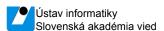
# Teoretické základy informatických vied Základné pojmy

Mgr. Martin Bobák, PhD.



2022/2023

Pôvodný autor: doc. Mgr. Daniela Chudá, PhD. Teoretické základy informatických vied, FIIT STU, 2020.

### Základné informácia o predmete

Typ štúdia: bakalárske

Ročník: 1. (INFO-3r) a 2. (INFO-4r) Semester: Letný Trvanie: 12 týždňov

Počet hodín týždenne (prednášky - cvičenia): 2 - 2

Prednášajúci:

- Mgr. Martin Bobák, PhD., martin.bobak@stuba.sk, ÚI SAV
- Ing. Juraj Petrík, juraj.petrik@stuba.sk, 3.31

#### Cvičiaci:

- Mgr. Martin Bobák, PhD., martin.bobak@stuba.sk, Úl SAV
- Ing. Ivan Kapustík, ivan.kapustik@stuba.sk, 4.44
- Ing. Jakub Perdek, jakub.perdek@stuba.sk
- Ing. Juraj Petrík, juraj petrik@stuba.sk, 3.31
- Ing. Igor Stupavský, igor.stupavsky@stuba.sk, 3.31
- Ing. Juraj Vincúr, juraj.vincur@stuba.sk, 3.38-2 (3D LAB)

### Základné informácia o predmete

- Teória formálnych jazykov a automatov.
- Teória vyčísliteľ nosti a zložitosti.
- Cieľom predmetu je získať vedomosti o Chomského hierarchii jazykov a jej vzťahu k abstraktným výpočtovým modelom.
- Prehĺbiť a rozvíjať abstraktné logické myslenie a podnietiť schopnosť samostatného riešenia exaktne formulovaných úloh a problémov.
- Získať zručnosti v konštruovaní umelých gramatík, abstraktných automatov, Turingových a počítadlových strojov.

# Sylabus predmetu

- Teória množín
- Gramatiky a jazyky
- Konečné automaty
- Zásobníkové automaty
- Turingove stroje
- Vypočítateľ nosť

### Harmonogram

Týždeň	Prednáška	Cvičenie	Iné			
1	Množiny, Jazyky	Množiny, Jazyky				
2	Konečné automaty	Množiny, Jazyky				
3	Regulárne gramatiky, Regulárne výrazy	Konečné automaty				
4	Zásobníkové automaty	Regulárne gramatiky, Regulárne výrazy	TEST 1			
5	Bezkontextové gramatiky, Uzáverové vlastnosti	Konečné automaty, Zásobníkové automaty				
6	LOA, Turingové stroje, Kontextové gramatiky	Zásobníkové automaty, Bezkontextové gramatiky				
7	Vypočítateľnosť, UTS, Syntaktická analýza, Uzáverové vlastnosti	Turingové stroje, LOA	TEST 2 Zadávanie projektov TS a RAM			
8	RAM, zložitosti	T vypočítateľnosť, Kontextové gramatiky				
9	Počítadlové stroje	RAM				
10	Ekvivalencie výpočtových modelov	RAM, AM				
11	Využiteľnosť automatov a gramatík, Vypočítateľnosť	AM, projekty				
12	Kryptografia, Aplikácie, Zhrnutie	projekty	TEST 3			
	M. Bob	ák Teoretické základy ir	formatických vied			

#### Rozvrh

Deň	8.00-8.50	9.00-9.50	10.00-10.50	11.00-11.50	12.00-12.50	13.00-13.50	14.00-14.50	15.00-15.50	16.00-16.50	17.00-17.50	18.00-18.50	19.00-19.50
Po		Teoretic informatic J.	SA-FIIT-FIIT) / 2 cké základy ckých vied <sup>(1)</sup> Petrík ) (BA-FIIT-FIIT)	Teoretické základ	j (1) Dustík				1.38 (U20b) ( Teoretické základ vied J. Pe 1.39 (U20a) (	y informatických <sub>I</sub> (1) etrík	1.38 (U20b) ( Teoretické základ vied J. Pe 1.39 (U20a) (	y informatických (1,2) etrík
			cké základy		ty informatických				Teoretické základ		Teoretické základ	
			kých vied (1,4)		(1)				vied		vied	(1,5)
			(apustík	J. Pe					I. Stup	avský	I. Stup	avský
			SA-FIIT-FIIT) / 3	1.40 (U40) (I								
			cké základy ckých vied <sup>(1)</sup>	reoreticke zaklad vied	ly informatických							
			Bobák	M. B								
Ut				Teoretické základ	na) (BA-FIIT-FIIT) dy informatických j (7) obák							
St												
št		Teoretic informa	(BA-FIIT-FIIT) cké základy tických vied Perdek	1.37 (LOS) (I Teoretické záklac viec J. Pe	ly informatických j (8)	1.37 (LOS) (I Teoretické záklac viec I. Kaj	y informatických (9)	Teoretické základ vie I. Kar	ed	Teoretické záklac vi I Stu	ed	
								1.39 (U20a) ( Teoretické základ	BA-FIIT-FIIT)	1.39 (U20a) ( Teoretické základ	BA-FIIT-FIIT)	
								vied		vied		
								J. Vi.	ncúr	J. Vi	ncúr	
Pi												

### Podmienky absolvovania predmetu TZIV

- Priebežné hodnotenie počas semestra 60 bodov (minimálne 30 bodov)
  - Absolvovanie 3 priebežných písomiek z príkladov a teórie - 35 bodov (podľa harmonogramu, píšu sa na cvičeniach)
    - Pri ospravedlnenej neúčasti na písomke si študent písomku napíše na najbližšom cvičení alebo na konci semestra (podľa rozhodnutia cvičiaceho).
    - Je možné si písomku napísať aj na inom cvičení (v prípade nahradenia si cvičenia)
    - Povinné odovzdanie a odprezentovanie 2 praktických zadaní TS a RAM - 15 bodov
    - Domáce úlohy a úlohy na cvičeniach 10 bodov (minimálne 3 bodov)
      - Štandardne sa odovzdávajú do pripraveného miesta odovzdania
    - Aktivita na cvičeniach 5 bodov (bonusové body)

Skuska (minimum 50%)	40
Test 1	8
Test 2	12
Test 3	15
Projekt TS	10
Projekt RAM	5
Aktivita (bonus)	5
DU a body za cvicenia (minimum 30%)	10
Spolu	105

