

Databáze

Přednáška 1. Relace

verze z 17. září 2023

Jan Laštovička



KATEDRA INFORMATIKY
UNIVERZITA PALACKÉHO V OLOMOUCI

1 Úvod

2 Relace a relační proměnné

3 SQL

Databázové systémy:

- PostgreSQL (relační model)
- MongoDB (dokumentový model)
- Elasticsearch (fulltextové vyhledávání)

1 Úvod

2 Relace a relační proměnné

3 SQL



Vlastníme filmy:

- The Avengers (1998)
- The Avengers (2012)
- The Matrix (1999)

Vlastníme filmy:

- The Avengers (1998)
- The Avengers (2012)
- The Matrix (1999)

Databáze?

Atribut = vhodný řetězec písmen anglické abecedy a podtržítka

Například: `title`, `year`, `movie_id`

Atribut = vhodný řetězec písmen anglické abecedy a podtržítka

Například: `title`, `year`, `movie_id`

\mathcal{Y} ... množina všech atributů

Atribut = vhodný řetězec písmen anglické abecedy a podtržítka

Například: `title`, `year`, `movie_id`

\mathcal{Y} ... množina všech atributů

Pro každé $y \in \mathcal{Y}$ máme D_y

- spočetná množina
- **doména atributu** y

Atribut = vhodný řetězec písmen anglické abecedy a podtržítko

Například: `title`, `year`, `movie_id`

\mathcal{Y} ... množina všech atributů

Pro každé $y \in \mathcal{Y}$ máme D_y

- spočetná množina
- **doména atributu** y

Například:

- D_{title} ... množina všech řetězců nad anglickou abecedou
- D_{year} ... množina všech přirozených čísel



Relační schéma a n -tice



Relační schéma = konečná podmnožina \mathcal{Y}

Například: `{title, year}`

Relační schéma = konečná podmnožina \mathcal{Y}

Například: $\{\text{title}, \text{year}\}$

$R \dots$ relační schéma

$$t: R \rightarrow \bigcup_{y \in R} D_y$$

- pro každé $y \in R$ je $t(y) \in D_y$
- n -tice nad R

Například:

$\{\langle \text{title}, \text{'The Matrix'} \rangle, \langle \text{year}, 1999 \rangle\}$ je n -tice nad $\{\text{title}, \text{year}\}$

Relační schéma = konečná podmnožina \mathcal{Y}

Například: $\{\text{title}, \text{year}\}$

$R \dots$ relační schéma

$t: R \rightarrow \bigcup_{y \in R} D_y$

- pro každé $y \in R$ je $t(y) \in D_y$
- n -tice nad R

Například:

$\{\langle \text{title}, \text{'The Matrix'} \rangle, \langle \text{year}, 1999 \rangle\}$ je n -tice nad $\{\text{title}, \text{year}\}$

$\text{Tupl}(R) \dots$ množina všech n -tic nad R

Relace

$R \dots$ relační schéma

Relace nad $R =$ konečná podmnožina $\text{Tupl}(R)$



$R \dots$ relační schéma

Relace nad $R =$ konečná podmnožina $\text{Tupl}(R)$

Například:

```
{ { <title, 'The Avengers'>, <year, 1998> },  
  { <title, 'The Avengers'>, <year, 2012> },  
  { <title, 'The Matrix'>, <year, 1999> } }
```

je relace nad $\{\text{title}, \text{year}\}$.

R ... relační schéma

Relace nad R = konečná podmnožina $\text{Tupl}(R)$

Například:

$\{\{\langle \text{title}, 'The Avengers' \rangle, \langle \text{year}, 1998 \rangle\},$
 $\{\langle \text{title}, 'The Avengers' \rangle, \langle \text{year}, 2012 \rangle\},$
 $\{\langle \text{title}, 'The Matrix' \rangle, \langle \text{year}, 1999 \rangle\}\}$

je relace nad $\{\text{title}, \text{year}\}$.

