בסיסי נתונים

אילוצים ותלויות פונקציונאליות

*מבוסס על מדריך הלמידה של האוניברסיטה הפתוחה

מכון טכנולוגי חולון
Holon Institute of Technology

דר' ברוריה הברמן דר' מיכל גורדון, 2016

תוכן עניינים

- אילוצי מפתח
- הגדרת אילוצים ב SQL (ראינו ב SQL): אילוצי תחום, אילוצי זיקה, זיקה כתלויות הכלה, זיקה ב SQL, טיפול בהפרת אילוצים
- Triggers ,Assertions
- תלויות פונקציונאליות טריויאליות
- תלויות פונקציונאליות ומפתחות
- סגור של תלויות פונקציונאליות
- סגור של קבוצת תכונות
- אלגוריתם לחישוב סגור
- מציאת מפתחות קבילים
- כיסוי קנוני
- תלויות ו SQL

אילוצים

בבסיסי נתונים שוחחנו על שלושה רכיבים עיקריים:

1. מערכת מושגים לייצוג מידע (מודלים)
2. שפה לטיפול במידע (SQL)
3. **כלים לביטוי אילוצים**

אילוצים

משמשים ל:

- מתן תיאור שלם של המערכת
- השפעה על תהליך התיכון – לדוגמא מידת ריבוי
- הגדרת צירופי נתונים שגויים – בדיקת נתונים בהכנסה
- הגדרת תגובות לאירועים – בדיקת תקינות מורחבת

אילוצים

נחלק את האילוצים ל:

אילוצי תחום – ערך בתכונה הוא מתחום התכונה

אילוצי מפתחות – אין שתי שורות עם אותו מפתח

תלויות פונקציונאליות

אילוצי תחום

ניתן להגדיר תחומי משתנים בתבנית היחסים כדי לבדוק תקינות ערכים ושאילתות.

ב SQL – הגדרות תחום: [] לא חובה

Char[length]

Bit [length]

Numeric [precision[,scale]]

Integer

Smallint

Float [precision], Real , double [precision]

Date time [precision]

NOT NULL – אילוץ הגבלת ערכים ריקים

שלמות קישור *referential integrity*

ליחסים R ו S שיש ביניהן תכונות משותפות, שורות ב R שאינן מצטרפות בצירוף טבעי לאף שורה ב S נקראות **מדולדלות dangling**.

דוגמא: לאלגברה אין שורה מתאימה עם שם מרצה

Course Name	Course No.
Calculus	101
Intro to CS	435
Algebra	231
Statistics	466

Course No.	Credit points
101	4
435	6
466	5

שלמות קישור referential integrity

ניתן לשמור על שלמות דרך אילוצי מפתחות, ולמנוע שורות מדולדלות.

ב SQL:

לשמור על שלמות הקישור ע"י הגדרת foreign key ו primary key.

תכונה השייכת למפתח קביל חייבת להיות מוגדרת כ NOT NULL.

ראינו דוגמאות ב SQL.

הבטחת שלמות קישור – מונעת n-יות מדולדלות בצירוף טבעי.

ראינו אילוצי זיקה ב SQL

אילוצי מפתח

המפתח הוא אילוץ – קבוצת תכונות, כך שאין ביחס שתי שורות שונות עם צירוף ערכים זהה במפתח.

בטיפוסי קשרים:

ברבים-לרבים – צירוף המפתחות של הטיפוסים המשתתפים הוא מפתח קביל.

ברבים-ליחיד – המפתח של טיפוס הישויות המשתתף כרבים מספיק כמפתח קביל.

--> ניתן להסיק מהגדרת המפתח של היחס את מידת הריבוי של טיפוס הקשרים.

אילוצים של מידת ריבוי במודל היחסים

אילוצי ריבוי בטיפוס קשרים מיוצגים במודל היחסים בתבנית יחסים.

כלומר – מידת הריבוי של טיפוס הקשרים מבוטאות במודל היחסים **כאילוצי מפתח** על **תבניות היחסים** המייצגות את טיפוס הקשרים.

כאילוץ פנימי ליחס אחד.

אילוצי השתתפות

במודל ישויות הקשרים – כל ישות מטיפוס E המשתתף בקשר R חייבת להשתתף לפחות במופע אחד של הקשר.

במודל היחסים – כל שורה ביחס המייצג את E חייבת **להשתקף** לפחות בשורה אחת ביחס המייצג את הקשר R. להשתקף כלומר כל ערך שמופיע במפתח של E, יופיע גם בשורה כלשהי ב R.

זהו אילוץ **בין** יחסים. ראינו ב SQL דוגמא.

