Teoría de Lenguajes

Clase Teórica 1

Autómatas Finitos

Primer 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.

Ingredientes básicos

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

Consideramos la concatenación de símbolos y formamos cadenas o palabras.

Una cadena sobre el alfabeto Σ es una secuencia finita de elementos (cero o más) de un alfabeto.

Cadena nula λ. No tiene símbolos.

Ejemplos: Estas son algunas cadenas sobre $\Sigma = \{a, j, r\}$, a, j ,aaj, rja, $raja$, $jarra$, etc.

Dado alfabeto Σ, definimos la clausura de Kleene del alfabeto Σ, que denotamos Σ^*, como el mínimo conjunto tal que

- $\lambda \in \Sigma^*$
- Si $(a \in \Sigma$ y $\alpha \in \Sigma^*)$ entonces $a\alpha \in \Sigma^*$

Definimos la clausura positiva de un alfabeto Σ, Σ^+ como el mínimo conjunto tal que si $(a \in \Sigma$ y $\alpha \in \Sigma^*)$ entonces $a\alpha \in \Sigma^+$.

Ejemplo: Sea $\Sigma = \{a, j, r\}$, entonces $arj \in \Sigma^*$ porque $j\lambda = j \in \Sigma^*$, $rj = rj \in \Sigma^*$, y $arj = arj \in \Sigma^*$.

Un lenguaje sobre un alfabeto Σ es un conjunto de cadenas sobre un alfabeto Σ.

Ejemplos:
$\emptyset$
$\{\lambda\}$ (notar que es distinto de $\emptyset$)
$\{0, 01, 011, 0111, 01111, \ldots\ldots\}$, 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*
- $\delta : Q \times \Sigma \to Q$ *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.

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

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

Definición (cadena aceptada por un AFD)

Una cadena $x \in \Sigma^*$ es aceptada por un AFD $M = \langle Q, \Sigma, \delta, q_0, F \rangle$ si y solo si $\widehat{\delta}(q_0, x) \in F$.

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,

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

Definición (autómata finito no determinístico (AFND))

Es una 5-upla $\langle Q, \Sigma, \delta, q_0, F\rangle$ *donde* Q, Σ, q_0 *y* F *tienen el mismo significado que para el AFD, pero* $\delta : Q \times \Sigma \to \mathcal{P}(Q)$.

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

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

Muchas veces utilizaremos el símbolo δ para ambas funciones.

Definición (cadena aceptada por un AFND)

Una cadena x *es aceptada por un AFND* $M = \langle Q, \Sigma, \delta, q_0, F \rangle$ *si y solo si* $\widehat{\delta}(q_0, x) \cap F \neq \phi$.

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 cadenas de* Σ^* *aceptadas por* M*,*

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

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

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

□

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?