Systemy baz danych 1

Anna Fiedorowicz

WMIE

10.10.2019

Zależności wielowartościowe

Definicja 5.1

Niech U będzie zbiorem atrybutów oraz $X,Y \subseteq U$, $Z = U \setminus XY$.
Niech R będzie relacją typu U.
Mówimy, że **zależność wielowartościowa** $X \twoheadrightarrow Y$ **jest spełniona** w relacji R, jeżeli dla dowolnych pięciu krotek $x \in R[X]$, $y, y' \in R[Y]$, $z, z' \in R[Z]$ zachodzi, że

$$\text{jeżeli } x \bowtie y \bowtie z \in R \ \& \ x \bowtie y' \bowtie z' \in R,$$

$$\text{to } x \bowtie y' \bowtie z \in R \ \& \ x \bowtie y \bowtie z' \in R.$$

Fakt, że zależność wielowartościowa $X \twoheadrightarrow Y$ zachodzi w relacji R (tzn., jest spełniona w R) oznaczamy $R \models X \twoheadrightarrow Y$.

Zależności wielowartościowe

Twierdzenie 5.1

Niech $R = R(U)$ będzie daną relacją typu U, niech $X,Y \subseteq U$, $Z = U \setminus XY$. Wówczas zależność wielowartościowa $X \twoheadrightarrow Y$ jest spełniona w relacji R wtedy i tylko wtedy, gdy

$$R = R[XY] \bowtie R[XZ].$$

Twierdzenie 5.2

Niech $R = R(U)$ będzie relacją typu U, i $X,Y \subseteq U$, $Z = U \setminus XY$. Wówczas

$$R \models X \rightarrow Y \ \Rightarrow \ R \models X \twoheadrightarrow Y.$$

Uwaga 5.1

Zależności wielowartościowe postaci $X \twoheadrightarrow U$, $X \twoheadrightarrow \emptyset$ nazywamy **trywialnymi**, *ponieważ są spełnione w każdej relacji R.*

Jeżeli $Y = U \setminus X$, to zależność wielowartościowa $X \twoheadrightarrow Y$ jest trywialna.

Przykład

Niech dany będzie zbiór atrybutów $U = \{P, D, Z, R\}$ (gdzie **P**racownik, **D**ziecko, **Z**arobki, **R**ok) i relacja R typu U:

R :	P	D	Z	R
	Jones	Mary	3000	2000
	Jones	Alice	3000	2000
	Jones	Mary	3500	2001
	Jones	Alice	3500	2001
	Brown	Carol	4000	2000
	Brown	Carol	5000	2001

W R spełnione są zależności wielowartościowe:

$P \twoheadrightarrow D$ i $P \twoheadrightarrow ZR$.

Zależności wielowartościowe trywialne to np.:

$PD \twoheadrightarrow ZR$, $P \twoheadrightarrow DZR$.

Przykład

Niech dany będzie zbiór atrybutów $U = \{P, D, Z, R\}$ (gdzie **P**racownik, **D**ziecko, **Z**arobki, **R**ok) i relacja R typu U:

R : P	D	Z	R
Jones	Mary	3000	2000
Jones	Alice	3000	2000
Jones	Mary	3500	2001
Jones	Alice	3500	2001
Brown	Carol	4000	2000
Brown	Carol	5000	2001

W R spełnione są zależności wielowartościowe:

$P \twoheadrightarrow D$ i $P \twoheadrightarrow ZR$.

Po dekompozycji relacji R (zgodnie z Twierdzeniem 5.1 względem $E \twoheadrightarrow C$, otrzymamy projekcje.

R_1 : P	D
Jones	Mary
Jones	Alice
Brown	Carol

R_2 : P	Z	R
Jones	3000	2000
Jones	3500	2001
Brown	4000	2000
Brown	5000	2001

Aksjomaty zależności wielowartościowych

Niech U będzie zbiorem atrybutów. Niech $M \subseteq \{X \twoheadrightarrow Y : X, Y \subseteq U\}$, tzn., M jest podzbiorem zbioru wszystkich możliwych zależności wielowartościowych nad U.

Definicja 5.2

Oznaczmy przez M^+ najmniejszy zbiór zależności wielowartościowych spełniający
M^+ zawiera zbiór M (tzn., $M \subseteq M^+$),
M^+ jest zamknięty ze względu na następujące reguły wyprowadzeń, gdzie $\{X, Y, Z, W \subseteq U\}$.

(M0) $Y \subseteq X \implies X \twoheadrightarrow Y \in M^+$ (Zwrotność)

(M1) $X \twoheadrightarrow Y \in M^+ \implies X \twoheadrightarrow U - XY \in M^+$ (Dopełnialność)

(M2) $X \twoheadrightarrow Y \in M^+ \implies XZ \twoheadrightarrow YZ \in M^+$ (Poszerzalność)

(M3) $(X \twoheadrightarrow Y \in M^+ \wedge Y \twoheadrightarrow Z \in M^+) \implies X \twoheadrightarrow Z \in M^+$ (Przechodniość)

(M4) $(X \twoheadrightarrow Y \in M^+ \wedge YZ \twoheadrightarrow W \in M^+) \Rightarrow XZ \twoheadrightarrow W \in M^+$ (Pseudoprzechodniość)

(M5) $(X \twoheadrightarrow Y \in M^+ \wedge X \twoheadrightarrow Z \in M^+) \implies X \twoheadrightarrow YZ \in M^+$ (Addytywność)

(M6) $(X \twoheadrightarrow Y \in M^+ \wedge X \twoheadrightarrow Z \in M^+) \implies X \twoheadrightarrow Y \cap Z \in M^+$ (Dekompozycyjność)

Zbiór M^+ nazywamy **domknięciem** zbioru M (rodziną zależności wielowartościowych generowaną przez M).

