



WYKŁAD 2

Relacyjny i obiektowy model
danych

Plan wykładu

- Relacyjny model danych
- Obiektowy model danych
- Struktury danych
- Operacje
- Ograniczenia integralnościowe



Model danych (1)

Model danych jest abstrakcją, spójną wewnątrznie, logiczną definicją struktur danych, operatorów danych, łączących łącznie abstrakcyjną maszynę, z którą pracuje użytkownik

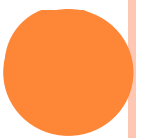
Implementacja modelu danych jest fizyczną realizacją na prawdziwej maszynie elementów maszyny abstrakcyjnej, które wspólnie składają się na model

Definiuje

- struktury danych
- operacje (manipulacje)
- ograniczenia integralnościowe, reguły poprawności danych

Relacyjny model danych

- relacje
- selekcja, projekcja, połączenie, operacje na zbiorach
- klucz podstawowy, klucz obcy, zawężenie dziedziny, unikalność, null



Model danych (2)

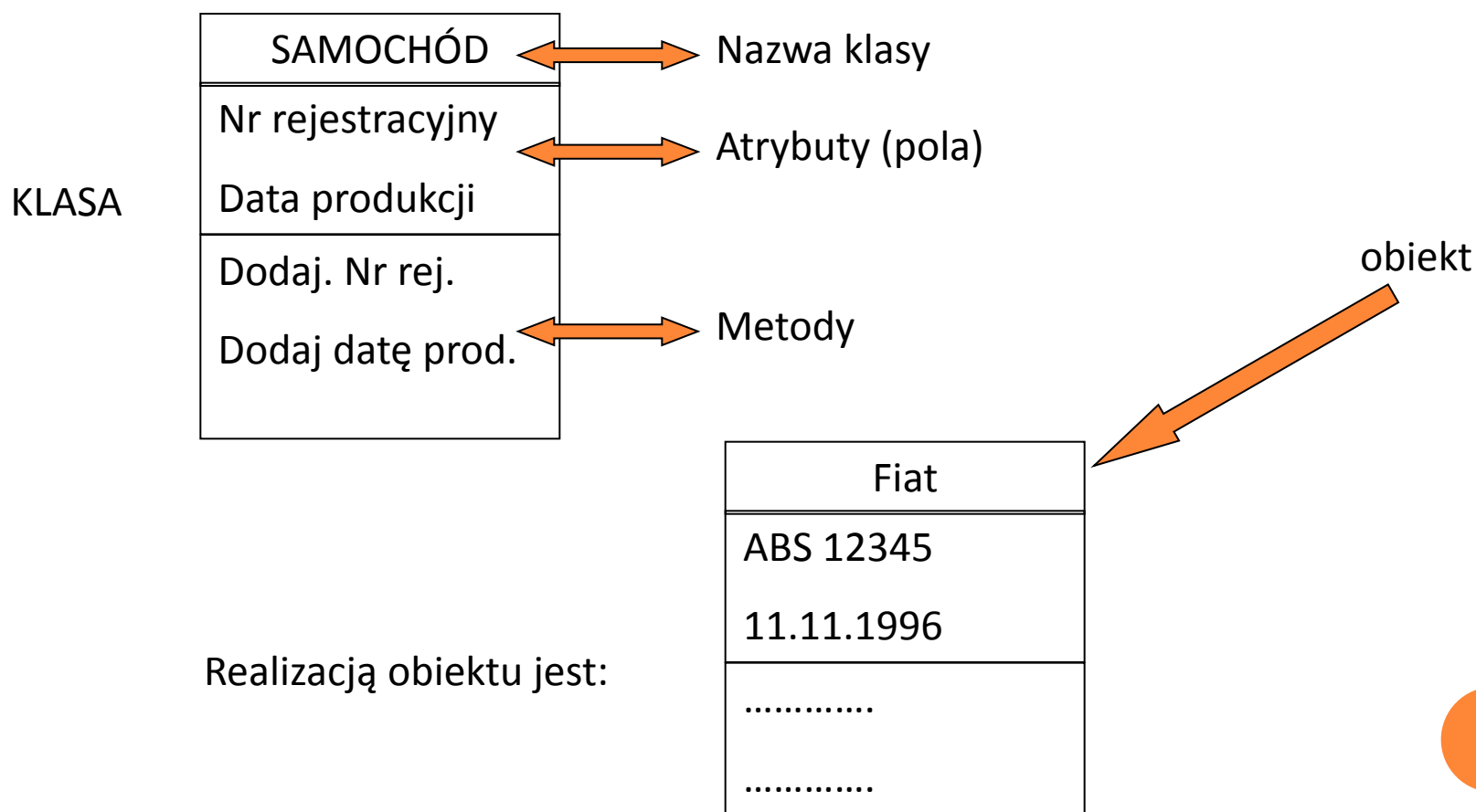
Podstawowe elementy obiektowego modelu danych

- Obiekt: stan i funkcjonalność
- Cechy obiektów: atrybuty i związki
- Funkcjonalność obiektu: metody
- Tożsamość obiektu
- Hermetyczność obiektów
- Klasa: typ danych i moduł programowy
- Dziedziczenie: współdzielenie implementacji i relacja podtypu
- Przeciążanie i dynamiczne wiązanie funkcjonalności obiektów



