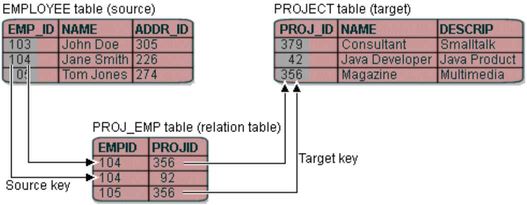
4 Relace a jejich použití

V lekci si podrobně rozebereme matematický aparát relací (a funkcí), kterému se v jeho abstraktní podobě mnoho pozornosti v dřívější výuce nevěnuje – na rozdíl od naivního pohledu na množiny a na „funkce“ ve významu analytických funkcí (jako $x+1$ či $\sin x$). Přitom na pojem relace velmi brzy narazí každý informatik už jen při studiu dat a databází a její abstraktní podobu bude potřebovat.



Stručný přehled lekce

* Co je relace a funkce. Reprezentace relací tabulkou a grafem.
* Základní vlastnosti binárních relací nad množinou.
* Inverze relace, skládání relací a jeho význam.

Zopakování kartézského součinu

Definice: *Kartézský součin* dvou množin A, B definujeme jako množinu všech uspořádaných dvojic ze složek z A a B

$$A \times B = \{(a,b) \mid a \in A, b \in B\}.$$

- Příklady $\{a,b\} \times \{a\} = \{(a,a),(b,a)\}$,
  $\{c,d\} \times \{a,b\} = \{(c,a),(c,b),(d,a),(d,b)\}$. $\square$

Definice *kartézského součinu* více množin: Pro každé $k \in \mathbb{N}$ definujeme

$$A_1 \times A_2 \times \cdots \times A_k = \{(a_1, a_2, \cdots, a_k) \mid a_i \in A_i \text{ pro každé } 1 \le i \le k\}.$$

- Například $\mathbb{Z}^3 = \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z} = \{(i,j,k) \mid i,j,k \in \mathbb{Z}\}$.

4.1 Relace a funkce nad množinami

Vedle množin jsou dalším důležitým základním „datovým typem“ matematiky relace, kterým vzhledem k jejich mnohotvárnému použití v informatice věnujeme významnou pozornost v této i příští lekci.

Definice 4.1. Relace mezi množinami $A_1, A_2, \cdots, A_k$, pro $k \in \mathbb{N}$, je libovolná podmnožina kartézského součinu

$$R \subseteq A_1 \times A_2 \times \cdots \times A_k \,. \Box$$

Pokud $A_1 = A_2 = \cdots = A_k = A$, hovoříme o *$k$-ární relaci R na A*. $\Box$

Příklady relací.

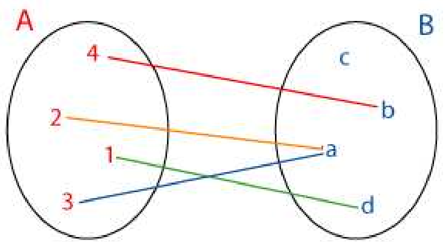
- $\{(1,a),(2,a),(2,b)\}$ je relace mezi $\{1,2,3\}$ a $\{a,b\}$.
- $\{(i, 2\cdot i) \mid i \in \mathbb{N}\}$ je *binární* relace na $\mathbb{N}$. $\Box$
- $\{(i,j,i+j) \mid i,j \in \mathbb{N}\}$ je *ternární* relace na $\mathbb{N}$.
- $\{3\cdot i \mid i \in \mathbb{N}\}$ je *unární* relace na $\mathbb{N}$. $\Box$
- Jaký význam vlastně mají unární a nulární relace na A?

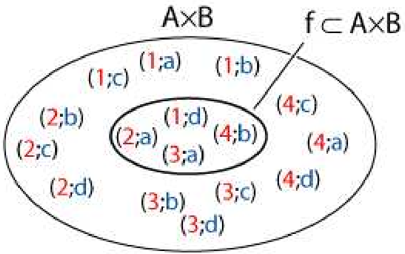
Funkce mezi množinami

Definice 4.2. **(Totální) funkce** z množiny A do množiny B je relace f mezi A a B taková, že pro každé $x \in A$ existuje právě jedno $y \in B$ takové, že $(x, y) \in f$. $\square$

Množina A se nazývá *definiční obor* a množina B *obor hodnot* funkce f.

