



Relacyjne Bazy Danych

Andrzej M. Borzyszkowski
PJATK/ Gdańsk

materiały dostępne elektronicznie
<http://szuflandia.pjwstk.edu.pl/~amb>

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

Algebra relacji

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

Relacyjna baza danych = relacje + operacje na relacjach

- Operacje działające na relacjach
 - operacje mogą być ze sobą składane tworząc dowolnie skomplikowane wyrażenia relacyjne tj. wyrażenia, których wartością jest relacja
 - czyli jest to algebra relacji
- Podstawowe operacje relacyjne:
 - obcięcie/wybór: wybiera pewne wiersze
 - rzut/projekcja: wybiera pewne atrybuty
 - zmiany nazwy atrybutów
 - złączenie: produkuje relację o atrybutach z dwu relacji, w tym wspólne atrybuty
- Operacje teoriomnogościowe: suma, przecięcie, różnica, iloczyn kartezjański

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

Operacja obcięcia

- Obcięcie – tylko wiersze spełniające warunek (*predicate*)
 - $\sigma[\text{miasto}='Gdańsk'](\text{Klient})$ (sigma)
 - w notacji bardziej przyjaznej **Klient WHERE miasto='Gdańsk'**
- $\sigma[\text{warunek}](\text{relacja})$
 - relacja jest dowolnym wyrażeniem algebry relacji
 - warunek jest wyrażeniem logicznym

2

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

nr	tytul	nazwisko	imie	kod_po	miasto	ulica_dom	telefon
5	Pan	Soroczyński	Jan	80-230	Gdańsk	Al. Hallera	58 309078
6	Pani	Niezabitowska-	Marzena	80-619	Gdańsk	Focha 39-41 m.66	58 309910
7	Pani	Kołąk	Agnieszka	80-832	Gdańsk	Wąwóz 4	NULL
9	NULL	Hałasa	Ewa	80-511	Gdańsk	Dywizjonu 303/303	58 348324
10	Pan	Sosnowy	Andrzej	80-266	Gdańsk	Leśna Góra 41h/088	58 346718
11	Pani	Songin	Barbara	80-376	Gdańsk	Grunwaldzka 1024/128	58 552874

4

Predykaty w operacji obciążenia

- Przykładowe postacie warunku obciążenia:
 - równość (i jej zaprzeczenie) $X1=X2, X1\neq C$
 - porządek (dla dziedzin uporządkowanych) $X1<X2, X1\leq X2$
 - należenie $X1 \text{ in } X2$
 - warunki złożone
 - **A WHERE c1 AND c2** koniunkcja
 - **A WHERE c1 OR c2** alternatywa
 - **A WHERE NOT c1** negacja
- Predykat stosowany jest do każdej krotki z osobna
 - nie można więc wyrażać go w zależności od wielu krotek
- $\sigma[\text{warunek-2}](\sigma[\text{warunek-1}](\text{relacja}))$ jest możliwym zastosowaniem obciążenia
 - wynik będzie ten samo co $\sigma[\text{warunek-2 and warunek-1}](\text{relacja})$

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

5

Iloczyn kartezjański

- Iloczyn kartezjański (produkt kartezjański, złączenie krzyżowe)
 - nagłówek: suma nagłówków $\{ X:DX, Y:DY \}$, które muszą mieć różne atrybuty
 - treść: wszystkie możliwe pary krotek $\{ X:x, Y:y \}$

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

7

Operacja rzutu

- Rzut – tylko wybrane atrybuty
 - $\pi[\text{nr,nazwisko,imie}](\text{Klient})$
 - w notacji ascii
Klient(nr,nazwisko,imie)
- Rzutem relacji A: $\pi[X_1, \dots, X_m](A)$ jest relacja z nagłówkiem $\{ X_1:D_{X1}, \dots, X_m:D_{Xm} \}$ i z treścią składającą się z tych krotek, dla których w relacji A występuje krotka, której fragmentem jest krotka z rzutu
- uwaga: relacja nie ma powtórzeń, fragmenty wszystkich krotek mogłyby wprowadzić powtórzenie

nr	nazwisko	imie
1	Kuśmerek	Małgorzata
2	Chodkiewicz	Jan
3	Szczęsna	Jadwiga
4	Łukowski	Bernard
5	Soroczyński	Jan
6	Niezabitowska-N	Marzena
7	Kołąk	Agnieszka
8	Kołąk	Agnieszka
9	Hałasa	Ewa
10	Sosnowy	Andrzej
11	Songin	Barbara
12	Wróblewicz-Terle	Urszula
13	Soroczyński	Bogdan
14	Miszke	Wojciech
15	Zaorski	Marcin
16	Wiśniewska	Grażyna
17	Wierciński	Henryk
18	Bazior	Gerard

