

Lenguajes Formales, Autómatas y Computabilidad

Expresiones Regulares

Segundo Cuatrimestre 2024

Bibliografía

Capítulo 3, *Introduction to Automata Theory, Languages and Computation*, J. Hopcroft, R. Motwani, J. Ullman, Second Edition, Addison Wesley, 2001.

En esta clase

- ▶ Definición de expresión regular
- ▶ Teorema: Para cada AFD M hay una expresión regular r tal que $r = \mathcal{L}(M)$.
- ▶ Teorema: Para cada expresión regular r hay un AFND- λ M con un solo estado final y sin transiciones de salida tal que $r = \mathcal{L}(M)$.

Definición

[Expresión regular] Dado un alfabeto Σ , una expresión regular denota un lenguaje sobre Σ :

- ▶ \emptyset es una expresión regular que denota el conjunto vacío \emptyset .
- ▶ λ es una expresión regular que denota el conjunto $\{\lambda\}$.
- ▶ para cada $a \in \Sigma$, a es una expresión regular que denota el conjunto $\{a\}$.
- ▶ si r y s denotan los lenguajes R y S entonces
 - $r \mid s$ denota $R \cup S$,
 - rs denota RS ,
 - r^* denota R^* , y
 - r^+ denota R^+ .

Ejemplo

- ▶ 00
- ▶ $(0 \mid 1)^*$
- ▶ $(0 \mid 1)^* 00 (0 \mid 1)^*$
- ▶ $(1 \mid 10)^*$
- ▶ $(0 \mid \lambda) (1 \mid 10)^*$

Definición (Autómata Finito No Determinístico con transiciones λ)

Un AFND- λ es una 5-upla $\langle Q, \Sigma, \delta, q_0, F \rangle$ donde

Q es conjunto de estados

Σ es el alfabeto de entrada

q_0 es estado inicial

F es conjunto de estados finales

$\delta : Q \times (\Sigma \cup \{\lambda\}) \rightarrow \mathcal{P}(Q)$.

Teorema

Dada una expresión regular r , existe un AFND- λ M con un solo estado final y sin transiciones a partir del mismo tal que $\mathcal{L}(M) = r$.

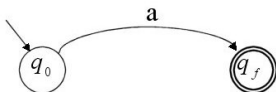
Demostración. Caso base: $r = \emptyset$,



Caso base: $r = \lambda$,



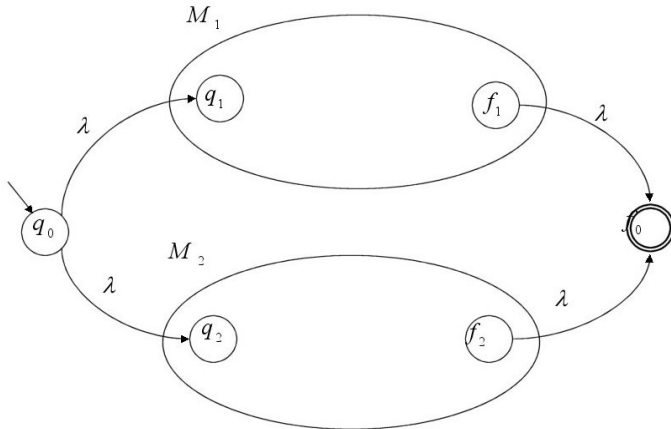
Caso base: $r = a$.



Caso inductivo: supongamos la expresión regular es $r_1|r_2, r_1r_2, r_1^*,$ ó r_2^+ y asumimos que vale la propiedad para r_1 y para r_2 .

Es decir, tanto para r_1 como para r_2 existen AFND- λ M_1 y M_2 con un solo estado final y sin transiciones a partir del mismo, tal que $\mathcal{L}(M_1) = r_1$ y $\mathcal{L}(M_2) = r_2$.

