Bazy danych i systemy zarządzania

Wykład IV

Operacje algebraiczne w relacyjnych bazach danych Część I

Definiowanie tablicy relacyjnej bazy danych

Definiowanie tabel

Aby zdefiniować tablicę reprezentującą relację w bazie danych należy określić następujące elementy:

1. **Schemat tabeli** postaci $R(A_1, A_2, \ldots, A_n)$, tzn.:
   - zdefiniować sekwencję atrybutów $A_1, A_2, \ldots, A_n$,
   - zdefiniować ich dziedziny $D_1, D_2, \ldots, D_n$ (określając typ, rozmiar, więzy spójności, etc. lub poprzez jawne (ekstensjonalne) zdefiniowania elementów każdego zbioru D_i, $i = 1, 2, \ldots, n$,
   - zdefiniować klucz lub indeks jednoznaczny (opcjonalnie).
2. **Krotki relacji (rekordy tabeli)**, co można uczynić:
   - *ekstensjonalnie*, np. wpisując lub wczytując ze zbioru,
   - *intensjonalnie*, definiując warunek selekcji oraz zbiór (generator) potencjalnych krotek (uniwersum), np. wczytać dane z odpowiedniego zbioru z wykorzystaniem filtru.

Predykat relacji

Predykatem relacji $R \subseteq D_1 \times D_2 \times \ldots \times D_n$ nazywamy funkcję postaci:

$$R \longrightarrow \{TRUE, FALSE\}$$

przy czym:
– $||R(d_1, d_2, \ldots, d_n)|| = TRUE$ wtw. gdy $(d_1, d_2, \ldots, d_n) \in R$, oraz
– $||R(d_1, d_2, \ldots, d_n)|| = FALSE$ wtw. gdy $(d_1, d_2, \ldots, d_n) \notin R$.
$||R(d_1, d_2, \ldots, d_n)||$ oznacza *wartość logiczną* formuły atomowej $R(d_1, d_2, \ldots, d_n)$.

Suma relacji zgodnych

Operację sumy (unii) relacji można przenieść bezpośrednio z algebry zbiorów; relacje są bowiem zbiorami. Jednak ze względu na specyficzną strukturę elementów tych zbiorów oraz wynikające z nich możliwości reprezentacji w tabelach, sensownie jest zdefiniować sumę dla *relacji zgodnych*, tzn. relacji o identycznym schemacie.

Niech $R(A_1, A_2, \ldots, A_n)$ oraz $S(A_1, A_2, \ldots, A_n)$ będą dowolnymi relacjami zgodnymi, określonymi w uniwersum $U = D_1 \times D_2 \times \ldots \times D_n$, $R \subseteq U$, $S \subseteq U$. Suma relacji zgodnych definiowana jest jako suma zbiorów $R \cup S$:

$$R \cup S = \{(d_1, d_2, \ldots, d_n) \in U : (d_1, d_2, \ldots, d_n) \in R \vee (d_1, d_2, \ldots, d_n) \in S\}.$$

Sumę relacji $R \cup S$ można też zadać intensjonalnie w U poprzez zdefiniowanie predykatu relacji jak następuje:

$$||R \cup S(d_1, d_2, \ldots, d_n)|| = ||R(d_1, d_2, \ldots, d_n)|| \vee ||S(d_1, d_2, \ldots, d_n)||.$$

Przykład

A	B	C
a_1	b_1	c_1
a_2	b_2	c_2
a_3	b_3	c_3
a_5	b_5	c_5

$\cup$

A	B	C
a_2	b_2	c_2
a_4	b_4	c_4
a_5	b_5	c_5
a_6	b_6	c_6
a_7	b_7	c_7

$=$

A	B	C
a_1	b_1	c_1
a_2	b_2	c_2
a_3	b_3	c_3
a_4	b_4	c_4
a_5	b_5	c_5
a_6	b_6	c_6
a_7	b_7	c_7

Zgodnie z prawami algebry zbiorów, duplikaty rekordów nie są uwzględniane (standardowo). Dla poprawnej realizacji operacji liczba i typy kolumn (atrybutów) powinny być identyczne. Liczba elementów relacji wynikowej $T = R \cup S$ spełnia warunek $card(T) \leq card(R) + card(S)$.