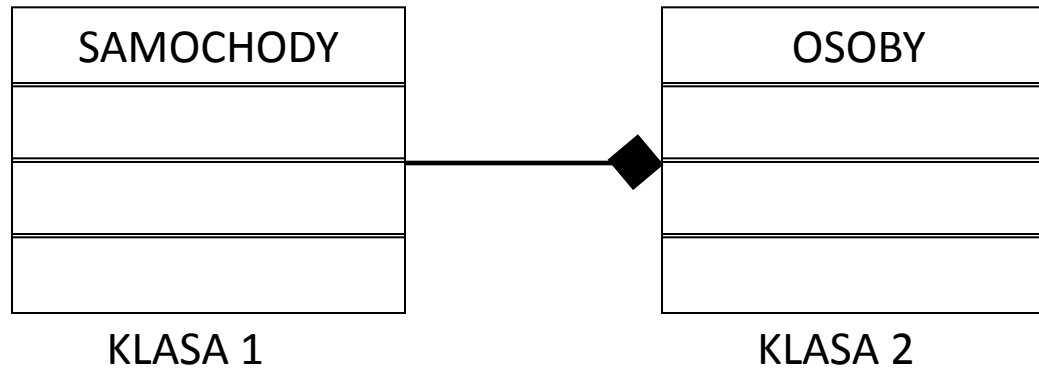
JĘZYK UML (UNIFIED MODELING LANGUAGE)

Zunifikowany język modelowania

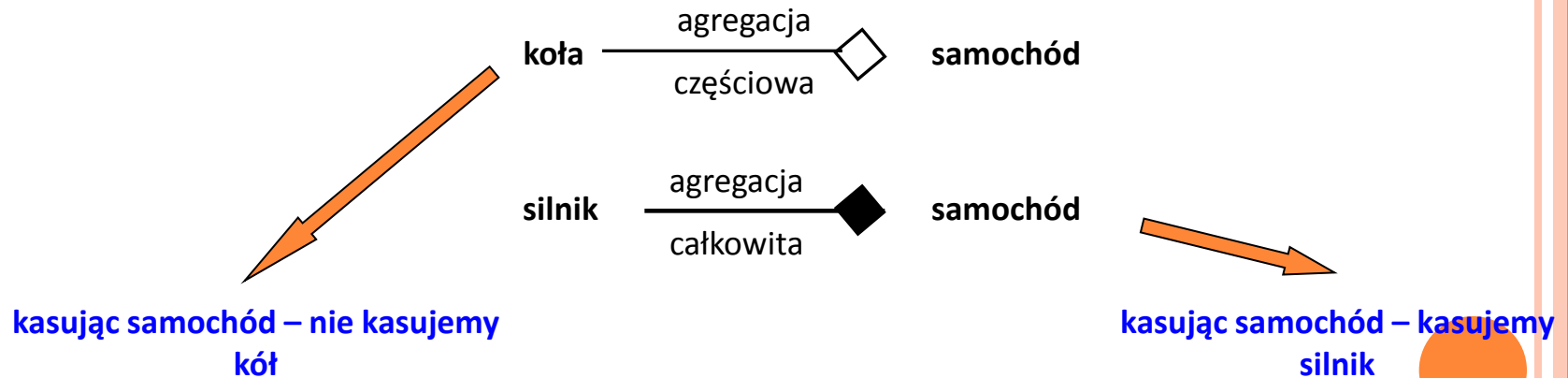


JĘZYK UML (UNIFIED MODELING LANGUAGE)

Zunifikowany język modelowania

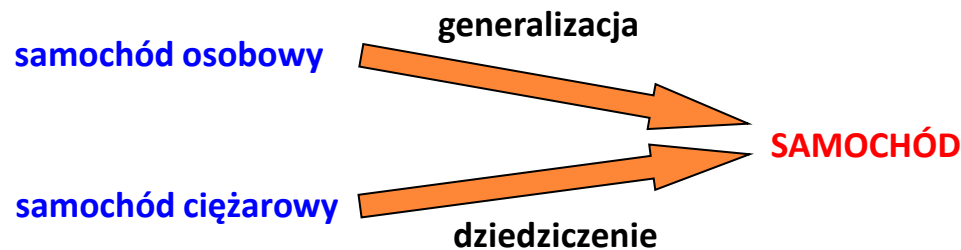


równoważny



JĘZYK UML (UNIFIED MODELING LANGUAGE)

Zunifikowany język modelowania



Samochód osobowy/ciężarowy – dziedziczy wszystkie cechy SAMOCHÓD

→ przewaga modelu OBIEKTOWEGO nad modelem RELACYJNYM

→ zlecenie częściowego wykonywania prac (tożsamość i hermetyczność)

