

Klasifikace gramatik, formálních jazyků a automatů přijímajících jazyky

ZPRACUJE: Mystik

Obsah

- 1 Základní pojmy
- 2 Gramatika
- 3 Chomského hierarchie klasifikace gramatik
 - 3.1 Typ 0 – obecné (neomezené) gramatiky
 - 3.2 Typ 1 – kontextové gramatiky
 - 3.3 Typ 2 – bezkontextové gramatiky
 - 3.4 Typ 3 – pravé/levé lineární gramatiky
 - 3.5 Speciální podtypy gramatik
 - 3.5.1 Pravé/levé regulární gramatiky
 - 3.5.2 Lineární gramatiky
 - 3.5.3 Deterministické bezkontextové gramatiky
 - 3.5.4 Rekursivní gramatiky

Základní pojmy

Abeceda (značíme S)

je neprázdná konečná množina, jejíž prvky nazýváme symboly, značíme S .

Řetězec / slovo / věta (značíme w)

konečná posloupnost symbolů abecedy

Alternativní definice:

- 1) prázdný řetězec je řetězec
- 2) je-li w řetězec a a symbol abecedy pak wa je řetězec
- 3) řetězce jsou vše, co lze získat jen a pouze aplikací pravidel 1) a 2)

Prázdný řetězec ε

řetězec neobsahující žádný symbol

Konkatenace (značíme xy)

připojení řetězce za jiný řetězec

- konkatenace je asociativní
- prázdný řetězec je neutrální prvek vzhledem ke konkatenaci

Iterace (značíme S^*)

množina všech řetězců nad abecedou S

(volný monoid nad abecedou S generovaný operací konkatenace)

Pozitivní iterace (značíme S^+)

iterace bez prázdného řetězce

Formální jazyk (značíme L)

jakákoli podmnožina $L \subseteq S^*$

Konkatenace jazyků

rozumíme binární operaci $L1 \cdot L2 = \{xy | x \in L1 \wedge y \in L2\}$

Iterace jazyka

$$L^0 = \{\varepsilon\}$$

$$L^n = L \cdot L^{n-1} \quad n \geq 1$$

$$L^* = \bigcup_{n \geq 0} L^n$$

$$L^+ = \bigcup_{n \geq 1} L^n$$

Gramatika

Gramatika (značíme G)

je čtveřice $G = (N, T, P, S)$, kde:

N je konečná množina **nonterminálních** symbolů (nonterminálů)

T je konečná množina **terminálních symbolů** (terminálů)

P je konečná množina **přepisovacích pravidel**

S je **počáteční (výchozí, startovací) nonterminál** $S \in N$

Přepisovací pravidlo

podmnožina kartézského součinu $(N \cup T)^* N (N \cup T)^* \times (N \cup T)^*$

- $(a, b) \in P$ zapisujeme ve tvaru $a \rightarrow b$
- kde a je **levá strana** a b je **pravá strana** pravidla.

Přímá derivace (značíme \Rightarrow)

binární relace mezi řetězci u, v která platí pokud můžeme řetězce zapsat jako $u = gad, v = gbd$ a $(a, b) \in P$

tj. pokud lze jeden řetězec vytvořit z druhého aplikací jednoho přepisovacího pravidla

Derivace \Rightarrow^+

tranzitivní uzávěr relace přímé derivace

tj. pokud lze vytvořit z druhého postupnou aplikací několika přepisovacích pravidel

Derivace \Rightarrow^*

tranzitivní a reflexivní uzávěr relace přímé derivace

tj. pokud lze vytvořit z druhého postupnou aplikací několika přepisovacích pravidel nebo pokud jsou řetězce stejné

Větná forma

řetězec terminálů a nonterminálů, které jsou derivací počátečního nonterminálu

Věta

větná forma obsahující pouze terminály

Jazyk generovaný gramatikou

množina všech vět, které patří do gramatiky

Chomského hierarchie klasifikace gramatik

Klasifikace je definována podle tvaru přepisovacích pravidel příslušných gramatik

Platí: $L_3 \subset L_2 \subset L_1 \subset L_0$

Typ 0 – obecné (neomezené) gramatiky

- Jazyky: **rekurzivně vyčíslitelné jazyky**
- Automaty: **Turingovy stroje**
- Tvar pravidel:

$$\alpha \rightarrow \beta$$

$$\alpha \in (N \cup \Sigma)^* N (N \cup \Sigma)^*$$

$$\beta \in (N \cup \Sigma)^*$$

Typ 1 – kontextové gramatiky

- Jazyky: **kontextové jazyky**
- Automaty: **lineárně omezené Turingovy stroje**
- Tvar pravidel:

$\alpha A \beta \rightarrow \alpha \gamma \beta$ (nebo $S \rightarrow \epsilon$ pokud $\epsilon \in L$ a S se nevyskytuje na pravé straně žádného pravidla.)

$$A \in N$$

$$\alpha, \beta \in (N \cup \Sigma)^*$$

$$\gamma \in (N \cup \Sigma)^+$$

- Nesmí dojít ke zkrácení větné formy.

Typ 2 – bezkontextové gramatiky

- Jazyky: **bezkontextové jazyky**
- Automaty: **zásobníkové automaty**
- Tvar pravidel:

$$A \rightarrow \gamma$$

$$A \in N$$

$$\gamma \in (N \cup \Sigma)^*$$

Typ 3 – pravé/levé linární gramatiky

- Jazyky: **regulární jazyky**
- Automaty: **konečné automaty**
- Tvar pravidel:

Pravá lineární: $A \rightarrow xB$ nebo $A \rightarrow x$

Levá lineární: $A \rightarrow Bx$ nebo $A \rightarrow x$

$$A, B \in N$$

$$x \in \Sigma^*$$

Speciální podtypy gramatik

Pravé/levé regulární gramatiky

- ekvivalentní gramatikám typu 3
- Tvar pravidel:

Pravá regulární: $A \rightarrow aB$ nebo $A \rightarrow x$

Levá regulární: $A \rightarrow Ba$ nebo $A \rightarrow x$

$$A, B \in N$$

$$a \in \Sigma$$

Lineární gramatiky

- podmnožina gramatik typu 2
- Tvar pravidel:

$$A \rightarrow xBy \text{ nebo } A \rightarrow x$$

$$A, B \in N$$

$$x, y \in \Sigma^*$$

Deterministické bezkontextové gramatiky

- podmnožina typu 2
- jazyky přijímané deterministickým zásobníkovým automatem

Rekurzivní gramatiky

- podmnožina typu 0
- přijímají je **úplné TS** (TS, které pro každý vstup rozhodnou - nikdy necyklí)
- všechny kontextové jazyky jsou rekurzivní, ale ne všechny rekurzivní jazyky jsou kontextové

Kategorie: Státnice 2011 | Teoretická informatika

Stránka byla naposledy editována 28. 5. 2011 v 08:47.