### Podmienky absolvovania predmetu TZIV

- Vykonanie záverečnej skúšky 40 bodov (minimálne 20 bodov)
- Získanie aspoň 56 percent z celkového hodnotenia (minimálne 56 bodov)
- Stupnica hodnotenia je štandardná
- Účasť na cvičeniach je povinná (t. j. 0 neospravedlnených absencií)

#### Akademická bezúhonnosť

- Odpisovanie je vedomé prezentovanie cudzej práce ako svoj vlastný výsledok. V tomto predmete sa nebude plagiát (odpísaná pasáž z diela iného autora, časť programu) tolerovať. Typickým príkladom plagiátu je použitie (častí) práce niekoho iného bez jej citovania. Autori projektu sú preto povinní uviesť v dokumentácií všetky zdroje informácií, ktoré použili pri vypracovaní projektu.
- Nedodržanie sa podľa Študijného a skúšobného poriadku bakalárskeho a inžinierskeho štúdia na FIIT STU posudzuje a rieši pred disciplinárnou komisiou.

# Študijná literatúra

- J. E. Hopcroft, R. Motwani, J. D. Ullman: Introduction to Automata Theory, Languages, and Computation. Addison-Wesley, tretie vydanie, 2007.
- P. Linz: An introduction to formal languages and automata. Jones & Bartlett Learning, piate vydanie, 2012.
- M. Sipser: Introduction to the Theory of Computation. Cengage Learning, 2012
- J.E. Hopcroft J.D. Ullman: Formálne jazyky a automaty. Alfa 1969.
- L'. Molnár a kol.: Gramatiky a jazyky. Alfa 1987.
- M. Nehéz, D. Chudá, I. Polický, M. Čerňanský: Teoretické základy informatiky
- Materiály v AIS

#### Definícia (Naivná definícia množiny)

Množina je súbor prvkov, ktoré majú spoločnú vlastnosť.

#### Zápis množiny:

- vymenovaním prvkov:  $A = \{a, b, c\}$
- charakterizáciou vlastností jej prvkov:

$$\forall x (x \in \{x | \varphi[x]\} \leftrightarrow \varphi[x])$$

Nech  $\varphi[x]$  je výroková formula, potom objekt x je prvkom množiny, ak má vlastnosť  $\varphi$  (napr.  $Parne = \{2n | n \in \mathbb{N}\}$ ).

 prázdna množina: (označenie ∅, {}) je množina, do ktorej nepatrí žiadny prvok (∀x : x ∉ ∅).
 Existuje práve jedna prázdna množina.

Russellov paradox: Množina všetkých množín, ktoré nie sú prvkami seba samých.

$$M = \{X | X \notin X\}$$

Čo platí pre množinu M?

- M ∈ M: Množina všetkých množín X, pre ktorú platí X ∉ X.
   Potom musí platiť, že M ∉ M SPOR.
- M ∉ M: Množina všetkých množín X, pre ktorú platí X ∉ X.
   Potom musí platiť, že M ∈ M SPOR.

Paradox holiča: V malom meste je jediný holič, ktorý holí práve tých mužov v meste, ktorí sa neholia sami. Také mesto však nemôže existovať, lebo tu opäť dochádza ku sporu: Holí holič sám seba? Sám seba má holiť práve vtedy, kedy sám seba holiť nebude.

Naivná teória množín je nekorektná (obsahuje viacero paradoxov). To viedlo k tomu, že teória množín sa zadefinovala axiomaticky (E. Zermelo a A. Fraenkel).

#### Definícia (Axiomatická definícia množiny)

Množina je matematický objekt, ktorý spĺňa axiómy teórie množín.

- Axióma existencie množín Existuje aspoň jedna množina.
- Axióma extenzionality Množiny, ktoré majú tie isté prvky, sa rovnajú.
- Schéma axióm vydelenia Z každej množiny je možné vydeliť množinu všetkých prvkov, ktoré spĺňajú danú vlastnosť.
- Axióma dvojice Ľubovoľné dve množiny určujú dvojprvkovú množinu.

- Axióma sumy

  Ku každej množine A je daná

  množina všetkých prvkov, ktoré
  patria do nejakého prvku množiny A.
- Axióma potencie Ku každej množine je daná množina všetkých podmnožín.
- Schéma axióm nahradenia Definovateľné zobrazenie zobrazuje množinu na množinu.
- Axióma nekonečna
   Existuje nekonečná množina.
- Axióma fundovanosti

#### Poznámky:

- "Všetko je množina."
- axiómy hovoria, ako sa množiny správajú t.j. charakterizujú ich vlastnosti (podobne ako pri teórií grúp, vektorových priestoroch . . . )
  - Formálne sú vlastnosti objektov vyjadrené pomocou predikátov.

#### Vlastnosti:

- Pre samotnú ZF teóriu sú postačujúce axiómy č. 2, 5, 6, 7, 8 a
   9, pričom ostatné axiómy sa z týchto dajú odvodiť.
- Všimnite si, že axióma č. 1 priamo vyplýva z axiómy č. 8.

#### Definícia (Definícia podmnožiny)

Množina A je **podmno**ž**inou** množiny B (označenie  $A \subseteq B$ ), ak každý prvok množiny A je aj prvkom množiny B.

$$A \subseteq B \iff \forall x[(x \in A) \Rightarrow (x \in B)]$$

Množina A je vlastnou podmnožinou množiny B, (označenie  $A \subset B$ ), ak platí:  $A \subseteq B \land A \neq B$ 

#### Poznámky:

- Platí, že každá množina je podmnožinou samej seba.
- Vzťah označovaný symbolom ⊆ sa nazýva inklúzia.
- Vzťah označovaný symbolom ⊂ sa nazýva vlastná inklúzia.
- $A = B \iff A \subseteq B \land B \subseteq A$
- Prázdna množina je podmnožinou ľubovoľnej množiny

#### Definícia (Definícia prieniku dvoch množín)

Prienikom množín A a B nazývame množinu

$$A \cap B = \{x | x \in A \land x \in B\}$$

Množiny C, D sa nazývajú disjunktné, ak  $C \cap D = \emptyset$ .