Caso $r = r_1 \mid r_2$. Por h.i. existen $M_1 = \langle Q_1, \Sigma_1, \delta_1, q_1, \{f_1\} \rangle$ y $M_2 = \langle Q_2, \Sigma_2, \delta_2, q_2, \{f_2\} \rangle$ tales que $\mathcal{L}(M_1) = r_1$ $\mathcal{L}(M_2) = r_2$. Sea $M = \langle Q_1 \cup Q_2 \cup \{q_0, f_0\}, \Sigma_1 \cup \Sigma_2, \delta, q_0, \{f_0\} \rangle$



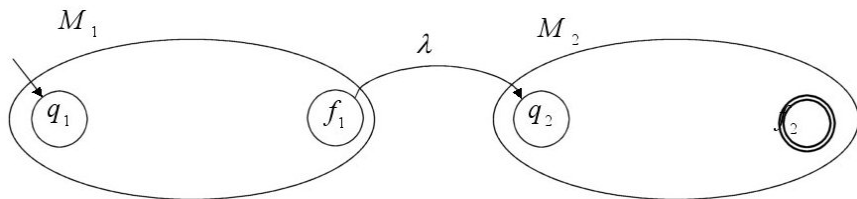
- ▶ $\delta(q_0, \lambda) = \{q_1, q_2\}$
- ▶ $\delta(q, a) = \delta_1(q, a)$ para $q \in Q_1 - \{f_1\}$ y $a \in \Sigma_1 \cup \{\lambda\}$
- ▶ $\delta(q, a) = \delta_2(q, a)$ para $q \in Q_2 - \{f_2\}$ y $a \in \Sigma_2 \cup \{\lambda\}$
- ▶ $\delta(f_1, \lambda) = \delta(f_2, \lambda) = \{f_0\}$.

Caso $r = r_1 r_2$.

Por h.i. existen $M_1 = \langle Q_1, \Sigma_1, \delta_1, q_1, \{f_1\} \rangle$ y $M_2 = \langle Q_2, \Sigma_2, \delta_2, q_2, \{f_2\} \rangle$, tales que $\mathcal{L}(M_1) = r_1$ y $\mathcal{L}(M_2) = r_2$ respectivamente.

Entonces podemos construir el autómata

$$M = \langle Q_1 \cup Q_2, \Sigma_1 \cup \Sigma_2, \delta, q_1, \{f_2\} \rangle$$



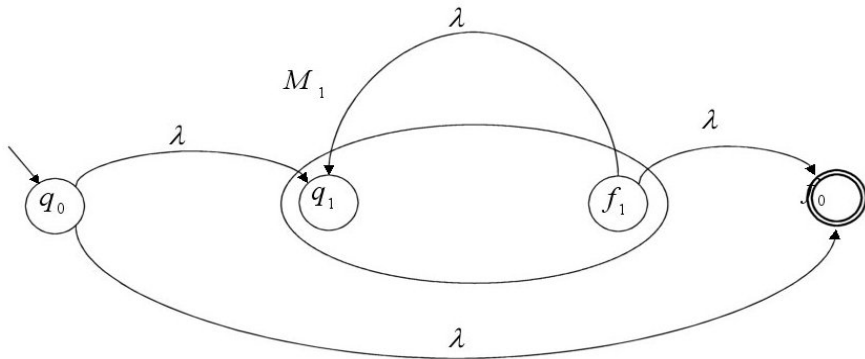
- ▶ $\delta(q, a) = \delta_1(q, a)$ para $q \in Q_1 - \{f_1\}$ y $a \in \Sigma_1 \cup \{\lambda\}$
- ▶ $\delta(f_1, \lambda) = \{q_2\}$
- ▶ $\delta(q, a) = \delta_2(q, a)$ para $q \in Q_2 - \{f_2\}$ y $a \in \Sigma_2 \cup \{\lambda\}$

Caso $r = r_1^*$.

Por h.i. existe $M_1 = \langle Q_1, \Sigma_1, \delta_1, q_1, \{f_1\} \rangle$, tal que $\mathcal{L}(M_1) = r_1$

Entonces podemos construir el autómata

$$M = \langle Q_1 \cup \{f_0, q_0\}, \Sigma_1, \delta, q_0, \{f_0\} \rangle$$



- $\delta(q, a) = \delta_1(q, a)$ para $q \in Q_1 - \{f_1\}$ y $a \in \Sigma_1 \cup \{\lambda\}$
- $\delta(q_0, \lambda) = \delta(f_1, \lambda) = \{q_1, f_0\}$.

Caso $r = r_1^+$.

Dado que $r_1^+ = r_1 r_1^*$, queda demostrado por los casos anteriores.