Zobrazení tabulkou:

title	year
The Avengers	1998
The Avengers	2012
The Matrix	1999

R ... relační schéma

Relace nad R = konečná podmnožina $\text{Tupl}(R)$

Například:

$\{\{\langle \text{title}, 'The Avengers' \rangle, \langle \text{year}, 1998 \rangle\},$
 $\{\langle \text{title}, 'The Avengers' \rangle, \langle \text{year}, 2012 \rangle\},$
 $\{\langle \text{title}, 'The Matrix' \rangle, \langle \text{year}, 1999 \rangle\}\}$

je relace nad $\{\text{title}, \text{year}\}$.

Zobrazení tabulkou:

title	year	nebo	year	title
The Avengers	1998		1999	The Matrix
The Avengers	2012		1998	The Avengers
The Matrix	1999		2012	The Avengers

Relace



$R \dots$ relační schéma

Relace nad $R =$ konečná podmnožina $\text{Tupl}(R)$

Například:

$\{\{\langle \text{title}, 'The Avengers' \rangle, \langle \text{year}, 1998 \rangle\},$
 $\{\langle \text{title}, 'The Avengers' \rangle, \langle \text{year}, 2012 \rangle\},$
 $\{\langle \text{title}, 'The Matrix' \rangle, \langle \text{year}, 1999 \rangle\}\}$

je relace nad $\{\text{title}, \text{year}\}$.

Zobrazení tabulkou:

title	year	nebo	year	title
The Avengers	1998		1999	The Matrix
The Avengers	2012		1998	The Avengers
The Matrix	1999		2012	The Avengers

prázdná relace = prázdná množina

Charakteristická vlastnost relace



$\mathcal{D} \dots$ relace nad R

\mathcal{D} ... relace nad R

$V(t)$

- výroková forma
- obor hodnot proměnné t je $\text{Tupl}(R)$
- $t \in \mathcal{D}$, právě když $V(t)$ je pravdivé
- zkráceně: $\mathcal{D} = \{t \in \text{Tupl}(R) \mid V(t)\}$
- **charakteristická vlastnost** \mathcal{D} (zkráceně jen **vlastnost** určující \mathcal{D})

\mathcal{D} ... relace nad R

$V(t)$

- výroková forma
- obor hodnot proměnné t je $\text{Tupl}(R)$
- $t \in \mathcal{D}$, právě když $V(t)$ je pravdivé
- zkráceně: $\mathcal{D} = \{t \in \text{Tupl}(R) \mid V(t)\}$
- **charakteristická vlastnost** \mathcal{D} (zkráceně jen **vlastnost** určující \mathcal{D})

Například: „Vlastním film $t(\text{title})$ vytvořený roku $t(\text{year})$.“

Zkráceně: „Vlastním film title vytvořený roku year .“

title	year
The Avengers	1998
The Avengers	2012
The Matrix	1999

Vždy modelujeme jen část reality.

Například: vlastněné filmy

Vždy modelujeme jen část reality.

Například: vlastněné filmy

Zjednodušení charakteristických vlastností relací.

Vždy modelujeme jen část reality.

Například: vlastněné filmy

Zjednodušení charakteristických vlastností relací.

Například místo:

„Vlastním film `title` vydaný roku `year`.“

Vždy modelujeme jen část reality.

Například: vlastněné filmy

Zjednodušení charakteristických vlastností relací.