Neformálně řečeno, ve funkci f je každé „vstupní" hodnotě x přiřazena jednoznačně „výstupní" hodnota y. (V obecné relaci počty „přiřazených" dvojic neomezujeme. . .)





Značení: Místo $(x,y) \in f$ píšeme obvykle $f(x) = y$.

Zápis $f : A \to B$ říká, že f je funkce s def. oborem A a oborem hodnot B. □

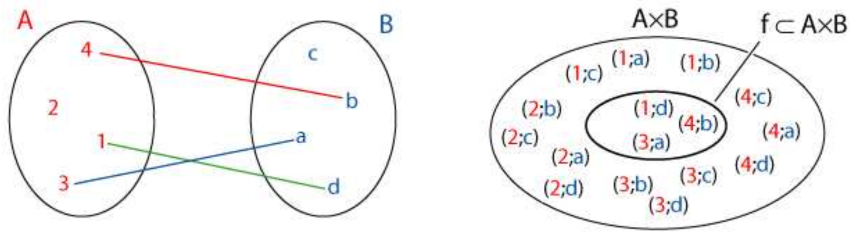
Funkcím se také říká *zobrazení*.

Příklady funkcí jsou třeba následující.

- Definujeme funkci $f : \mathbb{N} \to \mathbb{N}$ předpisem $f(x) = x + 8$.
  Pak $f = \{(x, x+8) \mid x \in \mathbb{N}\}$. □
- Definujeme funkci $plus : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ předpisem $plus(i,j) = i + j$.
  Pak $plus = \{(i, j, i+j) \mid i, j \in \mathbb{N}\}$.

Parciální (částečné) funkce

Definice: Pokud naši Definici 4.2 upravíme tak, že požadujeme pro každé $x \in A$ nejvýše jedno $y \in B$ takové, že $(x, y) \in f$, obdržíme definici *parciální funkce* z A do B. □



V parciální funkci f nemusí být pro některé „vstupní“ hodnoty x funkční hodnota definována (viz například $f(2)$ v uvedeném obrázku).

Pro *nedefinovanou* hodnotu používáme znak $\perp$.

Následuje několik příkladů parciálních funkcí.

- Definujeme parciální funkci $f : \mathbb{Z} \to \mathbb{N}$ předpisem

$$f(x) = \begin{cases} 3 + x & \text{jestliže } x \geq 0, \\ \perp & \text{jinak.} \end{cases}$$

Tj. $f = \{(x, 3 + x) \mid x \in \mathbb{N}\}$. $\square$

- Také funkce $f : \mathbb{R} \to \mathbb{R}$ daná běžným analytickým předpisem

$$f(x) = \log x$$

je jen parciální – není definována pro $x \leq 0$. $\square$

- Co je relace, přiřazující lidem v ČR jejich (česká) rodná čísla?

4.2 Reprezentace konečných relací

Ukládání dat — především sleduje vztahy mezi objekty (stejně jako relace).

$\rightsquigarrow$ relační databáze jako obecná ukázka použití relace.

Příklad 4.3. *Tabulka relační databáze prezentuje obecnou relaci.*

Definujme následující množiny („elementární typy“)

- ZNAK $= \{a, \cdots, z, A, \cdots, Z, mezera\}$,
- CISLICE $= \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. $\square$

Dále definujeme tyto množiny („odvozené typy“)

- JMENO $=$ ZNAK^{15}, PRIJMENI $=$ ZNAK^{20}, VEK $=$ CISLICE^{3},
- ZAMESTNANEC „$\in$“ JMENO $\times$ PRIJMENI $\times$ VEK. $\square$

Relaci „typu“ ZAMESTNANEC pak lze reprezentovat tabulkou:

JMENO	PRIJMENI	VEK
Jan	Novák	42
Petr	Vichr	28
Pavel	Zíma	26
Stanislav	Novotný	52

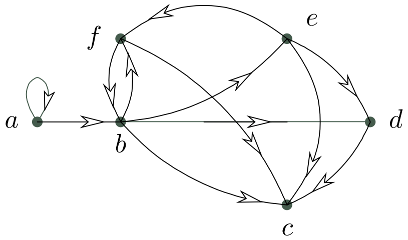
$\square$

Reprezentace binárních relací na množině

Značení: Binární relaci $R \subseteq M \times M$ lze jednoznačně znázornit jejím *grafem*:

- Prvky M znázorníme jako body v rovině.
- Prvek $(a, b) \in R$ znázorníme jako *orientovanou hranu* („šipku") z a do b. Je-li $a = b$, pak je touto hranou „smyčka" na a. □

