

Wykład 1

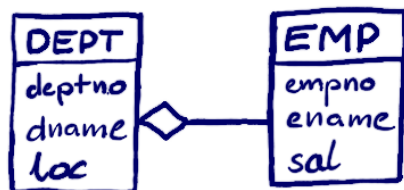
Co to są bazy danych?

Skąd się wziął SQL?

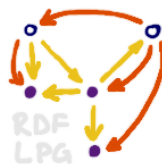
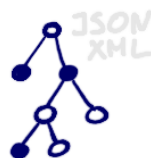
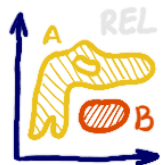
Co to są bazy danych?



modelowanie
pojęciowe



modelowanie
logiczne



$\forall x \forall y A(x,y) \Rightarrow \exists z B(x,z)$

$o \rightarrow \bullet \mid \bullet \mid \epsilon$
 $\bullet \rightarrow o^*$

$o \subseteq A \rightarrow \bullet$
 $o \subseteq B \rightarrow \bullet$

reprezentacja
fizyczna



(0,1) (0,2) (1,1) (1,2) ...

łatwo dodać, trudniej znaleźć

(0,1); (1,2) (2,0); (2,0) ...

trudno dodać, łatwiej znaleźć



modelowanie
przypadków
użycia



Kto zarabia najwięcej?
Co jest stolicą Tanzanii?
Jak dojechać z Bari do Bari?

Na urlopie zabierz premię.
Astana to teraz Nur-Sultan.
Spnedaj bilet Lublin - Rzeszów.

pisanie zapytań
i transakcji

$\{x: (7,x) \in A\}$

// ● / ○ /

MATCH $x: \bullet \rightarrow y: \bullet$
RETURN y

$A := A \setminus \{(3,4)\}$
 $B := B \cup A$

TRANSAKCJE

plany
zapytań

```

for i=1 to 100 do
  begin
  ...
  end;
  
```

realizacja
transakcji

```

for i=1 to 100 do
  for j=1 to 100 do
    begin
    ...
    end;
  end;
end;
  
```

Relacyjna rewolucja

1959-69: COBOL

- Bezpośrednia praca z fizyczną reprezentacją danych.
- Zapytania zależne od reprezentacji: przeglądanie plików, chodzenie po wskaźnikach.

1970: Edgar F. Codd

- Ustalmy jeden wspólny model danych: relacje.
- Metadane (schemat bazy): zbiór nazw relacji z arnościami.

$\{ \text{Film}^{(4)}, \text{Seans}^{(3)}, \text{Kino}^{(2)} \}$

- Dane (instancja bazy): zawartość relacji, tzn. zbiory krotek.

$I = \{ \text{Film}^I, \text{Seans}^I, \text{Kino}^I \}$

$\text{Film}^I = \{ (\text{Drogówka}, 2012, \text{Smarnowski}, \text{Topa}),$
 $(\text{Star Wars}, 1999, \text{Lucas}, \text{Ford}), \dots \}$

$\text{Seans}^I = \{ (\text{Femina}, \text{Drogówka}, 17:00),$
 $(\text{Muranów}, \text{Star Wars}, 23:15), \dots \}$

$\text{Kino}^I = \{ (\text{Femina}, \text{Solidarności}),$
 $(\text{Muranów}, \text{Andersa}), \dots \}$

Jak pisać zapytania?

Logika pierwszego rzędu (FO)

Formuły atomowe: $R(x_1, x_2, \dots, x_n)$, $x=y$, gdzie $R^{(n)}$ jest w schemacie (sygnaturze).

Spójniki: $\varphi \wedge \psi$, $\varphi \vee \psi$, $\neg \varphi$, $\varphi \rightarrow \psi$, $\varphi \leftrightarrow \psi$.

Kwantyfikatory: $\exists x. \varphi(x)$, $\forall x. \varphi(x)$.

$$\varphi(x, y) = \exists z \text{ Seans}(z, \text{Dragówka}, y) \wedge \text{Kino}(z, x)$$

$$\begin{aligned}\varphi^I &= \{ (a, b) \in \text{adom}(I)^2 \mid I \models \varphi(a, b) \} \\ &= \{ (\text{Solidarności}, 17:00) \}\end{aligned}$$

$\text{adom}(I)$ to dziedzina aktywna I ,
czyli wszystkie wartości użyte w I .