Różnica relacji zgodnych

Operację różnicy (odejmowania) relacji można przenieść bezpośrednio z algebry zbiorów (relacje są zbiorami). Jednak ze względu na specyficzną strukturę elementów tych zbiorów oraz wynikające z nich możliwości reprezentacji w tabelach, sensownie jest zdefiniować różnicę dla *relacji zgodnych*, tzn. relacji o identycznym schemacie.

Niech $R(A_1, A_2, \ldots, A_n)$ oraz $S(A_1, A_2, \ldots, A_n)$ będą dowolnymi relacjami zgodnymi, określonymi w uniwersum $U = D_1 \times D_2 \times \ldots \times D_n$, $R \subseteq U$, $S \subseteq U$. Różnica relacji zgodnych definiowana jest jako różnica zbiorów $R \setminus S$:

$$R \setminus S = \{(d_1, d_2, \ldots, d_n) \in U : (d_1, d_2, \ldots, d_n) \in R \wedge (d_1, d_2, \ldots, d_n) \notin S\}.$$

Różnicę relacji $R \setminus S$ można też zadać intensjonalnie w U poprzez zdefiniowanie predykatu relacji jak następuje:

$$||R \setminus S(d_1, d_2, \ldots, d_n)|| = ||R(d_1, d_2, \ldots, d_n)|| \wedge \neg ||S(d_1, d_2, \ldots, d_n)||.$$

Przykład

A	B	C
a_1	b_1	c_1
a_2	b_2	c_2
a_3	b_3	c_3
a_4	b_4	c_4
a_5	b_5	c_5
a_6	b_6	c_6
a_7	b_7	c_7

$\setminus$

A	B	C
a_1	b_1	c_1
a_2	b_2	c_2
a_3	b_3	c_3
a_5	b_5	c_5
a_8	b_8	c_8
a_9	b_9	c_9

$=$

A	B	C
a_4	b_4	c_4
a_6	b_6	c_6
a_7	b_7	c_7

Dla poprawnej realizacji operacji odejmowania liczba i typy kolumn (atrybutów) powinny być identyczne. Wynikiem odejmowania $R \setminus S$ jest relacja pusta wtw. gdy $R \subseteq S$. Liczba elementów relacji wynikowej $T = R \setminus S$ spełnia warunek $card(T) \leq card(R)$.

Iloczyn algebraiczny (przecięcie) relacji zgodnych

Operację iloczynu algebraicznego (zwykłego) (przecięcia) relacji można przenieść bezpośrednio z algebry zbiorów (relacje są zbiorami). Jednak ze względu na specyficzną strukturę elementów tych zbiorów oraz wynikające z nich możliwości reprezentacji w tabelach sensownie jest zdefiniować iloczyn dla *relacji zgodnych*, tzn. relacji o identycznym schemacie.

Niech $R(A_1, A_2, \ldots, A_n)$ oraz $S(A_1, A_2, \ldots, A_n)$ będą dowolnymi relacjami zgodnymi, określonymi w uniwersum $U = D_1 \times D_2 \times \ldots \times D_n$, $R \subseteq U$, $S \subseteq U$. Iloczyn relacji definiowany jest jako iloczyn zbiorów $R \cap S$, tzn.:

$$R \cap S = \{(d_1, d_2, \ldots, d_n) \in U : (d_1, d_2, \ldots, d_n) \in R \wedge (d_1, d_2, \ldots, d_n) \in S\}.$$

Iloczyn relacji $R \cap S$ można też zadać intensjonalnie w U poprzez zdefiniowanie predykatu relacji jak następuje:

$$||R(\cap S d_1, d_2, \ldots, d_n)|| = ||R(d_1, d_2, \ldots, d_n)|| \wedge ||S(d_1, d_2, \ldots, d_n)||.$$

