

PRACTICA 7: Caracterización de lenguajes regulares

7.1. Introducción

El objetivo de esta práctica es diseñar un programa que permita caracterizar el lenguaje regular reconocido por un autómata finito en el sentido de determinar si se trata de un lenguaje vacío, finito o infinito.

7.2. El lenguaje universal es un conjunto infinito numerable

Dado un alfabeto arbitrario, Σ , el lenguaje universal sobre este alfabeto Σ^* es infinito numerable; es decir, se puede establecer una relación uno a uno entre \mathbb{N} , el conjunto de los números naturales y Σ^* .

Veamos cómo establecer esta biyección. Sobre Σ definimos una relación de orden total (una ordenación de los símbolos de Σ) a la que llamaremos orden lexicográfico (coloquialmente a veces se le conoce como orden alfabético), de modo que los símbolos de Σ estarán ordenados según esta relación: $\Sigma = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$. Las cadenas de Σ^* pueden ordenarse en orden ascendente de longitud: primero la cadena de longitud cero (ϵ), luego las cadenas de longitud uno (las que tienen un único símbolo) $(\sigma_1, \sigma_2, \dots, \sigma_n)$, a continuación las cadenas de longitud dos, y así sucesivamente. Las cadenas que tienen el mismo número de símbolos las ordenaremos utilizando el orden lexicográfico.

Por ejemplo si $\Sigma = \{a, b\}$ entonces $\Sigma^* = \{\epsilon, a, b, aa, ab, ba, bb, aaa, aab, \dots\}$

De este modo hemos establecido la biyección entre las cadenas de Σ^* y el conjunto \mathbb{N} , de modo que Σ^* es un conjunto enumerable.

7.3. Caracterización de lenguajes regulares

Consideremos un autómata finito $M \equiv (\Sigma, Q, q_0, F, \delta)$. Sea k el número de estados del autómata: $k = |Q|$.

Teorema 1 *El lenguaje que reconoce el autómata, $L(M)$ es distinto de vacío si y sólo si M reconoce una cadena w de longitud $|w| < k$*

Teorema 2 *El lenguaje que reconoce M es infinito si y sólo si M reconoce una cadena $w \in \Sigma^*$ tal que $k \leq |w| < 2k$*

7.4. Práctica

Dado un autómata finito M que se leerá desde un fichero (el autómata puede ser un DFA o un NFA), realizar un programa que caracterice $L(M)$ indicando si es vacío, finito o infinito. Para ello el programa generará en orden lexicográfico ascendente cadenas pertenecientes a Σ^* (se sobreentiende que el alfabeto Σ se deduce de la especificación del autómata en el fichero).

El programa leerá por línea de comandos el nombre del fichero que contiene la definición del autómata (NFA o DFA).

- Si $L(M)$ es \emptyset , el programa lo indicará.
- Si $L(M)$ no es vacío, el programa mostrará al menos una cadena perteneciente a $L(M)$.
- Si $L(M)$ es infinito, el programa lo indicará.