→ łatwo rozbudować model OBIEKTOWY

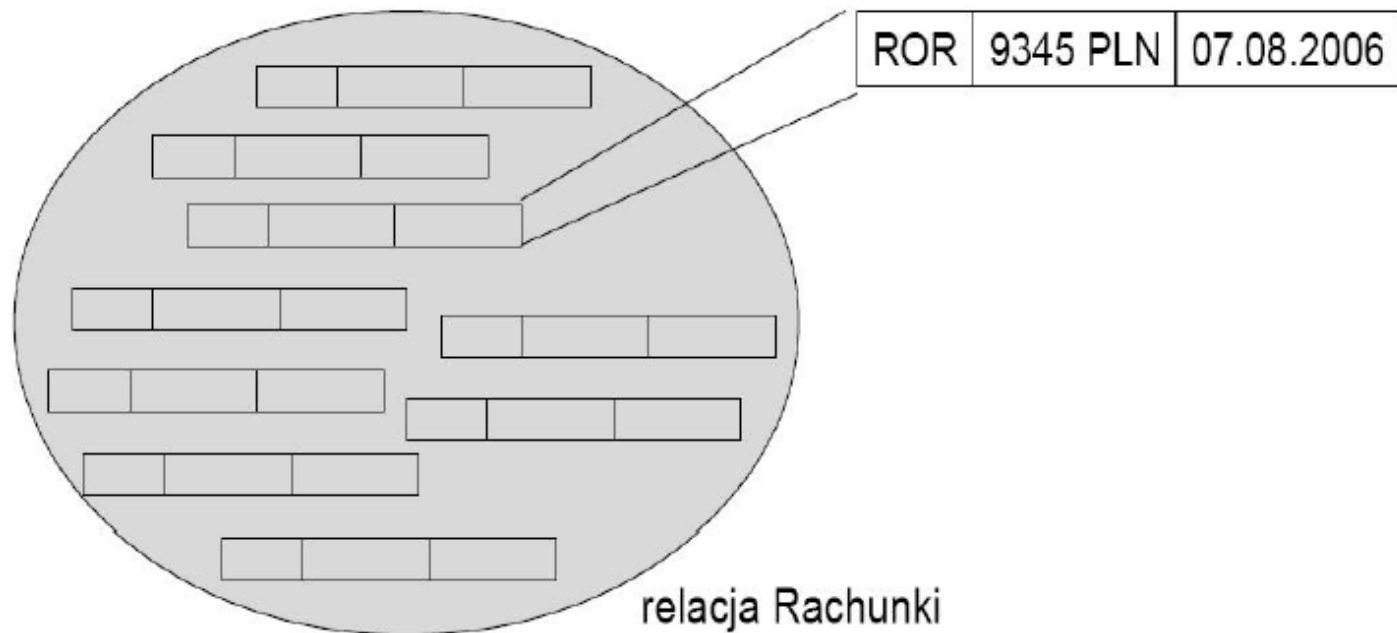


Struktury danych (1)

1. Baza danych jest zbiorem relacji
2. Schemat relacji R , oznaczony przez $R(A_1, A_2, \dots, A_n)$, składa się z nazwy relacji R oraz listy atrybutów A_1, A_2, \dots, A_n
3. Liczbę atrybutów składających się na schemat relacji R nazywamy stopniem relacji
4. Każdy atrybut A_i schematu relacji R posiada domenę, oznaczoną jako $\text{dom}(A_i)$
5. Domena definiuje zbiór wartości atrybut relacji poprzez podanie typu danych
6. Relacją r o schemacie $R(A_1, A_2, \dots, A_n)$, oznaczoną $r(R)$, nazywamy zbiór n -tek (krotek) postaci $r=\{t_1, t_2, \dots, t_m\}$.
7. Pojedyncza krotka t jest uporządkowaną listą n wartości $t=\langle v_1, v_2, \dots, v_n \rangle$, gdzie v_i , $1 \leq i \leq n$, jest elementem $\text{dom}(A_i)$ lub specjalną wartością pustą (NULL)
8. i -ta wartość krotki t , odpowiadająca wartości atrybutu A_i , będzie oznaczana przez $t[A_i]$
9. Relacja $r(R)$ jest relacją matematyczną stopnia n zdefiniowaną na zbiorze domen $\text{dom}(A_1), \text{dom}(A_2), \dots, \text{dom}(A_n)$ będącą podzbiorem iloczynu kartezyjskiego domen definiujących R : $r(R) \subseteq \text{dom}(A_1) \times \text{dom}(A_2) \times \dots \times \text{dom}(A_n)$

Struktury danych (2)

Relacja jest zbiorem krotek (k-wartości), które są listami wartości



relacja \neq zmienna relacyjna



STRUKTURA DANYCH

$R(A_1, A_2, A_3, \dots, A_n) \rightarrow$ schemat relacji R

Stopień relacji

Domena:
 → string
 → char

MODEL LOGICZNY

Studenci kursu BD1 (nr_id, Imię, Nazwisko, Grupa)

dom (Nazwisko) = łańcuch 80 znaków

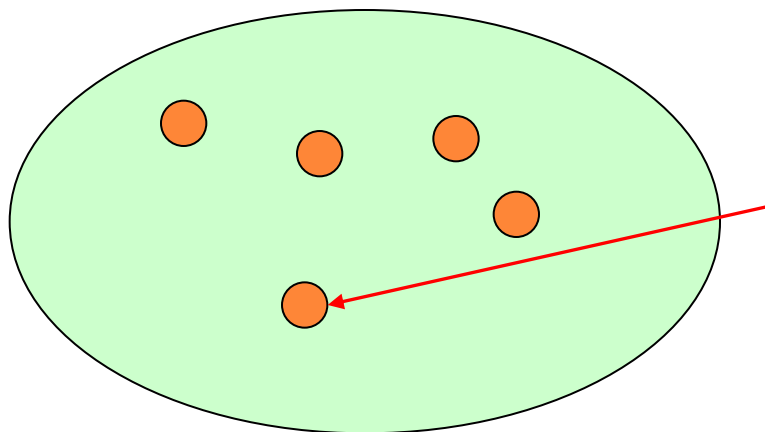
Ograniczenie
integralnościowe



SCHEMAT RELACJI

→ jest w postaci zbioru:

Studenci kursu BD1 (nr_id, Imię, Nazwisko, Grupa)



Każdy element tego zbioru – **opisany przez 4 atrybuty**

Każda instancja tego elementu - **to KROTKA**

KROTKA → na poziomie modelowania danych logicznych



SCHEMAT RELACJI

→ jest w postaci zbioru:

Studenci kursu BD1 (nr_id, Imię, Nazwisko, Grupa)

nr_id	Imię	Nazwisko	Grupa
1	Jan	Kowalski	PIH
....

Diagram illustrating a relation schema for "Studenci kursu BD1" with attributes: nr_id, Imię, Nazwisko, Grupa. A specific tuple t1 is highlighted in light blue, containing values (1, Jan, Kowalski, PIH). Red arrows point from the attribute names to the corresponding columns, and from the tuple values to the corresponding cells. The tuple is labeled "KROTKA t1". Below the table, the attributes are labeled V1, V2, V3, and V4, with red arrows pointing from the tuple values to these labels.

