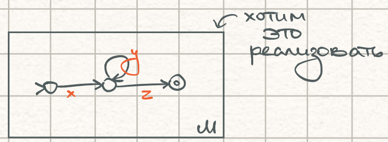
Теорема о накачке для регулярных языков

Теорема о накачке

Пусть $L \subseteq \Sigma^*$ - рег. язык. Тогда $\exists n_0 \geq 1$ т., ч. $\forall w \in L$ $\quad |w| \geq n_0 \Rightarrow \exists x, y, z \in \Sigma^*$ т., ч.:

$$\begin{cases} w = xyz \\ y \neq e \\ |xy| \leq n_0 \\ xy^kz \in L \quad \forall k \geq 0 \end{cases}$$ — **св-во накачивания**

Д-во: Т.к. L - рег. язык, то он явл. автоматным. Пусть $M = (Q, \Sigma, s, F, \delta)$ - ДКА, для ктр $L = L(M)$



Рассм. $n_0 = |Q|$ и пусть $w \in L = L(M)$ т., ч. $|w| \geq n_0$, т.е. $w = a_1a_2 \dots a_{n_0}a_{n_0+1} \dots a_m$ для нктр $m \geq n_0$

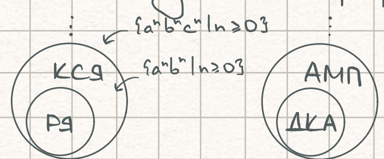
Имеем для $q_1 = s$ $\quad (q_1, a_1a_2 \dots a_m) \vdash_M (q_2, a_2 \dots a_m) \vdash_M \dots$
$\dots \vdash_M (q_{n_0}, a_{n_0} \dots a_m) \vdash_M (q_{n_0+1}, a_{n_0+1} \dots a_m) \vdash_M \dots \vdash_M$
$\vdash_M (q_m, a_m) \vdash_M (q_{m+1}, e)$, этот шаг всегда делается где $q_{m+1} \in F$

Т.к. $|Q| = n_0$, то $\exists i, j \leq n_0 + 1$ т., ч. $i < j$ и $q_i = q_j$

Полагаем $x = a_1 \dots a_{i-1}$ $\quad y = a_i \dots a_{j-1}$ $\quad z = a_j \dots a_m$. Тогда $xy^kz \in L = L(M)$ $\forall k \geq 0$

Контекстно-свободные грамматики и языки

Вспомним кусочек иерархии:



Опр **Контекстно-свободной грамматикой** наз. упоряд. четвёрка $G = (V, \Sigma, S, R)$, где V - конечное мн-во символов

$\Sigma \subseteq V$ - мн-во терминальных символов (символы из $V \setminus \Sigma$ наз. нетерминальными)

$S \in V \setminus \Sigma$ - начальный (стартовый) нетерм. символ

$R \subseteq (V \setminus \Sigma) \times V^*$ - конечное мн-во правил или продукций

Правило вида $(A, v) \in R$ обозн. $A \to_G v$ (или просто $A \to v$)

Опр Пусть $G = (V, \Sigma, S, R)$ - КС-грамматика и пусть $u, v \in V^*$. Говорим, что **u порождает v** в G (за 1 шаг) и обозн. $u \Rightarrow_G v$, если $u = xAy$ $\quad v = xzy$ для нктр $(A \to_G z) \in R$, $x, y, z \in V^*$, $A \in V \setminus \Sigma$

Опр Бинарное отношение $\Rightarrow_G^*$ опр-ся как рефл. и транзит. замыкание $\Rightarrow_G$:
$u \Rightarrow_G^* v \overset{def}{\Leftrightarrow}$ либо $u = v$, либо $\exists n \geq 0$ и $w_1, \dots, w_n$ т., ч. $u \Rightarrow_G w_1 \Rightarrow_G \dots \Rightarrow_G w_n \Rightarrow_G v$

Опр Для КС-грамм. $G = (V, \Sigma, S, R)$ язык $L(G) = \{w \in \Sigma^* \mid S \Rightarrow_G^* w\}$ наз. **языком, порождённым грамматикой** G