Przykład

A	B	C
a_1	b_1	c_1
a_2	b_2	c_2
a_3	b_3	c_3
a_4	b_4	c_4
a_5	b_5	c_5
a_6	b_6	c_6
a_7	b_7	c_7

$\cap$

A	B	C
a_1	b_1	c_1
a_3	b_3	c_3
a_5	b_5	c_5
a_6	b_6	c_6
a_8	b_8	c_8
a_9	b_9	c_9

$=$

A	B	C
a_1	b_1	c_1
a_3	b_3	c_3
a_5	b_5	c_5
a_6	b_6	c_6

Dla poprawnej realizacji operacji iloczynu liczba i typy kolumn (atrybutów) powinny być identyczne. Wynikiem operacji iloczynu jest relacja pusta wtw. gdy żaden rekord nie jest wspólny dla obu relacji. Liczba elementów relacji wynikowej $T = R \cap S$ spełnia warunek $card(T) \leq min(card(R), card(S))$. Iloczyn relacji można wyznaczyć korzystając z operacji różnicy wg wzoru: $R \cap S = R \setminus (R \setminus S)$.

Dopełnienie relacji

Operację dopełnienia (uzupełnienia) relacji można przenieść bezpośrednio z algebry zbiorów (relacje są zbiorami), o ile określone jest uniwersum U, takie że $R \subseteq U$. Dopełnienie relacji R jest relacją o identycznym schemacie.

Niech $R(A_1, A_2, \ldots, A_n)$ będzie relacją określoną w uniwersum $U = D_1 \times D_2 \times \ldots \times D_n$, $R \subseteq U$. Dopełnienie relacji R definiowane jest jako relacja $\overline{R}$ będąca uzupełnieniem R do uniwersum U, tzn.:

$$R = \{(d_1, d_2, \ldots, d_n) \in U : (d_1, d_2, \ldots, d_n) \notin R\}.$$

Dopełnienie relacji R można też zadać intensjonalnie w U poprzez zdefiniowanie predykatu relacji jak następuje:

$$||\overline{R}(d_1, d_2, \ldots, d_n)|| = \neg||R(d_1, d_2, \ldots, d_n)||.$$

W praktyce, jeżeli dziedziny atrybutów $A_1, A_2, \ldots, A_n$ nie są zadane jawnie, przyjmuje się, że są one wyznaczone jako dziedziny relacji ze względu na odpowiednie argumenty, tzn. zbiory wszystkich wartości występujących w danej kolumnie tabeli.

Przykład

$$\overline{\begin{array}{|c|c|c|}\hline A & B & C \\ \hline a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ \hline \end{array}} = \begin{array}{|c|c|c|}\hline A & B & C \\ \hline a_1 & b_1 & c_2 \\ a_1 & b_2 & c_1 \\ a_1 & b_2 & c_2 \\ a_2 & b_1 & c_2 \\ a_2 & b_2 & c_1 \\ a_2 & b_1 & c_1 \\ \hline \end{array}$$

Wynikiem operacji dopełnienia jest relacja pusta wtw. $R = U$. Liczba elementów relacji dopełnienia może być relatywnie duża i wynosi $card(U) - card(R)$, tzn. $card(D_1) \cdot card(D_2) \cdot \ldots \cdot card(D_n) - card(R)$.

Iloczyn kartezjański

Operację iloczynu kartezjańskiego relacji można przenieść bezpośrednio z algebry zbiorów (relacje są zbiorami). Jednak ze względu na specyficzną strukturę elementów tych zbiorów oraz przyjęty sposób reprezentacji w tabelach wygodnie jest posłużyć się schematami relacji. Iloczyn kartezjański definiuje się dla *dowolnych* relacji (nie ma wymagania zgodności relacji).

