשה.
- הטבלה החדשה תהיה מורכבת מהשדות הכפולים + המפתח של הטבלה המקורית.
- מה יהיה המפתח של אותה טבלה?
- כאשר אנחנו נכשלים בהגעה ל-1NF המפתח של הטבלה החדשה יהיה מורכב מיותר משדה אחד!
- שדה אחד יהיה המפתח של הטבלה ש"נכשלה" (המקורית).
- השדה הנוסף יגיע מהשדות הכפולים. שדה זה יהיה שדה שהשדות הכפולים תלויים בו ו\או במפתח הטבלה המקורית.
- המפתח: מס' הזמנה + מס' מוצר.
- השדות: שם המוצר, כמות, מחיר מוצר.

נרמול - 1st Normal Form

- בסיס הנתונים החדש (במצב 1NF) ייראה כך :

<u>מס'</u> <u>הזמנה</u>	תאריך הזמנה	מספר לקוח	שם לקוח	כתובת לקוח
<u>מס'</u> <u>הזמנה</u>	<u>מס'</u> <u>מוצר</u>	שם מוצר	כמות	מחיר מוצר

- תלויות בטבלה הכחולה :
- $\{\text{מספר לקוח}\} \leftarrow \{\text{שם הלקוח, כתובת הלקוח}\}$
- תלויות בטבלה האדומה :
- $\{\text{מספר המוצר}\} \leftarrow \{\text{שם המוצר, מחיר מוצר}\}$
- $\{\text{מס' הזמנה, מס' מוצר}\} \leftarrow \{\text{כמות}\}$

נרמול - 2nd Normal Form

<u>מס'</u> <u>הזמנה</u>	תאריך הזמנה	מספר לקוח	שם לקוח	כתובת לקוח
----------------------------	----------------	--------------	------------	---------------

• תלויות בטבלה הכחולה :

- {מספר לקוח} ← {שם הלקוח, כתובת הלקוח}

<u>מס'</u> <u>הזמנה</u>	<u>מס'</u> <u>מוצר</u>	שם מוצר	כמות	מחיר מוצר
----------------------------	---------------------------	------------	------	--------------

• תלויות בטבלה האדומה :

- {מספר המוצר} ← {שם המוצר, מחיר מוצר}
- {מס' הזמנה, מס' מוצר} ← {כמות}

- **2NF – הגדרה:** היחס נמצא ב-1NF ואין תלויות חלקיות על המפתח הראשי. השדות בטבלה שאינם חלק מהמפתח הראשי צריכים להיות תלויים בכל שדות המפתח הראשי של אותה הטבלה (תלות פונקציונאלית מלאה). אם יש לנו שדה שאינו במפתח שתלוי רק בחלק מהשדות המוגדרות כמפתח – אנחנו לא ב-2NF.
- ניתן לראות כי השדות "שם מוצר" ו-"מחיר מוצר" תלויים ב-"מס' מוצר" בלבד ואינם תלויים כלל במספר ההזמנה.
- **בעיה:** נניח שיש לנו מוצר שאף אחד לא הזמין אותו. איפה נוכל לתעד את עובדת קיום המוצר ב-DB שלנו? אם נרצה לקבל רשימה של כל המוצרים שאף אחד לא הזמין, המידע בעצם לא קיים בשום מקום ב-DB שלנו.
- יכולנו ליצור "הזמנת דמה" לשמירת פרטים אלו, אבל אז התוכנה הייתה צריכה לדעת להתעלם מהזמנה זו ואנחנו רוצים לפתור את הבעיה ברמת ה-DB.