תלויות פונקציונאליות functional dependencies

סוג של אילוצים על היחסים החוקיים.

הכללה של מושג מפתח על.

מוגדרים תיאורטית במושגי מודל היחסים.

אפשר לבטא אילוצים שביטאנו במודל ER, וגם אילוצים נוספים.

תלויות פונקציונאליות functional dependencies

הגדרה:

R תבנית יחסים, A, ו B קבוצות תכונות מ R.

התלות הפונקציונאלית A→B היא כלל הקובע שלכל שתי שורות עם ערכים זהים בתכונות A, חייבים להיות ערכים זהים בתכונות B.

כלומר: B **תלוי פונקציונאלית ב** A, אם לכל ערך של A, מתאים ערך אחד ויחיד של B בכל היחסים שהם מופעים של התבנית R.

פורמאלית: לכל יחס r(R) ולכל שתי שורות $t_1, t_2 \in r$ שבהן $t_1[A]=t_2[A]$, צריך להתקיים $t_1[B]=t_2[B]$

תלויות פונקציונאליות functional dependencies

ביחס הנ"ל

מתקיים A1→A3

אבל **לא** מתקיים A3→A1. למה?

A1	A2	A3
11	22	33
11	15	33
14	22	56
14	33	56
23	33	56

תלות פונקציונאלית טריוויאלית A→A

דוגמא: תלויות פונקציונאליות

Worker(workerId, departmentNo, manager)

כלומר: שורה מסוג (w,d,m) היא מספר מזהה של עובד, מספר מזהה של מחלקה, ומנהל של העובד.

תלויות החלות על התבנית מוסיפות מידע ונותנות תשובות לשאלות כמו:

1. האם עובד יכול לעבוד ביותר ממחלקה אחת?
2. האם לכל מחלקה מנהל אחד?
3. האם אדם יכול לנהל יותר ממחלקה אחת?
4. האם עובד יכול להיות מנוהל ע"י יותר ממנהל אחד באותה מחלקה?

התשובות ישתנו לפי התלויות הפונקציונאליות.

למה זה טוב?

Worker(workerId, departmentNo, manager)

עד היום לא ביטאנו אילוצים כאלה. מעכשיו נוכל, ונוכל גם לחשב מה אפשרי, מה ניתן להסיק, ולחשב מפתחות.

דוגמא: תלויות פונקציונאליות

Worker(workerId, departmentNo, manager)

מערכת תלויות פונקציונאליות לדוגמא: (מניחים כי **רק** התלויות שניתנות חלות)

workerId → departmentNo

לא יתכנו שתי שורות עם ערכים זהים ב workerId אך ערכים שונים ב departmentNo. כלומר לכל מזהה עובד ביחס, מתאים מספר מחלקה אחד ויחיד.

workerId → manager

לכל עובד יש מנהל יחיד.

דוגמא: תלויות פונקציונאליות

Worker(workerId, departmentNo, manager)

workerId → departmentNo

workerId → manager

1. האם עובד יכול לעבוד ביותר ממחלקה אחת? **לא**
2. האם לכל מחלקה מנהל אחד? **לא, כי לא נאמר כלום**
3. האם אדם יכול לנהל יותר ממחלקה אחת? **כן, כי לא נאמר כלום**
4. האם עובד יכול להיות מנוהל ע"י יותר ממנהל אחד באותה מחלקה? **לא**

דוגמא 2:

Worker(workerId, departmentNo, manager)

workerId → departmentNo

עובד אינו יכול לעבוד ביותר ממחלקה אחת

departmentNo → manager

לכל מחלקה מנהל יחיד.

.

דוגמא 2:

Worker(workerId, departmentNo, manager)

workerId → departmentNo

departmentNo → manager

1. האם עובד יכול לעבוד ביותר ממחלקה אחת? **לא**
2. האם לכל מחלקה מנהל אחד? **כן**
3. האם אדם יכול לנהל יותר ממחלקה אחת? **כן – אין תלות מגבילה**
4. האם עובד יכול להיות מנוהל ע"י יותר ממנהל אחד באותה מחלקה?

לא מהיסק.

לכל מחלקה מנהל יחיד, לכן לכל עובד באותה מחלקה מנהל יחיד, אחרת היינו מקבלים סתירה ל departmentNo → manager

דוגמא 2:

Worker(workerId, departmentNo, manager)

workerId → departmentNo

departmentNo → manager

אפשר להסיק מכלל הטרנזיטיביות:

workerId → manager

דוגמא 3:

Worker(workerId, departmentNo, manager)

workerId, departmentNo → manager

לצירוף מסוים של עובד ומחלקה מתאים רק מנהל אחד.