Niech $R(A_1, A_2, \ldots, A_n)$ oraz $S(B_1, B_2, \ldots, B_m)$ będą dowolnymi relacjami, $R \subseteq U$, $S \subseteq V$. Iloczyn kartezjański relacji $T = R \times S$ definiowany jest jako relacja o schemacie $T(A_1, A_2, \ldots, A_n, B_1, B_2, \ldots, B_m)$ taka, że:

$$R \times S = \{(d_1, \ldots, d_n, d_{n+1}, \ldots, d_{n+m}) : (d_1, \ldots, d_n) \in R \wedge (d_{n+1}, \ldots, d_{n+m}) \in S\}.$$

Iloczyn kartezjański relacji $R \times S$ można też zadać intensjonalnie w $U \times V$ poprzez zdefiniowanie predykatu relacji jak następuje:

$$||T(d_1, \ldots, d_n, d_{n+1}, \ldots, d_{n+m})|| = ||R(d_1, \ldots, d_n)|| \wedge ||S(d_{n+1}, \ldots, d_{n+m})||.$$

Przykład

A	B	C
a_1	b_1	c_1
a_2	b_2	c_2
a_3	b_3	c_3

$\times$

B	C	D
b_1	c_1	d_1
b_2	c_2	d_2
b_3	c_3	d_3

$=$

A	B	C	B	C	D
a_1	b_1	c_1	b_1	c_1	d_1
a_1	b_1	c_1	b_2	c_2	d_2
a_1	b_1	c_1	b_3	c_3	d_3
a_2	b_2	c_2	b_1	c_1	d_1
a_2	b_2	c_2	b_2	c_2	d_2
a_2	b_2	c_2	b_3	c_3	d_3
a_3	b_3	c_3	b_1	c_1	d_1
a_3	b_3	c_3	b_2	c_2	d_2
a_3	b_3	c_3	b_3	c_3	d_3

Liczba elementów relacji wynikowej $T = R \times S$ spełnia warunek $card(T) = card(R) \cdot card(S)$.

Projekcja

Projekcja (*rzutowanie*) jest jedną z podstawowych operacji algebry relacji. Dla intuicji, projekcja oznacza rzutowanie na wybrane kolumny, a więc ograniczenie informacji zadawanej w tabeli reprezentującej relację do wybranych kolumn. Innymi słowy, wybrane kolumny tabeli są zachowane, a inne – usuwane (pomijane; czasowo lub trwale).

Niech $\mathbf{A} = \{A_1, A_2, \ldots, A_n\}$ będzie wybranym zbiorem atrybutów, a R będzie relacją o schemacie $R(A_1, A_2, \ldots, A_n)$.

Niech $\mathbf{A}' = \{A^1, A^2, \ldots, A^k\}$ będzie zbiorem atrybutów wybranych ze zbioru $\mathbf{A}$, $\mathbf{A}' \subseteq \mathbf{A}$.

Niech P będzie relacją o schemacie $P(A^1, A^2, \ldots, A^k)$.

Operację projekcji π_P definiujemy jako operację postaci $\pi_P : \ R \longrightarrow P$ przekształcającą relację R w pewną relację P, taką, że:

$$\pi_P((d_1, d_2, \ldots, d_n)) = (d^1, d^2, \ldots, d^k),$$

gdzie, że $d^i = d_j$ dla $A^i = A_j$, $i = 1, 2, \ldots, k$, $j \in \{1, 2, \ldots, n\}$; $(d_1, d_2, \ldots, d_n) \in R$, $(d^1, d^2, \ldots, d^k) \in P$.

Oznaczenia:

Dla reprezentacji operacji projekcji relacji R zdefiniowanej jak wyżej stosowane są następujące oznaczenia:

- $\pi_{A^1, A^2, \ldots, A^k}(R)$ – rzut relacji R na atrybuty $A^1, A^2, \ldots, A^k$,
- $\pi_{i_1, i_2, \ldots, i_k}(R)$ – rzut relacji R na składowe $i_1, i_2, \ldots, i_k$,
- $R[A^1, A^2, \ldots, A^k]$ – rzut relacji R na atrybuty $A^1, A^2, \ldots, A^k$.

