Úvod do informatiky

přednáška čtvrtá

Miroslav Kolařík

Zpracováno dle učebního textu R. Bělohlávka:
Úvod do informatiky, KMI UPOL, Olomouc 2008.

Obsah

Obsah

Pojem relace je matematickým protějškem běžně používaného pojmu **vztah**. Např. dvě přímky v rovině mohou být ve vztahu "býti rovnoběžné". Dva členové rodiny mohou být ve vztahu "být sourozencem". Číslo 3 je ve vztahu "být menší" s číslem 5 ne však s číslem 2. Tři body v rovině mohou být ve vztahu "ležet na jedné přímce".

Všimněme si, čím je vztah určen. Za prvé je to tzv. **arita** vztahu, tj. číslo udávající počet objektů, které do vztahu vstupují. Za druhé jsou to množiny, jejichž prvky do vztahu vstupují.

Základním pojmem je pojem **uspořádané *n*-tice prvků**. Uspořádaná n-tice objektů $x_1, \ldots, x_n$ (v tomto pořadí) se označuje $\langle x_1, \ldots, x_n \rangle$. Prvek x_i ($1 \leq i \leq n$) se nazývá ***i*-tá složka** dané n-tice. Rovnost definujeme tak, že $\langle x_1, \ldots, x_n \rangle = \langle y_1, \ldots, y_m \rangle$, právě když $n = m$ a $x_1 = y_1, \ldots, x_n = y_n$.

Pojem relace je matematickým protějškem běžně používaného pojmu **vztah**. Např. dvě přímky v rovině mohou být ve vztahu "býti rovnoběžné". Dva členové rodiny mohou být ve vztahu "být sourozencem". Číslo 3 je ve vztahu "být menší" s číslem 5 ne však s číslem 2. Tři body v rovině mohou být ve vztahu "ležet na jedné přímce".

Všimněme si, čím je vztah určen. Za prvé je to tzv. **arita** vztahu, tj. číslo udávající počet objektů, které do vztahu vstupují. Za druhé jsou to množiny, jejichž prvky do vztahu vstupují.

Základním pojmem je pojem **uspořádané *n*-tice prvků.** Uspořádaná n-tice objektů $x_1, \ldots, x_n$ (v tomto pořadí) se označuje $\langle x_1, \ldots, x_n \rangle$. Prvek x_i ($1 \leq i \leq n$) se nazývá ***i*-tá složka** dané n-tice. Rovnost definujeme tak, že $\langle x_1, \ldots, x_n \rangle = \langle y_1, \ldots, y_m \rangle$, právě když $n = m$ a $x_1 = y_1, \ldots, x_n = y_n$.

Definice

Kartézský součin množin $X_1, \ldots, X_n$ je množina $X_1 \times \cdots \times X_n$ definovaná předpisem

$$X_1 \times \cdots \times X_n = \{\langle x_1, \ldots, x_n \rangle | x_1 \in X_1, \ldots, x_n \in X_n\}.$$

Je-li $X_1 = \cdots = X_n = X$ pak $X_1 \times \cdots \times X_n$ značíme také X^n (***n*-tá kartézská mocnina množiny** X).

Například pro množiny $A = \{x, y\}$, $B = \{1, 2, 3\}$ je
$A \times B = \{\langle x, 1\rangle, \langle x, 2\rangle, \langle x, 3\rangle, \langle y, 1\rangle, \langle y, 2\rangle, \langle y, 3\rangle\}$,
$A^2 = \{\langle x, x\rangle, \langle x, y\rangle, \langle y, x\rangle, \langle y, y\rangle\}$,
$B \times A = \{\langle 1, x\rangle, \langle 1, y\rangle, \langle 2, x\rangle, \langle 2, y\rangle, \langle 3, x\rangle, \langle 3, y\rangle\}$.
Všimněme si, že obecně $A \times B \neq B \times A$.

Uspořádanou 1-tici $\langle x \rangle$ obvykle ztotožňujeme s prvkem x (tj. $\langle x \rangle = x$). Potom tedy X^1 je vlastně množina X.