דוגמא 3:

Worker(workerId, departmentNo, manager)

workerId, departmentNo → manager

1. האם עובד יכול לעבוד ביותר ממחלקה אחת?
2. האם לכל מחלקה מנהל אחד?
3. האם אדם יכול לנהל יותר ממחלקה אחת?
4. האם עובד יכול להיות מנוהל ע"י יותר ממנהל אחד באותה מחלקה?

לא

חשוב: אין מגבלה על מספר מחלקות לעובד, או על מספר מנהלים במחלקה מסוימת. ההגבלה היא במספר המנהלים הקשורים **לצירוף** של עובד ומחלקה.

דוגמא 3:

Worker(workerId, departmentNo, manager)

workerId, departmentNo → manager

לצירוף מסוים של עובד ומחלקה מתאים רק מנהל אחד.

למשל: עובד המשתתף בפרוייקט אחד במחלקה אחת ובפרוייקט אחר במחלקה אחרת. לכל פרוייקט מנהל אחד, אך במחלקה אחת עובדים על יותר מפרוייקט אחד. כך במחלקה אחת כמה מנהלים, אבל לעובד במחלקה יש מנהל אחד.

מ X,Y → Z **לא ניתן להסיק** X→Z ולא ניתן להסיק Y→Z

שאלה:

Worker(workerId, departmentNo, manager)

איך נבטא תלות שאדם יכול לנהל מחלקה אחת לכל היותר?

manager → departmentNo

תלויות פונקציונליות ומפתחות

- תכונה או קבוצת תכונות, שכל תכונות התבנית תלויות בה פונקציונלית היא מפתח של התבנית.
- כך גם בכיוון ההפוך: כל התכונות של תבנית תלויות פונקציונלית בכל תכונה או קבוצת תכונות שהיא מפתח של התבנית.

- P **מפתח על** של R אם ורק אם $P \rightarrow R$

דוגמא: תלויות פונקציונליות במערכת הבנק

Branch Scheme:

Branch-name→branch-city

Branch-name→assets

Customer Scheme:

Name→ city

Name → street

Borrow Scheme:

loan-number→amount

loan-number →branch-name

Deposit Scheme:

account-number →balance

account-number → branch-name

סגור של קבוצות תלויות פונקציונאלית

אם R תבנית, ו F קבוצה של תלויות פונקציונאליות.

ה**סגור של F** מסומן F^+ והוא קבוצת **כל** התלויות הפונקיונאליות הנובעות לוגית מן הקבוצה F.

בהינתן סכימה R וקבוצה F של תלויות פונקציונאליות החלות על R. תלות פונקציונאלית f על R **נובעת לוגית** מ F אם כל יחס r בסכמה R המקיים את F, בהכרח מקיים את f.

דוגמא: עבור הסכמה R=(X,Y,Z) והתלויות X→Y ו Z→Y, אז בהכרח X→Z (לפי הגדרת תלות פונקציונאלית). לכן X→Z נובעת לוגית מהתלויות האחרות.

כללי היסק - כללי ארמסטרונג (אכסיומות):

רפלקסיביות (reflexivity rule):

אם X קבוצת תכונות ו $Y \subseteq X$ אזי $X \rightarrow Y$.

הכללה (augmentation):

אם Z קבוצת תכונות ו $X \rightarrow Y$ אז גם $ZX \rightarrow ZY$

טרנזיטיביות (Transitivity):

אם $X \rightarrow Y$ ו $Y \rightarrow Z$ אז $X \rightarrow Z$

הכללים **נאותים** (לא ניתן ליצור מהם כללים שאינם נובעים לוגית מ F)

ושלמים (ניתן ליצור מהם כל תלות פונקציונאלית בסגור)

חישוב סגור ע"י כללי ארמסטרונג:

איתחול $F^+ = F$

כל זמן שנוספו תלויות ל F^+ המשך:

הפעל את כלל הרפלקסיביות וההכללה עבור כל f in F, והוסף את התוצאה ל F^+.

הפעל את כלל הטרנזיטיביות עבור כל f , g in F עבורם ניתן, והוסף את התוצאה ל F^+.

דוגמא לפי ארמסטרונג לחישוב תלויות פונקציונליות

$F=\{A \rightarrow B, BC \rightarrow D\}$

נחשב סגור של תלויות:

צעד 1: $F^+=\{A \rightarrow B, BC \rightarrow D\}$

צעד 2: $AC \rightarrow BC$ מהכללה

צעד 3: $AC \rightarrow D$ מטרנזיטיביות של $AC \rightarrow BC$ and $BC \rightarrow D$

צעד 4: איחוד כשצד ימין זהה $ABC \rightarrow D$

$F^+=\{A \rightarrow B, BC \rightarrow D, AC \rightarrow BC, AC \rightarrow D, ABC \rightarrow D\}$

חישוב סגור ע"י כללי ארמסטרונג:

איתחול $F^+ = F$

כל זמן שנוספו תלויות ל F^+ המשך:

הפעל את כלל הרפלקסיביות וההכללה עבור כל f in F, והוסף את התוצאה ל F^+.