  Przykładowo, niech R będzie relacją o schemacie $R(A, B, C, D, E)$; wówczas

  $$\pi_{3,5,1}(R) = \pi_{C,E,A}(R) = R[C, E, A]$$

  Logicznie $||R(d_1, d_2, \ldots, d_n)[A^1, A^2, \ldots, A^k]||$ $(||R(d^1, d^2, \ldots, d^k)||)$wtw. $\exists x_1, x_2, \ldots, x_{n-k} R(x_1, \ldots, x_{i_1}, d^1, x_{i_1+1}, \ldots, x_{i_2}, d^2, \ldots, x_{i_k}, d^k, \ldots, x_{n-k})$.

Projekcja: przykład, własności, problemy

Przykład: Niech R będzie relacją o schemacie $R(A, B, C, D, E, F)$:

A	B	C	D	E	F
a_1	b_1	c_1	d_1	e_1	f_1
a_1	b_1	c_1	d_2	e_2	f_2
a_1	b_1	c_1	d_3	e_3	f_3
a_2	b_2	c_2	d_1	e_1	f_1
a_2	b_2	c_2	d_2	e_2	f_2

Projekcją (rzutem) relacji R na atrybuty E, C, A jest relacja $\pi_{E,C,A}(R)$ ($\pi_{5,3,1}(R)$, $R[E, C, A]$) równa:

E	C	A
e_1	c_1	a_1
e_2	c_1	a_1
e_3	c_1	a_1
e_1	c_2	a_2
e_2	c_2	a_2

Własności operacji projekcji:

- zmienia schemat relacji (ogranicza liczbę atrybutów i ich kolejność,
- tracona jest informacja o wartościach niektórych atrybutów,
- mogą powstać duplikaty (rekordy nierozróżnialne),
- liczba rekordów nie zmienia się (o ile nie usuniemy duplikatów (opcja `DISTINCT` w klauzuli `SELECT`; wówczas liczba rekordów maleje).

Operacja projekcji może prowadzić do utraty możliwości rozróżniania niektórych rekordów; może ona stanowić pewien operator *abstrakcji* (stosowany celowo); ale może też prowadzić do niezamierzonej utraty informacji.

Antyprojekcja

Antyprojekcja jest stosunkowo rzadko spotykaną operacją algebry relacji. Niech R będzie relacją o schemacie $R(A_1, A_2, \ldots, A_n, B)$, gdzie B jest pewnym wyróżnionym atrybutem (może on występować na dowolnej pozycji).

Niech $D_1, D_2, \ldots, D_n$ będą dziedzinami atrybutów ze względu na odpowiednie atrybuty (argumenty), tzn.:

$$D_i = \pi_i(R),$$

oraz niech $U = D_1 \times D_2 \times \ldots \times D_n$ będzie uniwersum względem atrybutów $A_1, A_2, \ldots, A_n$.

Niech $D_B = \pi_B(R)$ będzie zbiorem wszystkich wartości atrybutu B występujących w relacji R.

Operacja antyprojekcji relacji R której wynikiem jest relacja $R]B[$ (antyprojekcja R na atrybut B) jest zdefiniowana jako:

$$R]B[= \{b \in B : \ \forall (d_1, d_2, \ldots, d_n) \in U, \ (d_1, d_2, \ldots, d_n, b) \in R\}.$$

Przykład: Niech $R(A_1, A_2, B)$ będzie relacją zdefiniowaną jak następuje:

A_1	A_2	B
a_1	a_1	b_1
a_2	a_1	b_1
a_1	a_2	b_1
a_2	a_2	b_1
a_1	a_1	b_2
a_2	a_1	b_2
a_1	a_2	b_2

Wynikiem operacji antyprojekcji R na B jest relacja $R]B[$ jak poniżej:

B
b_1

Logicznie $||R]B[(b)||$ wtw. gdy $\forall (d_1, d_2, \ldots, d_n) \in U$, $(d_1, d_2, \ldots, d_n, b) \in R$.

