Lenguajes Formales, Autómatas y Computabilidad

Clase Teórica
Autómatas Finitos y Gramáticas Regulares

Segundo cuatrimestre 2024

Bibliografía

Capítulo 2, *Introduction to Automata Theory, Languages and Computation*, J. Hopcroft, R. Motwani, J. Ullman, Second Edition, Addison Wesley, 2001.

En esta clase

- Definición de autómata finito determinístico (AFD)
- Definición de autómata finito no-determinístico (AFND)
- Teorema: Para todo AFND hay un AFD que reconoce el mismo lenguaje.
- Teorema: Para cada gramática regular hay un AFND que acepta el lenguaje generado por la gramática,
- Teorema: Para cada AFD hay una gramática que genera el mismo lenguaje que el acpetado por el AFD.

Definición (autómata finito determinístico (AFD))

Es una 5-upla $\langle Q, \Sigma, \delta, q_0, F \rangle$ *donde*

- Q *es un conjunto finito de estados*
- Σ *el alfabeto de entrada*
- $\delta : Q \times \Sigma \to Q$ *es la función de transición*
- $q_0 \in Q$ *es el estado inicial*
- $F \subseteq Q$ *es el conjunto de estados finales*

Definición (función de transición generalizada $\widehat{\delta}$)

Definimos $\widehat{\delta} : Q \times \Sigma^* \to Q$,

- $\widehat{\delta}(q, \lambda) = q$
- $\widehat{\delta}(q, xa) = \delta\left(\widehat{\delta}(q, x), a\right)$, *con* $x \in \Sigma^*$ *y* $a \in \Sigma$.

Notar que $\widehat{\delta}(q, a) = \delta\left(\widehat{\delta}(q, \lambda), a\right) = \delta(q, a)$.
Muchas veces usaremos símbolo δ para ambas funciones.

Definición (lenguaje aceptado por un AFD)

El lenguaje aceptado por un AFD $M = \langle Q, \Sigma, \delta, q_0, F\rangle$, *al que denotamos* $\mathcal{L}(M)$, *es el conjunto de cadenas de* Σ^* *aceptadas por* M,

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

Veremos a los autómatas finitos como funciones tales que para cada cadena dan un valor booleano: aceptación o no aceptación,

$$M : \Sigma^* \to \{0, 1\}$$

Autómata finito no determinístico

Definición (autómata finito no determinístico (AFND))

Es una 5-upla $\langle Q, \Sigma, \delta, q_0, F \rangle$ *donde*

- Q *es un conjunto finito de estados*
- Σ *el alfabeto de entrada*
- $\delta : Q \times \Sigma \to \mathcal{P}(Q)$
- $q_0 \in Q$ *es el estado inicial*
- $F \subseteq Q$ *es el conjunto de estados finales*

Definición (función de transición generalizada $\widehat{\delta}$)

Definimos $\widehat{\delta} : Q \times \Sigma^* \to \mathcal{P}(Q)$,

- $\widehat{\delta}(q, \lambda) = \{q\}$
- $\widehat{\delta}(q, xa) = \left\{ p \in Q : \exists r \in \widehat{\delta}(q, x) \text{ y } p \in \delta(r, a) \right\}$, *con* $x \in \Sigma^*$ *y* $a \in \Sigma$.

Notar que

$$\begin{aligned}
\widehat{\delta}\left(q, \lambda a\right) &= \left\{p \in Q : \exists r \in \widehat{\delta}\left(q, \lambda\right) \text{ y } p \in \delta\left(r, a\right)\right\} \\
&= \left\{p \in Q : \exists r \in \{q\} \text{ y } p \in \delta\left(r, a\right)\right\} \\
&= \left\{p \in Q : p \in \delta\left(q, a\right)\right\} \\
&= \delta\left(q, a\right).
\end{aligned}$$

Muchas veces utilizaremos el símbolo δ para ambas funciones.

