1 Úvod

Poznámka (Historie)

- První formalizace pojmu algoritmus – Ada, Countess of Lovelace 1852.
- Intenzivnější vývoj s rozvojem počítačů ve 2. čtvrtině 20. století.
- Co stroje umí a co ne? – Church, Turing, Kleene (konečné automaty / neuronové sítě), Post, Markov, Chomsky (zásobníkové automaty a formální teorie konečných automatů, zkoumal Angličtinu).

Poznámka (Cíl)
Osvojit si abstraktní model počítače, vnímat jak drobné změny v definici vedou k velmi rozdílným důsledkům. Zažít skutečnost alg. nerozhodnutelných problémů a připravit se na přednášku o složitosti a NP-úplnosti.

Poznámka (Praktické využití)
Korektnost algoritmů, zpracování přirozeného jazyka, lexikální a syntaktická analýza v překladačích. Návrh, popis a verifikace hardwaru (automaty, integrované obvody, stroje). Vyhledávání v textu atd.

2

Definice 2.1 (Deterministický konečný automat (DFA))

Deterministický konečný automat $A = (Q, \Sigma, \delta, q_0, F)$ sestává z: konečné množiny stavů (Q), konečné neprázdné množiny vstupních symbolů (abecedy, Σ), přechodové funkce, tj. zobrazení $Q \times \Sigma \to Q$ (značí se hranami grafu, δ), počátečního stavu (vede do něj šipka 'odnikud', $q_0 \in Q$) a neprázdné množiny (přijímajících) stavů (značí se dvojitým kruhem / šipku 'ven', $F \subseteq Q$).

Úmluva
Přidáváme 0-2 stavy: fail (pokud je nějaký přechod nedefinován, vede sem a všechno z fail vede do fail) a final (pokud je F prázdné, všechny šipky z něj vedou zpět do něj).

Definice 2.2 (Slovo, jazyk)

Mějme neprázdnou množinu symbolů Σ. Slovo je konečná (i prázdná) posloupnost symbolů $s \in \Sigma$, prázdné slovo se značí λ nebo ε.

Množinu všech slov v abecedě Σ značíme Σ^* a množinu všech neprázdných Σ^+.

Jazyk $L \subseteq \Sigma^*$ je množina slov v abecedě Σ.

Definice 2.3 (Operace: zřetězení, mocnina, délka slova)

Nad slovy Σ^* definujeme operace: Zřetězení slov $u.v$ nebo uv, mocnina (počet opakování) u^n ($u^0 = \lambda$, $u^1 = u$, $u^{n+1} = u^n.u$), délka slova $|u|$ ($|\lambda| = 0$, $|auto| = 4$), počet výskytů $s \in \Sigma$ ve slově u značíme $|u|_s$ ($|zmrzlina|_z = 2$).

Definice 2.4 (Rozšířená přechodová funkce)

Mějme přechodovou funkci $\delta : Q \times \Sigma \to Q$. Rozšířenou přechodovou funkci $\delta^* : Q \times \Sigma^* \to Q$ (tranzitivní uzávěr δ) definujeme induktivně: $\delta^*(q, \lambda) = q$ a $\delta^*(q, wx) = \delta(\delta^*(q, w), x)$ pro $x \in \Sigma$ a $w \in \Sigma^*$.

Definice 2.5 (Jazyky rozpoznatelné konečnými automaty, regulární jazyky)

Jazykem rozpoznávaným (akceptovaným, přijímaným) konečným automatem A nazveme jazyk $L(A) = \{w | w \in \Sigma^* \wedge \delta^*(q_0, w) \in F\}$.

Jazyk je rozpoznatelný konečným automatem, jestliže existuje konečný automat A takový, že $L = L(A)$.

Třídu jazyků rozpoznatelných konečnými automaty označíme $\mathcal{F}$ a nazveme ji regulární jazyky.

Věta 2.1 (!Iterační (pumpin) lemma pro regulární jazyky)

Mějme regulární jazyk L. Pak existuje konstanta $n \in \mathbb{N}$ (závislá na L) tak, že každé $w \in L; |w| \geq n$ můžeme rozdělit na tři části, $w = xyz$, že $y \neq \lambda \wedge |xy| \leq n \wedge \forall k \in \mathbb{N}_0$, slovo xy^kz je také v L.

Důkaz

Mějme regulární jazyk L, pak existuje DFA A s n stavy, že $L = L(A)$. Vezmeme libovolné slovo $a_1 a_2 \dots a_n \dots a_m = w \in L$ délky $m \geq n$, $a_i \in \Sigma$. Následně definujeme $\forall i : p_i = \delta^*(q_0, a_1 a_2 \dots a_i)$. Platí $p_0 = q_0$. Z Dirichletova principu se některý stav opakuje. Vezmeme první takový, tj. $(\exists i, j)(0 \leq i < j \leq n \wedge p_i = p_j)$.

Definujeme $x = a_1 a_2 \dots a_i$, $y = a_{i+1} a_{i+2} \dots a_j$ a $z = a_{j+1} a_{j+2} \dots a_m$, tj. $w = xyz$, $y \neq \lambda$, $|xy| \leq n$. □