

Езици и автомати

Иво Стратев

24 ноември 2017 г.

Регулярни езици

Основни

Нека Σ е азбука. Тогава \emptyset , $\{\varepsilon\}$, $\forall a \in \Sigma \{a\}$ - са регулярни езици

Операции

Обединение

Ако L_1 , L_2 са регулярни, то и $L_1 \cup L_2 = \{\omega \mid \omega \in L_1 \vee \omega \in L_2\}$ е регулярен.

Конкатенация

Ако L_1 , L_2 са регулярни, то и $L_1 \cdot L_2 = \{\omega_1 \omega_2 \mid \omega_1 \in L_1 \wedge \omega_2 \in L_2\}$ е регулярен.

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

$$\forall n \in \mathbb{N} \ L^{n+1} = L^n \cdot L$$

Звезда на Клини

Ако L е регулярен, то и $L^* = \bigcup_{n \in \mathbb{N}} L^n$ е регулярен.

Позитивна обвивка

Ако L е регулярен, то и $L^+ = L \cdot L^* = L \cdot \bigcup_{n \in \mathbb{N}} L^n = \bigcup_{n \in \mathbb{N}^+} L^n$ е регулярен.

Дефиниция

Един език е регулярен, ако се получава от основните с помощта на краен брой приложения на операциите: обединение, конкатенация и звезда на Клини.

Регулярни изрази

Базови

Символите \emptyset , ε , a са регулярни изрази ($\forall a \in \Sigma$)

"Производни"

Ако r_1 и r_2 са регулярни изрази, то и $r_1 + r_2$, $r_1 \cdot r_2$ и r_1^* също са регулярни изрази.

Език на регулярен израз

$$\mathcal{L}(\emptyset) = \emptyset$$

$$\mathcal{L}(\varepsilon) = \{\varepsilon\}$$

$$\forall a \in \Sigma \quad \mathcal{L}(a) = \{a\}$$

Нека r_1 и r_2 са регулярни изрази и $L_1 = \mathcal{L}(r_1)$, $L_2 = \mathcal{L}(r_2)$. Тогава:

$$L_1 \cup L_2 = \mathcal{L}(r_1) \cup \mathcal{L}(r_2) = \mathcal{L}(r_1 + r_2)$$

$$L_1 \cdot L_2 = \mathcal{L}(r_1) \cdot \mathcal{L}(r_2) = \mathcal{L}(r_1 \cdot r_2)$$

$$L_1^* = \mathcal{L}(r_1)^* = \mathcal{L}(r_1^*)$$

Недетерминирани крайни автомати

Определение

Недетерминиран краен автомат представлява:

$N = (\Sigma, Q, s, \Delta, F)$, където:

Σ - крайно множество от букви (азбука)

Q - крайно множество от състояния

$s \in Q$ - начално състояние

$\Delta : Q \times \Sigma \rightarrow 2^Q$ - функция на преходите. Тя е тотална. Ако за някоя двойка (q, a) няма преход в автомата, то $\Delta(q, a) = \emptyset$;

$F \subseteq Q$ - множество от финални състояния

Дефиниция за Δ^*

$$\Delta^* : Q \times \Sigma^* \rightarrow 2^Q :$$

$$\Delta^*(q, \varepsilon) = \{q\}$$

$$\Delta^*(q, a\alpha) = \bigcup_{p \in \Delta(q, a)} \Delta^*(p, \alpha)$$

Език на недетерминирания автомат $\Delta^*, \mathcal{L}(N)$

$$\mathcal{L}(N) = \{\omega \in \Sigma^* \mid \Delta^*(s, \omega) \in F\}$$

"Действия" с недетерминирани автомати

Нека Σ е азбука.