#### Definícia (Definícia rozdielu dvoch množín)

Rozdielom množín A a B nazývame množinu

$$A \setminus B = \{x | x \in A \land x \notin B\}$$

#### Definícia (Definícia zjednotenia dvoch množín)

Zjednotením množín A a B nazývame množinu

$$A \cup B = \{x | x \in A \lor x \in B\}$$

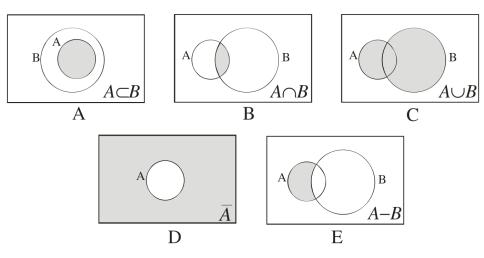
#### Definícia (Definícia zjednotenia systému množín)

Majme množinu I indexov systému množím  $\{A_i|i\in I\}=\{A_{i_1},\ldots,A_{i_k}\}$ . Potom jeho zjednotenie nazývame množinu

$$\bigcup_{i=1}^k A_i = \{x | \exists i \in I \land x \in A_i\}$$

- **1** Komutatívny zákon  $A \cup B = B \cup A$   $A \cap B = B \cap A$
- **2** Asociatívny zákon  $(A \cup B) \cup C = A \cup (B \cup C)$   $(A \cap B) \cap C = A \cap (B \cap C)$
- **3** Distributívny zákon  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$   $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- $\begin{array}{c}
  \mathbf{0} & \mathbf{Identita} \\
  A \cup \emptyset = A \\
  A \cap U = A
  \end{array}$

- **3** Komplement  $A^{C} = U \setminus A$   $A \cup A^{C} = U$   $A \cap A^{C} = \emptyset$   $(A^{C})^{C} = A$   $\emptyset^{C} = U$   $U^{C} = \emptyset$
- De Morganove zákony  $(A \cup B)^C = A^C \cap B^C$  $(A \cap B)^C = A^C \cup B^C$



Zdroj: Vladimír Kvasnička, Jiří Pospíchal: Algebra a diskrétna matematika, STU, 2008.

#### Definícia (Definícia karteziánskeho súčinu množín)

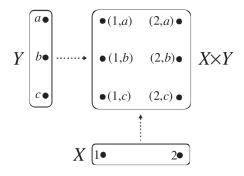
Nech A, B sú ľubovoľné množiny. Karteziánskym súčinom množín A a B nazývame množinu

$$A \times B = \{(a,b)|a \in A \land b \in B\}$$

Analogickým spôsobom možno definovať karteziánsky súčin k množín - jeho prvkami sú usporiadané k-tice.

#### Poznámky:

- Prvkami množiny A × B sú usporiadané dvojice prvkov, ktoré budeme označovať (a, b).
- Ak  $a \neq b$ , potom  $(a, b) \neq (b, a)$ .
- Neusporiadané dvojice prvkov a, b budeme označovať {a, b}.



Zdroj: Vladimír Kvasnička, Jiří Pospíchal:Algebra a diskrétna matematika, STU, 2008.

#### Definícia (Definícia binárnej relácie)

Nech sú A, B ľubovoľné množiny. Binárnou reláciou R z množiny A do množiny B nazveme ľubovoľnú podmnožinu karteziánskeho súčinu  $A \times B$ .

$$R \subset A \times B$$

#### Poznámky:

- Prvkami relácie sú usporiadané dvojice.
- Ak  $(a, b) \in R$ , tak používame tiež zápis aRb. Najdôležitejšou operáciou na binárnych reláciách je kompozícia relácií.

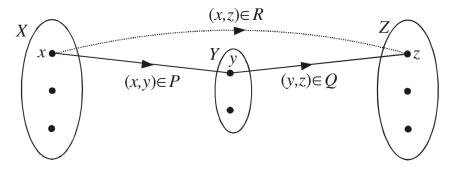
#### Definícia (Definícia kompozície relácií)

Nech  $R_1$  a  $R_2$  sú binárne relácie. Kompozícia (binárnych) relácií  $R_1 \circ R_2$  je definovaná nasledovne:

$$R_1 \circ R_2 = \{(x,z) | \exists y : (x,y) \in R_1 \land (y,z) \in R_2\}$$

#### Poznámky:

- Ak R ⊆ A × A pre nejakú množinu A, tak hovoríme, že R je relácia na množine A.
- Identická relácia  $I_A = \{(a,a)|a \in A\}$  je príkladom relácie na množine.
- Pre reláciu R na množine A poznáme napríklad operácie mocniny, reflexívneho a tranzitívneho uzáveru.



Zdroj: Vladimír Kvasnička, Jiří Pospíchal: Algebra a diskrétna matematika, STU, 2008.

#### Definícia (Definícia zobrazenia)

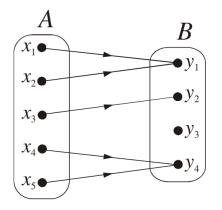
Binárna relácia  $f \subset A \times B$  sa nazýva zobrazenie (funkcia) z množiny A do množiny B, ak platí:

$$\forall x \in A, \forall y_1 \in B, \forall y_2 \in B : \\ [(x,y_1) \in f \land (x,y_2) \in f \rightarrow y_1 = y_2] \ (jednoznačná) \\ \forall x \in A, \exists y \in B : (x,y) \in f \ (všade \ definovaná)$$

- Ak  $f \subseteq X \times Y$  je zobrazenie, tak ho zapisujeme ako  $f \cdot X \rightarrow Y$
- Ak  $(x, y) \in f$  pre  $x \in X, y \in Y$ , tak to zapisujeme ako f(x) = f(x)у.

Zobrazenie  $f: X \to Y$  sa nazýva:

- injektívne, ak  $\forall x_1, x_2 \in X, x_1 \neq x_2$  potom platí, že  $f(x_1) \neq f(x_2)$
- surjektívne, ak  $\forall y \in Y, \exists x \in X$  potom platí, že f(x) = y
- bijektívne, ak je injektívne a zároveň surjektívne.



Zdroj: Vladimír Kvasnička, Jiří Pospíchal: Algebra a diskrétna matematika, STU, 2008.

#### Definícia (Definícia kompozície zobrazení)

 $Ak\ f:A\to B\ a\ g:B\to C$ , potom pre ich kompozíciu platí  $f\circ g:A\to C$ . Formálne pre l'ubovolné prvky  $x\in A\ a\ z\in C$  platí:

$$(f \circ g)(x) = z \iff \exists y \in B(f(x) = y \land g(y) = z) \iff gf(x) = z$$

Je to dôsledok toho, že jednoznačné a všade definované relácie sú uzavreté na operáciu kompozície.