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

6

Operacja złączenia

- Złączenie zapewnia integrację danych
 - dane z kilku tabel są zbierane w jednej tabeli wynikowej
- Łączy ze sobą krotki z różnych relacji
 - krotki dobierane są na podstawie pasujących wartości odpowiednich atrybutów
 - najczęściej klucz obcy jednej relacji i klucz kandydujący drugiej relacji
- Operacja daje się wyrazić za pomocą kolejno
 - iloczynu kartezjańskiego (wszystkie pary krotek)
 - obciążenia (wybór par pasujących do siebie)
 - rzutu (wyrugowanie powtarzających się atrybutów)
- Oznaczenia matematyczne: $R \bowtie_{[\text{warunek złączenia}]} S$

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

8

Złączenie, przykład

nr	nazwisko	imie
3	Szczęsna	Jadwiga
4	Łukowski	Bernard
5	Soroczyński	Jan
6	Niezabitowska-Nasiadko	Marzena
7	Kołąk	Agnieszka
8	Kołąk	Agnieszka

klient_nr	data_zlozenia
3	21.02.2021
3	23.03.2021
3	13.03.2021
5	4.05.2021
6	1.02.2021
6	22.03.2021
8	7.04.2021
8	12.01.2021

nr	nazwisko	imie	data_zlozenia
3	Szczęsna	Jadwiga	21.02.2021
3	Szczęsna	Jadwiga	23.03.2021
3	Szczęsna	Jadwiga	13.03.2021
5	Soroczyński	Jan	4.05.2021
6	Niezabitowska-Nasiadko	Marzena	1.02.2021
6	Niezabitowska-Nasiadko	Marzena	22.03.2021
8	Kołąk	Agnieszka	7.04.2021
8	Kołąk	Agnieszka	12.01.2021

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

9

Uogólniony rzut (i zmiana nazwy)

- Zmiana nazwy
 - całej relacji: $\rho[S](R)$ (rho)
 - poszczególnych atrybutów w relacji: $\rho[S(B1, \dots, Bn)](R)$
 - tutaj zakładamy, że nowy schemat S ma tyle samo atrybutów i o tych samych dziedzinach
 - w praktyce będzie notacja wymieniająca tylko nowe nazwy
- Rzut w sensie ogólniejszym
 - nagłówek: może zawierać również wyrażenia dla atrybutów relacji (i nazwy dla nich)
 - treść: krotki, do których zastosowano to wyrażenie
 - np. jeśli w nagłówku są atrybuty cena i koszt, to w rzucie może się pojawić nowy atrybut zysk zdefiniowany jako różnica cena-koszt

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

11

Dodatkowe operacje relacyjne – funkcje agregujące (grupowanie)

- Funkcja agregująca przekształca zbiór wartości w pojedynczą wartość
 - avg: średnia
 - min, max, sum
 - count: liczba elementów
- Relacja jest grupowana w/g równych wartości niektórych atrybutów
 - do każdej grupy stosowana jest funkcja agregująca
 - szczególny przypadek: cała relacja jest jedną grupą

ilu	miasto
3	Sopot
1	Gardeja Pierwsza
6	Gdańsk
1	Kwidzyn
1	Gdakowo
4	Gdynia
1	Dziewięć Włók
1	Prabuty

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

10

Dodatkowe operacje relacyjne – złączenie zewnętrzne

- Złączenie naturalne „gubi” te krotki z jednej relacji, które nie pasują do drugiej z relacji
 - może to być pożądanym rezultatem, np. szukamy tylko klientów, którzy coś zakupili
 - ale może prowadzić do utraty informacji, np. o klientach, którzy nie dokonali zakupów
- Złączenie zewnętrzne
 - krotki z jednej z relacji nie pasujące do żadnej krotki z drugiej uzupełniane są wartością NULL
 - można „chronić” tylko jedną z relacji przed utratą informacji: złączenie lewe i prawe

