

Задача

Дано регулярное выражение α в обратной польской записи, задающее язык L , и слово $u \in \{a, b, c\}^*$. Найти длину самого длинного префикса u , принадлежащего L .

Описание алгоритма

Построим по регулярному выражению α недетерминированный конечный автомат $M = \langle Q, \Sigma, \Delta, q_0, F \rangle$.

Пусть $S = \{ \langle q, k \rangle \in Q \times \{0, 1, \dots, |u|\} \mid \exists w \in \Sigma^k : \langle q_0, w \rangle \vdash \langle q, \varepsilon \rangle \}$. Тогда $\langle q, k \rangle \in S \iff (k = 0 \wedge q = q_0) \vee (\exists a \in \Sigma \exists q_1 \in S : \langle q_1, k-1 \rangle \in S \wedge (\langle q_1, a \rangle \rightarrow q) \in \Delta)$.

Ответ – $\max\{k \mid \exists q \in F : \langle q, k \rangle \in S\}$. S можно найти с помощью поиска в ширину за время $O(|Q||u|) = O(|\alpha||u|)$.