Interpretacja operacji antyprojekcji

Operacja antyprojekcji pozwala odpowiadać na pytania typu:

Którzy dostawcy dostarczają **wszystkie** *produkty (występujące w tabeli)?*

Przykład: Niech będzie dana tabela dostawców wraz ze specyfikacją dostarczanego produktu:

Dostawca	*Element*
Budex	*cegła*
Budex	*pustak*
Budex	*cement*
Budex	*piasek*
Matbud	*cegła*
Matbud	*pustak*
Matbud	*cement*
Matbud	*gips*
Matbud	*piasek*
Matbud	*żwir*
Probud	*cegła*
Probud	*cement*
Probud	*piasek*
Probud	*żwir*

Wynikiem antyprojekcji tej tabeli na atrybut *Dostawca* jest tabela:

Dostawca
Matbud

Antyprojekcja pozwala odpowiadać na pytania zawierające kwantyfikator ogólny ($\forall$).

Iloraz

Operacja ilorazu umożliwa *dzielenie* relacji przez relację.

Niech R będzie relacją o schemacie $R(A_1, A_2, \ldots, A_n)$, a S relacją o schmemacie $S(B_1, B_2, \ldots, B_k)$, przy czym zakładamy że $n > k$ oraz $S \neq \emptyset$.

Wynikiem operacji dzielenia relacji R przez relację S jest relacja $R \div S$ będąca zbiorem wszystkich $(n-k)$-krotek t, takich, że dla **wszystkich** k-krotek $s \in S$, krotka $ts \in R$.

Niech $\pi_{1,2,\ldots,n-k}(R) = T$.
T jest zbiorem wszystkich początkowych ciągów krotek relacji R (o długości $n-k$). Wówczas relacja

$$(T \times S) \setminus R$$

jest zbiorem wszystkich n-krotek nie należących do R, utworzonych poprzez konkatenację pierwszych $n-k$ składowych z relacji R z **wszystkimi** krotkami z S.

Niech

$$V = \pi_{1,2,\ldots,n-k}((T \times S) \setminus R);$$

V jest zbiorem wszystkich $n-k$-krotek v, takich, że dla pewnej krotki $s \in S$, konkatenacja $vs \notin R$. Zatem

$$T \setminus V = R \div S,$$

lub:

$$R \div S = \pi_{1,2,\ldots,n-k}(R) \ \setminus \ \pi_{1,2,\ldots,n-k}((\pi_{1,2,\ldots,n-k}(R) \times S) \setminus R)$$

Operacja dzielenia relacji R przez relację S służy do znajdowania wszystkich początkowych ciągów z relacji R o długości $n-k$ (wszystkich fragmentów lub początków), które w R są skonkatenowane ze **wszystkimi** krotkami **zadanymi w relacj**i S.

Iloraz: przykład i interpretacja

Operacja dzielenia relacji R przez relację S ma istotne znaczenie praktyczne i stanowi rozszerzenie operacji antyprojekcji. W przypadku operacji antyprojekcji, można było uzyskać odpowiedź na pytanie typu:

Którzy dostawcy dostarczają **wszystkie** *produkty?*

Natomiast w przypadku operacji ilorazu można uzyskać odpowiedź na pytanie typu:

Którzy dostawcy dostarczają produkty z **określonego zbioru***?*

Interesujący nas zbiór zadawany jest właśnie relacją S. Poniżej przedstawiono prosty przykłd.

Przykład:

Dostawca	*Element*
Budex	*cegła*
Budex	*pustak*
Budex	*cement*
Budex	*piasek*
Matbud	*pustak*
Matbud	*cement*
Matbud	*gips*
Matbud	*piasek*
Probud	*cegła*
Probud	*cement*
Probud	*piasek*
Probud	*żwir*

$\div$

Element
cegła
cement
piasek

$=$

Dostawcya
Budex
Probud

Operacja dzielenia pozwala zatem na realizację zapytań formalizowalnych logicznie z zastosowaniem kwantyfikatora ogólnego ($\forall$).

