Lenguajes Formales, Autómatas y Computabilidad

Expresiones Regulares

Segundo Cuatrimestre 2025

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 expresion regular r tal que $r = \mathcal{L}(M)$.
- Teorema: Para cada expresion 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:

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

Un AFND-λ es una 5-upla $M = \langle Q, \Sigma, \delta, q_0, F \rangle$ donde
Q es conjunto de estados
Σ es el alfabeto de entrada
q_0 es estado inicial
F es conjunto de estados finales
$\delta : Q \times (\Sigma \cup \{\lambda\}) \to \mathcal{P}(Q)$.
El lenguaje aceptado por M

$$\mathcal{L}(M) = \left\{ x \in \Sigma^* : \widehat{\overline{\delta}}(q_0, x) \cap F \neq \phi \right\}.$$

donde
$\overline{\delta} : Q \times \Sigma \to \mathcal{P}(Q)$, $\overline{\delta}(q, a) = Cl_\lambda \Big(\bigcup_{p \in Cl_\lambda(q)} \delta(p, a) \Big)$
$Cl_\lambda(q) = \{p \in Q : (q, p) \in R^*\}$ con $R = \{(q, p) : p \in \delta(q, \lambda)\}$.
$Cl_\lambda(P) = \bigcup_{p \in P} Cl_\lambda(p)$.

$\widehat{\overline{\delta}} : Q \times \Sigma^* \to \mathcal{P}(Q), \quad x \in \Sigma^*, a \in \Sigma$
$\widehat{\overline{\delta}}(q, \lambda) = Cl_\lambda(q), \quad \widehat{\overline{\delta}}(q, xa) = \bigcup_{p \in \widehat{\overline{\delta}}(q, x)} \overline{\delta}(p, a).$

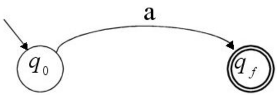
Teorema

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: $r = \emptyset$,

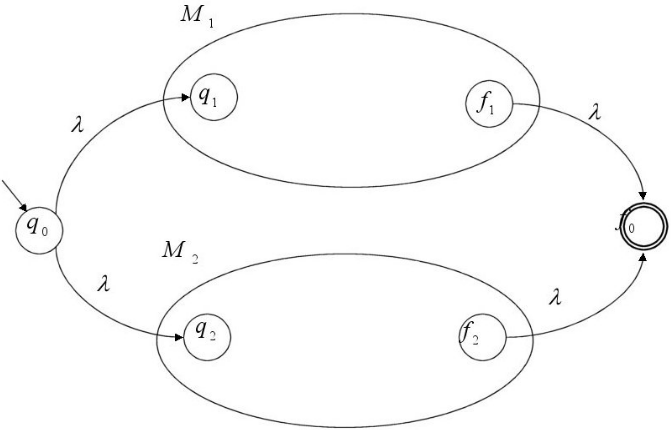
Caso base: $r = \lambda$,

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 $\mathcal{L}(M_2) = r_2$.

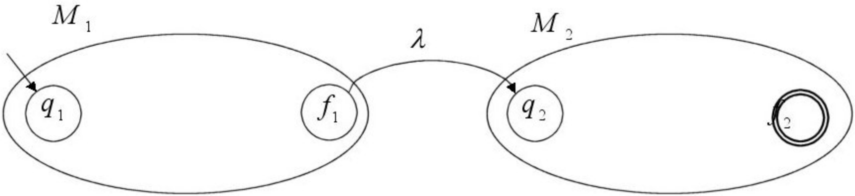
Caso $r = r_1 \mid r_2$. Por h.i. existen $M_1 = \langle Q_1, \Sigma_1, \delta_1, q_1, \{f_1\} \rangle$ y $M_2 = \langle Q_2, \Sigma_2, \delta_2, q_2, \{f_2\} \rangle$ tales que $\mathcal{L}(M_1) = r_1$ $\mathcal{L}(M_2) = r_2$. Sea $M = \langle Q_1 \cup Q_2 \cup \{q_0, f_0\}, \Sigma_1 \cup \Sigma_1,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



- $\delta(q_0, \lambda) = \{q_1, q_2\}$
- $\delta(q, a) = \delta_1(q, a)$ 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 $r = r_1 r_2$.