נרמול - 2nd Normal Form

• תלויות בטבלה הכחולה :

• {מספר לקוח} ← {שם הלקוח, כתובת הלקוח}

<u>מס'</u> <u>הזמנה</u>	תאריך הזמנה	מספר לקוח	שם לקוח	כתובת לקוח
----------------------------	----------------	--------------	------------	---------------

• תלויות בטבלה האדומה :

• {מספר המוצר} ← {שם המוצר, מחיר מוצר}

• {מס' הזמנה, מס' מוצר} ← {כמות}

<u>מס'</u> <u>הזמנה</u>	<u>מס'</u> <u>מוצר</u>	שם מוצר	כמות	מחיר מוצר
----------------------------	---------------------------	------------	------	--------------

• כרגיל – נצטרך ליצור טבלה חדשה.

• המפתח לטבלה החדשה יהיה אותו חלק של המפתח שגרם לבעיה בטבלה הישנה.

• במקרה שלנו – מס' המוצר.

• שאר השדות בטבלה ילכו אחרי המפתח שלהם :

• "שם מוצר" ו-"מחיר מוצר" תלויים ב-"מס' מוצר" ולכן ילכו לטבלה החדשה.

• "כמות" תלוי ב-"מס' הזמנה" וב-"מס' מוצר" ולכן יישאר בטבלה הישנה.

נרמול - 2nd Normal Form

- בסיס הנתונים החדש (במצב 2NF) ייראה כך:

<u>מס'</u> <u>הזמנה</u>	תאריך הזמנה	מספר לקוח	שם לקוח	כתובת לקוח
----------------------------	----------------	--------------	------------	---------------

<u>מס'</u> <u>הזמנה</u>	<u>מס'</u> <u>מוצר</u>	כמות	<u>מס'</u> <u>מוצר</u>	שם מוצר	מחיר מוצר
----------------------------	---------------------------	------	---------------------------	------------	--------------

- תלויות בטבלה הכחולה:

• {מספר לקוח} ← {שם הלקוח, כתובת הלקוח}

נרמול - 3rd Normal Form

- תלויות בטבלה הכחולה:
- {מספר לקוח} ← {שם הלקוח, כתובת הלקוח}

מס' הזמנה	תאריך הזמנה	מספר לקוח	שם לקוח	כתובת לקוח
מס' הזמנה	מס' מוצר	כמות	שם מוצר	מחיר מוצר

- **3NF – הגדרה:** היחס נמצא ב-2NF ואין תלויות טרנזיטיביות במפתח הראשי. אין ביחס שדות שאינם חלק מהמפתח, התלויים בשדות אחרים שאינם מפתח.
- ניתן לראות בטבלה הכחולה כי השדות "שם הלקוח" ו-"כתובת הלקוח", שאינם חלק מהמפתח, תלויים בשדה "מס' הלקוח" שאינו המפתח של הטבלה.
- **בעיה:** דיי דומה לבעיה שהייתה לנו בעבר. נניח שיש לנו לקוח שלא הזמין שם מוצר. איפה נוכל לתעד את עובדת קיום הלקוח ב-DB שלנו?
בעיה נוספת היא עדכון פרטי הלקוח. אם נרצה לעדכן למשל את כתובת הלקוח, במבנה הנוכחי נצטרך לעשות זאת בכל אחת מההזמנות שלו.

נרמול - 3rd Normal Form

- תלויות בטבלה הכחולה:
- {מספר לקוח} ← {שם הלקוח, כתובת הלקוח}

מס' הזמנה	תאריך הזמנה	מספר לקוח	שם לקוח	כתובת לקוח
מס' הזמנה	מס' מוצר	כמות	שם מוצר	מחיר מוצר

- ניצור טבלה חדשה:
- המפתח החדש יהיה השדה שלא היה מפתח שבו היה תלוי השדה השני שלא היה מפתח (או חלק מהמפתח).
- במקרה שלנו – מס' הלקוח.
- שאר השדות הולכים אחרי המפתח שלהם:
- "תאריך הזמנה" ו"מספר לקוח" תלויים ב"מס' הזמנה" ולכן יישארו בטבלה שלהם.
- "שם הלקוח" ו"כתובת הלקוח" תלויים ב"מס' הלקוח" ולכן יעברו לטבלה החדשה.