Definición (lenguaje aceptado por un AFND)

El lenguaje aceptado por AFND $M = \langle Q, \Sigma, \delta, q_0, F\rangle$, al que denotamos $\mathcal{L}\left(M\right)$, es el conjunto de cadenas de Σ^* aceptadas por M,

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

Podemos extender la función de transición aún más, haciendo que mapee conjuntos de estados y cadenas en conjuntos de estados.

Definición (función de transición de conjuntos de estados)

Función de transición $\delta : \mathcal{P}(Q) \times \Sigma^* \to \mathcal{P}(Q)$ *dada por*

$$\delta(P, x) = \bigcup_{q \in P} \delta(q, x).$$

Es trivial ver que, para todo AFD existe un AFND equivalente.
Lo que no es tan obvio es que lo recíproco también es cierto:
para cada AFND existe un AFD equivalente.

Teorema (Equivalencia entre AFND y AFD (Rabin & Scott, 1959))

Dado un AFND $M = \langle Q, \Sigma, \delta, q_0, F \rangle$, *existe un AFD* $M' = \langle Q', \Sigma, \delta', q_0', F' \rangle$ *tal que* $\mathcal{L}(M) = \mathcal{L}(M')$.

Demostración del teorema

Construimos un AFD $M' = \langle Q', \Sigma, \delta', q_0', F' \rangle$

$Q' = \{[q_1, \ldots, q_i], q_1, \ldots, q_i \in Q\}$ (son los elementos de $\mathcal{P}(Q)$)

$F' = \{[q_1, \ldots, q_i] \in Q' : \{q_1, \ldots, q_i\} \cap F \neq \phi\}$

$q_0' = [q_0]$

$\delta'([q_1, \ldots, q_j], a) = [p_1, \ldots, p_i]$ si y solo si $\delta(\{q_1, \ldots, q_j\}, a) = \{p_1, \ldots, p_i\}$.

Demostremos que para toda cadena $x \in \Sigma^*$,
$\delta'(q_0', x) = [q_1, \ldots, q_i]$ si y solo si $\delta(q_0, x) = \{q_1, \ldots, q_i\}$.

Demostración por inducción en la longitud de la cadena.
Escribimos $|x|$ para la longitud de la cadena x.

Caso Base: $|x| = 0$, o sea $x = \lambda$.
Por definición de $\widehat{\delta}$,

$$\delta'(q_0', \lambda) = [q_0] \quad \text{y} \quad \delta(q_0, \lambda) = \{q_0\},$$

por lo que $\delta'(q_0', \lambda) = [q_0]$ si y solo si $\delta(q_0, \lambda) = \{q_0\}$.

Caso inductivo: suponemos que vale para x tal que $|x| = n$, es decir suponemos $\delta'(q'_0, x) = [p_1, \ldots, p_k]$ si y solo si $\delta(q_0, x) = \{p_1, \ldots, p_k\}$.

Veamos que vale para xa, con $a \in \Sigma$.

$$\delta'(q'_0, xa) = \delta'(\delta'(q'_0, x), a) = [r_1, \ldots, r_i]$$

si y solo si

(por definición de δ' en AFD M')

$\exists [p_1, \ldots, p_k]$,

$\delta'(q'_0, x) = [p_1, \ldots, p_k]$ y $\delta'([p_1, \ldots, p_k], a) = [r_1, \ldots, r_i]$

si y solo si

(por HI y y por definición de δ en AFND M)

$\exists \{p_1, \ldots, p_k\}$,

$\delta(q_0, x) = \{p_1, \ldots, p_k\} \wedge \delta(\{p_1, \ldots, p_k\}, a) = \{r_1, \ldots, r_i\}$

si y solo si

por def δ en AFND M,

$\delta(q_0, xa) = \delta(\delta(q_0, x), a) = \{r_1, \ldots, r_i\}$

Concluimos,
$\delta'(q'_0, xa) = [r_1, \ldots, r_i]$ si y solo si $\delta(q_0, xa) = \{r_1, \ldots, r_i\}$.

