

Lenguajes Formales, Autómatas y Computabilidad

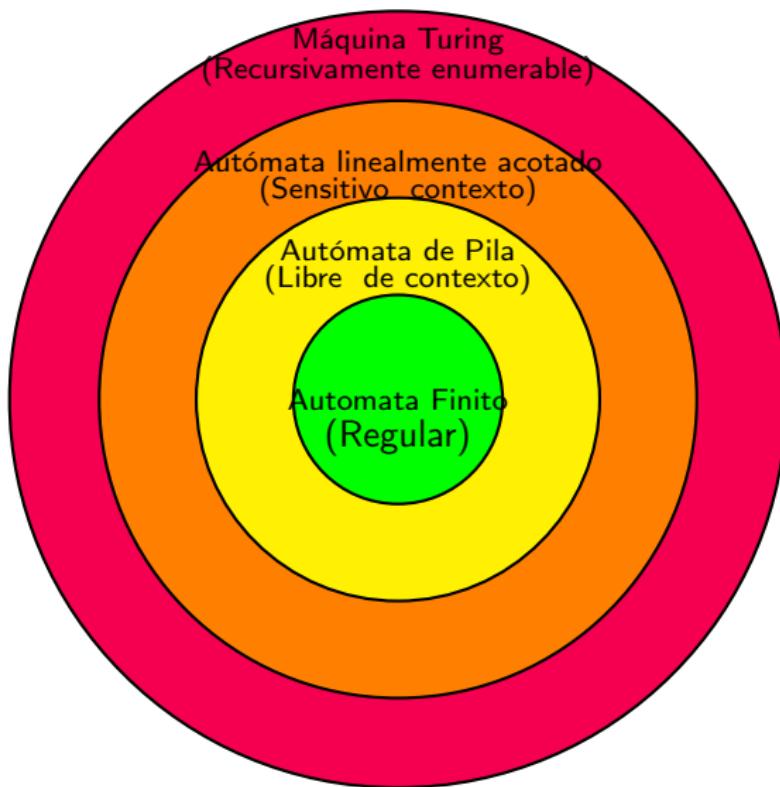
Clase Teórica
Autómatas Finitos

Segundo cuatrimestre 2025

Bibliografía

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

Jerarquía de Autómatas y los lenguajes que aceptan



Alfabetos y Lenguaje

Un alfabeto es un conjunto finito, no vacío, de símbolos.

Consideramos la concatenación de símbolos.

Una palabra sobre el alfabeto Σ es la concatenación de los elementos de una secuencia finita de símbolos. También la llamamos cadena o string.

La palabra nula λ . No tiene símbolos.

Ejemplos: Estas son algunas palabras sobre $\Sigma = \{a, j, r\}$,
 $\lambda, a, j, r, aa, aj, ar, ja, jj, aaj, rja, raja, jarra$, etc.

El conjunto de todas las palabras sobre un alfabeto

Dado alfabeto Σ , escribimos:

$$\Sigma^0 = \{\lambda\}$$

$$\Sigma^1 = \Sigma$$

$$\Sigma^2 = \Sigma\Sigma = \{ab : a \in \Sigma, b \in \Sigma\}$$

$$\Sigma^3 = \Sigma\Sigma\Sigma = \{abc : a \in \Sigma, b \in \Sigma, c \in \Sigma\}$$

...

La clausura de Kleene del alfabeto Σ

$$\Sigma^* = \bigcup_{i \geq 0} \Sigma^i$$

La clasura positiva del alfabeto Σ ,

$$\Sigma^+ = \bigcup_{i \geq 1} \Sigma^i$$

Ejemplos:

Supongamos $\Sigma = \{a, b\}$. Entonces $\Sigma^2 = \{aa, ab, ba, bb\}$,

$\Sigma^3 = \{aaa, aab, aba, abb, baa, bab, bba, bbb\}$, ...

Σ^* es el conjunto de todas las palabras sobre el alfabeto *Sigma*, de todas las longitudes, 0, 1, 2,

El conjunto de todas las palabras sobre un alfabeto

¿Cuál es la cardinalidad del conjunto Σ^i ?

El conjunto Σ tiene $|\Sigma|$ elementos.

Para cada $i \geq 0$, Σ^i tiene $|\Sigma|^i$ palabras.

El conjunto Σ^* tiene una cantidad infinita numerable de palabras.

Hay tantas palabras como números naturales

Teorema

$|\Sigma^*|$ es igual al la cardinalidad de \mathbb{N} .

Un orden entre pares de elementos es una relación antisimétrica y transitiva.

Definición (Orden longitud-lexicográfico en Σ^*)

Definimos el orden $\prec \subset \Sigma^* \times \Sigma^*$:

Asumimos un orden lexicográfico entre los elementos del alfabeto. Lo extendemos a un lexicográfico entre todas las palabras de la misma longitud. Y ahora lo extendemos a un orden total de todas las palabras: Las palabras de menor longitud son menores que las de mayor longitud.

Por ejemplo para $\Sigma = \{a, b, c\}$,

$\lambda \prec a \prec b \prec c \prec aa \prec ab \prec ac \prec ba \prec bb \prec bc \prec ca \prec cb \prec cc \prec aaa \prec aab \prec aac \prec aba \prec \dots$

Demostración del Teorema.