$t1(V2) = \text{Jan}$

Każdy atrybut relacji ma unikalną nazwę (w ramach tej samej relacji nie może się powtórzyć)

Własności atrybutów są ATOMOWE – 1NF/1PN

Atrybuty JEDNOWARTOŚCIOWE



Imię Nazwisko (atr. 2-wartościowy (0PN))



Baza danych

Podsumowanie definicyjne

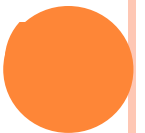
- Baza danych = zbiór relacji
- Schemat bazy danych = zbiór schematów relacji
- Schemat relacji = zbiór {atrybut, dziedzina, [ograniczenia integralnościowe]}
- Relacja = zbiór krotek
- Krotka = lista wartości atomowych



Relacje - cechy

Charakterystyka relacji

- Każdy atrybut relacji ma unikalną nazwę
- Porządek atrybutów w relacji nie jest istotny
- Porządek krotek w relacji nie jest istotny i nie jest elementem definicji relacji
- Wartości atrybutów są atomowe (elementarne)
- Relacja nie zawiera rekordów powtarzających się (z definicji zbiorów)



Relacje - klucze

Unikalność krotek relacji - klucze

Ograniczenie na unikalność krotek relacji

1. Każdy podzbiór S atrybutów relacji R , taki że dla każdych dwóch krotek ze zbioru $r(R)$ zachodzi $t_1[S] \neq t_2[S] \rightarrow$ superkluczem (super key) R
2. Superklucz - cały schemat relacji
3. Superklucz może posiadać nadmiarowe atrybuty
4. Kluczem K schematu relacji R nazywamy superklucz schematu R o takiej własności, że usunięcie dowolnego atrybutu A z K powoduje, że $K' = K - A$ nie jest już superkluczem
5. Klucz jest minimalnym superkluczem zachowującym własność unikalność krotek relacji
6. Schemat relacji może posiadać więcej niż jeden klucz

Wyróżniony klucz \rightarrow klucz podstawowy (primary key)

Pozostałe klucze \rightarrow klucze wtórne lub kandydujące (foreign key)



Ograniczenia integralnościowe (1)

1. Mechanizm (reguła), który gwarantuje że dane wpisane do relacji spełnią nałożone na nie warunki (czuwa nad tym SZBD)
2. Definiuje się na poziomie
 - pojedynczego atrybutu
 - całej relacji
3. Rodzaje
 - klucz podstawowy (primary key)
 - klucz obcy (foreign key)
 - unikalność (unique)
 - zawężenie domeny/dziedziny (check)
 - wartość pusta/niepusta (NULL/NOT NULL)



Ograniczenia integralnościowe (2)


Klucz podstawowy relacji (primary key - PK)

- atrybut (lub zbiór atrybutów), którego wartość jednoznacznie identyfikuje krotkę
- wartość ta jest unikalna w obrębie całej relacji i jest niepusta

Przykłady:

- IdPrac, adres e-mail, NIP, PESEL, nr dowodu, nr paszportu

relacja Pracownicy



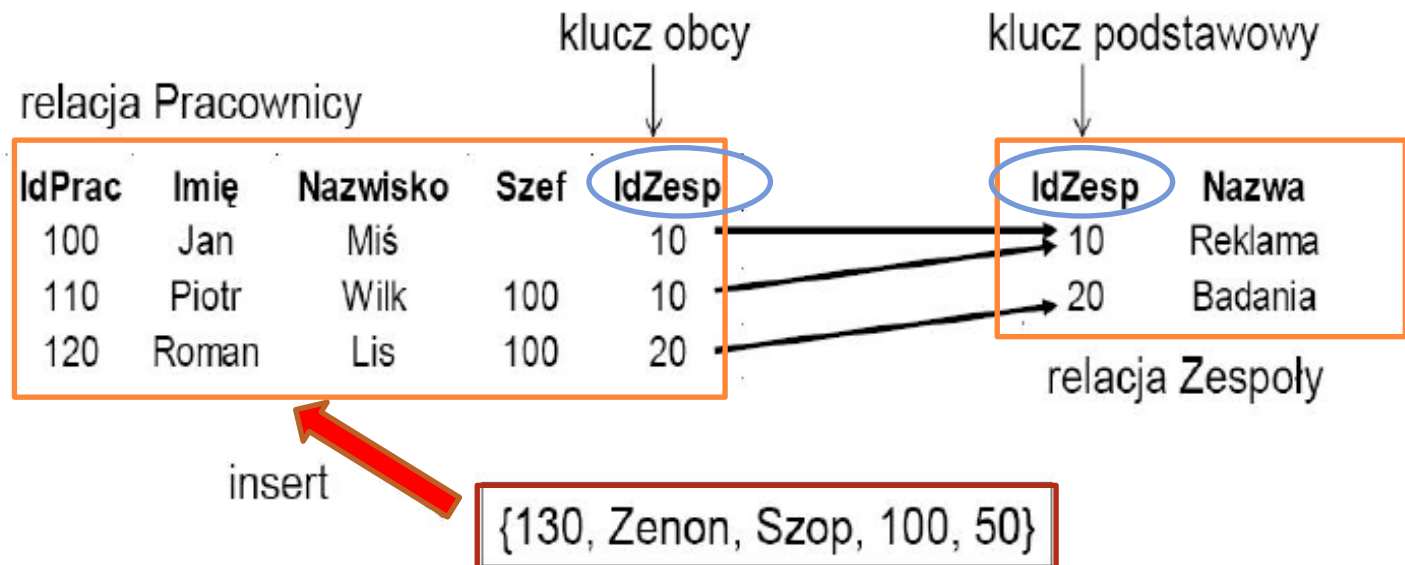
IdPrac	Imię	Nazwisko	Szef
100	Jan	Miś	
110	Piotr	Wilk	100
120	Roman	Lis	100



Ograniczenia integralnościowe (3)

Klucz obcy relacji (foreign key - FK)

- atrybut (lub zbiór atrybutów), który wskazuje na klucz podstawowy
- służy do reprezentowania powiązań między danymi (łączenia relacji)



Naruszono w tym przypadku ograniczenie integralnościowe klucza obcego.