Nos queda probar que $\mathcal{L}(M) = \mathcal{L}(M')$.

$$x \in \mathcal{L}(M)$$

si y solo si

$$\delta(q_0, x) = \{q_1, \ldots, q_i\} \wedge \{q_1, \ldots, q_i\} \cap F \neq \phi$$

si y solo si

$$\delta'(q_0', x) = [q_1, \ldots, q_i] \wedge [q_1, \ldots, q_i] \in F'$$

si y solo si

$$x \in \mathcal{L}(M').$$

□

Recordemos que una gramática $G = \langle V_N, V_T, P, S\rangle$ es regular si todas las producciones son de la forma $A \to \lambda$, o $A \to a$ o $A \to aB$

Teorema

Dada una gramática regular $G = \langle V_N, V_T, P, S\rangle$ *existe un AFND* $M = \langle Q, \Sigma, \delta, q_0, F\rangle$ *tal que* $\mathcal{L}(G) = \mathcal{L}(M)$.

Demostración del teorema

Definamos M de la siguiente manera:

- $Q = V_N \cup \{q_f\}$, para mayor claridad, llamaremos q_A al estado correspondiente al no terminal A
- $\Sigma = V_T$
- $q_0 = q_S$
- $q_B \in \delta(q_A, a) \Leftrightarrow A \rightarrow aB \in P$
- $q_f \in \delta(q_A, a) \Leftrightarrow A \rightarrow a \in P$
- $q_A \in F \Leftrightarrow A \rightarrow \lambda \in P$
- $q_f \in F$.

Lema

Para todo $w \in V_T^*$*, Si* $A \stackrel{*}{\Rightarrow} wB$ *entonces* $q_B \in \delta(q_A, w)$*.*

Demostración del Lema.

Por inducción en la longitud de w.
Caso base $|w| = 0$, es decir $w = \lambda$. Como $A \stackrel{*}{\Rightarrow} A$ y $q_A \in \delta(q_A, \lambda)$,

$$A \stackrel{*}{\Rightarrow} A \Leftrightarrow q_A \in \delta(q_A, \lambda).$$

Caso $|w| = n+1, n \geq 0$, es decir, $w = \alpha a$ con $\alpha = n$. Asumamos h.i. para longitud n, es decir, vale para α.

$$\begin{aligned} A \stackrel{*}{\Rightarrow} \alpha a B \quad & \Leftrightarrow \exists C \in V_N, A \stackrel{*}{\Rightarrow} \alpha C \wedge C \rightarrow aB \in P \\ & \Leftrightarrow \ \exists q_C \in Q, q_C \in \delta(q_A, \alpha) \wedge q_B \in \delta(q_C, a) \ \text{ por h.i.} \\ & \Leftrightarrow q_B \in \delta(\delta(q_A, \alpha), a) \\ & \Leftrightarrow q_B \in \delta(q_A, \alpha a) \end{aligned}$$

$\square$

Continuación de la demostración del Teorema

$$
\begin{aligned}
wa \in \mathcal{L}(G) \Leftrightarrow & S \overset{*}{\Rightarrow} wa \\
& \Leftrightarrow \left(\exists A \in V_N, S \overset{*}{\Rightarrow} wA \land A \to a \in P\right) \lor \\
& \quad \left(\exists B \in V_N, S \overset{*}{\Rightarrow} waB \land B \to \lambda \in P\right) \text{ (únicas dos formas)} \\
& \Leftrightarrow \left(\exists q_A \in Q, q_A \in \delta(q_S, w) \land q_f \in \delta(q_A, a)\right) \lor \\
& \quad \left(\exists q_B \in Q, q_B \in \delta(q_S, wa) \land q_B \in F\right) \text{ (por el Lema)} \\
& \Leftrightarrow q_f \in \delta(q_S, wa) \lor \left(\exists q_B \in Q, q_B \in \delta(q_S, wa) \land q_B \in F\right) \\
& \Leftrightarrow wa \in \mathcal{L}(M).
\end{aligned}
$$

$$
\begin{aligned}
\lambda \in \mathcal{L}(G) \Leftrightarrow & S \overset{*}{\Rightarrow} \lambda \\
& \Leftrightarrow S \to \lambda \in P \\
& \Leftrightarrow q_S \in F \\
& \Leftrightarrow \lambda \in \mathcal{L}(M).
\end{aligned}
$$

□

Teorema

Dado un AFD $M = \langle Q, \Sigma, \delta, q_0, F \rangle$ *existe una gramática regular* $G = \langle V_n, V_T, P, S \rangle$ *tal que* $\mathcal{L}(G) = \mathcal{L}(M)$.