הפעל את כלל הטרנזיטיביות עבור כל f , g in F עבורם ניתן, והוסף את התוצאה ל F^+.

- החישוב סופי כי עבור סכמה עם n תכונות, ניתן לחשב לכל היותר 2^n*2^n לפי מספר תת הקבוצות האפשריות בכל צד של תלות פונקציונלית.

מכון טכנולוגי חולון
Holon Institute of Technology

כללים נוספים (שניתן להסיק מכללי ארמסטרונג):

איחוד Union:

אם X→Y ו X→Z אז X→YZ (קיצור ל X→Y,Z)

פירוק Decomposition:

אם X→YZ אז X→Y ו X→Z

טרנזיטיביות למחצה Psedutransitivity:

אם X→Y ו **W**Y→Z אז **W**X→Z

<u>הערה</u>**: לא ניתן לפרק צד שמאל של תלות.**

כלומר, אם CD→B , לא ניתן להסיק את זוג התלויות הבא: B → C ו- D→B.

מצד שני, **אפשר לאחד צדי שמאל של שתי תלויות בעלות צד ימין זהה**. כלומר, אם B → C ו- D→B , אזי CD→B.

דוגמא:

הוכח את **כלל הטרנזיטיביות** למחצה מכללי ארמסטרונג.

צריך להוכיח שאם $X \rightarrow Y$ ו- $WY \rightarrow Z$, אזי $WX \rightarrow Z$.

נפעיל את כלל ההכללה עם קבוצת התכונות W, על התלות הפונקציונלית
$X \rightarrow Y$ ונקבל כי $WX \rightarrow WY$.

נפעיל את כלל הטרנזיטיביות על זוג התלויות:
$WX \rightarrow WY$ (הוכחנו) ו- $WY \rightarrow Z$ (נתון) ונקבל כי: $WX \rightarrow Z$

דוגמא:

הוכח את **כלל הפירוק** באמצעות כללי ארמסטרונג.

צריך להוכיח שאם $X \rightarrow YZ$ אזי $X \rightarrow Y$ ו- $X \rightarrow Z$.

(1) $X \rightarrow YZ$ (נתון).
(2) $YZ \rightarrow Y$ (כלל הרפלקסיביות, תלות טריוויאלית).
(3) $X \rightarrow Y$ (לפי 1,2, טרנזיטיביות).
(4) $YZ \rightarrow Z$ (כלל הרפלקסיביות, תלות טריוויאלית).
(5) לכן $X \rightarrow Z$ (לפי 1,4, טרנזיטיביות).

סגור של ת"פ

לא אפקטיבי לחשב סגור של ת"פ.

אבל הגיוני לשאול אם תלות פונקיונאלית נמצאת בסגור של F, ואז צריך לבדוק אם ניתן להסיק אותה.

אפשר גם לשאול האם הסגור של שתי קבוצות של ת"פ שווה.

$F1^+ = F2^+$

איך נבדוק?

סגור של ת"פ

לא אפקטיבי לחשב סגור של ת"פ.

אבל הגיוני לשאול אם תלות פונקיונאלית נמצאת בסגור של F, ואז צריך לבדוק אם ניתן להסיק אותה.

אפשר גם לשאול האם הסגור של שתי קבוצות של ת"פ שווה.

$F1^+ = F2^+$

איך נבדוק?

צריך להראות שכל f_i מ F1 ניתן להסיק מ F2, והפוך, שכל f_j מ F2, ניתן להסיק מ F1.

סגור של קבוצת תכונות

אם A קבוצת תכונות, ו F קבוצת תלויות פונקציונאליות.

הסגור של A ביחס ל F (מסומן A^+) הוא קבוצת כל התכונות הנקבעות פונקציונלית ע"י A בהינתן F.

אלגוריתם:

אתחל CA=A

כל זמן שיש שינויים ב CA בצע:

לכל תלות פונקציונלית D→B שנמצאת ב F

אם B נמצאת ב CA, אז CA=CA U D.