Například místo:

`„Vlastním film title vydaný roku year.“`

uvedeme jen:

`„Film title byl vydaný roku year.“`

identifikátor (zkráceně **id**)

- jednoznačně určuje prvek reálného světa
- přirozené číslo, náhodný řetězec, ...
- prvek ztotožňujeme s jeho identifikátorem

identifikátor (zkráceně `id`)

- jednoznačně určuje prvek reálného světa
- přirozené číslo, náhodný řetězec, ...
- prvek ztotožňujeme s jeho identifikátorem

Například: „Film `movie_id` má název `title` a byl vydán roku `year`.“

identifikátor (zkráceně `id`)

- jednoznačně určuje prvek reálného světa
- přirozené číslo, náhodný řetězec, ...
- prvek ztotožňujeme s jeho identifikátorem

Například: „Film `movie_id` má název `title` a byl vydán roku `year`.“

<code>movie_id</code>	<code>title</code>	<code>year</code>
1	The Avengers	1998
2	The Avengers	2012
3	The Matrix	1999

Entity a vztahy



Relace modelují

- entity (například filmy nebo herce)
- vztahy mezi nimi (například role herce ve filmu)

Relace modelují

- entity (například filmy nebo herce)
- vztahy mezi nimi (například role herce ve filmu)

Příklad vztahu mezi entitami:

actor_name	movie_title	movie_year
Keano Reeves	The Matrix	1999
Keano Reeves	Dracula	1992
Laurence Fishburne	The Matrix	1999
Gary Oldman	Dracula	1992
Anthony Hopkins	The Elephant Man	1980

„Zajímá mě herec `actor_name`,
který hrál ve mnou vlastněném filmu
`movie_title` z roku `movie_year`.“

Zkráceně: „Herec `actor_name` hrál ve filmu `movie_title` z roku `movie_year`.“

Relační proměnné



relační proměnná

- proměnná
- hodnota je relace
- typ je relační schéma
- charakteristická vlastnost

Relační proměnné



relační proměnná

- proměnná
- hodnota je relace
- typ je relační schéma
- charakteristická vlastnost

základní relace

- relační proměnná
- hodnotu lze přímo měnit

relační proměnná

- proměnná
- hodnota je relace
- typ je relační schéma
- charakteristická vlastnost

základní relace

- relační proměnná
- hodnotu lze přímo měnit

Například základní relace movie nad $\{\text{title}, \text{year}\}$

Vlastnost: „Film title je z roku year.“

Hodnota:

title	year
The Avengers	1998
The Avengers	2012
The Matrix	1999

1 Úvod

2 Relace a relační proměnné

3 SQL

Odlišnost od relací:

- určují pořadí sloupců
- určují pořadí řádků
- mohou obsahovat duplicitní řádky
- buňka nemusí obsahovat hodnotu
- názvy sloupců nemusí být jedinečné

Odlišnost od relací:

- určují pořadí sloupců
- určují pořadí řádků
- můžou obsahovat duplicitní řádky
- buňka nemusí obsahovat hodnotu
- názvy sloupců nemusí být jedinečné

SQL tabulky chápané jako relace:

- názvy sloupců jsou jedinečné
- neobsahují duplicitní řádky
- nespolehneme se na pořadí řádků
- všechny buňky mají hodnoty

Odlišnost od relací:

- určují pořadí sloupců
- určují pořadí řádků
- můžou obsahovat duplicitní řádky
- buňka nemusí obsahovat hodnotu
- názvy sloupců nemusí být jedinečné

SQL tabulky chápané jako relace:

- názvy sloupců jsou jedinečné
- neobsahují duplicitní řádky
- nespolehneme se na pořadí řádků
- všechny buňky mají hodnoty

Jiné tabulky používat nebudeme.

... pojmenovávají domény atributů

Například: text, integer a mnoho dalších

... pojmenovávají domény atributů

Například: `text`, `integer` a mnoho dalších

Atribut (sloupec) SQL tabulky má doménu určenou skalárním typem.

... pojmenovávají domény atributů

Například: `text`, `integer` a mnoho dalších

Atribut (sloupec) SQL tabulky má doménu určenou skalárním typem.

Omezení: Stejné atributy v různých SQL tabulkách musí být stejného skalárního typu.

... pojmenovávají domény atributů

Například: `text`, `integer` a mnoho dalších

Atribut (sloupec) SQL tabulky má doménu určenou skalárním typem.