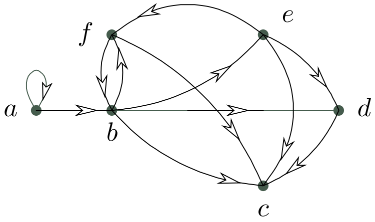
Například mějme $M = \{a, b, c, d, e, f\}$ a $R = \{(a, a), (a, b), (b, c), (b, d), (b, e), (b, f), (d, c), (e, c), (f, c), (e, d), (e, f), (f, b)\}$, pak:



Pozor, nejedná se o „grafy funkcí" známé třeba z matematické analýzy. □

V případě, že M je nekonečná nebo „velká", může být reprezentace R jejím grafem nepraktická (záleží také na míře „pravidelnosti" R).

Značení: Binární relaci $R \subseteq M \times M$ lze jednoznačně zapsat také pomocí *matice* relace – matice $\boldsymbol{A}$ typu $M \times M$ s hodnotami z $\{0, 1\}$.



$$\rightarrow \qquad \begin{array}{c} a \\ b \\ c \\ d \\ e \\ f \end{array} \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{pmatrix} = \boldsymbol{A}$$

A závěrem se můžeme opět vrátit k reprezentaci tabulkou...

JMENO	PRIJMENI
Jan	Novák
Petr	Vichr
Pavel	Zíma
Stanislav	Novotný

4.3 Vlastnosti binárních relací na množině

Definice 4.4. Nechť $R \subseteq M \times M$. Binární relace R je

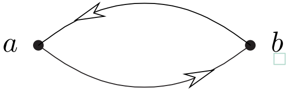
- *reflexivní*, právě když pro každé $a \in M$ platí $(a, a) \in R$;

□

- *ireflexivní*, právě když pro každé $a \in M$ platí $(a, a) \notin R$;

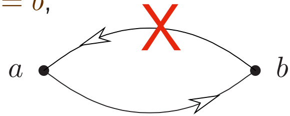
□

- *symetrická*, právě když pro každé $a, b \in M$ platí, že jestliže $(a, b) \in R$, pak také $(b, a) \in R$;

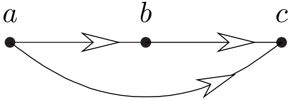


□

- *antisymetrická*, právě když pro každé $a, b \in M$ platí, že jestliže $(a, b), (b, a) \in R$, pak $a = b$;



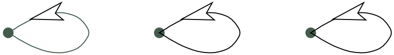
- *tranzitivní*, právě když pro každé $a, b, c \in M$ platí, že jestliže $(a, b), (b, c) \in R$, pak také $(a, c) \in R$.



Následují dva základní typy binárních relací; kde R je

- relace *ekvivalence*, právě když je R reflexivní, symetrická a tranzitivní;
- *částečné uspořádání*, právě když je R reflexivní, antisymetrická a tranzitivní (často říkáme jen *uspořádání*).

Pozor, může být relace symetrická i antisymetrická zároveň? Ano!



Ukázkové binární relace

Příklad 4.5. *Několik příkladů relací definovaných v přirozeném jazyce.*

Nechť M je množina všech studentů 1. ročníku FI. Uvažme postupně relace $R \subseteq M \times M$ definované takto

- $(x,y) \in R$ právě když x a y mají stejné rodné číslo; □
- $(x,y) \in R$ právě když x má stejnou výšku jako y (dejme t. na celé mm); □
- $(x,y) \in R$ právě když výška x a y se neliší více jak o 2 mm; □
- $(x,y) \in R$ právě když x má alespoň takovou výšku jako y; □
- $(x,y) \in R$ právě když x má jinou výšku než y (dejme tomu na celé mm); □
- $(x,y) \in R$ právě když x je zamilován(a) do y. □

Které z nich tedy jsou ekvivalencí nebo uspořádáním? □

Příklad 4.6. *Jaké vlastnosti mají následující relace?*

- Buď $R \subseteq \mathbb{N} \times \mathbb{N}$ definovaná takto $(x, y) \in R$ právě když x dělí y. □ (Částečné uspořádání, ale ne každá dvě čísla jsou porovnatelná.) □
- Buď $R \subseteq \mathbb{N} \times \mathbb{N}$ definovaná takto $(x, y) \in R$ právě když x a y mají stejný zbytek po dělení číslem 5. □(Ekvivalence.) □
- Nechť $F = \{f \mid f : \mathbb{N} \to \mathbb{N}\}$ je množina funkcí. Buď $R \subseteq F \times F$ definovaná takto $(f, g) \in R$ právě když $f(x) < g(x)$ pro všechna x. □(Ireflexivní, antisymetrická a tranzitivní, ale ne reflexivní – není uspořádáním.) □