Definice

Kartézský součin množin $X_1, \dots, X_n$ je množina $X_1 \times \cdots \times X_n$ definovaná předpisem

$$X_1 \times \cdots \times X_n = \{\langle x_1, \dots, x_n\rangle | x_1 \in X_1, \dots, x_n \in X_n\}.$$

Je-li $X_1 = \cdots = X_n = X$ pak $X_1 \times \cdots \times X_n$ značíme také X^n (***n*-tá kartézská mocnina množiny** X).

Například pro množiny $A = \{x, y\}$, $B = \{1, 2, 3\}$ je
$A \times B = \{\langle x, 1\rangle, \langle x, 2\rangle, \langle x, 3\rangle, \langle y, 1\rangle, \langle y, 2\rangle, \langle y, 3\rangle\}$,
$A^2 = \{\langle x, x\rangle, \langle x, y\rangle, \langle y, x\rangle, \langle y, y\rangle\}$,
$B \times A = \{\langle 1, x\rangle, \langle 1, y\rangle, \langle 2, x\rangle, \langle 2, y\rangle, \langle 3, x\rangle, \langle 3, y\rangle\}$.
Všimněme si, že obecně $A \times B \neq B \times A$.

Uspořádanou 1-tici $\langle x\rangle$ obvykle ztotožňujeme s prvkem x (tj. $\langle x\rangle = x$). Potom tedy X^1 je vlastně množina X.

Můžeme přistoupit k definici pojmu relace.

Definice

Nechť $X_1, \dots, X_n$ jsou množiny. **Relace** R mezi $X_1, \dots, X_n$ je libovolná podmnožina kartézského součinu $X_1 \times \cdots \times X_n$.

Číslu n říkáme **arita** relace R, R se nazývá n-**ární**. Pro $n = 1, 2, 3, 4$ se místo n-ární používá také unární, binární, ternární, kvaternární. To, že R je unární relace v X vlastně znamená, že $R \subseteq X$.

Můžeme přistoupit k definici pojmu relace.

Definice

Nechť $X_1, \dots, X_n$ jsou množiny. **Relace** R mezi $X_1, \dots, X_n$ je libovolná podmnožina kartézského součinu $X_1 \times \cdots \times X_n$.

Číslu n říkáme **arita** relace R, R se nazývá n-**ární**. Pro $n = 1, 2, 3, 4$ se místo n-ární používá také unární, binární, ternární, kvaternární. To, že R je unární relace v X vlastně znamená, že $R \subseteq X$.

Příklad. V jisté rodině žije s rodiči syn, dcera a babička těchto dětí z matčiny strany. (Mezi těmito osobami existují různé vztahy neboli relace.) Pro konkrétnost si představme, že nejmladší v rodině je dcera, nejstarší babička a otec je starší než matka. Každou n-ární relaci R v rodině $X = \{o, m, d, s, b\}$ můžeme popsat pomocí množiny uspořádaných n-tic osob $\langle x_1, x_2, \dots, x_n \rangle \in R^n$, které jsou v příslušné relaci. Tak například binární relace R: "osoba x_1 je mladší než osoba x_2", je charakterizována množinou uspořádaných dvojic $R = \{\langle d, s\rangle, \langle d, m\rangle, \langle d, o\rangle, \langle d, b\rangle, \langle s, m\rangle, \langle s, o\rangle, \langle s, b\rangle, \langle m, o\rangle, \langle m, b\rangle, \langle o, b\rangle\}$, binární relaci P: "osoba x_1 je stejného pohlaví jako osoba x_2", odpovídá množina uspořádaných dvojic
$P = \{\langle d, d\rangle, \langle d, m\rangle, \langle m, d\rangle, \langle d, b\rangle, \langle b, d\rangle, \langle m, b\rangle, \langle b, m\rangle, \langle m, m\rangle, \langle b, b\rangle, \langle s, s\rangle, \langle s, o\rangle, \langle o, s\rangle, \langle o, o\rangle\}$
a ternární relaci T: "x_1 je dítětem x_2 a x_3", kde x_2 je otec a x_3 je matka, odpovídá množina usp. trojic $T = \{\langle d, o, m\rangle, \langle s, o, m\rangle\}$.