Omezení: Stejné atributy v různých SQL tabulkách musí být stejného skalárního typu.

Skalární typ atributů můžeme zobrazit v tabulce:

title text	year integer
The Avengers	1998
The Avengers	2012
The Matrix	1999


```
CREATE TABLE relation (  
    attribute1 scalar_type1,  
    attribute2 scalar_type2,  
    ⋮  
);
```

- deklaruje základní relaci *relation* nad $\{attribute1, attribute2, \dots\}$
- doména atributu *attribute_i* je dána skalárním typem *scalar_type_i*
- hodnotou je prázdná relace

```
CREATE TABLE relation (  
    attribute1 scalar_type1,  
    attribute2 scalar_type2,  
    ⋮  
);
```

- deklaruje základní relaci *relation* nad $\{attribute1, attribute2, \dots\}$
- doména atributu *attribute_i* je dána skalárním typem *scalar_type_i*
- hodnotou je prázdná relace

Příklad:

```
CREATE TABLE movie (  
    title text,  
    year integer,  
);
```


Zjištění hodnoty relační proměnné



Zjištění hodnoty relační proměnné



Příkaz

```
TABLE relation;
```

zobrazí hodnotu relační proměnné *relation*.

Příkaz

```
TABLE relation;
```

zobrazí hodnotu relační proměnné *relation*.

Příklad:

```
# TABLE movie;  
title | year  
-----+-----  
(0 rows)
```

Přidání n -tic do relační proměnné



Přidání n -tic do relační proměnné



- *relation*... relační proměnná nad $R = \{y_1, \dots, y_n\}$
- pořadí atributů y_1, \dots, y_n
- t_1, \dots, t_m jsou n -tice nad R

Přidání n -tic do relační proměnné



- *relation*... relační proměnná nad $R = \{y_1, \dots, y_n\}$
- pořadí atributů y_1, \dots, y_n
- t_1, \dots, t_m jsou n -tice nad R

Příkaz

```
INSERT INTO relation VALUES  
    (value11, ..., value1n),  
    ⋮  
    (valuem1, ..., valuemn);
```

nastaví hodnotu proměnné *relation* na $\mathcal{D} \cup \{t_1, \dots, t_m\}$

- $value_{ij} = t_i(y_j)$
- \mathcal{D} ... hodnota proměnné *relation*
- \mathcal{D} a $\{t_1, \dots, t_m\}$ jsou disjunktní


```
# INSERT INTO movie VALUES  
  ( 'The Matrix', 1999 ),  
  ( 'The Avengers', 2012 ),  
  ( 'The Avengers', 1998 );
```



```
# INSERT INTO movie VALUES
  ( 'The Matrix', 1999 ),
  ( 'The Avengers', 2012 ),
  ( 'The Avengers', 1998 );
```

```
# TABLE movie;
  title      | year
-----+-----
The Matrix   | 1999
The Avengers | 2012
The Avengers | 1998
(3 rows)
```

Další příklad



```
INSERT INTO movie VALUES ('A Space Odyssey', 1968);
```

```
INSERT INTO movie VALUES ('A Space Odyssey', 1968);
```

```
# TABLE movie;
```

title	year
The Matrix	1999
The Avengers	2012
The Avengers	1998
A Space Odyssey	1968

(4 rows)

```
INSERT INTO movie VALUES ('A Space Odyssey', 1968);
```

```
# TABLE movie;
```

title	year
The Matrix	1999
The Avengers	2012
The Avengers	1998
A Space Odyssey	1968

(4 rows)

Připomínka: na pořadí n -tic se nelze spolehnout

Příkaz

```
DROP TABLE relation;
```

zruší základní relaci *relation*

Příkaz

```
DROP TABLE relation;
```

zruší základní relaci *relation*

Příklad:

```
DROP TABLE movie;
```