□

Co v případě, že naše relace některou z poptávaných vlastností nemá, ale nám by se ta vlastnost hodila? To řeší tzv. uzávěry relací: □

- Reflexivní uzávěr přidá do relace všechny dvojice (x, x) nad množinou.
- Symetrický uzávěr zahrne do relace všechny obrácené dvojice k existujícím dvojicím, neboli všechna chybějící (x, y) pokud $(y, x) \in R$. □
- Tranzitivní uzávěr nelze popsat až tak jednoduše, ale stručně řečeno přidává do relace všechny ty dvojice (x, y), pro které se lze (v grafu relace) dostat z x do y „po šipečkách“.

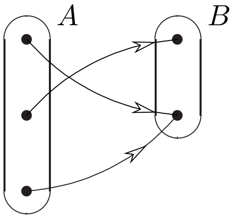
4.4 Inverzní relace a skládání relací

Ve výkladu se nyní vrátíme k obecnému pojetí relací mezi množinami.

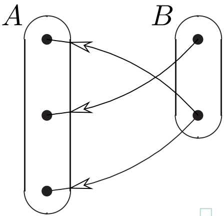
Definice: Nechť $R \subseteq A \times B$ je binární relace mezi A a B. *Inverzní relace* k relaci R se značí R^{-1} a je definována takto:

$$\boldsymbol{R^{-1} = \{(b,a) \mid (a,b) \in R\}}$$

$\boldsymbol{R}$:



$\boldsymbol{R^{-1}}$:



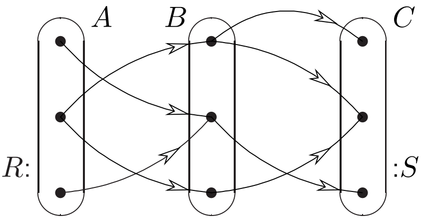
□

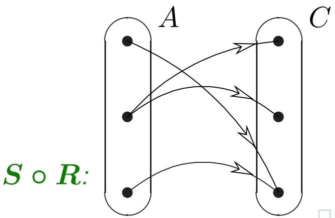
R^{-1} je tedy relace mezi B a A.

Definice skládání

Definice 4.7. **Složení (kompozice)** relací R a S.
Nechť $R \subseteq A \times B$ a $S \subseteq B \times C$ jsou binární relace. *Složení* relací R a S (v tomto pořadí!) je relace $S \circ R \subseteq A \times C$ definovaná takto:

$$\boldsymbol{S \circ R = \{(a,c) \mid \text{existuje } b \in B \text{ takové, že } (a,b) \in R, (b,c) \in S\}}$$





Složení relací čteme „R *složeno s* S" nebo (pozor na pořadí!) „S *po* R".

Několik matematických příkladů skládání relací následuje zde.

- Je-li
  - $A = \{a, b\}, \quad B = \{1, 2\}, \quad C = \{X, Y\}$,
  - $R = \{(a, 1), (b, 1), (b, 2)\}, \quad S = \{(1, X)\}$,

  pak složením vznikne relace
  - $S \circ R = \{(a, X), (b, X)\}$. □
- Složením funkcí $h(x) = x^2$ a $f(x) = x + 1$ na $\mathbb{R}$ vznikne funkce

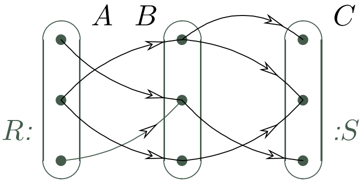
$$(f \circ h)(x) = f(h(x)) \;=\; x^2 + 1.\square$$

- Složením těchže funkcí „naopak“ ale vznikne funkce

$$(h \circ f)(x) = h(f(x)) \;=\; (x + 1)^2.\square$$

Poznámka: Nepříjemné je, že v některých oblastech matematiky (například v algebře při skládání zobrazení) se setkáme s právě opačným zápisem skládání, kdy se místo $S \circ R$ píše $R \cdot S$ nebo jen RS.