Demostración del Teorema

Definimos gramática $G = (V_N, V_T, P, S)$, donde $V_N = Q$ y llamaremos A_p al no terminal correspondiente a $p \in Q$; $S = A_{q_0}$; $V_T = \Sigma$ y el conjunto P es:

$$\begin{aligned} A_p \to aA_q \in P \Leftrightarrow & \;\; \delta\left(p, a\right) = q \\ A_p \to a \in P \Leftrightarrow & \;\; \delta\left(p, a\right) = q \in F \\ S \to \lambda \in P \Leftrightarrow & \;\; q_0 \in F \end{aligned}$$

Lema

$\delta(p, w) = q$ *si y solo si* $A_p \overset{*}{\Rightarrow} wA_q$.

Demostración del Lema.

Por inducción en la longitud de w.
Para $w = \lambda$, es cierto que $\delta(p, \lambda) = p$ y que $A_p \overset{*}{\Rightarrow} A_p$, por lo tanto $\delta(p, \lambda) = p \Leftrightarrow A_p \overset{*}{\Rightarrow} A_p$.
Asumamos h.i. vale para α de longitud n, con $n \geq 0$, y veamos que vale para para $w = \alpha a$.

$$\begin{aligned}\delta(p, \alpha a) = q &\Leftrightarrow \exists r \in Q, \delta(p, \alpha) = r \wedge \delta(r, a) = q \\ &\Leftrightarrow \exists A_r, A_p \overset{*}{\Rightarrow} \alpha A_r \wedge A_r \to aA_q \in P \text{ por h.i.} \\ &\Leftrightarrow A_p \overset{*}{\Rightarrow} \alpha a A_q.\end{aligned}$$

$\square$

Continuación de la demostración del Teorema

Asumamos Lema: $\delta(p, w) = q$ si y solo si $A_p \overset{*}{\Rightarrow} wA_q$.

$$\begin{aligned}
wa \in \mathcal{L}(M) \Leftrightarrow & \delta(q_0, wa) \in F \\
\Leftrightarrow & \exists p \in Q, \delta(q_0, w) = p \wedge \delta(p, a) \in F \\
\Leftrightarrow & \exists A_p, A_{q_0} \overset{*}{\Rightarrow} wA_p \wedge A_p \to a \in P \\
\Leftrightarrow & A_{q_0} \overset{*}{\Rightarrow} wa \\
\Leftrightarrow & wa \in \mathcal{L}(G) \\
\lambda \in \mathcal{L}(M) \Leftrightarrow & \ \ q_0 \in F \\
\Leftrightarrow & S \to \lambda \in P \\
\Leftrightarrow & S \overset{*}{\Rightarrow} \lambda \\
\Leftrightarrow & \lambda \in \mathcal{L}(G).
\end{aligned}$$

□

Ejercicios

1. Dado un AFD $M = \langle Q, \Sigma, \delta, q_0, F \rangle$, y sea a in símbolo de Σ. Construir otro AFD M' que acepte el lenguaje $L = \{ax \in \Sigma^* : x \in \mathcal{L}(M)\}$.
2. Indicar Verdadero o Falso y justificar

   Si $M = \langle Q, \Sigma, \delta, q_0, F \rangle$ es un AFD entonces reconoce al menos $|Q|$ palabras distintas, es decir $\#\mathcal{L}(M) \geq |Q|$.

   Si $M = \langle Q, \Sigma, \delta, q_0, F \rangle$ es AFND entonces todas las palabras de $\mathcal{L}(M)$ tienen longitud menor or igual que $|Q|^2$.
3. ¿Cuántos AFD hay con $|Q| = 2$ y $|\Sigma| = 3$?
4. ¿qué pasa si revierto todas las flechas de un AFD ?
5. ¿qué pasa si invierto estados finales con no finales de un AFND?