Ograniczenia integralnościowe (4)

Klucz obcy relacji (foreign key) – formalna reguła

Dane są relacje R1 i R2. Podzbiór FK atrybutów relacji R1 nazywany jest kluczem obcym R1 jeżeli:

- atrybuty w FK mają taką samą domenę jak atrybuty klucza podstawowego PK relacji R2
- dla każdej krotki $t1$ relacji R1 istnieje dokładnie jedna krotka $t2$ relacji R2, taka że $t1 [FK] = t2 [PK]$, lub $t1 [FK] = \text{null}$

Klucz obcy (ograniczenie referencyjne) gwarantuje, że rekordy z tabeli R1 występują w kontekście związanego z nim rekordu z tabeli R2



Ograniczenia integralnościowe (5)

Zawężenie dziedziny (ograniczenie domeny) atrybutu (check)

- ograniczenie dozwolonych wartości do pewnego podzbioru przez wyrażenie logiczne określające przedział lub za pomocą wyliczeniowej listy wartości

Przykłady:

- płeć: K, M, nieznana, N/A (zgodnie ze standardem ISO)
- pensja: wartości dodatnie
- kolor oczu: niebieskie, szare, piwne



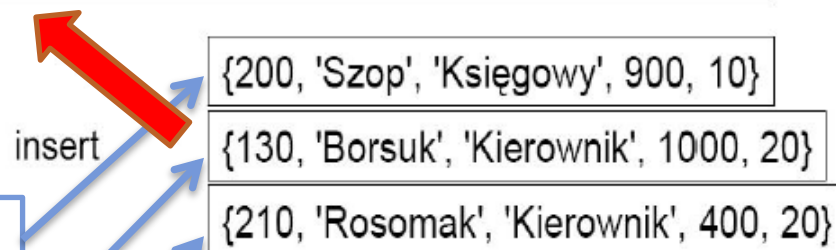
Ograniczenia integralnościowe (5)

Zawężenie dziedziny - przykład

- Etat - dziedzina: {'Analityk', 'Dyrektor', 'Referent', 'Kierownik', 'Sekretarka'}
- Płaca - dziedzina: $\text{płaca} > 500$
- IdPrac - klucz podstawowy

relacja Pracownicy

IdPrac	Nazwisko	Etat	Płaca	Szef	IdZesp
120	Kowalski	Analityk	850	100	10
100	Tarzan	Dyrektor	1700		10
130	Nowak	Referent	600	100	10
110	Józek	Kierownik	1200	100	20
140	Nowacki	Analityk	800	110	20
150	Bunio	Sekretarka	700	100	10



Naruszenie integralności etatu

Naruszenie integralności PK

Naruszenie integralności płacy



Operatory algebry relacyjnej (1)

Założenia podstawowe

1. Operatory relacyjne mają charakter ogólny (ang. *generic*), co oznacza, że mogą być stosowane do wszystkich możliwych relacji. Nie istnieje potrzeba tworzenia specjalnej relacji złączenia relacji A i B a innej dla relacji B i C.
2. Operatory działają w trybie samego odczytu: odczytują operandy i zwracają wynik, lecz nie modyfikują niczego w samych operandach. Nie ingerują w zmienne relacyjne, lecz operują na relacjach
3. Operacje INSERT, DELETE i UPDATE nie należą do algebry relacyjnej ponieważ modyfikują zawartość zmiennej relacyjnej



Operatory algebry relacyjnej (2)

selekcja

projekcja

iloczyn
kartyzański

a
b
c

x
y

a	x
a	y
b	x
b	y
c	x
c	y

iloczyn krzyżowy, złączenie krzyżowe

a	x
a	y
a	z
b	x
c	y

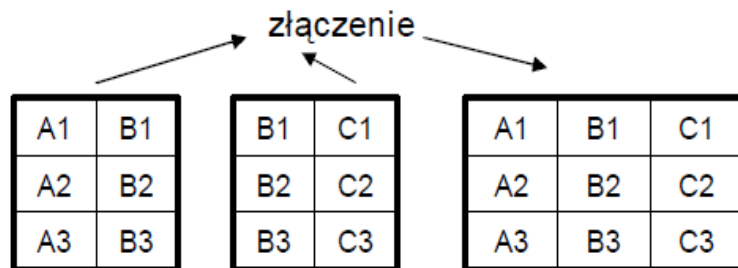
x
y

dzielenie

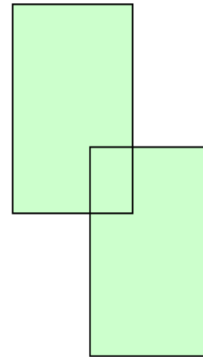
a



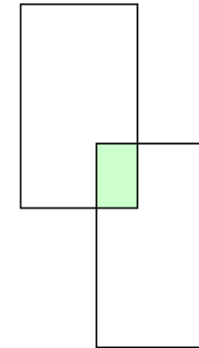
Operatory algebry relacyjnej (3)



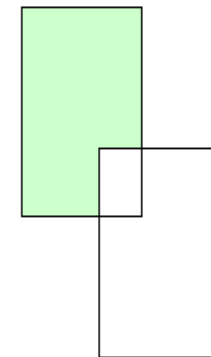
unia



część wspólna



różnica



Operatory algebry relacyjnej (4)

Selekcja

Przeznaczenie:

- wyodrębnienie podzbioru krotek relacji, które spełniają warunek selekcji

Notacja: $\sigma_{\langle \text{warunek selekcji} \rangle} (\langle \text{Nazwa relacji} \rangle)$

- warunek selekcji jest zbiorem predykatów postaci
 - $\langle \text{atrybut} \rangle \langle \text{operator relacyjny} \rangle \langle \text{litera} \rangle$
 - $\langle \text{atrybut} \rangle \langle \text{operator relacyjny} \rangle \langle \text{atrybut} \rangle$
- predykaty są łączone operatorami logicznymi: AND lub OR