Algebra relacji (RA)

Wyrażenia atomowe: $\{ (a_1, a_2, \dots, a_n) \}$, R , gdzie R jest w schemacie.

Selekcja: $\sigma_{i=j}(S) = \{ (a_1, \dots, a_n) \mid (a_1, \dots, a_n) \in S, a_i = a_j \}$

$\sigma_{i=c}(S) = \{ (a_1, \dots, a_n) \mid (a_1, \dots, a_n) \in S, a_i = c \}$

Rzut (permutacja, duplikacja): $\pi_{i_1, i_2, \dots, i_k}(S) = \{ (a_{i_1}, a_{i_2}, \dots, a_{i_k}) \mid (a_1, a_2, \dots, a_n) \in S \}$

Produkt (kartezjański): $S \times T = \{ (a_1, \dots, a_n, b_1, \dots, b_m) \mid (a_1, \dots, a_n) \in S, (b_1, \dots, b_m) \in T \}$

Suma: $S \cup T = \{ (a_1, a_2, \dots, a_n) \mid (a_1, a_2, \dots, a_n) \in S \text{ lub } (a_1, a_2, \dots, a_n) \in T \}$

Różnica: $S \setminus T = \{ (a_1, a_2, \dots, a_n) \mid (a_1, a_2, \dots, a_n) \in S, (a_1, a_2, \dots, a_n) \notin T \}$

Wyrażenie $E = \pi_{5,3} \sigma_{1=4} \sigma_{2='Dragówka'}(Seans \times Kino)$ na I daje relację $E^I = \pi_{5,3} \sigma_{1=4} \sigma_{2='Dragówka'}(Seans^I \times Kino^I)$.

SELECT adres, godzina **FROM** Seans, Kino

WHERE Seans.kino = Kino.nazwa **AND** Seans.film = "Dragówka";

Twierdzenie Codd (1972)

$$FO = RA$$

Tzn. w logice pierwszego rzędu i w algebrze relacji
da się wyrazić dokładnie te same zapytania.

Od RA do FO

E	φ_E
R	$R(x_1, \dots, x_n)$
$\{(a_1, \dots, a_n)\}$	$x_1 = a_1 \wedge \dots \wedge x_n = a_n$
$\sigma_{i=j}(E)$	$\varphi_E(x_1, \dots, x_n) \wedge x_i = x_j$
$\pi_{i_1, \dots, i_k}(E)$	$\exists y_1 \dots \exists y_n \varphi_E(y_1, \dots, y_n) \wedge x_1 = y_{i_1} \wedge \dots \wedge x_k = y_{i_k}$
$E \times F$	$\varphi_E(x_1, \dots, x_n) \wedge \varphi_F(x_{n+1}, \dots, x_{n+m})$
$E \cup F$	$\varphi_E(x_1, \dots, x_n) \vee \varphi_F(x_1, \dots, x_n)$
$E - F$	$\varphi_E(x_1, \dots, x_n) \wedge \neg \varphi_F(x_1, \dots, x_n)$

Od FO do RA

φ	E_φ
$R(x_3, x_3, x_1)$	$\pi_{3,1} \sigma_{1=2} R$ <p>powtórzenia zmiennych porządkowanie kolumn wg. numerów zmiennych</p>
$x_3 = x_1$	$\sigma_{1=2}(\text{adom} \times \text{adom})$
$\varphi(x_1, x_2) \vee \psi(x_1, x_3)$	$E_\varphi \times \text{adom} \vee \pi_{1,3,2}(E_\psi \times \text{adom})$ <p>x_3 x_2</p>
$\neg \varphi(x_1, x_3)$	$\text{adom} \times \text{adom} - E_\varphi$
$\exists x_3 \varphi(x_1, x_3, x_4)$	$\pi_{1,3} E_\varphi$

$$\text{adom} = \bigcup_{R^{(n)} \in \Sigma} \bigcup_{i=1}^n \pi_i(R)$$