Indicar Verdadero o Falso, justificar

1. $r|s = s|r$
2. $(r^*)^* = r^*$
3. $\emptyset^* = \lambda$
4. $(r|s)^* = r^*|s^*$
5. $(rs|r)^*r = r(sr|r)^*$

Teorema

Dado un AFD $M = \langle \{q_1, \dots, q_n\}, \Sigma, \delta, q_1, F \rangle$ existe una expresión regular r tal que $\mathcal{L}(M) = r$.

Demostración. Denotemos con $R_{i,j}^k$ el conjunto de cadenas de Σ^* que llevan al autómata M desde el estado q_i al estado q_j pasando por estados cuyo índice es, a lo sumo, k . Definamos $R_{i,j}^k$ en forma recursiva:

$$R_{i,j}^k = R_{i,k}^{k-1} (R_{k,k}^{k-1})^* R_{k,j}^{k-1} \cup R_{i,j}^{k-1} \text{ para } k \geq 1$$

$$R_{i,j}^0 = \begin{cases} \{a : \delta(q_i, a) = q_j\}, a \in \Sigma & \text{si } i \neq j \\ \{a : \delta(q_i, a) = q_j\} \cup \{\lambda\}, a \in \Sigma & \text{si } i = j \end{cases}$$

Caso base: $k = 0$. Debemos dar r_{ij}^0 , tal que $r_{ij}^0 = R_{ij}^0$.

$R_{i,j}^0$ es el conjunto de cadenas de un solo caracter o λ .

Por lo tanto, $r_{i,j}^0$ es:

- ▶ $a_1 \mid \dots \mid a_p$, con a_1, \dots, a_p símbolos de Σ , si $\delta(q_i, a_s) = q_j$ para $s = 1, \dots, p$ y $q_i \neq q_j$.
- ▶ $a_1 \mid \dots \mid a_p \mid \lambda$, con a_1, \dots, a_p símbolos de Σ , si $\delta(q_i, a_s) = q_j$ para $s = 1, \dots, p$ y además $q_i = q_j$.
- ▶ \emptyset , si no existe ningún a_i que una q_i y q_j y $q_i \neq q_j$.
- ▶ λ , si no existe ningún a_i que una q_i y q_j y además $q_i = q_j$.

Caso inductivo. Por hipótesis inductiva, tenemos:

$$r_{ik}^{k-1} = R_{ik}^{k-1}, r_{kk}^{k-1} = R_{kk}^{k-1}, r_{kj}^{k-1} = R_{kj}^{k-1} \text{ y } r_{ij}^{k-1} = R_{ij}^{k-1}.$$

Definimos $r_{ij}^k = r_{ik}^{k-1} (r_{kk}^{k-1})^* r_{kj}^{k-1} \mid r_{ij}^{k-1}$ y verificamos que

$$\begin{aligned} r_{ij}^k &= \left(r_{ik}^{k-1} (r_{kk}^{k-1})^* r_{kj}^{k-1} \mid r_{ij}^{k-1} \right) \\ &= \left(r_{ik}^{k-1} (r_{kk}^{k-1})^* r_{kj}^{k-1} \right) \cup \left(r_{ij}^{k-1} \right) \\ &= \left(r_{ik}^{k-1} \right) \left(r_{kk}^{k-1} \right)^* \left(r_{kj}^{k-1} \right) \cup \left(r_{ij}^{k-1} \right) \\ &= R_{ik}^{k-1} \left(R_{kk}^{k-1} \right)^* R_{kj}^{k-1} \cup R_{ij}^{k-1} \\ &= R_{ij}^k. \end{aligned}$$

Entonces, como $Q = \{q_1, \dots, q_n\}$ y q_1 es el estado inicial de M ,

$$\begin{aligned} \mathcal{L}(M) &= R_{1j_1}^n \cup \dots \cup R_{1j_m}^n, \text{ con } F = \{q_{j_1}, \dots, q_{j_m}\} \\ &= r_{1j_1}^n \cup \dots \cup r_{1j_m}^n \\ &= r_{1j_1}^n \mid \dots \mid r_{1j_m}^n \end{aligned}$$

Concluimos $\mathcal{L}(M) = r_{1j_1}^n \mid \dots \mid r_{1j_m}^n$.