Опр Язык $L \subseteq \Sigma^*$ наз. **контекстно-свободным**, если $\exists$ КС-грамм. G т., ч. $L = L(G)$

Ex Рассм. КС-грамм. $G=(V,\Sigma,S,R)$, где $V=\{S,a,b\}$, $\Sigma=\{a,b\}$, $R=\{S\to aSb,\ S\to e\}$
Тогда $L(G)=\{a^nb^n \mid n\geq 0\}$ – это КСЯ, но не РЯ (следует из т. о накачке)

Теорема
Всякий РЯ явл-ся КСЯ

Д-во: Т.к. из регул-ти следует автоматность, пусть $M=(Q,\Sigma,s,F,\delta)$ – ДКА, распозн. данный РЯ $L\subseteq\Sigma^*$
Рассм. КС-грамм. $G_M=(Q\cup\Sigma,\Sigma,s,R_M)$, где
$R_M=\{q\to ap \mid \delta(q,a)=p\}\cup\{q\to e \mid q\in F\}$

Такая грамматика наз. праворегулярной

Тогда $L(G_M)=L(M)=L$

Теорема св-во замкнутости класса КСЯ
Класс КСЯ замкнут отн-но объед-я, конкатенации и итерации ("звёздочки" Клини):

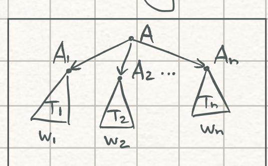
Д-во: Пусть $L_1=L(G_1)$, $L_2=L(G_2)$ для нктр КС-грамм. $G_i=(V_i,\Sigma_i,S_i,R_i)$ $i=1,2$
Можно считать, что $(V_1\setminus\Sigma_1)\cap(V_2\setminus\Sigma_2)=\varnothing$
- $L_1\cup L_2=L(G_\cup)$, где $G_\cup=(V_1\cup V_2\cup\{S\},\Sigma_1\cup\Sigma_2,S,R_\cup)$, где S – новый символ, $S\notin V_1\cup V_2$, $R_\cup=R_1\cup R_2\cup\{S\to S_1,\ S\to S_2\}$
- $L_1\circ L_2=L(G_\circ)$, где $G_\circ=(V_1\cup V_2\cup\{S\},\Sigma_1\cup\Sigma_2,S,R_\circ)$, где $S\notin V_1\cup V_2$, $R_\circ=R_1\cup R_2\cup\{S\to S_1S_2\}$
- $(L_1)^*=L(G_*)$, где $G_*=(V_1\cup\{S\},\Sigma_1,S,R_*)$, где $S\notin V_1$, $R_*=R_1\cup\{S\to SS_1,\ S\to e\}$

Зам Класс КСЯ не замкнут отн-но пересеч-я и доп-я (д-во позже)

Деревья разбора (parse trees)

Опр Пусть $G=(V,\Sigma,S,R)$ – нктр КС-грамм. Мн-во **деревьев разбора** для G, а также понятия результата, корня, листьев и высоты дерева разбора опр-ся индуктивно:

- $\boxed{\cdot a}$ $(a\in\Sigma)$ – дерево разбора для G (рез. = корень = лист = a, высота = 0)
- $\boxed{A \downarrow e}$ $((A\to e)\in R)$ – дерево разбора для G (рез = лист = e, корень = A, высота = 1)
- Пусть $T_1,\dots,T_n$ – деревья разбора для G с корнями $A_1,\dots,A_n\in V$, результатами $w_1,\dots,w_n\in\Sigma^*$, мн-вами листьев $X_1,\dots,X_n$ и высотами $h_1,\dots,h_n$ и пусть $(A\to A_1,\dots,A_n)\in R$. Тогда дерево:



явл-ся деревом разбора для G, у ктр корень = A, результат = $w_1w_2\dots w_n$, мн-во листьев = $X_1\cup X_2\cup\dots\cup X_n$, высота = $\max\{h_1,\dots,h_n\}+1$