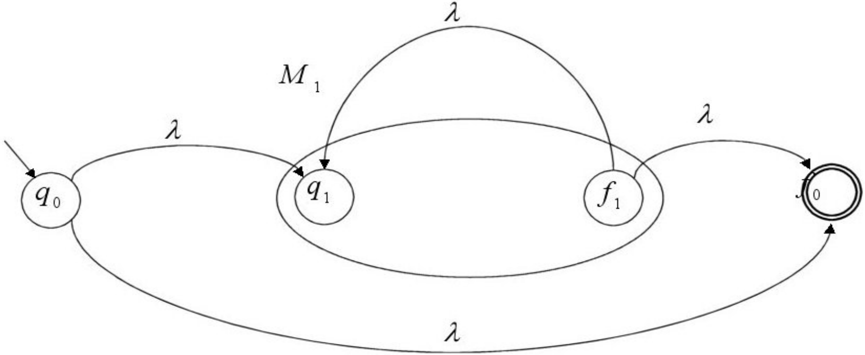
Por h.i. existen $M_1 = \langle Q_1, \Sigma_1, \delta_1, q_1, \{f_1\} \rangle$ y $M_2 = \langle Q_2, \Sigma_2, \delta_2, q_2, \{f_2\} \rangle$, tales que $\mathcal{L}(M_1) = r_1$y $\mathcal{L}(M_2) = r_2$ respectivamente.
Entonces podemos construir el AFND-λ
$M = \langle Q_1 \cup Q_2, \Sigma_1 \cup \Sigma_2, \delta, q_1, \{f_2\} \rangle$



- $\delta(q, a) = \delta_1(q, a)$ 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 AFND-λ
$M = \langle Q_1 \cup \{f_0, q_0\}, \Sigma_1, \delta, q_0, \{f_0\} \rangle$



- $\delta(q, a) = \delta_1(q, a)$ 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 $M = \langle \{q_1, \ldots, q_n\}, \Sigma, \delta, q_1, F \rangle$ *existe una expresión regular* r *tal que* $\mathcal{L}(M) = r$*.*

Demostración.
Sea AFD $M = (Q, \Sigma, \delta, q_1, F)$ donde $Q = \{q_1, \dots q_n\}$ $F = q_{j_1}, \dots q_{j_m}$.
Denotemos con $R^k_{i,j}$ 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 $R^k_{i,j}$ en forma recursiva:

$$R^k_{i,j} = R^{k-1}_{i,k} \left(R^{k-1}_{kk}\right)^* R^{k-1}_{k,j} \cup R^{k-1}_{i,j} \text{ para } k \geq 1$$

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

Notemos que $\mathcal{L}(M) = R^n_{1,j_1} \cup \dots R^n_{1,j_m}$
Demostarremos que hay una expresión regualar para expresar $\mathcal{L}(M)$.

Demostratermos que para cada $R^k_{i,j}$ hay una expresión regular, por inducción en k.

Caso base: $k = 0$. Debemos dar r^0_{ij}, tal que $r^0_{ij} = R^0_{ij}$.
$R^0_{i,j}$ es el conjunto de cadenas de un solo caracter o λ.
Por lo tanto, $r^0_{i,j}$ es:

- $a_1 \mid \ldots \mid a_p$,con $a_1, \ldots, a_p$ símbolos de Σ, si $\delta\left(q_{i,} a_s\right) = q_j$ para $s = 1, \ldots, p$ y $q_i \neq q_j$.
- $a_1 \mid \ldots \mid a_p \mid \lambda$,con $a_1, \ldots, a_p$ símbolos de Σ, si $\delta\left(q_{i,} a_s\right) = q_j$ para $s = 1, \ldots, p$ y además $q_i = q_j$.
- $\varnothing$, 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,

$$(h.i) \qquad r_{i,j}^{k-1} = R_{ij}^{k-1},$$

para todos los i, j entre 1 y n.
Definimos

$$r_{ij}^{k} = r_{ik}^{k-1} \left(r_{kk}^{k-1}\right)^{*} r_{kj}^{k-1} \mid r_{ij}^{k-1}$$

y verificamos

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

Dado que $Q = \{q_1, \ldots q_n\}$ y q_1 es el estado inicial de M,

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

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

Ejercicios

1. Dar un algoritmo que decida si dos expresiones regulares son equivalentes, es decir, denotan el mismo conjunto.
2. Dada una expresión regular r, dar un algoritmo para producir la expresión regular del complemento del conjunto denotado por r.
3. ¿Se puede adaptar la demostración del último Teorema y partir de un AFND, en vez de partir de un AFD?