#### Poznámky:

- $f \circ g = \{(x, z) \in A \times C | gf(x) = z\}$
- $\forall x \in A : (f \circ g)(x) = gf(x)$
- Všimnite si zápisy zloženého zobrazenia gf(x) a  $(f \circ g)(x)$ .
- Kompozícia dvoch injektívnych zobrazení je injektívne, dvoch surjektívnych zobrazení je surjektívne a dvoch bijektívnych zobrazení je bijektívne.

#### Definícia (Definícia potenčnej množiny)

Nech A je množina. Potenčná množina  $\mathcal{P}(A)$  je množina všetkých podmnožín množiny A:

$$\mathcal{P}(A) = \{X | X \subseteq A\}$$

Príklad:

$$A = \{a, b, c\}$$

$$P(A) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}\}$$

#### Poznámky:

ullet zvykne sa označovať dvoma spôsobmi:  $\mathcal{P}(A)$ , alebo  $2^A$ 

#### Definícia (Definícia množiny prirodzených čísiel)

Množina prirodzených čísel  $\mathbb N$  je najmenšia množina (vzhľadom na inklúziu), pre ktorú platí:

- $\bullet$   $0 \in \mathbb{N}$
- $ak \ x \in \mathbb{N}$ ,  $tak \ (x+1) \in \mathbb{N}$ .

(zjednodušenie Peanových axióm)

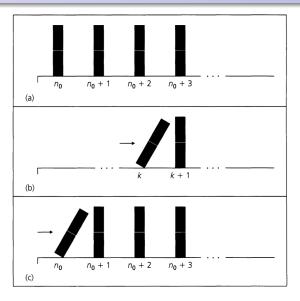
#### Veta (Veta o matematickej indukcii)

Nech P(x) je predikát vyjadrujúci nejakú vlastnosť čísel. Nech  $k \in \mathbb{N}$ . Predpokladajme, že platí:

- P(0),
- ak P(k), tak P(k + 1).

Potom pre všetky  $n \in \mathbb{N}$  platí P(n), teda vlastnosť P majú všetky prirodzené čísla.

$$\forall P \Big( P(0) \land \forall k \big( P(k) \rightarrow P(k+1) \big) \rightarrow \forall n \big( P(n) \big) \Big)$$



Zdroj: Grimaldi, R. P.: Discrete and Combinatorial Mathematics, Pearson Addison-Wesley, 2004.

Množina kladných celých čísel  $\mathbb{N}^+$ :

$$\mathbb{N}^+ = \mathbb{N} \setminus \{0\}$$

Množina (nezáporných) párnych čísel  $\mathbb{E}v$  (Even):

$$\mathbb{E}v = \{2 \cdot k | k \in \mathbb{N}\}\$$

Množina (kladných) nepárnych čísel  $\mathbb{O}dd$ :

$$\mathbb{O}dd = \{2 \cdot k + 1 | k \in \mathbb{N}\}\$$

Množina celých čísel  $\mathbb{Z}$ :

$$\mathbb{Z} = \mathbb{N} \cup \{-k | k \in \mathbb{N}\}\$$

Množina všetkých prvočísel  $\mathbb{P}$ r

Je to množina všetkých takých čísiel z  $\mathbb{N}^+$  , ktoré majú iba dva rôzne kladné delitele.

Množina racionálnych čísiel Q

Každé racionálne číslo sa dá zapísať v tvare zlomku, teda v tvare  $\frac{p}{q}$ , pričom  $p \in \mathbb{Z}, q \in \mathbb{N}^+$ . Avšak rôzne zlomky môžu vyjadrovať to isté číslo (napr.  $\frac{1}{5} = \frac{3}{15}$ ). Preto sa množina  $\mathbb{Q}$  definuje ako množina všetkých zlomkov v tzv. vykrátenom (resp. kanonickom) tvare.

Množina reálnych čísiel  $\mathbb R$ 

Množina reálnych čísiel vznikne tzv. zaplnením množiny čísiel. Okrem všetkých čísiel z  $\mathbb Q$  obsahuje aj ďalšie čísla, ako napr.  $\sqrt{2}$ ,  $\sqrt{3}$ ,  $\pi$  a iné.

### Mohutnosť množiny - Opakovanie

- mohutnosť (kardinalita) množiny A vyjadruje veľkosť (=počet prvkov) množiny A
- označuje sa symbolom |A|

#### Definícia (Definícia rovnosti mohutnosti dvoch množín)

Nech A, B sú množiny. Budeme hovoriť, že množiny A a B majú rovnakú mohutnosť, ak existuje bijektívne zobrazenie  $\varphi:A\to B$  Rovnosť dvoch množín budeme označovať ako |A|=|B|.

# Definícia (Definícia menšej alebo rovnakej mohutnosti dvoch množín)

Nech A, B sú množiny. Budeme hovoriť, že množina A má mohutnosť menšiu alebo rovnakú ako množina B, označujeme  $|A| \leq |B|$ , ak existuje injektívne zobrazenie  $\psi: A \to B$ . Množina A má mohutnosť menšiu ako množina B, označujeme |A| < |B|, ak  $|A| \leq |B|$  a neplatí |A| = |B|.

#### Definícia (Definícia konečnej množiny)

Množina A sa nazýva konečná, ak  $|A| < |\mathbb{N}|$ .

#### Vlastnosti:

- mohutnosť konečnej množiny je možné vyjadriť prirodzeným číslom. Nech  $k \in \mathbb{N}$ . Mohutnosť konečnej množiny A označíme číslom k, píšeme |A|=k, ak  $|A|=|\{1,\,2,\,\ldots,\,k\}|$ . Mohutnosť prázdnej množiny sa označuje nulou.
- $\bullet |A \times B| = |A| \cdot |B|$
- $|A \cup B| \le |A| + |B|$
- $|A \cap B| \leq \min(|A|, |B|)$
- $|A \setminus B| \leq |A|$
- $|P(A)| = 2^{|A|}$

- ako spočítať nekonečnú množinu?
- je prirodzených čísiel viacej ako celých čísiel, reálnych čísiel...?
- kardinálne číslo symbol, ktorým sa označuje mohutnosť množiny.
  - konečné množiny prirodzené číslo
  - nekonečné množiny nekonečné kardinály  $(|\mathbb{N}|=\aleph_0)$

- kardinalita definovaná ako relácia namiesto čísla
- |A| = |B| práve vtedy keď existuje **bijekcia**  $f : A \rightarrow B$ .