מאפיינים של הסגור

הסגור של קבוצת תכונות מכיל את הקבוצה עצמה: $X \subseteq X^+$

$R = R^+$, עבור המקרה הפרטי $R=X$

אם $Y \subseteq X$ אז גם $Y^+ \subseteq X^+$

הוכחה: אם $Y \subseteq X$ נובע כי $X \rightarrow Y$ לפי הגדרת הסגור $X \rightarrow X^+$ ומטרנזיטיביות

מקבלים $X^+ \rightarrow Y$, $Y \subseteq Y^+$ לכן $Y \rightarrow Y^+$ ומטרזיטיביות $X^+ \rightarrow Y^+$

דוגמא

Worker(workerId, deptNo, manager

workerId→ deptNo, deptNo → manager

נקבע מה הסגור של כל תת קבוצה של תכונות מ :Worker

$\{workerId\}^+ = ?$

$\{deptNo\}^+ = ?$

$\{manager\}^+ = ?$

$\{workerId, deptNo\}^+ = ?$

$\{workerId, manager\}^+ = ?$

$\{deptNo, manager\}^+ = ?$

$\{workerId, deptNo, manager\}^+ = ?$

כל קבוצת תכונות שהסגור שלה הוא Worker היא מפתח.

בדוגמה זו המפתחות הם: {employee_id}, {employee_id, dept_no }, {employee_id, manager} וכן {employee_id, dept_no, manage}.

{employee_id} הוא מפתח קביל, וכל האחרים הם מפתחות על, שכן הם מכילים את {employee_id}.

דוגמא

Worker(workerId, deptNo, manager

workerId→ deptNo, deptNo → manager

נקבע מה הסגור של כל תת קבוצה של תכונות מ Worker:

$\{workerId\}^+ = \{workerId, deptNo, manager\} = Worker$

$\{deptNo\}^+ = \{deptNo, manager\}$

$\{manager\}^+ = \{manager\}$

$\{workerId, deptNo\}^+ = Worker$

$\{workerId, manager\}^+ = Worker$

$\{deptNo, manager\}^+ = \{deptNo, manager\}$

$\{workerId, deptNo, manager\}^+ = Worker$

כל קבוצת תכונות שהסגור שלה הוא Worker היא מפתח.

בדוגמה זו המפתחות הם: {worker_id}, {worker_id, dept_no }, {worker_id, manager} וכן { worker_id, dept_no, manage}.

{worker_id} הוא מפתח קביל, וכל האחרים הם מפתחות על, שכן הם מכילים את {worker_id}.

דוגמא נוספת

Worker(workerId, deptNo, manager

workerId, deptNo → manager

נקבע מה הסגור של כל תת קבוצה של תכונות מ :Worker

$\{workerId\}^+ = \{workerId\}$

$\{deptNo\}^+ = \{deptNo\}$

$\{manager\}^+ = \{manager\}$

$\{workerId, deptNo\}^+ = Worker$

$\{workerId, manager\}^+ = \{workerId, manager\}$

$\{deptNo, manager\}^+ = \{deptNo, manager\}$

$\{workerId, deptNo, manager\}^+ = Worker$

כאן המפתח הקביל הוא {workerId, deptNo}.

דוגמא:

בהנתן $F=\{A \rightarrow CD, B \rightarrow E, C \rightarrow AE\}$

מהו A^+? (קיצור ל $\{A\}^+$)

CA= {A}

CA= {A} and {CD} = {ACD}

CA = {ACD} and {AE} = {ACDE}

אין יותר תלויות שתכונות בצידן השמאלי נמצאות בסגור, והתכונות בצידן הימני לא, לכן סיימנו.

דוגמא:

בהנתן $F=\{A\rightarrow CD, B\rightarrow E, C\rightarrow AE\}$

מהו $\{AB\}^+$?

CA= {AB}

CA= {AB} and {CD} = {ABCD}

CA = {ABCD} and {AE} = {ABCDE}

אין יותר תלויות שתכונות בצידן השמאלי נמצאות בסגור, והתכונות בצידן הימני לא, לכן סיימנו.

דוגמא:

נתונה תבנית $R = (A, B, C, G, H, I)$, וקבוצות התלויות הפונקציונליות $F = \{A \rightarrow B, A \rightarrow C, CG \rightarrow H, CG \rightarrow I, B \rightarrow H\}$. נרשום להלן כמה מהתלויות הפונקציונליות השייכות ל- F^+:

1. התלות $A \rightarrow H$ נובעת מהפעלת כלל הטרנזיטיביות על התלויות הפונקציונליות $A \rightarrow B$ ו- $B \rightarrow H$.
2. התלות $CG \rightarrow HI$ נובעת מהפעלת כלל האיחוד על התלויות הפונקציונליות $CG \rightarrow H$ ו- $CG \rightarrow I$.
3. $AG \rightarrow I$ נובעת מהפעלת כלל ההכללה וכלל הטרנזיטיביות. נראה זאת בשני שלבים.

