תלויות פונקציונליות Functional Dependencies

דוגמה

Flight	Class	Price	Airpl	Num

השדות Flight וClass מגדירים את הטיסה(הם מפתח ראשי) - כלומר אם הם שווים בשני השדות האלה הם שווים בכל שאר השדות - כלומר הם אותה שורה בדיוק!

הגדרה

נתון $R(S)$(יחס R עם סכימה S) ותתי קבוצות $X, Y \subset S$. נגדיר שמתקיים $X \to Y$, כלומר Y תלוי בX, אם

$$\forall_{t_1,t_2} t_1(X) = t_2(X) \Rightarrow t_1(Y) = t_2(Y)$$

דוגמה

Flight	Class	Price	Airpl	Num
123	bus	700	Boeing	250
123	ec	600	Boeign	250
124	bus	500	Airbus	300

- טיסה מתבצעת במטוס אחד, ולכן גם מדגם אחד, ולכן $\text{Flight} \to \text{Aripl}$
- מספר המקומות במטוס קבוע, ולכן $\text{Airpl} \to \text{Num}$
- מכיוון שהמטוס קבוע לטיסה, ומספר המקומות קבוע למטוס, מספר המקומות קבוע לטיסה - $\text{Flight} \to \text{Num}$
- הטיסה והמחלקה מגדירים את המחיר - $\text{Flight}, \text{Class} \to \text{Price}$
- אפשר גם לבודד את המחלקה לפי המחלקה והטיסה, לכן $\text{Flight}, \text{Price} \to \text{Class}$

אלו הם התלויות הפונקציונליות של הטבלה הזו לפי הלוגיקה. כשנרצה למדל את הטבלה, נפעיל הגיון אנושי ונבחר את המפתח $\text{key} : \text{Flight}, \text{Class}$.

אפשר גם לבחור Superkey כך ש$\text{SK} \supseteq \text{Key}$.

ניקח קבוצת תלויות אחרת:

$$F2\left\{\begin{array}{c}\text{Flight}\rightarrow\text{Airpl},\text{Num}\\ \text{Airpl}\rightarrow\text{Num}\\ \text{Flight},\text{Class}\rightarrow\text{Price}\end{array}\right.$$

יש לנו יחס R. נסמן $I_R - \text{instance}(state)$ of R. לR יכולים להיות המון instanceים, המון stateים. אם לכל state לגיטימי תלות פונקציונאלית f מתבצעת, אזי אומרים שR מקיים f - R satisfies f

נסמן $F1 \Leftarrow F2$ - $F1$ גורר $F2$. אם יחס R המקיים את התלויות של $F1$ מקיים גם את התלויות של $F2$.

גרירה בשני הכיוונים היא שקילות - $F1 \equiv F2$ - $F1$ שקול ל$F2$ אם $F1 \Leftarrow F2$ וגם $F2 \Leftarrow F1$. בדוגמה שלנו $F1 \not\equiv F2$. אבל ניתן לקחת

$$F3\left\{\begin{array}{c}\text{Flight}\rightarrow\text{Airpl},\text{Num}\\ \text{Airpl}\rightarrow\text{Num}\\ \text{Flight},\text{Class}\rightarrow\text{Price}\\ \text{Flight},\text{Price}\rightarrow\text{Class}\end{array}\right.$$

ואז $F1 \equiv F3$

כללים

1) Splitting/Combining Rule

נתון $R(S)$, $X, Y \subset S$, $Y = \{y_1, y_2, \dots y_n\}$. נגדיר:

$F1 = \{X \rightarrow Y\}$

$$F2 = (X \rightarrow y_1, \dots X \rightarrow y_n)$$

splitting $F1 \Rightarrow F2$

combining $F2 \Rightarrow F1$

בדוגמה שלנו

- splitting

$$\text{Flight}\rightarrow\text{Airpl}\Rightarrow\begin{array}{c}\text{Flight}\rightarrow\text{Airpl}\\ \text{Flight}\rightarrow\text{Num}\end{array}$$

- combining

$$\text{Flight}\rightarrow\text{Airpl}\Leftarrow\begin{array}{c}\text{Flight}\rightarrow\text{Airpl}\\ \text{Flight}\rightarrow\text{Num}\end{array}$$

סטנדרט standard

כאשר כותבים תלות פונקציונלית, נהוג שבצד ימין יהיה שדה אחד. למשל

$$\text{Flight}, \text{Class}, \text{Airpl} \to \text{Num}$$

הערה: זוהי תלות נכונה, אבל נעדיף לכתוב קבוצה מינימלית בצד שמאל -

$$\cancel{\text{Flight}}, \cancel{\text{Class}}, \text{Airpl} \to \text{Num}$$

2) Trivial FD

אילוץ טריוויאלי על יחס C Trivial constraint הוא תנאי שמתקיים תמיד:

$$\forall_{I_R} C(R) = \text{true}$$

התנאי תמיד מתקיים - לא תלוי בinstance.