Własności: operacja selekcji jest komutatywna:

$$\sigma_{\langle \text{war1} \rangle} (\sigma_{\langle \text{war2} \rangle} (R)) = \sigma_{\langle \text{war2} \rangle} (\sigma_{\langle \text{war1} \rangle} (R))$$



Operatory algebry relacyjnej (5)

Selekcja cd. - przykłady

S1: $\sigma_{\text{IdZesp}=10}(\text{Pracownicy})$

```
select IdPrac, Nazwisko, Etat, Szef,  
       Zatrudniony, Płaca, IdZesp  
from Pracownicy  
where IdZesp=10
```

S2: $\sigma_{\text{Płaca} > 7000}(\text{Pracownicy})$

```
select IdPrac, Nazwisko, Etat, Szef,  
       Zatrudniony, Płaca, IdZesp  
from Pracownicy  
where Płaca > 7000
```

S3: $\sigma_{(\text{IdZesp}=10 \text{ AND } \text{Płaca} > 7000) \text{ OR } (\text{IdZesp}=20 \text{ AND } \text{Płaca} > 8000)}(\text{Pracownicy})$

```
select IdPrac, Nazwisko, Etat, Szef,  
       Zatrudniony, Płaca, IdZesp  
from Pracownicy  
where (IdZesp=10 and Płaca>7000)  
or     (IdZesp=20 and Płaca>8000)
```

S4: $\sigma_{\text{Etat}='Księgowy' \text{ AND } (\text{Płaca} \geq 6000 \text{ AND } \text{Płaca} < 9000)}(\text{Pracownicy})$

```
select IdPrac, Nazwisko, Etat, Szef,  
       Zatrudniony, Płaca, IdZesp  
from Pracownicy  
where Etat='KSIĘGOWY'  
and     (Płaca>=6000 and Płaca<9000)
```



Operatory algebry relacyjnej (6)

Projekcja

Przeznaczenie:

- wyodrębnienie wybranych atrybutów relacji

Notacja: $\pi_{\langle \text{atrybuty} \rangle} (\langle \text{Nazwa relacji} \rangle)$

- atrybuty jest podzbiorem atrybutów ze schematu relacji

Własności: operacja projekcji nie jest komutatywna

Składanie operacji projekcji jest możliwe jeżeli lista2 zawiera wszystkie atrybuty lista1

$$\pi_{\langle \text{lista1} \rangle} (\pi_{\langle \text{lista2} \rangle} (R)) = \pi_{\langle \text{lista1} \rangle} (R)$$



Operatory algebry relacyjnej (7)

Projekcja - przykład

P1: $\pi_{\text{Nazwisko}}(\text{Pracownicy})$

```
select Nazwisko  
from Pracownicy
```

P2: $\pi_{\text{Nazwisko, Etat, Płaca}}(\text{Pracownicy})$

```
select Nazwisko, Etat, Płaca  
from Pracownicy
```



Operatory algebry relacyjnej (8)

Składanie operacji

Wynik danej operacji może być zbiorem wejściowym dla innej operacji

$$\sigma_{\text{IdZesp} = 10} (\text{Pracownicy}) \Rightarrow \text{PracZesp10}$$

$$\pi_{\text{IdPrac}, \text{Nazwisko}} (\text{PracownicyZesp10}) \Rightarrow \text{PracZesp10Wynik}$$



$$\text{PracZesp10Wynik} = \pi_{\text{IdPrac}, \text{Nazwisko}} (\sigma_{\text{IdZesp} = 10} (\text{Pracownicy}))$$

Złożenie
operacji



Operatory algebry relacyjnej (9)

Operacje na zbiorach 1

Kompatybilność relacji

- Dwie relacje: $R(A_1, \dots, A_n)$ i $S(B_1, \dots, B_n)$ są kompatybilne, jeżeli mają ten sam stopień i jeżeli $\text{dom}(A_i) = \text{dom}(B_i)$ dla $1 \leq i \leq n$

Operacje na zbiorach

- dla dwóch kompatybilnych relacji: $R(A_1, \dots, A_n)$ i $S(B_1, \dots, B_n)$



Operatory algebry relacyjnej (10)

Operacje na zbiorach 2

Suma:

- Wynikiem tej operacji, oznaczanej przez $R \cup S$, jest relacja zawierająca wszystkie krotki, które występują w R i wszystkie krotki, które występują w S , z wyłączeniem duplikatów krotek
- Operacja sumy jest operacją komutatywną: $R \cup S = S \cup R$

Iloczyn:

- Wynikiem tej operacji, oznaczonej przez $R \cap S$, jest relacja zawierająca krotki występujące zarówno w R i S
- Operacja iloczynu jest operacją komutatywną: $R \cap S = S \cap R$

Różnica:

- Wynikiem tej operacji, oznaczonej przez $R - S$, jest relacja zawierająca wszystkie krotki, które występują w R i nie występują w S
- Operacja różnicy nie jest operacją komutatywną: $R - S \neq S - R$



Operatory algebry relacyjnej (11)

Operacje na zbiorach - przykłady

Uczniowie

Imię	Nazwisko
Ala	Kusiak
Edek	Musiak
Adam	Zając
Olek	Struś
Ola	Buba

Instruktorzy

Imię	Nazwisko
Jan	Kuc
Edek	Musiak
Wacek	Misiek

```
select Imię, Nazwisko
from Uczniowie
UNION
select Imię, Nazwisko
from Instruktorzy;
```

```
select Imię, Nazwisko
from Uczniowie
INTERSECT
select Imię, Nazwisko
from Instruktorzy;
```

Uczniowie \cup Instruktorzy

Imię	Nazwisko
Ala	Kusiak
Edek	Musiak
Adam	Zając
Olek	Struś
Ola	Buba
Jan	Kuc
Wacek	Misiek