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

12

Złączenie zewnętrzne, przykład

nr	nazwisko	imie	data_zlozenia
1	Kuśmerek	Małgorzata	
2	Chodkiewicz	Jan	
3	Szczęsna	Jadwiga	21.02.2021
3	Szczęsna	Jadwiga	23.03.2021
3	Szczęsna	Jadwiga	13.03.2021
4	Łukowski	Bernard	
5	Soroczyński	Jan	4.05.2021
6	Niezabitowska-Nasiadko	Marzena	1.02.2021
6	Niezabitowska-Nasiadko	Marzena	22.03.2021
7	Kołąk	Agnieszka	
8	Kołąk	Agnieszka	7.04.2021

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

13

Złączenie, własności

- łączność: $(A \bowtie B) \bowtie C = A \bowtie (B \bowtie C)$
 - można opuszczać nawiasy
- przemienność: $A \bowtie B = B \bowtie A$
- jeśli relacje A i B nie mają wspólnych atrybutów, wówczas jest to iloczyn kartezjański relacji: każda krotka A jest skombinowana z każdą krotką B
- jeśli relacje A i B mają identyczne wszystkie atrybuty, wówczas jest to przecięcie relacji: tylko wspólne krotki

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

15

Złączenie naturalne, definicja

- Niech A i B będą relacjami o nagłówkach $\{X:DX, Y:DY\}$ oraz $\{Y:DY, Z:DZ\}$,
 - złączeniem naturalnym relacji A i B , $A \bowtie B$, jest relacja z nagłówkiem $\{X:DX, Y:DY, Z:DZ\}$ i z treścią składającą się z takich krotek $\{X:x, Y:y, Z:z\}$, że krotka $\{X:x, Y:y\}$ należy do relacji A , a krotka $\{Y:y, Z:z\}$ należy do relacji B
- Najczęstszy przypadek: wspólne atrybuty Y stanowią klucz kandydujący w jednej relacji i klucz obcy w drugiej
 - w istniejących językach NIE zakłada się, że atrybuty o tej samej nazwie oznaczają to samo
 - złączenie wymaga jawnego podania nazw atrybutów, które chcemy utożsamić (t.j. warunku łączącego)

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

14

Operacje teoriomnogościowe

- Przekrój (iloczyn, część wspólna)
 - nagłówek: równy *wspólnemu* nagłówkowi obu relacji (dopuszczalne są lekkie odchylenia, dziedziny atrybutów nie muszą być identyczne, ale zgodne, np. liczbowe)
 - treść: wszystkie krotki należące do obu relacji
- Suma (unia)
 - nagłówek: równy *wspólnemu* nagłówkowi obu relacji
 - treść: wszystkie krotki należące do co najmniej jednej z relacji
- Różnica
 - nagłówek: równy *wspólnemu* nagłówkowi obu relacji
 - treść: krotki z jednej relacji nie należące do drugiej
 - np. studenci, którzy nie zaliczyli egzaminu

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

16

Rachunek krotek

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

17

Rachunek krotek

- Zmienne krotkowe t
 - $\{ \langle t.A1, \dots, t.An \rangle \mid \Phi(t) \}$ lub $\{ t \mid \Phi(t) \}$
 - $\langle t.A1, \dots, t.An \rangle$ jest złożeniem kilku atrybutów w krotkę
 - predykat Φ jest zbudowany z następujących elementów
 - należenie do relacji $t \in r$, zapisywane również $r(t)$
 - warunki na wartości atrybutów (podobnie jak w obciążeniu)
 - kwantyfikatory „dla każdego” $\forall t$ oraz „istnieje” $\exists t$
 - elementy połączone są spójnikami koniunkcji, alternatywy oraz negacji
- Przykład,
 - relacja $\sigma[\text{miasto}='Gdańsk'](\text{Klient})$ jest równa $\{ t \mid \text{Klient}(t) \text{ AND } t.\text{miasto}='Gdańsk' \}$
 - relacja $\pi[\text{nr}, \text{nazwisko}, \text{imie}](\text{Klient})$ jest równa $\{ \langle t.\text{nr}, t.\text{nazwisko}, t.\text{imie} \rangle \mid \text{Klient}(t) \}$

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

18

Rachunek krotek, c.d.