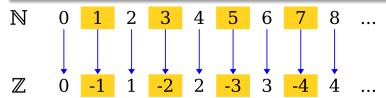
#### Vlastnosti:

- |A| = |A|. Definujme f ako f(x) = x
- Ak platí, že |A| = |B|, potom |B| = |A|.
- Ak platí, že |A| = |B| a zároveň |B| = |C|, potom |A| = |C|.

#### Veta

Mohutnosť množiny prirodzených čísel je rovnaká ako mohutnosť množiny celých čísel, teda platí nasledujúca rovnosť:

$$|\mathbb{N}| = |\mathbb{Z}|$$



Definujme  $f: \mathbb{N} \to \mathbb{Z}$ 

$$f(n) = \begin{cases} \frac{n}{2} & n \in \mathbb{E}v \\ -\frac{n+1}{2} & n \in \mathbb{O}dd \end{cases}$$

- Podobne vieme ukázať:
  - $\mathbb{N} = \mathbb{N}^+$
  - $\mathbb{N} = \mathbb{E} \mathbf{v}$
  - $\mathbb{N} = \mathbb{O}dd$
- Menšia množina nemusí mať menej prvkov napriek tomu, že množiny  $\mathbb{E} v$  a  $\mathbb{O} dd$  sú vlastnými podmnožinami množiny  $\mathbb{N}$ , všetky tri majú rovnakú mohutnosť. Podobne  $\mathbb{N}^+ \subset \mathbb{N}$ , pričom obidve množiny majú rovnakú mohutnosť.

#### Lema

 $\mathbb{P}r \subset \mathbb{N}^+ \subset \mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}.$ 

#### Veta

Nech A, B sú disjunktné nekonečné množiny, pre ktoré platí:  $|\mathbb{N}| = |A| = |B|$ . Potom platí:

$$|\mathbb{N}| = |A \cup B|.$$

Dôkaz: Z predpokladu  $|\mathbb{N}|=|\mathsf{A}|$  vyplýva, že existuje bijektívne zobrazenie  $\varphi_1:\mathbb{N}\to A$ , ktoré každé prirodzené číslo zobrazí na nejaký prvok z množiny A. Konkrétne, prirodzené číslo  $n\in\mathbb{N}$  sa zobrazí na prvok  $\varphi_1(n)\in A$ . Podobne, z rovnosti  $|\mathbb{N}|=|\mathsf{B}|$  vyplýva, že existuje bijektívne zobrazenie  $\varphi_2:\mathbb{N}\to B$ , ktoré prirodzené číslo  $m\in\mathbb{N}$  zobrazí na prvok  $\varphi_2(m)\in B$ . Keďže množiny A a B sú disjunktné, môžeme definovať nové zobrazenie  $\psi:\mathbb{N}\to(A\cup B)$  nasledujúcim spôsobom:

$$\psi(n) = \begin{cases} \varphi_1(\frac{n}{2}) & n \in \mathbb{E}v \\ \varphi_2(\lfloor \frac{n}{2} \rfloor) & n \in \mathbb{O}dd \end{cases}$$

#### Veta

Nech A, B sú disjunktné nekonečné množiny, pre ktoré platí:  $|\mathbb{N}|=|A|=|B|$ . Potom platí:

$$|\mathbb{N}|=|A\cup B|.$$

Teda platí:  $\psi(0)=\varphi_1(0),\ \psi(2)=\varphi_1(1),\ \psi(4)=\varphi_1(2),\ \psi(6)=\varphi_1(3),\ \text{atd}'.$  To znamená, že v zobrazení  $\psi$  sa napr. číslo 4 zobrazí na ten prvok z množiny A, na ktorý bolo v zobrazení  $\varphi_1$  zobrazené číslo 2.

$$\psi(1) = \varphi_2(0), \ \psi(3) = \varphi_2(1), \ \psi(5) = \varphi_2(2), \ \psi(7) = \varphi_2(3), \ \text{atd}'.$$
 Zobrazenie  $\psi$  je bijektívne. Tým sme dokázali, že platí  $|\mathbb{N}| = |A \cup B|.$ 

**Poznámka:** Tvrdenie platí aj vtedy, ak v predpokladoch vynecháme podmienku disjunktnosti množín A a B.

#### Veta

Množina  $P = \mathbb{N} \times \mathbb{N}$  všetkých usporiadaných dvojíc prirodzených čísel je spočítateľná, teda platí nasledujúca nerovnosť:

$$|\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$$

**Dôkaz:** Myšlienka dôkazu je založená na nasledujúcej úvahe. Prvkami množiny  $\mathbb{N} \times \mathbb{N}$  sú usporiadané dvojice prirodzených čísel, ktoré usporiadame do dvojrozmernej tabuľky, ohraničenej ľavým a horným okrajom.

$\mathbb{N} \times \mathbb{N}$	0	1	2	3	
0	(0,0)	(0,1)	(0,2)	(0,3)	
1	(1,0)	(1, 1)		(1,3)	
2	(2,0)	(2,1)	(2,2)	(2,3)	
3	(3,0)	(3,1)	(3,2)	(3, 3)	
:	:	:	:	:	·

Zdroj: Daniela Chudá:Teoretické základy informatických vied, FIIT STU, 2020.

#### Veta

Množina  $P = \mathbb{N} \times \mathbb{N}$  všetkých usporiadaných dvojíc prirodzených čísel je spočítateľná, teda platí nasledujúca nerovnosť:

$$|\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$$

Prvky v tejto tabuľke očíslujeme prirodzenými číslami tak, aby uvedené očíslovanie reprezentovalo bijektívne zobrazenie  $f: \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ . Dvojiciam v tabuľke budeme priraďovať prirodzené čísla v smere diagonály, zhora-dole a sprava-doľava tak, ako je to naznačené nižšie

$\mathbb{N}$	0	1	2	3	
0	0	1	3	6	
1	2	4	7	11	
2	5	8	12	17	
3	9	13	18	24	
i	:	:	:	÷	٠٠.

Zdroj: Daniela Chudá:Teoretické základy informatických vied, FIIT STU, 2020.