Tvrzení 4.8. *Nechť* $R \subseteq A \times B$ *a* $S \subseteq B \times C$ *jsou binární relace. Pak inverzí složené relace* $S \circ R$ *je relace*



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

4.5 Skládání relací „v praxi“

Příklad 4.9. *Skládání v relační databázi studentů, jejich předmětů a fakult.*

Mějme dvě binární relace – jednu R přiř. studentům MU kódy jejich zapsaných předmětů, druhou S přiř. kódy předmětů jejich mateřským fakultám.

R :

student (učo)	předmět (kód)
121334	MA010
133935	M4135
133935	IA102
155878	M1050
155878	IB000

S :

předmět (kód)	fakulta MU
MA010	FI
IB000	FI
IA102	FI
M1050	PřF
M4135	PřF

Jak z těchto „tabulkových“ relací zjistíme, kteří studenti mají zapsané předměty na kterých fakultách (třeba na FI)?

Jedná se jednoduše o složení relací $S \circ R$. V našem příkladě třeba:

$S \circ R$:

student (učo)	fakulta MU
121334	FI
133935	FI
133935	PřF
155878	FI
155878	PřF

□

Zobecněné skládání relací

Definice: (**skládání relací vyšší arity**): Mějme relace $T \subseteq K_1 \times K_2 \times \cdots \times K_k$ a $U \subseteq L_1 \times L_2 \times \cdots \times L_\ell$, přičemž pro nějaké $m < \min(k,\ell)$ platí $L_1 = K_{k-m+1}, L_2 = K_{k-m+2}, \dots, L_m = K_k$. Pak relaci T lze *složit* s relací U na zvolených m složkách $L_1, \dots, L_m$ („překrytí“) s použitím Definice 4.7 takto:

- Položme $A = K_1 \times \cdots \times K_{k-m}$, $B = L_1 \times \cdots \times L_m$ a $C = L_{m+1} \times \cdots \times L_\ell$.
- Příslušné relace pak jsou
  - $R = \{(\vec{a}, \vec{b}) \in A \times B \mid (a_1, \dots a_{k-m}, b_1, \dots b_m) \in T\}$ a
  - $S = \{(\vec{b}, \vec{c}) \in B \times C \mid (b_1, \dots b_m, c_{m+1}, \dots c_\ell) \in U\}$.
- Nakonec přirozeně položme $U \circ_m T \simeq S \circ R$, takže vyjde $U \circ_m T =$

$$\left\{ (\vec{a}, \vec{c}) \mid \text{ ex. } \vec{b} \in B, \text{ že } (a_1, \dots a_{k-m}, b_1, \dots b_m) \in T \text{ a } (b_1, \dots b_m, c_{m+1}, \dots c_\ell) \in U \right\}.$$

Schematicky pro snadnější orientaci:

$$\begin{array}{rllll}
T \subseteq & K_1 \times \cdots \times K_{k-m} \times & K_{k-m+1} \times \cdots \times K_k & & \\
U \subseteq & & L_1 \times \cdots \times L_m & \times L_{m+1} \times \cdots \times L_\ell & \\
U \circ_m T \subseteq & \underbrace{K_1 \times \cdots \times K_{k-m}}_{A} \times & \underbrace{\qquad\qquad\qquad}_{B} & \times \underbrace{L_{m+1} \times \cdots \times L_\ell}_{C} &
\end{array}$$

Příklad 4.10. *Skládání v relační databázi pasažérů a letů u let. společností.*

Podívejme se na příklad hypotetické rezervace letů pro cestující, relace T. Jak známo (tzv. codeshare), letecké společnosti si mezi sebou „dělí“ místa v letadlech, takže různé lety (podle kódů) jsou ve skutečnosti realizovány stejným letadlem jedné ze společností. To zase ukazuje relace U.

T :

pasažér	datum	let
Petr	5.11.	OK535
Pavel	6.11.	OK535
Jan	5.11.	AF2378
Josef	5.11.	DL5457
Alena	6.11.	AF2378

U :

datum	let	letadlo
5.11.	OK535	ČSA
5.11.	AF2378	ČSA
5.11.	DL5457	ČSA
6.11.	OK535	AirFrance
6.11.	AF2378	AirFrance

Ptáme-li se nyní, setkají se Petr a Josef na palubě stejného letadla? Případně, čí letadlo to bude? Odpovědi nám dá složení relací $U \circ_2 T$, jak je popsáno výše.

$U \circ_2 T$:

pasažér	letadlo
Petr	ČSA
Josef	ČSA
Pavel	AirFrance
...	...

□