דוגמה: $\text{Flight}, \text{Class} \to \text{Flight}$

בהגדרה: נתון $R(S)$, $X, Y \subset S$, $X \subseteq Y$. אזי $X \to Y$ נקראת תלות טריוויאלית

$$\{X \to Y\} \equiv \{X \to Z\}, Z = Y - X$$

זה טריוויאלי ש$X \to Z$, כי $Y - Z \subseteq X$ ו$X \to Y$

בדוגמה שלנו

$$\text{Flight}, \text{Class} \to \text{Class}, \text{Price}$$

ברור ש$\text{Flight}, \text{Class} \to \text{Class}$ - זה ממש מפריע כאן. לכן נוריד את זה:

$$\text{Flight}, \text{Class} \to \text{Price}$$

אמרנו שיכולה להיות גרירה $F2 \Rightarrow F1$ - אם אם $F1(R)$ אז $F2(R)$. האם ניתן להוכיח את זה בצורה פורמלית?

יהי $R(X, Y, Z)$ (יחס R וסכימה X, Y, Z). נניח שמתקיים

$$F1 = \{X \to Y, Y \to Z\} \qquad F2 = \{X \to Z\}$$

רוצים להוכיח

$$F1 \Rightarrow F2$$

הוכחה

$$\left.\begin{array}{c} t_1 = (x, y_1, z_1) \\ t_2 = (x, y_2, z_2) \\ X \to Y \end{array}\right] \Rightarrow y_1 = y_2 = y$$

$$\left.\begin{array}{c} t_1 = (x, y, z_1) \\ t_2 = (x, y, z_2) \\ Y \to Z \end{array}\right] \Rightarrow z_1 = x_2 = z$$

$$\left.\begin{array}{c} t_1 = (x, y, z) \\ t_2 = (x, y, z) \end{array}\right] \Rightarrow X \to Z$$

עיקרון כללי General Principle

Closure of a set of attributes

תהי X – set of attr קבוצת שדות, $X \subseteq S$, וF – set of FDs. נרצה לחשב את

$$X^+ - \text{closure of } X \text{ under } F$$

המוגדר ע"י:
X^+ הוא קבוצת y_i כך ש$F \Rightarrow X \to y_i$. כלומר $X^+ \subseteq S$ מקסימלי המקיים

$$F \Rightarrow X \to X^+$$

במילים

בהינתן קבוצת תלויות פונקציונאליות F וקבוצת שדות X, נרצה למצוא את כל השדות התלויים בX תחת F.

אלגוריתם

% Init
`while(`$\exists\,(f = X' \to Y') \in F, |Y'| > 1$`)`
    `split` f;
X^+=X;
% Body
`while(`$\exists\,(f = X' \to y') \in F, X' \subseteq X^+, y' \notin X^+$`)`
    X^+=$X^+ \cup \{y'\}$
`return` X^+

דוגמה

$$R\left(A,B,C,D,E,G\right)$$

$$F=\{A,B\rightarrow C;B,C\rightarrow A,D;D\rightarrow E;C,G\rightarrow B\}$$

$$X=\{A,B\}$$

בשלב הראשון נפצל את F, כך שבצד ימין יהיה תמיד רק איבר אחד, ומשם כל פעם מוסיפים עוד שדות שתלויים בX^{+}

stage	X^{+}	F	new X^{+}
init	A,B	$A,B\rightarrow C$ $B,C\rightarrow A$ $B,C\rightarrow D$ $D\rightarrow E$ $C,G\rightarrow B$	
1	A,B,C	$B,C\rightarrow A$ $B,C\rightarrow D$	B,C,D
2	A,B,C,D	D→E	A,B,C,D,E
3	A,B,C,D,E	C,G→B	

שימושים Applications

1)

$$\begin{gathered}X\rightarrow Y\Leftarrow F\\ \Updownarrow\\ Y\subset X^{+}\end{gathered}$$

כדי לבדוק אם תלות מתבצעת, נוכל להוכיח בעזרת כללים כלשהם שהיא אכן מתבצעת, אבל זה מאוד קשה ומסובך.

2)

$$\begin{gathered}X^{+}=S\\ \Downarrow\\ X\text{ is Superkey}\end{gathered}$$

ואז אפשר לנסות להוריד תכונות מX ולראות אם מגיעים למפתח.