בשלב ראשון, נפעיל את כלל ההכללה, עם קבוצת התכונות G, על התלות הפונקציונלית $A \rightarrow C$, ונקבל כי $AG \rightarrow CG$.

בשלב שני, נפעיל את כלל הטרנזיטיביות על תוצאה זו:

$AG \rightarrow CG$ ועל $CG \rightarrow I$, ונקבל כי $AG \rightarrow I$.

כללי אצבע לחישוב מפתחות

כדי לקצר את תהליך מציאת המפתחות אפשר להשתמש בכמה שיקולים מקדימים:

- תכונה שאינה תלויה פונקציונלית באף תכונה אחרת חייבת להיות כלולה בכל מפתח קביל. לכן יש לבדוק רק תת קבוצות המכילות אותה.
- אם מצאנו שתכונה מסויימת או קבוצת תכונות היא מפתח קביל – אין צורך לבדוק את התת-קבוצות המכילות אותה.
- אם הסגור של קבוצת תכונות X מכיל מפתח – גם X הוא מפתח.

דוגמא:

הוכח כי:

- לתבנית R=(A, B, C, D, E) עם קבוצת התלויות הפונקציונליות {A→CD, B→E, C→AE} יש אפוא שני מפתחות קבילים: {AB}, {BC}.

B אינה תלויה באף תלות אחרת, לא מופיעה בצד ימין. לכן היא כלולה.

B אינה מפתח קביל, לכן נבדוק תת קבוצות המכילות אותה. AB מפתח קביל.

נבדוק תת-קבוצות המכילות את B ולא את AB. נבדוק את BC, מוסיפים את E, ואת AE, והגענו למפתח לכן נעצור.

תת קבוצות BD, BE, BDE אין צורך לבדוק כי E ו D אינן מופיעים בצד שמאל ולכן לא יוכלו לתרום לסגור אלא את עצמן.

בדיקה האם תלות פונקציונלית שייכת לסגור של קבוצת תלויות

חישוב הסגור של קבוצת תכונות מאפשר לבדוק האם תלות פונקציונלית מסוימת נובעת מקבוצת תלויות נתונה, בלי להשתמש בכללי היסק.

דוגמה: נתונה תבנית R=(A, B, C, D) עם קבוצת תלויות פונקציונליות החלות עליה, F:

$F = \{AB \rightarrow D, C \rightarrow B, CD \rightarrow A, AD \rightarrow C\}$

האם התלות $AC \rightarrow D$ נמצאת ב- F^+?

בדיקה האם תלות פונקציונלית שייכת לסגור של קבוצת תלויות

כדי לענות על שאלה זו יש לחשב את הסגור של קבוצת התכונות {AC} תחת F, ולראות האם $D \subseteq \{AC\}^+$

בשלב הראשון: $\{AC\}^+ = \{AC\}$

בשלב השני מוסיפים את B על פי התלות $C \rightarrow B$ $\{AC\}^+ = \{AC\} \cup \{B\} = \{ACB\}$

בשלב השלישי אפשר כבר להוסיף את D ע"פ $AB \rightarrow D$ $\{AC\}^+ = \{ACB\} \cup \{D\} = R$

ברגע שהוספנו את D לסגור של {AC} אנו יודעים כי התלות $AC \rightarrow D$ חלה על R. שימוש בכללי ההיסק על F כדי להגיע למסקנה זו קשה הרבה יותר.

כיסוי קנוני (כיסוי מינימלי)

יחס המקיים קבוצת תלויות פונקציונליות F מקיים גם את כל התלויות הפונקציונליות ב- F^+. לכן, כדי לבדוק האם יחס מקיים את האילוצים המוגדרים עליו על ידי התלויות הפונקציונליות, מספיק לבדוק אם הוא מקיים את F. למעשה אם נוכל למצוא קבוצת תלויות מצומצמת עוד יותר מ-F, שהסגור שלה הוא כל F^+, נוכל להסתפק בבדיקת התלויות בקבוצה זו.

הרעיון: נפשט את בדיקת קיום הת"פ ע"י הגדרת יצוג שקול ל F אבל פשוט יותר (כשלשניהם אותו סגור).

כיסוי קנוני (כיסוי מינימלי) הגדרה

יהא F קבוצה נתונה של תלויות פונקציונליות. **כיסוי קנוני של F** הוא קבוצת תלויות פונקציונליות F_c, המקיימת:

- $F_c^+ = F^+$. (אפשר לבטא זאת גם כך: כל התלויות ב-F נובעות לוגית מ- F_c, וכל התלויות ב- F_c נובעות לוגית מ- F);
- אף אחת מהתלויות ב- F_c אינה מכילה **תכונה עודפת**;
- אין בתלויות של F_c שתי תלויות עם אותה קבוצת תכונות בצד שמאל.