Нека $N_1 = (\Sigma, Q_1, s_1, \Delta_1, F_1)$ и $N_2 = (\Sigma, Q_2, s_2, \Delta_2, F_2)$ - К.Н.А и $Q_1 \cap Q_2 = \emptyset$

Нека $N = (\Sigma, Q, s, \Delta, F)$

Конкатенация, $\mathcal{L}(N) = \mathcal{L}(N_1) \cdot \mathcal{L}(N_2)$

$$Q = Q_1 \cup Q_2$$

$$s = s_1$$

$$\Delta(q, a) = \begin{cases} \Delta_1(q, a), & q \in Q_1 \setminus F_1 \\ \Delta_2(q, a), & q \in Q_2 \\ \Delta_1(q, a) \cup \Delta_2(s_2, a), & q \in F_1 \end{cases}$$

$$F = \begin{cases} F_1 \cup F_2, & s_2 \in F_2 \\ F_2, & s_2 \notin F_2 \end{cases}$$

Обединение, $\mathcal{L}(N) = \mathcal{L}(N_1) \cup \mathcal{L}(N_2)$

$$s \notin Q_1 \cup Q_2$$

$$Q = Q_1 \cup Q_2 \cup \{s\}$$

$$\Delta(q, a) = \begin{cases} \Delta_1(q, a), & q \in Q_1 \\ \Delta_2(q, a), & q \in Q_2 \\ \Delta_1(s_1, a) \cup \Delta_2(s_2, a), & q = s \end{cases}$$

$$F = \begin{cases} F_1 \cup F_2 \cup \{s\}, & s_1 \in F_1 \vee s_2 \in F_2 \\ F_1 \cup F_2, & s_1 \notin F_1 \wedge s_2 \notin F_2 \end{cases}$$

Позитивна обвивка, $\mathcal{L}(N) = \mathcal{L}(N_1)^+$

$$Q = Q_1$$

$$s = s_1$$

$$\Delta(q, a) = \begin{cases} \Delta_1(q, a), & q \in Q_1 \setminus F_1 \\ \Delta_1(q, a) \cup \Delta_1(s, a), & q \in F_1 \end{cases}$$

$$F = F_1$$

Звезда на Клини, $\mathcal{L}(N) = \mathcal{L}(N_1)^* = \{\varepsilon\} \cup \mathcal{L}(N_1)^+$

$$E = (\Sigma, \{s_\varepsilon\}, s_\varepsilon, \Delta_\varepsilon, \{s_\varepsilon\})$$

$$s_\varepsilon \notin Q_1$$

$$\Delta_\varepsilon(s_\varepsilon, \varepsilon) = \{s_\varepsilon\}$$

$$\forall a \in \Sigma \quad \Delta_\varepsilon(s_\varepsilon, a) = \emptyset$$

$$Q = Q_1 \cup \{s_\varepsilon\}$$

$$s = s_\varepsilon$$

$$\Delta(q, a) = \begin{cases} \Delta_1(q, a), & q \in Q_1 \setminus F_1 \\ \Delta_1(q, a) \cup \Delta_1(s, a), & q \in F_1 \\ \Delta_1(s, a), & q = s_\varepsilon \end{cases}$$

$$F = F_1 \cup \{s_\varepsilon\}$$

Лема за покачването (Pumping Lemma)

Нека $L \subseteq \Sigma^*$ е безкраен регулярен език.

$$\exists p \in \mathbb{N}^+ : \forall \omega \in L : |\omega| \geq p, \exists x, y, z \in \Sigma^* :$$

$$\omega = xyz \wedge |xy| \leq p \wedge |y| \geq 1 \wedge (\forall i \in \mathbb{N} \ xy^i z \in L)$$

Следствие (Контрапозиция на лемата за покачването)

Нека $L \subseteq \Sigma^*$ е безкраен език.

Ако е изпълнено, че:

$$\forall p \in \mathbb{N}^+ \exists \omega \in L : |\omega| \geq p, \forall x, y, z \in \Sigma^* :$$

$$\omega = xyz \wedge |xy| \leq p \wedge |y| \geq 1 \wedge (\exists i \in \mathbb{N} xy^i z \notin L).$$

Тогава L не е регулярен.