נרמול - 3rd Normal Form

- בסיס הנתונים החדש (במצב 3NF) ייראה כך:

<u>מספר</u> לקוח	שם לקוח	כתובת לקוח
---------------------	------------	---------------

<u>מס' הזמנה</u>	תאריך הזמנה	מספר לקוח
------------------	----------------	--------------

<u>מס' מוצר</u>	שם מוצר	מחיר מוצר
-----------------	------------	--------------

<u>מס' הזמנה</u>	<u>מס' מוצר</u>	כמות
------------------	-----------------	------

נרמול – הגדרה כללית יותר

- הערה חשובה:

- בהגדרות שראינו עבור 2NF ו-3NF, כאשר חיפשנו תלויות חלקיות וטרנזיטיביות (בהתאמה) חיפשנו אותם מול המפתח הראשי. הגדרה כללית יותר היא שלא יהיו תלויות כאלה גם עבור כל אחד מהמועמדים למפתח (ולא רק עבור המפתח הראשי). כלומר:

- **2NF – הגדרה כללית:** היחס נמצא ב-1NF וכל שדה, שאינו חלק ממועמד למפתח כלשהו, תלוי תלות פונקציונאלית מלאה במועמד למפתח כלשהו. (אין תלות חלקית של שדה שאינו חלק ממועמד למפתח במועמד למפתח כלשהו)

- **3NF – הגדרה כללית:** היחס נמצא ב-2NF ואין שדה, שאינו חלק ממועמד למפתח כלשהו, שתלוי תלות טרנזיטיבית במועמד למפתח כלשהו.

- הגדרה זו מייצרת תהליך נרמול מסובך וארוך יותר (משום שאנחנו צריכים לחפש (ולפתור) תלויות חלקיות וטרנזיטיביות בכל המועמדים למפתח ולא רק במפתח הראשי), אם כי היא תפתור בעיות (יתירות) שעלולות שלא להתגלות במידה ונתייחס רק למפתח הראשי (ולכן היא גם נכונה יותר). עדיין – לעיתים נקבל תוצאה זהה עבור שני ההגדרות. בכל מקרה, הפעולה של פיצול הטבלאות נעשית בצורה דומה גם כאשר אנחנו מתייחסים למועמדים למפתח ולא רק למפתח הראשי.

נרמול - Boyce-Codd Normal Form

- דטרמיננטה:

- נתון: $R(S)$ (היחס R עם הסכמה S).
 $D \subset S$ היא דטרמיננטה אם קיים $X \subset S$ כך שהתלות $D \rightarrow X$ היא תלות פונקציונאלית מלאה.

- דוגמא:

- יש לנו את היחס $R(A,B,C)$.
התלויות שניתנו לנו הם: $\{A, B\} \rightarrow \{C\}$; $\{C\} \rightarrow \{B\}$.
 - יש לנו ביחס 2 דטרמיננטות: $\{A, B\}$, $\{C\}$.

- BCNF - הגדרה: בסיס הנתונים נמצא במצב שבו בכל היחסים כל הדטרמיננטות הן מועמדות למפתח (או לסופר-מפתח) (כלומר, מכילות תת-קבוצה שהיא מפתח).

נרמול - Boyce-Codd Normal Form

- **BCNF - הגדרה:** יחס שבו כל הדטרמיננטות שלו הן מועמדות ל(סופר-)מפתח.
- דוגמא:
 - יש לנו את היחס $R(A,B,C)$.
התלויות שניתנו לנו הם: $\{A, B\} \rightarrow \{C\}$; $\{C\} \rightarrow \{B\}$.
 - יש לנו ביחס 2 דטרמיננטות: $\{A, B\}$, $\{C\}$.
- ניתן לראות שהמפתח הוא $\{A, B\}$, כלומר $\{C\}$ היא דטרמיננטה שאינה מועמדת ל(סופר-)מפתח.
- הפתרון, כרגיל – טבלה חדשה:
 - עבור $X, D \subset S$ כאשר $D \rightarrow X$ FFD: $D \rightarrow X$ היא דטרמיננטה שאינה מפתח וגוררת בצורה מלאה את X אזי היחס R יתפרק ל- $R_1(D, X), R_2(S - X)$
 - במקרה שלנו: $R_1(C, B), R_2(A, C)$
- יש להמשיך את התהליך עד שבסיס הנתונים הוא ב-BCNF.