Uogólniona suma relacji

Zdefiniujemy uogólnioną sumę relacji (niekoniecznie zgodnych), czyli relacji o różnych schematach. Niech $R(A_1, A_2, \ldots, A_n)$ oraz $S(B_1, B_2, \ldots, B_m)$ będą dowolnymi relacjami, $\mathbf{A} = \{A_1, A_2, \ldots, A_n\}$ oraz $\mathbf{B} = \{B_1, B_2, \ldots, B_m\}$. Niech $\mathbf{B}' = \mathbf{B} \setminus \mathbf{A}$, tzn. zbiór $\mathbf{B}'$ zawiera wszystkie atrybuty zbioru $\mathbf{B}$ nie występujące w $\mathbf{A}$. Niech $\mathbf{B}' = \{B^1, B^2, \ldots, B^k\}$, $k \leq m$.

Niech $R + S$ będzie nową relacją o schemacie zdefiniowanym przez $(R + S)(A_1, A_2, \ldots, A_n, B^1, B^2, \ldots, B^k)$. Relacja $R + S$ będzie nazywana *uogólnioną sumą* relacji R i S, jeżeli spełniony jest warunek:

$$\pi_{A_1,A_2,\ldots,A_n}(R + S) = R$$

oraz

$$\pi_{B_1,B_2,\ldots,B_m}(R + S) = S$$

Przykład:

A	B
a_1	b_1
a_2	b_2
a_3	b_3

$+$

B	C
b_1	c_1
b_2	c_2
b_3	c_3

$=$

A	B	C
a_1	b_1	c_1
a_1	b_1	c_2
a_1	b_1	c_3
a_1	b_2	c_2
a_1	b_3	c_3
a_2	b_1	c_1
a_2	b_2	c_1
a_2	b_2	c_2
a_2	b_2	c_3
a_2	b_3	c_3
a_3	b_1	c_1
a_3	b_2	c_2
a_3	b_3	c_1
a_3	b_3	c_2
a_3	b_3	c_3

Interpretacja sumy uogólnionej: przykład

Uogólniona suma relacji jest relacją o schemacie zawierającym wszystkie atrybuty obu relacji (powtarzające się atrybuty występują tylko raz) oraz wszystkie połączenia krotek relacji R i S, takie, że obcięcie danej krotki do zboru atrybutów wybranej relacji należy do tej relacji.

Uogólniona suma relacji odpowiada na pytanie typu:

> *Jakich odbiorców mogłyby (potencjalnie) znaleźć produkty dostarczane przez pewnych dostawców oraz którzy dostawcy mogliby (potencjalnie) dotarczać produkty wymagane przez pewnych odbiorców?*

Przykład:

Dostawca	*Produkt*
Budex	*cegła*
Budex	*pustak*
Budex	*cement*
Budex	*piasek*
Matbud	*pustak*
Matbud	*cement*
Matbud	*gips*
Matbud	*piasek*

+

Produkt	*Odbiorca*
cegła	*Wykbud*
pustak	*Wykbud*
cement	*Rembud*
cement	*Wykbud*
piasek	*Rembud*
piasek	*Wykbud*
gips	*Rembud*

=

Dostawca	*Produkt*	*Odbiorca*
Budex	*cegła*	*Rembud*
Budex	*pustak*	*Rembud*
Budex	*cement*	*Rembud*
Budex	*piasek*	*Rembud*
Matbud	*pustak*	*Rembud*
Matbud	*cement*	*Rembud*
Matbud	*gips*	*Rembud*
Budex	*gips*	*Rembud*
Matbud	*piasek*	*Rembud*
Budex	*cegła*	*Wykbud*
Budex	*pustak*	*Wykbud*
Budex	*cement*	*Wykbud*
Budex	*piasek*	*Wykbud*
Matbud	*pustak*	*Wykbud*
Matbud	*cement*	*Wykbud*
Matbud	*gips*	*Wykbud*
Matbud	*piasek*	*Wykbud*