תכונה עודפת בתלות (בצד שמאל או בצד ימין) אם אפשר למחוק אותה מן התלות ועדיין לקבל את אותו סגור.

כיסוי קנוני (כיסוי מינימלי) הגדרה

באופן פורמלי: יהא F קבוצת תלויות פונקציונליות, ויהא X→Y תלות ב- F.

- תכונה $A \in X$ היא תכונה עודפת (בצד שמאל) אם אפשר להסיק מ- F את התלות (X-A)→Y.
- תכונה $B \in Y$ היא תכונה עודפת (בצד ימין) אם אפשר להסיק מקבוצת התלויות: {X→(Y-B)} U {X→Y} - F את התלות X→Y.
  (כלומר, אם אפשר לשחזר את התלות X→Y לאחר שמוחקים ממנה את התכונה B).

מכון טכנולוגי חולון
Holon Institute of Technology

כיסוי קנוני

האלגוריתם לחישוב **כיסוי קנוני** של קבוצת תלויות F הוא כדלקמן:

1. אתחול: $F_c := F$.
2. אחד ב- F_c תלויות שהן זהות בצד שמאל על ידי איחוד צידי ימין שלהן;
3. בדוק האם יש ב- F_c תלות המכילה תכונה עודפת. אם כן, מחק תלות זו מ- F_c, והוסף במקומה את התלות המתקבלת ממנה **לאחר** מחיקת התכונה העודפת.
4. חזור על שלבים 2 ו- 3 כל עוד אחד מהם לפחות גורם לשינוי ב- F_c.
5. F_c הוא הכיסוי הקנוני.

דוגמא:

יהא $R=(A, B, C, D, E)$ תבנית יחסים, ויהא F קבוצת תלויות פונקציונליות החלה על R: $F= \{A \rightarrow B, CE \rightarrow A, B \rightarrow C, AC \rightarrow D\}$. נבחן האם יש תכונה עודפת באחת התלויות ב- F. שאלה זו רלבנטית רק לגבי התלויות $CE \rightarrow A$ ו- $AC \rightarrow D$ (הן היחידות שיש להן יותר מתכונה אחת בצד כלשהו).

דוגמא:

יהא $R=(A, B, C, D, E)$ תבנית יחסים, ויהא F קבוצת תלויות פונקציונליות החלה על R: $F= \{A\rightarrow B, CE\rightarrow A, B\rightarrow C, AC\rightarrow D\}$.

- כדי לבדוק אם C עודפת ב- $CE\rightarrow A$ יש לבדוק אם $E\rightarrow A$ נובעת מ- F. לשם כך יש לחשב את E^+, ולראות אם הוא מכיל את A. אבל $E^+ = E$, שכן E לבדה אינה מופיעה באף תלות. לכן C אינה עודפת בתלות זו.
- כדי לבדוק אם E עודפת ב- $CE\rightarrow A$ יש לבדוק אם $C\rightarrow A$ נובעת מ- F. גם כאן, $C^+ = C$, ולכן E אינה עודפת בתלות זו.

דוגמא המשך:

- כדי לבדוק אם A עודפת ב- AC→D יש לבדוק אם C→D נובעת מ- F. לשם כך יש לחשב את C^+ תחת F, ולראות אם הוא מכיל את D. כבר ראינו בבדיקה הקודמת כי $C^+ = C$, ולכן A אינה עודפת בתלות זו.
- כדי לבדוק אם C עודפת ב- AC→D יש לבדוק אם A→D נובעת מ- F. לשם כך יש לחשב את הסגור של A, ולראות אם הוא מכיל את D. הפעם חישוב הסגור אינו טריביאלי.

דוגמא המשך:

$F= \{A \rightarrow B, CE \rightarrow A, B \rightarrow C, AC \rightarrow D\}$

בשלב הראשון: $\{A\}^+ = \{A\}$

בשלב השני מוסיפים B ע"פ התלות $A \rightarrow B$: $\{A\}+= \{A\}U\{B\}=\{AB\}$

בשלב השלישי מוסיפים את C, על פי התלות $B \rightarrow C$:

$\{A\}+=\{AB\} \cup \{C\}=\{ABC\}$

בשלב האחרון משתמשים בתלות $AC \rightarrow D$ וכך מוסיפים את D:

$\{A\}+=\{ABC\} \cup \{D\}=\{ABCD\}$

התלות $A \rightarrow D$ נובעת אפוא מ- F, ולכן C תכונה עודפת בתלות $AC \rightarrow D$

כיסוי קנוני (דוגמא)