נרמול - 4th Normal Form

- תלות רב ערכית (לא טריוויאלית):

- נתון $R(S)$, כאשר $\{X, Y\} \subset S$ (כלומר, ישנם ערכים נוספים מעבר ל- X, Y (אחרת אין תלות רב-ערכית לא טריוויאלית)).
תלות רב ערכית בין X ל- Y (נסמן זאת כך: $X \twoheadrightarrow Y$) היא כאשר לכל ערך ב- X יש קב' ערכים ב- Y וקבוצת ערכים ב- $S-X-Y$ אך קב' הערכים ב- Y אינה תלויה כלל בקב' הערכים ב- $S-X-Y$.

- דוגמא:

- נניח שיש לנו טבלה של קורסים המפרטת את המרצים ואת הספרים בקורס. לכל קורס יש מס' מרצים וכן מס' ספרים.
היחס יהיה (מס' קורס, מס' מרצה, מס' ספר).
היחס אכן נמצא ב- $BCNF$. אין לנו אף דטרמיננטה שאינה מפתח.
נשים לב שלכל ערך של מס' קורס יש קב' של מרצים אפשריים.
בנוסף, לכל ערך של מס' קורס יש קב' של ספרים אפשריים.
לכן ניתן לסמן: מס' קורס \twoheadrightarrow מס' מרצה, וכן מס' קורס \twoheadrightarrow מס' ספר.
שימו לב שהשדה "מס' מרצה" (Y) אינו תלוי בשדה "מס' ספר" ($S-X-Y$)
וכן "מס' ספר" (Y) אינו תלוי בשדה "מס' מרצה" ($S-X-Y$).
במצב זה יש לנו תלות רב ערכית.

- 4NF - הגדרה: בסיס הנתונים במצב שבו אין לנו תלות רב ערכית באף יחס.

נרמול - 4th Normal Form

- 4NF - הגדרה: בסיס הנתונים במצב שבו אין לנו תלות רב ערכית באף יחס.

- דוגמא:

- נניח שיש לנו טבלה של קורסים המפרטת את המרצים ואת הספרים בקורס. לכל קורס יש מס' מרצים וכן מס' ספרים.

היחס יהיה (מס' קורס, מס' מרצה, מס' ספר).

היחס אכן נמצא ב-BCNF. אין לנו אף דטרמיננטה שאינה מפתח.

נשים לב שלכל ערך של מס' קורס יש קב' של מרצים אפשריים.

בנוסף, לכל ערך של מס' קורס יש קב' של ספרים אפשריים.

לכן ניתן לסמן: מס' קורס \leftarrow מס' מרצה, וכן מס' קורס \leftarrow מס' ספר.

שימו לב שהשדה "מס' מרצה" (Y) אינו תלוי בשדה "מס' ספר" ($S-X-Y$)

וכן "מס' ספר" (Y) אינו תלוי בשדה "מס' מרצה" ($S-X-Y$).

במצב זה יש לנו תלות רב ערכית.

- הפתרון:

ע"מ לפתור מצב זה, עלינו לפרק את הטבלה ל-2 טבלאות שונות:

היחס $R(X, Y, Z)$ (כאשר יש תלות רב ערכית בין X ל- Y ו- Z הם שאר השדות)

יתפרק ל- $R_1(X, Z)$, $R_2(X, Y)$.

במקרה שלנו: (מס' קורס, מס' מרצה), (מס' קורס, מס' ספר)

- יש להמשיך את התהליך עד שבסיס הנתונים הוא ב-4NF.