Pojem relace má ústřední roli v tzv. **relačním databázovém modelu**. Tzv. relační pohled na databáze spočívá v tom, že databázi chápeme jako relaci. Například databázi znázorněnou v následující tabulce, která obsahuje v řádcích vybrané

název	rok výroby	barva	cena
⋮	⋮	⋮	⋮
Škoda Favorit	1993	bílá	10 000
Škoda Felicia	1996	modrá	19 000
Škoda Felicia	1997	červená	31 000
Škoda Forman	1993	bílá	5 000
⋮	⋮	⋮	⋮

informace o autech fiktivního autobazaru, můžeme chápat jako 4-ární relaci R mezi množinami (těm se v databázích říká **domény**) $D_1 = \{\dots,$ Škoda Favorit, Škoda Felicia, Škoda Forman, $\dots\}$, $D_2 = \{n \in \mathbb{N} \mid 1900 \leq n \leq 2010\}$, $D_3 = \{\dots,$ bílá, červená, modrá, $\dots\}$, $D_4 = \{z \in \mathbb{Z} \mid 0 \leq z \leq 10000000\}$. Tedy $R \subseteq D_1 \times D_2 \times D_3 \times D_4$.

Relace jsou dány záznamy (řádky) v databázi, takže například $\langle$ Škoda Favorit, 1993, bílá, 10 000 $\rangle \in R$ apod.

V relačních databázích jsou zavedeny i jiné operace než ty, které zavedeme my (například SELECT). Tyto operace slouží k manipulaci a zpřístupňování dat v databázi a lze se s nimi setkat v každé učebnici databázových systémů. Samozřejmě, že lze s databázemi provádět i základní množinové operace: sjednocení, průnik a rozdíl.

Obsah

Relace jsou množiny. Proto s nimi lze provádět množinové operace ($\cap, \cup, \setminus$) a lze na ně aplikovat vztah rovnosti a inkluze ($=, \subseteq$).

S binárními relacemi lze provádět i další operace. Začneme tzv. inverzní relací. **Inverzní relací** k relaci $R \subseteq X \times Y$ je relace R^{-1} mezi Y a X definovaná předpisem:

$$R^{-1} = \{\langle y, x\rangle \mid \langle x, y\rangle \in R\}.$$

Další operací je tzv. skládání. Je-li R relací mezi množinami X a Y a S relací mezi množinami Y a Z, pak **složením relací** R a S je relace $R \circ S$ mezi X a Z definovaná předpisem

$$R \circ S = \{\langle x, z\rangle \mid \exists y \in Y : \langle x, y\rangle \in R \text{ a } \langle y, z\rangle \in S\}.$$

Vztahy a operace s (binárními) relacemi

Relace jsou množiny. Proto s nimi lze provádět množinové operace ($\cap, \cup, \setminus$) a lze na ně aplikovat vztah rovnosti a inkluze ($=, \subseteq$).

S binárními relacemi lze provádět i další operace. Začneme tzv. inverzní relací. **Inverzní relací** k relaci $R \subseteq X \times Y$ je relace R^{-1} mezi Y a X definovaná předpisem:

$$R^{-1} = \{\langle y, x\rangle \mid \langle x, y\rangle \in R\}.$$

Další operací je tzv. skládání. Je-li R relací mezi množinami X a Y a S relací mezi množinami Y a Z, pak **složením relací** R a S je relace $R \circ S$ mezi X a Z definovaná předpisem

$$R \circ S = \{\langle x, z\rangle \mid \exists y \in Y : \langle x, y\rangle \in R \text{ a } \langle y, z\rangle \in S\}.$$

Vztahy a operace s (binárními) relacemi

Relace jsou množiny. Proto s nimi lze provádět množinové operace ($\cap, \cup, \setminus$) a lze na ně aplikovat vztah rovnosti a inkluze ($=, \subseteq$).

S binárními relacemi lze provádět i další operace. Začneme tzv. inverzní relací. **Inverzní relací** k relaci $R \subseteq X \times Y$ je relace R^{-1} mezi Y a X definovaná předpisem:

$$R^{-1} = \{\langle y, x\rangle \mid \langle x, y\rangle \in R\}.$$

Další operací je tzv. skládání. Je-li R relací mezi množinami X a Y a S relací mezi množinami Y a Z, pak **složením relací** R a S je relace $R \circ S$ mezi X a Z definovaná předpisem

$$R \circ S = \{\langle x, z\rangle \mid \exists y \in Y : \langle x, y\rangle \in R \text{ a } \langle y, z\rangle \in S\}.$$

Věta

Pro relace $R \subseteq X \times Y$, $S \subseteq Y \times Z$, $T \subseteq Z \times U$ platí

a) $(R^{-1})^{-1} = R$

b) $(R \circ S)^{-1} = S^{-1} \circ R^{-1}$

c) $R \circ (S \circ T) = (R \circ S) \circ T$.