Релация на Майхил-Нероуд и минимален автомат

Релация на Майхил-Нероуд

Нека Σ е азбука и $L \subseteq \Sigma^*$

$$R_L = \{(x, y) \in \Sigma^* \times \Sigma^* \mid \forall z \in \Sigma^* xz \in L \iff yz \in L\}$$

$$\Sigma^*/R_L = \{[\alpha]_{R_L} \mid \alpha \in \Sigma^*\}$$

Теорема за съществуване на минимален автомат

Ако съществува $n \in \mathbb{N} : n = |\Sigma^*/R_L|$. Тогава L е регулярен и съществува минимален Д.К.А $M = (\Sigma, Q_M, s_M, \delta_M, F_M)$, такъв че $\mathcal{L}(M) = L$.

$$Q_M = \Sigma^*/R_L$$

$$s_M = [\varepsilon]_L$$

$$\delta_M([\alpha]_{R_L}, a) = [\alpha.a]_L$$

$$F_M = \{[\alpha]_{R_L} \mid \alpha \in L\} = \{q \in Q_M \mid q \subseteq L\}$$

Лема δ_M е коректно дефинирана (δ_M задава функция)

$$\text{Лема } \delta_M^*([\varepsilon]_{R_L}, \omega) = [\omega]_{R_L}$$

$$\text{Лема } \mathcal{L}(M) = L$$

Еквивалентност на състояния на автомат и минимален автомат

Нека Σ е азбука и $A = (\Sigma, Q, s, \delta, F)$ е свързан, тотален Д.К.А

Еквивалентност на състояния на автомат

$$R_{\equiv} = \{(p, q) \in Q \mid \forall \omega \in \Sigma^* \delta^*(p, \omega) \in F \iff \delta^*(q, \omega) \in F\}$$

Минимален Д.К.А построен по състояния на класовете на еквивалентност на R_{\equiv}

$$\text{Нека } A_{\equiv} = (\Sigma, Q_{\equiv}, s_{\equiv}, \delta_{\equiv}, F_{\equiv})$$

$$Q_{\equiv} = Q/R_{\equiv} = \{[q]_{R_{\equiv}} \mid q \in Q\}$$

$$s_{\equiv} = [s]_{R_{\equiv}}$$

$$\delta_{\equiv}([q]_{R_{\equiv}}, a) = [\delta(q, a)]_{R_{\equiv}}$$

$$F_{\equiv} = \{[q]_{R_{\equiv}} \mid q \in Q, [q]_{R_{\equiv}} \subseteq F\} = \{[q]_{R_{\equiv}} \mid q \in Q, [q]_{R_{\equiv}} \cap F \neq \emptyset\}$$

Лема δ_{\equiv} е коректно дефинирана (δ_{\equiv} задава функция)

$$\text{Лема } \delta_{\equiv}^*([s]_{R_{\equiv}}, \omega) = [\delta^*(s, \omega)]_{R_{\equiv}} \implies \mathcal{L}(A_{\equiv}) = L$$

Лема A_{\equiv} е с минимален брой състояния

Задачи:

Ако $L \subseteq \{a, b\}^*$ е регулярен език.

То езикът $\{(ab)^n(ba)^k \mid n, k \in \mathbb{N}\}$ е регулярен, защото очевидно се описва от регулярния израз $(ab)^*(ba)^*$.

Ако D е краен език над $\{a, b\}^*$, то $L \cup \overline{D}$ винаги е регулярен, защото щом D е регулярен, то и \overline{D} е регулярен, следователно и $L \cup \overline{D}$ е регулярен.

Ако $K \subseteq \{a, b\}^*$ не е регулярен. То $L \cap K$ може да е не регулярен, но може и да е регулярен, например ако $L \cap K = \emptyset$ то той ще е регулярен.

Използвана литература

Записки по "Езици, автомати, изчислимост" на главен асистент д-р Стефан Вътев от Факултет по Математика и Информатика на Софийски университет "Св. Климент Охридски"

Уводни записки от курса по Езици, автомати и изчислимост четен на спец. Информатика през зимния семестър на 2017г във ФМИ на СУ "Св. Климент Охридски" от доц. д-р Александра Соскова

Записки по темата "Регулярни езици" от курса по Езици, автомати и изчислимост четен на спец. Информатика през зимния семестър на 2017г във ФМИ на СУ "Св. Климент Охридски" от доц. д-р Александра Соскова