Reguły dla zależności wielowartościowych i funkcyjnych

Następujące reguły wyprowadzeń zachodzą pomiędzy zależnościami funkcyjnymi i wielowartościowymi:

(FM1) $X \to Y \in F^+ \implies X \twoheadrightarrow Y \in M^+$ (replikacja)

(tzn., z zależności funkcyjnej można wyprowadzić odpowiadająca jej zależność wielowartościową)

(FM2) $(X \twoheadrightarrow Z \in M^+ \wedge Y \twoheadrightarrow V \in M^+ \wedge V \subseteq Z \wedge Y \cap Z = \emptyset) \implies X \to V \in F^+$ (scalenie).

Używając aksjomatów (F1)–(F3), (M0)–(M6) i reguł (FM1), (FM2), można otrzymać różnego typu domknięcia zbiorów zależności funkcyjnych F i wielowartościowych M:

- Za pomocą aksjomatów Armstronga (F1)–(F3) dostajemy domknięcie F^+ zbioru zależności funkcyjnych F;
- Za pomocą aksjomatów (M0)–(M6) dostajemy domknięcie M^+ zbioru zależności wielowartościowych M;
- Za pomocą aksjomatów Armstronga (F1)–(F3), aksjomatów (M0)–(M6) i reguł (FM1), (FM2) dostajemy domknięcie $(F \cup M)^+$ obu zbiorów.

Ogólne schematy relacyjne

Niech U będzie zbiorem atrybutów, F będzie zbiorem zależności funkcyjnych nad U i niech M będzie zbiorem tych zależności wielowartościowych nad U, które nie są zależnościami funkcyjnymi.

Definicja 5.3

Parę uporządkowaną $\mathbf{R} = (U, F \cup M)$ nazywamy ogólnym schematem relacyjnym o zbiorze atrybutów U, zbiorze zależności funkcyjnych F i zbiorze zależności wielowartościowych M.

Definicja 5.4

Niech R będzie relacją. Mówimy, że R jest przypadkiem (instancją) ogólnego schematu relacyjnego $\mathbf{R} = (U, F \cup M)$, jeżeli R jest typu U oraz wszystkie zależności funkcyjne ze zbioru F i wszystkie zależności wielowartościowe ze zbioru M są spełnione w R.

Czwarta postać normalna

Definicja 5.5 (Czwarta postać normalna (ozn. 4PN lub 4NF))

Niech $\mathbf{R} = (U, F \cup M)$ będzie ogólnym schematem relacyjnym. Mówimy, że schemat $\mathbf{R}$ jest w czwartej postaci normalnej (4PN), jeżeli z faktu że istnieje nietrywialna zależność wielowartościowa $X \twoheadrightarrow Y \in M^+$, gdzie $Y \subseteq U \setminus X$, wynika istnienie zależności funkcyjnej $X \to U \in F^+$.

Uwaga 5.2

Jeżeli ogólny schemat relacyjny $\mathbf{R}$ jest w 4PN, to

- *wszystkie nietrywialne zależności wielowartościowe w $\mathbf{R}$ są funkcyjne;*
- *dla każdej nietrywialnej zależności wielowartościowej $X \twoheadrightarrow Y$ w $\mathbf{R}$, zbiór X jest kluczem lub nad-kluczem (tzn., nadzbiorem klucza) schematu $\mathbf{R}$.*

Czwarta postać normalna

Definicja 5.6 (Czwarta postać normalna (ozn. 4PN lub 4NF))

Niech $\mathbf{R} = (U, F \cup M)$ będzie ogólnym schematem relacyjnym. Mówimy, że schemat $\mathbf{R}$ jest w czwartej postaci normalnej (4PN), jeżeli z faktu że istnieje nietrywialna zależność wielowartościowa $X \twoheadrightarrow Y \in M^+$, gdzie $Y \subseteq U \setminus X$, wynika istnienie zależności funkcyjnej $X \to U \in F^+$.

Uwaga 5.3

Zachodzą poniższe zależności:

$$4PN \subset 3.5PN \subset 3PN \subset 2PN.$$

Wszystkie relacje zawierania są właściwe.

Ćwiczenia

Zadanie 1. Sprowadź poniższe ogólne schematy relacyjne do 4PN.

1. $\mathbf{R} = (U, F \cup M) = (\{A, B, C, D\}, \{A \twoheadrightarrow BCD, B \to AC, C \to D\})$.
2. $\mathbf{S} = (W, F \cup M) = (\{A, B, C, D, E\}, \{D \twoheadrightarrow CE, A \to BC, C \to DE\})$.
3. $\mathbf{T} = (V, F \cup M) = (\{A, B, C, D, E\}, \{A \to BCDE, AB \twoheadrightarrow CD\})$.
4. $\mathbf{Z} = (X, F \cup M) = (\{A, B, C, D, E\}, \{A \to B, C \to DE\})$.

Zadanie 2. Niech $R = (U, F)$ będzie danym schematem relacyjnym i niech $X, Y, Z, V, W \subset U$. Uzasadnij, że jeżeli $X \to YZ \in F^+$ i $V \to W \in F^+$, to $XV \to YZW \in F^+$.

Zadanie 3. Niech $R = (U, F \cup M)$ będzie danym ogólnym schematem relacyjnym i niech $X \subset U$, $Y = U \setminus X$. Uzasadnij, że $X \twoheadrightarrow Y \in M^+$.