Uczniowie \cap Instruktorzy

Imię	Nazwisko
Edek	Musiak

Uczniowie - Instruktorzy

Imię	Nazwisko
Ala	Kusiak
Adam	Zając
Olek	Struś
Ola	Buba

Instruktorzy - Uczniowie

Imię	Nazwisko
Jan	Kuc
Wacek	Misiek

```
select Imię, Nazwisko
from Uczniowie
MINUS
select Imię, Nazwisko
from Instruktorzy;
```

```
select Imię, Nazwisko
from Instruktorzy
MINUS
select Imię, Nazwisko
from Uczniowie;
```


Operatory algebry relacyjnej (12)

Iloczyn kartezjański

- Dane są dwie relacje: $R(A_1, \dots, A_n)$ i $S(B_1, \dots, B_m)$ - wynikiem iloczynu kartezjańskiego relacji R i S , oznaczonym przez $R \times S$, jest relacja Q stopnia $n+m$ i schemacie: $Q(A_1, \dots, A_n, B_1, \dots, B_m)$
- Krotkom w relacji Q odpowiadają wszystkie kombinacje krotek z relacji R i S
- Jeżeli relacja R ma N krotek, a relacja S ma M krotek, to relacja Q będzie miała $N \cdot M$ krotek



Operatory algebry relacyjnej (13)

Iloczyn kartezjański - przykład

Pracownicy

Imię	Nazwisko
Ala	Kusiak
Edek	Musiał
Adam	Zajac

Zespoły

Nazwa	Lokalizacja
Reklama	Krucza 10
Badania	Piotrowo 3A

Pracownicy x Zespoły

Imię	Nazwisko	Nazwa	Lokalizacja
Ala	Kusiak	Reklama	Krucza 10
Edek	Musiał	Reklama	Krucza 10
Adam	Zajac	Reklama	Krucza 10
Ala	Kusiak	Badania	Piotrowo 3A
Edek	Musiał	Badania	Piotrowo 3A
Adam	Zajac	Badania	Piotrowo 3A



Operatory algebry relacyjnej (14)

Połączenie – złączenie

Przeznaczenie:

- łączenie na podstawie warunku połączeniowego wybranych krotek z dwóch relacji w pojedynczą krotkę

Notacja:

operacja połączenia relacji $R(A_1, \dots, A_n)$ i $S(B_1, \dots, B_m)$, jest oznaczona jako: $R \leftrightarrow \langle \text{warunek połączeniowy} \rangle S$

- warunek połączeniowy jest zbiorem predykatów połączonych operatorami logicznymi AND
- predykaty są postaci: $A_i \theta B_j$
 - A_i i B_j są atrybutami połączeniowymi
 - A_i jest atrybutem R , B_j jest atrybutem S
 - $\text{dom}(A_i) = \text{dom}(B_j)$,
 - θ jest operatorem relacyjnym ze zbioru $\{ =, \neq, <, \leq, >, \geq \}$



Operatory algebry relacyjnej (15)

Połączenie – złączenie

- Ogólna postać operacji połączenia (theta join) – $R \theta S$
- Połączenie równościowe (equi join) – θ jest operatorem $=$
- Połączenie nierównościowe (non-equi join) – θ jest operatorem różnym od $=$
- Połączenie naturalne (natural join)
 - połączenie równościowe
 - jeden z atrybutów połączeniowych jest usunięty ze schematu relacji wynikowej
 - oznaczane jako: $R * S$
 - atrybuty połączeniowe w obu relacjach muszą mieć taką samą nazwę



Operatory algebry relacyjnej (15)

Połączenie - przykład

Pracownicy

IdPrac	Imię	Nazwisko	Szef	IdZesp
100	Jan	Miś		10
110	Piotr	Wilk	100	10
120	Roman	Lis	100	20

Zespoly


IdZesp	Nazwa
10	Reklama
20	Badania

```
select *
from pracownicy p join zespoly z
on p.id_zesp=z.id_zesp
```

połączenie równościowe
(niestandardowe)

```
select nazwisko, nazwa
from pracownicy p join zespoly z
on p.id_zesp=z.id_zesp
```

połączenie równościowe
(standardowe)

Pracownicy  Pracownicy

Szef=IdPrac

IdPrac	Imię	Nazwisko	Szef	IdZesp	IdPrac	Imię	Nazwisko	Szef	IdZe
110	Piotr	Wilk	100	10	100	Jan	Miś		10
120	Roman	Lis	100	20	100	Jan	Miś		10

połączenie naturalne
(standardowe)