#### Veta

Množina  $P=\mathbb{N}\times\mathbb{N}$  všetkých usporiadaných dvojíc prirodzených čísel je spočítateľná, teda platí nasledujúca nerovnosť:

$$|\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$$

Platí: f(0, 0) = 0, f(0, 1) = 1, f(1, 0) = 2, f(0, 2) = 3, f(1, 1) = 4, f(2, 0) = 5, atď. Funkčný predpis potom možno vyjadriť nasledovne:

$$f(m,n)=\frac{(m+n)(m+n+1)}{2}+m$$

Takto definované zobrazenie je bijektívne, a teda požadovaná rovnosť je dokázaná.

#### Veta

Mohutnosť množiny prirodzených čísel je menšia ako mohutnosť množiny reálnych čísel, teda platí nasledujúca nerovnosť:

$$|\mathbb{N}| < |\mathbb{R}|$$

**Dôkaz** (Sporom): Predpokladajme, že platí  $|\mathbb{N}|=|\mathbb{R}|$  a budeme sa snažiť dostať spor, z ktorého by potom hneď vyplynulo tvrdenie vety. Ak by platilo  $|\mathbb{N}|=|\mathbb{R}|$ , tak potom aj  $|(0,1)|=|\mathbb{N}|$ . To by znamenalo, že interval (0,1) možno zoradiť do postupnosti:

#### Veta

Mohutnosť množiny prirodzených čísel je menšia ako mohutnosť množiny reálnych čísel, teda platí nasledujúca nerovnosť:

$$|\mathbb{N}| < |\mathbb{R}|$$

kde  $a_{i,k}$  je k-ta cifra desatinného rozvoja čísla  $a_i$ . Ukážeme, že existuje číslo b = 0,  $b_1b_2$  ..., ktoré zrejme patrí do intervalu (0, 1), ale nenachádza sa v postupnosti  $a_1, a_2, \ldots$  Toto číslo zostrojíme pomocou diagonalizácie nasledovným spôsobom.

#### Veta

Mohutnosť množiny prirodzených čísel je menšia ako mohutnosť množiny reálnych čísel, teda platí nasledujúca nerovnosť:

$$|\mathbb{N}| < |\mathbb{R}|$$

Stačí vziať  $b_1 \neq a_{1,1}$ ,  $b_2 \neq a_{2,2}$ , ...,  $b_i \neq a_{i,i}$ ... Keďže b sa v i-tej cifre líši od i-teho čísla v postupnosti, nemôže sa rovnať žiadnemu číslu z tejto postupnosti.

$$b_b = \begin{cases} 1 & \text{ak } a_{k,k} \neq 1 \\ 9 & \text{ak } a_{k,k} = 1 \end{cases}$$

#### Veta

Mohutnosť množiny prirodzených čísel je menšia ako mohutnosť množiny reálnych čísel, teda platí nasledujúca nerovnosť:

$$|\mathbb{N}| < |\mathbb{R}|$$

Zdroj: Daniela Chudá: Teoretické základy informatických vied, FIIT STU, 2020.

Potom číslo b patrí do intervalu (0,1), teda existuje také m, že  $a_m=b$ . Špeciálne, musí byť  $a_{m,m}=b_m$ . Ale číslo b bolo zostrojené tak, aby pre každé k bolo  $b_k\neq a_{k,k}$ , teda aj  $b_m\neq a_{m,m}$ . To je hľadaný spor.  $\square$ 

#### Veta

Mohutnosť množiny prirodzených čísel je menšia ako mohutnosť množiny reálnych čísel, teda platí nasledujúca nerovnosť:

$$|\mathbb{N}| < |\mathbb{R}|$$

**Poznámka:** Ukázali sme, že neexistuje bijekcia  $f: |\mathbb{N}| \to |(0,1)|$ . Pomocou diagonalizácie sme ukázali, že žiadna funkcia  $f: |\mathbb{N}| \to |(0,1)|$  nie je surjektívna. To sme ukázali tak, že sme našli prvok  $x \in (0,1)$ , na ktorý sa f nezobrazí (nech je f akákoľ vek funkcia).

## Spočítateľné a nespočítateľné množiny - Opakovanie

### Definícia (Definícia spočítateľ nej a nespočítateľ nej množiny)

Množina A sa nazýva spočítateľná, ak  $|A| \leq |\mathbb{N}|$  (t.j.  $|A| \leq \aleph_0$ ). Množina, ktorá nie je spočítateľná, sa nazýva nespočítateľná.

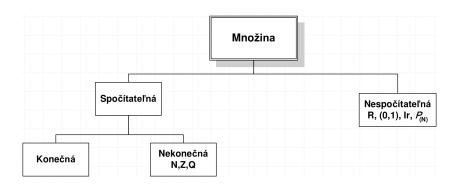
#### Veta

Množina  $\mathbb{R} \setminus \mathbb{Q}$  iracionálnych čísel je nespočítateľná.

**Dôkaz** (Sporom): Predpokladajme, že množina  $\mathbb{R}\setminus\mathbb{Q}$  je spočítateľná. V tom prípade je buď konečná alebo nekonečná spočítateľná. Predpokladajme najprv, že množina  $\mathbb{R}\setminus\mathbb{Q}$  je konečná. Keďže vieme, že platí  $|\mathbb{Q}|=\aleph_0$ , existuje bijektívne zobrazenie  $(\mathbb{R}\setminus\mathbb{Q})\cup\mathbb{Q}\to\mathbb{N}$ . To znamená, že množina  $(\mathbb{R}\setminus\mathbb{Q})\cup\mathbb{Q}=\mathbb{R}$  je nekonečná spočítateľná, čo je spor.

V prípade, keby bola množina  $\mathbb{R}\setminus\mathbb{Q}$  nekonečná spočítateľná, dostaneme taktiež spor. Množina  $\mathbb{Q}$  je nekonečná spočítateľná, pričom množiny  $\mathbb{Q}$  a  $\mathbb{R}\setminus\mathbb{Q}$  sú disjunktné. Potom ich zjednotenie, teda množina  $\mathbb{R}$ , musí byť spočítateľná. To je opäť spor. Tým je tvrdenie dokázané.  $\square$ 

## Spočítateľné a nespočítateľné množiny - Opakovanie



Zdroj: Daniela Chudá: Teoretické základy informatických vied, FIIT STU, 2020.