K důkazu c):

$\langle x, u \rangle \in R \circ (S \circ T) \Leftrightarrow$

$\Leftrightarrow \exists y \in Y$: $\langle x, y \rangle \in R$ a $\langle y, u \rangle \in S \circ T \Leftrightarrow$

$\Leftrightarrow \exists y \in Y$: $\langle x, y \rangle \in R$ a $\exists z \in Z$: $\langle y, z \rangle \in S$ a $\langle z, u \rangle \in T \Leftrightarrow$

$\Leftrightarrow \exists y \in Y$ a $\exists z \in Z$: $\langle x, y \rangle \in R$, $\langle y, z \rangle \in S$, $\langle z, u \rangle \in T \Leftrightarrow$

$\Leftrightarrow \exists z \in Z$: $\langle x, z \rangle \in R \circ S$ a $\langle z, u \rangle \in T \Leftrightarrow$

$\Leftrightarrow \langle x, u \rangle \in (R \circ S) \circ T$.

Věta

Pro relace $R \subseteq X \times Y$, $S \subseteq Y \times Z$, $T \subseteq Z \times U$ platí

a) $(R^{-1})^{-1} = R$

b) $(R \circ S)^{-1} = S^{-1} \circ R^{-1}$

c) $R \circ (S \circ T) = (R \circ S) \circ T$.

K důkazu c):

$\langle x, u \rangle \in R \circ (S \circ T) \Leftrightarrow$

$\Leftrightarrow \exists y \in Y$: $\langle x, y \rangle \in R$ a $\langle y, u \rangle \in S \circ T \Leftrightarrow$

$\Leftrightarrow \exists y \in Y$: $\langle x, y \rangle \in R$ a $\exists z \in Z$: $\langle y, z \rangle \in S$ a $\langle z, u \rangle \in T \Leftrightarrow$

$\Leftrightarrow \exists y \in Y$ a $\exists z \in Z$: $\langle x, y \rangle \in R$, $\langle y, z \rangle \in S$, $\langle z, u \rangle \in T \Leftrightarrow$

$\Leftrightarrow \exists z \in Z$: $\langle x, z \rangle \in R \circ S$ a $\langle z, u \rangle \in T \Leftrightarrow$

$\Leftrightarrow \langle x, u \rangle \in (R \circ S) \circ T$.

Přirozených způsobů jak skládat relace existuje více. Předpokládejme opět, že $R \subseteq X \times Y$, $S \subseteq Y \times Z$. Pak $R \lhd S$, $R \rhd S$ a $R \Box S$ jsou relace mezi X a Z definované předpisy

$$
\begin{aligned}
R \lhd S &= \{\langle x, z\rangle \,|\, \forall y \in Y : \text{ pokud } \langle x, y\rangle \in R, \text{ pak } \langle y, z\rangle \in S\}; \\
R \rhd S &= \{\langle x, z\rangle \,|\, \forall y \in Y : \text{ pokud } \langle y, z\rangle \in S, \text{ pak } \langle x, y\rangle \in R\}; \\
R \Box S &= \{\langle x, z\rangle \,|\, \forall y \in Y : \langle x, y\rangle \in R, \text{ právě když } \langle y, z\rangle \in S\}.
\end{aligned}
$$

Zřejmě relace $\Box$ je podmnožinou relace $\lhd$ i $\rhd$. Platí, že $\Box = \lhd \cap \rhd$.