Definimos una biyección $f : \mathbb{N} \rightarrow \Sigma^*$, $f(i) =$ la i -ésima palabra en el orden \prec longitud lexicográfico sobre Σ^* .

□

Lenguaje sobre un alfabeto

Un lenguaje L sobre un alfabeto Σ es un conjunto de palabras sobre Σ . Es decir, $L \subseteq \Sigma^*$.

Ejemplos:

\emptyset

$\{\lambda\}$ (notar que es distinto de \emptyset)

$\{0, 01, 011, 0111, 01111, \dots\}$, es un lenguaje sobre $\Sigma = \{0, 1\}$.

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
- ▶ $\delta : Q \times \Sigma \rightarrow Q$ 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 $\widehat{\delta} : Q \times \Sigma^* \rightarrow Q$,

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

Notar que $\widehat{\delta}(q, a) = \delta(\widehat{\delta}(q, \lambda), a) = \delta(q, a)$.

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 palabras 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 palabra dan un valor booleano: aceptación o no aceptación,

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

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
- ▶ $\delta : Q \times \Sigma \rightarrow \mathcal{P}(Q)$
- ▶ $q_0 \in Q$ es el estado inicial
- ▶ $F \subseteq Q$ es el conjunto de estados finales

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

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

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

Notar que

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

Definición (lenguaje aceptado por un AFND)

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

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

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

Definición (función de transición de conjuntos de estados)

Función de transición δ -extendida : $\mathcal{P}(Q) \times \Sigma \rightarrow \mathcal{P}(Q)$,

$$\delta\text{-extendida}(P, a) = \bigcup_{q \in P} \delta(q, a).$$

Y definimos $\widehat{\delta}$ -extendida : $\mathcal{P}(Q) \times \Sigma^ \rightarrow \mathcal{P}(Q)$,*

$$\widehat{\delta}\text{-extendida}(P, x) = \bigcup_{q \in P} \widehat{\delta}(q, x).$$

Es trivial ver que, para todo AFD existe un AFND que reconoce el mismo lenguaje.

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

Dado AFND $M = \langle Q, \Sigma, \delta, q_0, F \rangle$ donde

$\delta : Q \times \Sigma \rightarrow \mathcal{P}(Q)$, y su extensión es δ -extendida : $\mathcal{P}(Q) \times \Sigma \rightarrow \mathcal{P}(Q)$, construimos un AFD $M' = \langle Q', \Sigma, \delta', q'_0, F' \rangle$, con $\delta' : Q' \times \Sigma \rightarrow Q'$,

$$Q' = \mathcal{P}(Q)$$

$$\delta' = \delta\text{-extendida}$$

$$q'_0 = \{q_0\}$$

$$F' = \{P \in Q' : P \cap F \neq \emptyset\}$$

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

Por definición AFD M' ,

$$x \in \mathcal{L}(M') \text{ si y solo si } \widehat{\delta}'(q'_0, x) \in F'$$

donde $F' = \{P \in Q' : P \cap F \neq \emptyset\}$

Y por definición de AFND M ,

$$x \in \mathcal{L}(M) \text{ si y solo si } \widehat{\delta}(q_0, x) \cap F \neq \emptyset.$$

Dado que $F' = \{P \in Q' : P \cap F \neq \emptyset\}$ tenemos

$S \in F'$ exactamente cuando $S \cap F \neq \emptyset$

Entonces basat con demostrar que para toda palabra $x \in \Sigma^*$,
 $\widehat{\delta}'(q'_0, x) = \widehat{\delta}(q_0, x)$.

Lo faremos por inducción en la estructura de la palabra x .

Caso Base: $x = \lambda$.

Por definición de $\widehat{\delta}'$ y $\widehat{\delta}$, $\widehat{\delta}'(q'_0, \lambda) = \{q_0\}$ y $\widehat{\delta}(q_0, \lambda) = \{q_0\}$,
por lo que $\widehat{\delta}'(q'_0, \lambda) = \widehat{\delta}(q_0, \lambda) = \{q_0\}$.

Caso inductivo

H1: suponemos que vale para $x \in \Sigma^*$ es decir suponemos
 $\widehat{\delta}'(q'_0, x) = \widehat{\delta}(q_0, x)$.

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

Por definición de $\widehat{\delta}'$ en AFD M' , existe $R \in Q'$ tal que

$$\widehat{\delta}'(q'_0, xa) = \delta'(\widehat{\delta}'(q'_0, x), a) = R$$

si y solo si

(por definición de $\widehat{\delta}'$ en AFD M')

$$\exists P, \widehat{\delta}'(q'_0, x) = P \text{ y } \delta'(P, a) = R$$

si y solo si

(por H1 y por definición de $\widehat{\delta}$ y δ -extendida en AFND M)

$$\exists P, \widehat{\delta}(q_0, x) = P \text{ y } \delta\text{-extendida}(P, a) = R$$

si y solo si

por def $\widehat{\delta}$ y δ -extendida en AFND M ,

$$\widehat{\delta}(q_0, xa) = \delta\text{-extendida}(\delta(q_0, x), a) = R.$$

Concluimos, $\delta'(q'_0, xa) = R$ si y solo si $\delta(q_0, xa) = R$.

□

Ejercicios

1. Dado un AFD $M = \langle Q, \Sigma, \delta, q_0, F \rangle$, y sea a un 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 o 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?