# Úvod do teórie formálnych jazykov

## Čo je to jazyk?

- súbor slov a metód skladania slov, ktorý používa a chápe dostatočne veľká ľudská spoločnosť (Webster - americký výkladový slovník angličtiny).
- sústava osobitných znakových hodnôt, ktorá slúži ako nástroj vyjadrovania, dorozumievania a ukladania poznatkov v istom spoločenstve (Slovník súčasného slovenského jazyka)
- matematický systém (J.E. Hopcroft J.D. Ullman: Formálne jazyky a automaty. Alfa 1969.)

# Úvod do teórie formálnych jazykov

Na abstraktnej úrovni rozlišujeme 3 základné typy jazykov:

- prirodzené (lingvistické)
- umelé (napr. programovacie)
- formálne

Jazyk je možné skúmať z dvoch hľadísk:

- syntax (štruktúra a stavba)
- sémantika (význam a obsah)

# Úvod do teórie formálnych jazykov

### Reprezentácia jazyka:

- Matematický opis jazyka množina.
- Systematické generovanie slov z jazyka (generovanie jazyka) gramatika.
- Zostrojenie algoritmu, ktorý určí, či dané slovo patrí do jazyka (rozpoznanie jazyka) – automat

### Teória jazykov:

 štúdium množín znakov, reťazcov, ich reprezentácií, štruktúr a vlastností

### Abeceda

### Definícia (Definícia abecedy jazyka)

**Abeceda** je konečná neprázdna množina symbolov (písmen). Označuje sa  $\Sigma$ .

#### Príklad:

• 
$$\Sigma = \{a, b, c\}$$

### Slovo

### Definícia (Definícia slova nad abecedou $\Sigma$ )

**Slovo** (veta, reťazec) nad abecedou  $\Sigma$  je konečná postupnosť symbolov zo  $\Sigma$ . Označuje sa malým písmenom (najčastejšie u, v, w, x, y).

$$u = a_1 a_2 ... a_n$$

 $kde \ \forall i \in \mathbb{N}^+ : a_i \in \Sigma.$ 

Prázdnu postupnosť (prázdne slovo) označujeme  $\varepsilon$ .

### Príklady:

- u = aa
- v = ababac
- $\mathbf{w} = \varepsilon$

### Podslovo

### Definícia (Definícia podslova)

**Podslovo** u slova  $v = a_1 a_2 ... a_n$  je súvislá podpostupnosť  $u = a_i a_{i+1} ... a_j$  pričom platí  $1 \le i \le j \le n$ . Špeciálne podslová:

- prefix (počiatočné podslovo, predpona) platí i = 1.
   Slovo u je prefixom slova v, ak platí že v = uw pre nejaké slovo w.
- sufix (koncové podslovo, prípona) platí j = n.
   Slovo u je sufixom slova v, ak platí že v = wu pre nejaké slovo w.
- infix (podslovo, časť slova) ak platí že  $v = w_1 u w_2$  pre nejaké slová  $w_1 a w_2$ .

Prázdne slovo  $\varepsilon$  je prefixom a zároveň sufixom každého slova.

# Operácie nad slovami

- Rovnosť slov
- Zreťazenie slov
- Mocnina slova
- Zrkadlový obraz
- Dĺžka slova
- Počet výskytov symbolu
- Homomorfizmus

### Rovnosť slov

### Definícia (Definícia rovnosti slov)

Nech k,  $l \in \mathbb{N}^+$  a  $u = a_1...a_k$ ,  $v = b_1...b_l$  sú slová. Potom platí, že slovo u sa rovná slovu v (zapisujeme u = v), ak k = l a zároveň  $\forall 1 \leq i \leq k$  :  $a_i = b_i$ .

Naopak, slová u, v sa nerovnajú (zapisujeme u  $\neq$  v), ak k  $\neq$  l alebo  $\exists 1 \leq j \leq k : a_i \neq b_i$ .

### Príklady:

- $\varepsilon$ aa = aa
- $abbbaaaa = ab^3a^4$

### Zreťazenie slov

### Definícia (Definícia zreťazenia slov)

Nech  $u = u_1 u_2 ... u_n, v = v_1 v_2 ... v_m$ , potom

 $u \cdot v = u_1 u_2 \dots u_n v_1 v_2 \dots v_m.$ 

Nech u, v, w sú slová nad abecedou Σ. Potom operácia zreťazenia spĺňa nasledujúce vlastnosti:

- (uzavretosť) slovo uv je tiež nad abecedou Σ
- (asociatívnosť) (uv)w = u(vw)
- (neutrálny prvok)  $u\varepsilon = \varepsilon u = u$

Znak zreťazenia (bodku) budeme väčšinou vynechávať.

#### Príklad:

•  $u = aabb, v = ccc, w = \varepsilon$  potom uv = aabbccc, vw = ccc

### Mocnina slova

### Definícia (Definícia mocniny slova)

Nech w je ľubovoľné slovo nad abecedou  $\Sigma$  a  $i \in \mathbb{N}^+$ , potom platí:

- $w^0 = \varepsilon$
- $w^{i} = w^{i-1}w$

(je zrejmé, že platí  $w^1 = w$ .)

### Príklady:

- $a^2b^1e^0s^3 = aabsss$
- $(ab)^2 \neq a^2b^2$

# Zrkadlový obraz slova

### Definícia (Definícia zrkadlového obrazu slova)

Nech w je ľubovoľné slovo, potom pre jeho zrkadlový obraz platí:

- $\bullet \ \varepsilon^R = \varepsilon$
- $(au)^R = u^R a$

Ak pre slovo w platí  $w^R = w$ , takéto slovo sa nazýva palindróm.