Obsah

Binární relace lze znázorňovat tabulkami. Např. relace $R = \{\langle 1,\alpha\rangle, \langle 1,\delta\rangle, \langle 2,\beta\rangle, \langle 2,\gamma\rangle, \langle 3,\alpha\rangle\}$ mezi množinami $X = \{1,2,3\}$ a $Y = \{\alpha,\beta,\gamma,\delta,\varepsilon\}$ je znázorněna v následující tabulce:

R	α	β	γ	δ	ε
1	$\times$			$\times$	
2		$\times$	$\times$		
3	$\times$				

Tedy, je-li $\langle x,y\rangle \in R$, je v průsečíku řádku x a sloupce y symbol $\times$, jinak tam není nic.

Poznámka. Chceme-li matematické pojmy zpracovávat v počítači, je třeba je vhodným způsobem reprezentovat (uložit je v paměti počítače tak, aby výpočty s danými pojmy byly rychlé).

Binární relace lze znázorňovat tabulkami. Např. relace $R = \{\langle 1,\alpha\rangle, \langle 1,\delta\rangle, \langle 2,\beta\rangle, \langle 2,\gamma\rangle, \langle 3,\alpha\rangle\}$ mezi množinami $X = \{1,2,3\}$ a $Y = \{\alpha,\beta,\gamma,\delta,\varepsilon\}$ je znázorněna v následující tabulce:

R	α	β	γ	δ	ε
1	$\times$			$\times$	
2		$\times$	$\times$		
3	$\times$				

Tedy, je-li $\langle x,y\rangle \in R$, je v průsečíku řádku x a sloupce y symbol $\times$, jinak tam není nic.

Poznámka. Chceme-li matematické pojmy zpracovávat v počítači, je třeba je vhodným způsobem reprezentovat (uložit je v paměti počítače tak, aby výpočty s danými pojmy byly rychlé).

I. Reprezentace maticí (tabulkou)

Matice typu $m \times n$ je obdélníkové schéma o m řádcích a n sloupcích, ve kterém se na každém místě odpovídajícím nějakému řádku a nějakému sloupci nachází nějaká (nejčastěji číselná) hodnota. Označme takovou matici M. Pro každé $i \in \{1, \dots, m\}$ a $j \in \{1, \dots, n\}$ označme m_{ij} prvek matice z průsečíku řádku i a sloupce j.

Matice (a tabulky) představují základní způsob reprezentace binárních relací. Nechť R je relace mezi množinami $X = \{x_1, \dots, x_m\}$ a $Y = \{y_1, \dots, y_n\}$. Relaci R reprezentujeme maticí (tabulkou), ve které se na místě odpovídajícímu řádku i a sloupci j nachází hodnota, která určuje zda dvojice $\langle x_i, y_j \rangle$ je v relaci R. Obvykle se používá 1 (popř. $\times$) k označení $\langle x_i, y_j \rangle \in R$ a 0 (popř. prázdné místo) k označení $\langle x_i, y_j \rangle \notin R$.

Matice M_R reprezentující relaci $R \subseteq \{x_1, \dots, x_m\} \times \{y_1, \dots, y_n\}$ je definována předpisem

$$m_{ij} = 1 \quad \text{je-li} \quad \langle x_i, y_j \rangle \in R,$$

$$m_{ij} = 0 \quad \text{je-li} \quad \langle x_i, y_j \rangle \notin R.$$

M_R se nazývá **matice relace** R. Naopak také, každá binární matice M typu $m \times n$, tj. matice s hodnotami 0 a 1, reprezentuje relaci mezi $X = \{x_1, \dots, x_m\}$ a $Y = \{y_1, \dots, y_n\}$.

Příklad

Tabulková a maticová reprezentace relace R mezi množinami $X = \{\alpha, \beta, \gamma, \delta\}$ a $Y = \{1, 2, 3\}$, kde $R = \{\langle\alpha, 1\rangle, \langle\alpha, 3\rangle, \langle\beta, 1\rangle, \langle\beta, 2\rangle, \langle\gamma, 3\rangle, \langle\delta, 2\rangle, \langle\delta, 3\rangle\}$:

R	1	2	3
α	$\times$		$\times$
β	$\times$	$\times$	
γ			$\times$
δ		$\times$	$\times$

$$M_R = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

Poznámka. Maticová reprezentace je názorná. Její nevýhodou je velká paměťová složitost. (Např. matice 1000×1000 má milion políček; má-li jen 3000 jedniček, uchovává se zbytečně i 997000 nul. Pro takové případy se používají jiné reprezentace).

