Автомат Антимирова

Лучшая команда разработчиков по ТФЯ

2022 г.

Частичные производные

$\alpha_c(R)$ — это регулярное выражение R' такое, что если $w \in \mathscr{L}(R')$, то $cw \in \mathscr{L}(R)$. Обратное не обязательно выполняется. Вычислить частичные производные можно по следующему рекурсивному алгоритму.

$$\alpha_c(c) = \{\varepsilon\}$$

$$\alpha_c(c') = \varnothing$$

$$\alpha_c(\varepsilon) = \varnothing$$

$$\alpha_c(r_1 r_2) = \begin{cases} \{r r_2 | r \in \alpha_c(r_1)\} \cup \alpha_c(r_2) & \text{если} \varepsilon \in \mathscr{L}(r_1) \\ \{r r_2 | r \in \alpha_c(r_1)\} & \text{иначе} \end{cases}$$

$$\alpha_c(\perp) = \varnothing$$

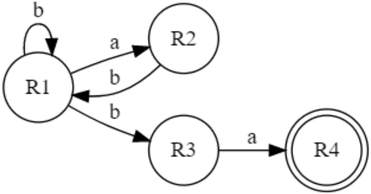
$$\alpha_c(r_1 | r_2) = \alpha_c(r_1) \cup \alpha_c(r_2)$$

$$\alpha_c(r^*) = \{r' r^* | r' \in \alpha_c(r)\}$$

Автомат Антимирова аналогичен автомату Брзозовски, но состояния представляют собой элементы α_w, а не δ_w. Упрощать по ACI состояния не требуется — их множество и так конечно.

Пример автомата Антимирова

Положим $\mathsf{R}_1 = (\mathtt{ab} \mid \mathtt{b})^*\mathtt{ba}$.
Тогда (производные, равные пустому множеству, здесь опущены):
$\alpha_{\mathtt{a}}(\mathsf{R}_1) = \{\mathtt{b}(\mathtt{b} \mid \mathtt{b})^*\mathtt{ba}\}$ – положим $\mathsf{R}_2 = \mathtt{b}(\mathtt{ab} \mid \mathtt{b})^*\mathtt{ba}$
$\alpha_{\mathtt{b}}(\mathsf{R}_1) = \{(\mathtt{ab} \mid \mathtt{b})^*\mathtt{ba}, \mathtt{a}\}$ – положим $\mathsf{R}_3 = \mathtt{a}$
$\alpha_{\mathtt{b}}(\mathsf{R}_2) = \{(\mathtt{ab} \mid \mathtt{b})^*\mathtt{ba}\}$ – тут ничего нового
$\alpha_{\mathtt{a}}(\mathsf{R}_3) = \{\varepsilon\}$ — положим $\mathsf{R}_4 = \varepsilon$
Соответствующий автомат имеет состояния R_i и один недетерминированный переход.



Дополнительные сведения

Связь с автоматом Томпсона

Автомат Антимирова может быть получен из автомата Томпсона путем последовательного применения к нему следующих операций:

$$\mathtt{RemEps}(\mathtt{DeAnnote}(\mathtt{Minimize}(\mathtt{RemEps}(\mathtt{Annote}(\mathtt{Thompson}(r))))))$$

Теорема

Пусть r – взвешенное регулярное выражение над K. Если K является $null - k$-замкнутым для r, то $\mathtt{Antimirov}(r)$ может быть вычислен за $O(m \log m + mn)$ путем применения удаления ε-переходов и минимизации.

Мы будем говорить, что K является $null - k$-замкнутым для α, если $\exists k \geqslant 0$, такое, что для каждого подтерма β^* $null(\beta)^* = \oplus_{i=0}^{k} null(\beta)^i$.