### Príklady:

- $abbac^R = cabba$
- $(a^3b)^R = ba^3$

## Dĺžka slova

### Definícia (Definícia dĺžky slova)

**Dĺžka slova** w je dĺžka postupnosti symbolov, ktorá ho tvorí. Označuje sa ako |w|.

#### Vlastnosti:

- Ak  $w = a_1 a_2 ... a_n$ , potom |w| = n.
- $|\varepsilon| = 0$
- $|v^R| = |v|$
- $|u \cdot v| = |u| + |v|$

# Počet výskytov symbolu

### Definícia (Definícia počtu výskytov symbolu)

**Počet výskytov symbolu** a v slove w sa označuje ako  $\sharp_a w$  a je definovaný nasledovne:

• 
$$\sharp_a \varepsilon = 0$$

## Homomorfizmus

### Definícia (Definícia homomorfizmu)

**Homomorfizmus** h nad slovami je zobrazenie z množiny všetkých slov nad abecedou  $\Sigma_1$  do množiny všetkých slov nad abecedou  $\Sigma_2$   $(h: \Sigma_1^* \to \Sigma_2^*)$ , pričom pre  $\forall u, v \in \Sigma_1^*$  platí:

$$h(u \cdot v) = h(u) \cdot h(v)$$

#### Vlastnosti:

- zachováva operáciu zreťazenia
- $h(\varepsilon) = \varepsilon$

$$h: \{0,1\}^* \to \{a,b\}^*$$
  
 $h(0) = ab$   
 $h(1) = ba$   
 $h(0011) = ababbaba$ 

## Jazyk

### Definícia (Definícia jazyka)

**Jazyk** nad abecedou  $\Sigma$  je ľubovoľná množina slov nad abecedou  $\Sigma$ . Označujú sa väčšinou písmenom L (s indexom, ak je to potrebné).

#### Mohutnosť:

- jazyky konečné
- jazyky nekonečné (väčšinou spočítateľné)

### Množinové operácie:

- zjednotenie, prienik, rozdiel, rovnosť, nerovnosť, podmnožina
- zreťazenie jazykov, mocnina jazyka, Kleeneho uzáver (iterácia), kladný uzáver, doplnok jazyka, homomorfizmus jazyka

## Zreťazenie jazykov

### Definícia (Definícia zreťazenia jazykov)

Nech  $L_1$  a  $L_2$  sú jazyky. **Zreťazenie jazykov** (označuje sa ·) je definované nasledovne:

$$L_1 \cdot L_2 = L_1 L_2 \{uv | u \in L_1, v \in L_2\}$$

Vlastnosti:

- $L_1L_2 \neq L_2L_1$
- $L \cdot \emptyset = \emptyset \cdot L = \emptyset$
- $L \cdot \{\varepsilon\} = \{\varepsilon\} \cdot L = L$

# Mocnina jazyka

### Definícia (Definícia mocniny jazyka)

Nech L je jazyk. **Mocnina jazyka** L (označuje sa L<sup>i</sup> , kde  $i \in \mathbb{N}$ ) je definovaná nasledovne:

- $L^0 = \{ \varepsilon \}$
- $L^i = L^{i-1} \cdot L$  pre  $i \in \mathbb{N}^+$

# Kleeneho uzáver (iterácia) jazyka

### Definícia (Definícia iterácie jazyka)

Nech L je jazyk. **Kleeneho iterácia** alebo uzáver jazyka L (označuje sa  $L^*$ ) je definovaná nasledovne:

$$L^* = \bigcup_{i=0}^{\infty} L^i$$

# Kleeneho kladný uzáver (iterácia) jazyka

### Definícia (Definícia kladnej iterácie jazyka)

Nech L je jazyk. **Kleeneho kladná iterácia** alebo kladný uzáver jazyka L (označuje sa  $L^+$ ) je definovaná nasledovne:

$$L^+ = \bigcup_{i=1}^{\infty} L^i$$

Vlastnosti:

• 
$$L^* = L^+ \cup \{\varepsilon\}$$

# Doplnok (komplement) jazyka

### Definícia (Definícia doplnku jazyka)

Nech L je jazyk a  $\Sigma_L$  je najmenšia abecedu,nad ktorou je L. **Doplnok jazyka** alebo komplement jazyka L (označuje sa  $L^C$ ) je definovaný nasledovne:

$$L^C = \Sigma_I^* \setminus L$$

# Homomorfizmus jazyka

### Definícia (Definícia homomorfizmu jazyka)

Nech  $\Sigma_1$  a  $\Sigma_2$  sú abecedy, L je jazyk nad  $\Sigma_1$  a  $h: \Sigma_1^* \to \Sigma_2^*$  je homomorfizmus (nad slovami). **Homomorfizmus jazyka** (označuje sa h(L)) je definovaný nasledovne:

$$h(L) = \{h(w)|w \in L\}$$

# Triedy jazykov

### Definícia (Definícia tried jazykov)

Jazyk sa nazýva kontextový (bezkontextový, resp. regulárny), ak je generovaný kontextovou (bezkontextovou, resp. regulárnou) gramatikou. Jazyk sa nazýva rekurzívne vyčísliteľný, ak je generovaný frázovou gramatikou.

# Triedy jazykov

### Tried jazykov sa označujú nasledovne:

- L<sub>RE</sub> trieda rekurzívne vyčísliteľných jazykov, (ang. recursive enumerable), generovaných frázovou gramatikou (Turingov stroj)
- $\mathcal{L}_{CS}$  trieda kontextových jazykov, (ang. context sensitive), generovaných kontextovou gramatikou (lineárne ohraničený automat)
- L<sub>CF</sub> trieda bezkontextových jazykov (ang. context free), generovaných bezkontextovou gramatikou (zásobníkový automat)
- $\mathcal{R}$  trieda regulárnych jazykov (ang. regular), generovaných regulárnou gramatikou (konečný automat)

## Hierarchia tried jazykov

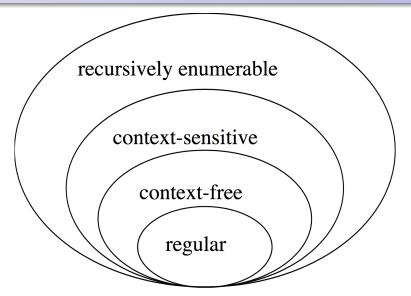
### Definícia (Definícia hierarchie tried jazykov)

Medzi triedami jazykov **Chomského hierarchie** platia nasledujúce vzťahy  $\mathcal{L}_{RE} \supsetneq \mathcal{L}_{CSE} \supsetneq \mathcal{L}_{CF} \supsetneq \mathcal{R}$ .

Vyššie definovaná trieda kontextových jazykov (budeme sa im venovať neskôr) neumožňuje vygenerovať prázdne slovo, rozšírime ju nasledovne:

$$\mathcal{L}_{\textit{CSE}} = \{L, L \cup \{\varepsilon\} | L \in \mathcal{L}_{\textit{CS}}\}$$

## Hierarchia tried jazykov



Zdroj: https://en.wikipedia.org/wiki/Chomsky\_hierarchy

# Ďakujem vám za pozornosť.

## Dotazník k prednáške:

https://forms.gle/gPpbqYUYSnGfbG9X9