Pro binární matice můžeme zavést operace, které odpovídají operacím s relacemi. Mějme binární matice M, N typu $m \times n$ a matici K typu $n \times k$. Definujme následující operace:

$$
\begin{aligned}
&M \vee N = P, \quad p_{ij} = \max\{m_{ij}, n_{ij}\};\\
&M \wedge N = P, \quad p_{ij} = \min\{m_{ij}, n_{ij}\};\\
&M - N = P, \quad p_{ij} = \max\{0, m_{ij} - n_{ij}\};\\
&M \cdot K = P, \quad p_{ij} = \max\{m_{il} \cdot k_{lj};\, l = 1, \ldots, n\};\\
&M^T, \quad m_{ij}^T = m_{ji}.
\end{aligned}
$$

Například operace $\vee$ přiřazuje maticím M a N matici P, jejíž každý prvek p_{ij} je roven maximu z hodnot m_{ij} a n_{ij}.

Operace s relacemi lze provádět pomocí vhodných operací s maticemi.

Věta

Pro relace $R, S \subseteq X \times Y$, $U \subseteq Y \times Z$ je

$$M_{R \cup S} = M_R \vee M_S;$$

$$M_{R \cap S} = M_R \wedge M_S;$$

$$M_{R - S} = M_R - M_S;$$

$$M_{R \circ U} = M_R \cdot M_U;$$

$$M_{R^{-1}} = (M_R)^T.$$

Důkaz je jednoduchý – stačí porovnat definice operací s maticemi a definice operací s relacemi.

II. Reprezentace grafem

Grafy představují další způsob reprezentace binárních relací, který je názorný. Graf binární relace R na množině X dostaneme tak, že každý prvek $x \in X$ znázorníme v rovině jako kroužek s označením daného prvku. Pokud $\langle x, y \rangle \in R$, nakreslíme z kroužku odpovídajícího x do kroužku odpovídajícího y orientovanou čáru (se šipkou).

Poznámka. Graf je jedním ze základních pojmů tzv. diskrétní matematiky. Zde nebyl definován přesně.

III. Reprezentace seznamem seznamů

Další způsob reprezentace pro uložení binární relace R na množině X (pro uložení binární relace v paměti počítače) je reprezentace seznamem seznamů. Tuto reprezentaci tvoří hlavní (spojový) seznam, ve kterém jsou uloženy všechny prvky množiny X. Z každého prvku $x \in X$ hlavního seznamu vede seznam obsahující právě ty $y \in X$, pro které $\langle x, y \rangle \in R$.

Poznámka. Reprezentace seznamem seznamů je paměťově úsporná a je vhodná pro počítačové zpracování.
(Např. u výše zmíněné relace R ($|R| = 3000$) na množině X s 1000 prvky bude mít hlavní seznam 1000 prvků + průměrně 3 políčka pro prvky z něho vedoucí; včetně ukazatelů potřebuje zhruba 2 krát 4000 paměťových buněk; přičemž maticová reprezentace jich potřebovala jeden milion.)

Poznámka. Spojový seznam je jednou ze základních datových struktur. Více o něm viz jakoukoli učebnici algoritmů a datových struktur.

III. Reprezentace seznamem seznamů

Další způsob reprezentace pro uložení binární relace R na množině X (pro uložení binární relace v paměti počítače) je reprezentace seznamem seznamů. Tuto reprezentaci tvoří hlavní (spojový) seznam, ve kterém jsou uloženy všechny prvky množiny X. Z každého prvku $x \in X$ hlavního seznamu vede seznam obsahující právě ty $y \in X$, pro které $\langle x, y \rangle \in R$.

Poznámka. Reprezentace seznamem seznamů je paměťově úsporná a je vhodná pro počítačové zpracování.
(Např. u výše zmíněné relace R ($|R| = 3000$) na množině X s 1000 prvky bude mít hlavní seznam 1000 prvků $+$ průměrně 3 políčka pro prvky z něho vedoucí; včetně ukazatelů potřebuje zhruba 2 krát 4000 paměťových buněk; přičemž maticová reprezentace jich potřebovala jeden milion.)

Poznámka. Spojový seznam je jednou ze základních datových struktur. Více o něm viz jakoukoli učebnici algoritmů a datových struktur.