4. Теорема о пересечении КС-языка с регулярным языком.

Теорема

Пересечение контекстно-свободного языка (КСЯ) с регулярным языком (РЯ) является контекстно-свободным языком (КСЯ).

Или в простой форме: КСЯ ∩ РЯ = КСЯ.

Док-во:

Док-во:

L - КСЯ $\Rightarrow \exists G = (\Sigma, \Gamma, P, S)$ - КСГ

$L(G) = L$

функции переходов

алфавит

конечные состояния

M - РЯ $\Rightarrow \exists A = (\Sigma, Q, \delta, q_0, F)$ - 1КА (детерминиров конечный автомат)

$L(A) = M$

мн-во сост

начальное сост

$\forall f_i \in F$ Рассмотрим автомат.

$A_i = (\Sigma, Q, \delta, q_0, f_i)$ разделили на кол-во выходных состояний.

$M = L(A) = \bigcup_i^n L(A_i)$

$L \cap M = L \cap \bigcup_i L(A_i) = \bigcup_i (L \cap L(A_i))$

$L \cap L(A_i)$ - КСЯ ?

Рассмотрим $A_f = (\Sigma, Q, \delta, q_0, f)$

$f \in F$ и пересечение $L \cap L(A_f)$

$H = (\Sigma, \Gamma', P', S')$

$\Gamma' = Q \times (\Gamma \cup \Sigma) \times Q$

$S' = (q_0, S, f)$

Построение P':

Вида:

1) $(A \to X_1 \ldots X_n) \in P \Rightarrow ((q, A, p) \to (q, X_1, r_1) \ldots$
$\ldots (r_{n-1}, X_n, p)) \in P' \quad \forall (q, p, r_1, \ldots, r_{n-1}) \in Q^{n+1}$

2) $(q, a, p) \in \delta \Rightarrow ((q, a, p) \to a) \in P'$

$L \cap L(A_f) = L(H)$?

$w \in L$, то $S \overset{*}{\underset{G}{\Rightarrow}} w = a_1 \ldots a_k$

Тогда $(q_0, S, f) \overset{*}{\underset{H}{\Rightarrow}} (q_0, a_1, r_1)(r_1, a_2, r_2) \ldots (r_{k-1}, a_k, f)$

$w \in L(A_f) \Rightarrow \exists$ путь из q_0 в f, помеч. $a_1, \ldots, a_k$

$w \in L(H) \Leftrightarrow \exists S \overset{*}{\underset{H}{\Rightarrow}} (q_0, a_1, r_1)(r_1, a_2, r_2) \ldots$
$(r_{k-1}, a_k, f) \overset{*}{\underset{H}{\Rightarrow}} a_1 \ldots a_k$

■

Суть доказательства и конструкция

Доказательство строится конструктивно: по КС-грамматике G, порождающей язык K, и детерминированному конечному автомату (ДКА) A, распознающему регулярный язык M, строится новая КС-грамматика H, порождающая их пересечение.

1. **Нетерминалы новой грамматики**: Представляют собой тройки вида (q, X, r), где q и r — состояния автомата, а X — символ исходной грамматики (терминал или нетерминал).
2. **Аксиома**: Новой аксиомой становится тройка (q_0, S, f), где q_0 — начальное состояние автомата, S — аксиома исходной грамматики, а f — заключительное состояние автомата.
3. **Правила вывода**:

- Если в исходной грамматике есть правило $A \to X_1 X_2 \dots X_k$, то в новую грамматику добавляются правила, моделирующие все возможные переходы автомата через эти символы: $(q, A, r) \to (q, X_1, p_1)(p_1, X_2, p_2) \dots (p_{k-1}, X_k, r)$ для всех комбинаций состояний.
- Для терминалов добавляются правила вида $(q, a, r) \to a$, если в автомате существует переход из состояния q в r по символу a.

Таким образом, вывод в грамматике H возможен тогда и только тогда, когда цепочка одновременно выводима в исходной КС-грамматике и «прочитывается» автоматом, переходя из начального состояния в конечное.

Контекст и значение

Важно понимать, что в общем случае класс КС-языков **не замкнут** относительно пересечения. Например, если взять два КС-языка $L_1 = a^n b^n a^m$ и $L_2 = a^m b^n a^n$, то их пересечением будет язык $a^n b^n a^n$, который, согласно теореме о накачке, не является контекстно-свободным. Однако пересечение с регулярным языком всегда сохраняет свойство контекстной свободности.

Практическое применение

Эта теорема часто используется для доказательства того, что какой-либо язык **не является контекстно-свободным**. Если при пересечении исследуемого языка с заведомо регулярным языком получается язык, не являющийся КС-языком (что можно проверить по лемме о накачке), то и исходный язык не был контекстно-свободным.