- Kwantyfikatory:
 - nazwiska klientów, którzy złożyli zamówienie:
 - $\{ t.\text{nazwisko} \mid \text{Klient}(t) \text{ AND } \exists u (\text{Zamówienie}(u) \text{ AND } u.\text{klient_nr}=t.\text{nr}) \}$
 - jest to de facto złączenie (a właściwie jego rzut)
- Nazwiska klientów, którzy zamówili każdy towar dostępny w ofercie
 - $\{ t.\text{nazwisko} \mid \text{Klient}(t) \text{ AND } \forall u (\text{Towar}(u) \Rightarrow \exists v (\text{Zamówienie}(v) \text{ AND } v.\text{klient_nr}=t.\text{nr} \text{ AND } \exists w (\text{Pozycja}(w) \text{ AND } w.\text{zamowienie_nr}=v.\text{nr} \text{ AND } w.\text{towar_nr}=u.\text{nr}))) \}$

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

19

Logika kwantyfikatorów

- NOT $(\exists u) \Phi(u) \Leftrightarrow (\forall u) \text{NOT } \Phi(u)$
- NOT $(\forall u) \Phi(u) \Leftrightarrow (\exists u) \text{NOT } \Phi(u)$
- W szczególności kwantyfikator uniwersalny \forall może zostać zastąpiony przez bardziej skomplikowane wyrażenie bez niego
 - $(\forall u) \Phi(u) \Leftrightarrow \text{NOT } (\exists u) \text{NOT } \Phi(u)$
 - „nieprawda, że istnieje kontrprzykład”
- Przydatne mogą też być prawa de Morgana
 - NOT $(\Phi(t) \text{ AND } \Psi(t)) \Leftrightarrow (\text{NOT } \Phi(t) \text{ OR } \text{NOT } \Psi(t))$
 - NOT $(\Phi(t) \text{ OR } \Psi(t)) \Leftrightarrow (\text{NOT } \Phi(t) \text{ AND } \text{NOT } \Psi(t))$
- Konstrukcje z kwantyfikatorem ogólnym powinny ograniczać zbiór potencjalnych wartości
 - tzn. dopuszczalne są jedynie konstrukcje $(\forall t) t \in r \Rightarrow$

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

20

Rachunek dziedzin

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

21

Rachunek dziedzin, QBE

- Język QBE (*query by example*) opracowany w IBM
 - nie ma potrzeby nazywać zmiennych, których nie potrzebujemy w wyniku
 - nazwy zmiennych często są przykładowymi wynikami (i są syntaktycznie wyróżnione)
 - wyświetl imiona i nazwiska klientów z Gdańska:
 $\{ \langle _Jan, _Kowalski \rangle \mid Klient(_, _, _Kowalski, _Jan, 'Gdańsk', _, _) \}$
 - wyświetl nazwy towarów zamawiane z Gdańska $\{ _donica \mid Klient(_17, _, _, _, 'Gdańsk', _, _) \text{ AND } Zamówienie(_44, _17, \dots) \text{ AND } Pozycja(_44, _32, _) \text{ AND } Towar(_32, _donica, \dots) \}$
- Język używany jest często w środowisku graficznym
 - nazwy potrzebnych zmiennych lub stałych wprowadzane są bezpośrednio do nagłówek tabeli

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

23

Rachunek dziedzin

- Zmienne dziedzinowe x_i , dla każdej dziedziny atrybutu
 - $\{ \langle x_1, \dots, x_n \rangle \mid \Phi(x_1, \dots, x_m) \}$ gdzie w formule może wystąpić więcej zmiennych niż chcemy mieć w wyniku
 - predykat Φ jest zbudowany analogicznie jak w rachunku krotek
- Przykład,
 - relacja $\sigma[\text{miasto}='Gdańsk'](\text{Klient})$ jest równa $\{ \langle nr, tytuł, nazwisko, \dots, miasto, kod_p, tel \rangle \mid Klient(nr, tytuł, nazwisko, \dots, miasto, kod_p, tel) \text{ AND } miasto='Gdańsk' \}$
 - relacja $\pi[nr, nazwisko, imie](\text{Klient})$ jest równa $\{ \langle nr, nazwisko, imie \rangle \mid \exists tytuł, miasto, kod_p, tel, \dots Klient(nr, tytuł, nazwisko, \dots, miasto, kod_p, tel) \}$

© Andrzej M. Borzyszkowski

Relacyjne Bazy Danych

22