נחשב את הכיסוי הקנוני של קבוצת התלויות מהדוגמה הקודמת:

$F = \{A \rightarrow B, CE \rightarrow A, B \rightarrow C, AC \rightarrow D\}$

- בצעד ראשון, $F_c := F$.
- בצעד שני יש לאחד תלויות עם צדי שמאל זהים, אך אין כאלה. F_c נשאר אפוא ללא שינוי.
- בצעד השלישי יש לבדוק האם יש תכונות עודפות. ביצענו בדיקה זו בדוגמא הקודמת, ולכן נשנה את F_c: $F_c = \{A \rightarrow B, CE \rightarrow A, B \rightarrow C, A \rightarrow D\}$.
- מכיוון שביצענו שינוי ב- F_c, יש לחזור שוב לצעדים 2 ו-3.
- הפעם יש ב- F_c שתי תלויות עם צידי שמאל זהים, ולכן נשנה שוב את F_c:

$F_c = \{A \rightarrow BD, CE \rightarrow A, B \rightarrow C\}$

כיסוי קנוני (דוגמא המשך)

- עתה יש ב- F_c תלות עם שתי תכונות בצד ימין. יש לבדוק אם אחת מהן עודפת. כדי לבדוק אם D למשל עודפת, יש לבדוק האם התלות A→D נובעת מקבוצת התלויות: {A→B, CE→A, B→C}. לשם כך יש לחשב את A+ תחת תלויות אלה. $A^+ = \{ABC\}$.
- אין הסגור מכיל את D, ולכן D אינה עודפת. באופן דומה אפשר לבדוק ולראות כי גם B אינה עודפת.
- יש לחזור לשלב 2. הפעם אין תלויות שאפשר לאחד, ולכן F_c אינו משתנה. גם שלב 3 אינו מניב שינוי, ולכן האלגוריתם עוצר. קיבלנו אפוא את הכיסוי הקנוני:

$$F_c = \{A \rightarrow BD,\ CE \rightarrow A,\ B \rightarrow C\}$$

מציאת מפתח

נמצא את המפתחות הקבילים של התבנית R, על-פי הכיסוי הקנוני אליו הגענו בדוגמא הקודמת.

ראשית אפשר לראות כי E אינה מופיעה בצד ימין של אף תלות, ולכן היא תהיה כלולה בכל מפתח קביל.

- E לבדה אינה מפתח, שכן $E^+ = E$. ננסה לצרף לה אפוא עוד תכונה. הסגור של A כולל את BD, על-פי התלות $A \rightarrow BD$, ומכאן שהוא כולל גם את C, על-פי התלות $B \rightarrow C$. מקבלים: $R = \{ABCDE\} = \{AE\}^+$, כלומר {AE} מפתח קביל של R.
- עתה ננסה לצרף ל- E את B. הסגור של B כולל את C (שכן $B \rightarrow C$), ומ- C ו- E ביחד מגיעים ל- A. מכיוון שהסגור כולל את A ואת E הוא בהכרח כולל את כל R. מקבלים אפוא $R = \{BE\}^+$, ולכן גם {BE} מפתח קביל.

מציאת מפתח

- באופן דומה אפשר לראות שגם {CE} מפתח. (CE→A, ולכן שוב, מ- A ו-C אפשר להגיע לכל R.)
- נותר לבדוק את {DE}, שכן כל קבוצת התכונות המכילה את E עם B,A או C מכילה כבר מפתח קביל. אבל $\{DE\}^+ = \{DE\}$, כלומר צירוף תכונות זה אינו מפתח.

קיבלנו כי ל- R שלושה מפתחות קבילים: {AE}, {BE} ו- {CE}.

תלויות פונקציונליות ו SQL

ת"פ הן אילוצים המשמשים לשיפור ניתוח מערכת.

ב SQL אין דרך לבטא ת"פ, אך משתמשים באילוצי מפתח שראינו, ואפשר תלויות אחרות לבטא כתנאי לוגי ע"י Assertions.

לדוגמא: ננסח deptNo → manager

נבדוק אם לא קיימות ביחס worker שתי שורות עם ערך זהה ב deptNo וערך שונה ב manager.

```
CREATE ASSERTION  singleMan CHECK (NOT EXISTS(
SELECT w1.dept_no, w2.dept_no FROM worker AS w1, worker AS w2
WHERE w1.dept_no=w2.dept_no AND w1.manager != w2.manager))
```

סיכום

- תלויות פונקציונאליות
- סגור של קבוצת תלויות
- סגור של קבוצת תכונות
- חישוב מפתחות
- כיסוי קנוני של קבוצת תלויות (תכונות עודפות)