Pracownicy * Zespoly

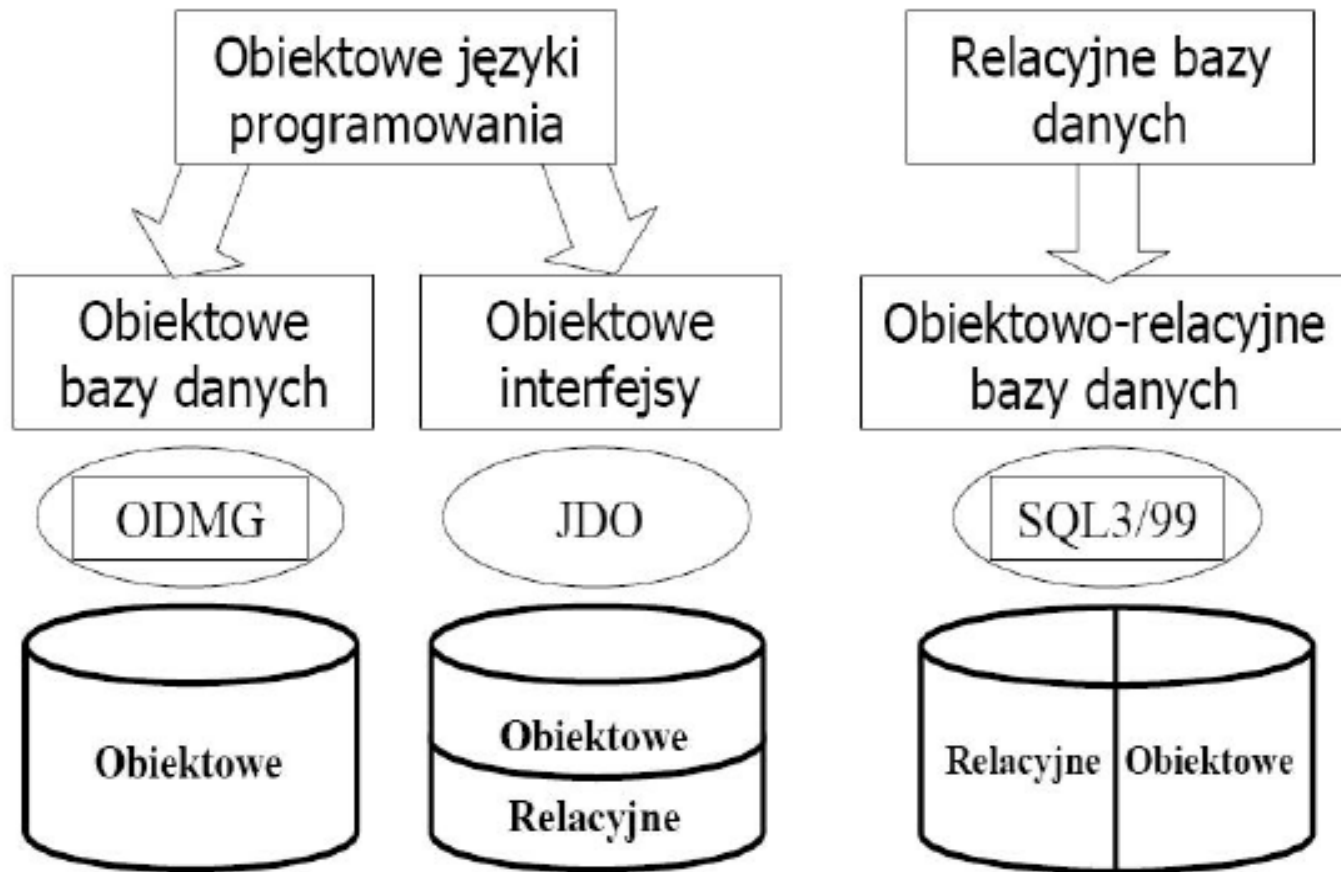
IdPrac	Imię	Nazwisko	Szef	IdZesp	Nazwa
100	Jan	Miś		10	Reklama
110	Piotr	Wilk	100	10	Reklama
120	Roman	Lis	100	20	Badania

```
select *
from pracownicy p natural join zespoly z
```



Obiektowy Model danych

Drogi rozwoju obiektowych baz danych



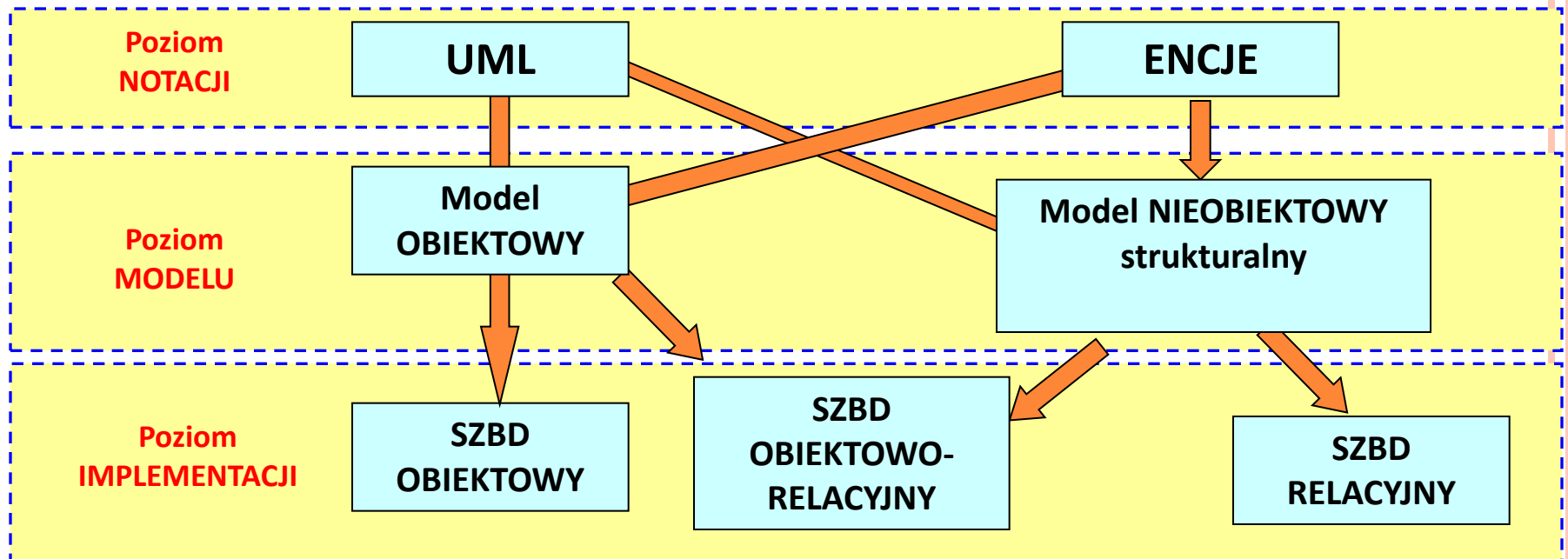
MODEL ERD

→ „model związków encji”

MODEL ER → schemat relacyjny

TRANSFORMACJA

MODEL KONCEPLUALNY → schemat relacyjny



Koniec wykładu 2

