Propiedades de los lenguajes regulares

Lema del bombeo. El lema del bombeo permite mostrar que no todos los lenguajes son regulares.

Lema (del bombeo)

Sea L un lenguaje regular. Entonces existe n constante tal que $\forall w \in L$ con $|w| \geq n$ se obtiene que

$$w = uvx \text{ con } |v| \geq 1 \text{ y } |uv| \leq n$$

y además

$$\forall i, \quad uv^i x \in L.$$

Dem.

Si L es finito entonces existe m la mayor de las longitudes de las palabras de L. Se pone $n = m + 1$ y el lema se cumple por vacuidad (no hay palabras, en L, de longitud $\geq n$).

Si L es infinito, entonces, por el teorema de Kleene, existe M un AFD tal que $L = L(M)$. Sea

$$M = (Q, \Sigma, s, F, \delta).$$

Definimos $n = |Q|$ el número de estados de M. Si $w \in L$ con $|w| \geq n$ entonces

$$w = a_1 \cdots a_{|w|-1} a_{|w|}$$

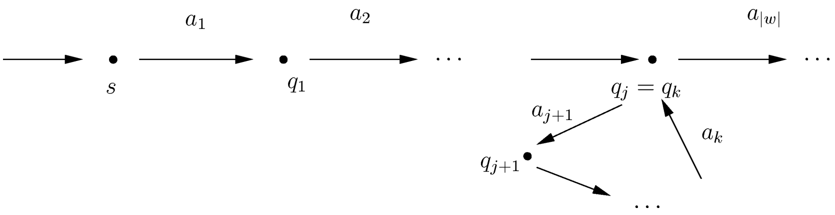
con cada $a_i \in \Sigma$. Definimos además $q_1, q_2, \ldots, q_{|w|}$ como los estados que resultan de las transiciones indicadas por w, esto es:

$$q_1 = \delta(s, a_1), q_2 = \delta(q_1, a_2), \ldots, q_{|w|} = \delta(q_{|w|-1}, a_{|w|}).$$

Puesto que $w \in L(M)$ entonces $q_{|w|} \in F$:

$$\longrightarrow \underset{s}{\bullet} \xrightarrow{a_1} \underset{q_1}{\bullet} \xrightarrow{a_2} \underset{q_2}{\bullet} \xrightarrow{a_3} \cdots \xrightarrow{a_{|w|}} \underset{q_{|w|}}{\bullet}$$

Como sólo hay n estados, los estados $s, q_1, \ldots, q_n, \ldots, q_{|w|}$ no pueden ser todos diferentes, pues en tal caso tendríamos $|w| + 1 > n$: más estados que n!. Por lo que algunos de éstos se repiten: existen j, k tales que $1 \leq j < k \leq n \leq |w|$ tales que $q_j = q_k$. Lo que obliga la aparición de un ciclo:



Definimos

- $u = a_1 \cdots a_j$
- $v = a_{j+1} \cdots a_k$
- $x = a_{k+1} \cdots a_{|w|}$

Entonces

$$w = uvx$$

Como v corresponde al ciclo, y este existe entonces $|v| > 1$. Además
$$|uv| = k \leq n.$$

Finalmente notemos que u lleva el estado inicial s al estado que se repite q_j, luego x continua con este estado hasta llevarlo al estado de aceptación $q_{|w|}$;

$$\begin{aligned} \delta(s, ux) &= \delta(\delta(s, u), x) \\ &= \delta(q_j, u) \\ &= \delta(q_k, x) \\ &= q_{|w|} \in F \end{aligned}$$

por lo que $ux \in L$. Similarmente, como v^2 lleva el estado q_j al estado
$$q_k$$
entonces uv^2x también se acepta. En general, si $i \geq 0$ entonces

$$\begin{aligned} \delta(s, uv^i x) &= \delta(\delta(\delta(s, u), v^i), x) \\ &= \delta(\delta(q_j, v^i), x) \\ &= \delta(q_k, x)) \\ &= q_{|w|} \in F, \end{aligned}$$

por lo que $uv^i x \in F$.

Ejemplo

Sea $\Sigma = \{0, 1\}$ y L el lenguaje de todas las cadenas con el mismo número de 0's que de 1's. Demostrar que L no es regular.

Dem. Por contradicción: si L fuera regular entonces existe n tal que se cumple el lema del bombeo. Esto es las palabras de longitud más grande o igual a n se descomponen en tres partes con ciertas características. En particular $q = 0^n 1^n \in L$ es una palabra de longitud $|w| = 2n > n$, luego tal se puede descomponer en tres partes u, v, x tales que

$$w = 0^n 1^n = uvx$$

con $|uv| \leq n$, $|v| > 1$ y $uv^i x \in L$, $\forall i \geq 0$. Como uv es la parte al principio de $0^n 1^n$ y $|uv| \leq n$ entonces uv está formada sólo por ceros, por lo que x, que ess la parte final de w, debe de tener todos los unos:

$$w = \underbrace{0 \cdots 0}_{uv} \underbrace{0 \cdots 0 1 \cdots 1}_{x}$$

Pero para $i = 0$, $ux \in L$, por lo que ux tiene n unos, de donde debe de tener n ceros que son los de u. Se deduce entonces que v no contribuye con ningún cero a $w = uvx$. Esto es $v = \epsilon$, lo cual contradice el elma del bombeo. Por lo tanto L no es regular.

Ejemplo

Sea

$$L = \{a^{i^2} \mid i \geq 1\}$$

entonces L no es un lenguaje regular.

Dem. Supongamos que L es regular. Existe una constante n tal que cualquier palabra $w \in L$ con $|w| \geq n$ se descompone según las características del lema del bombeo.

Por definición de L tenemos que $a^{n^2} \in L$. Luego, como $|a^{n^2}| = n^2 \geq n$, a^{n^2} se puede decomponer como

$$a^{n^2} = uvx$$

con $|uv| \leq n$ y $uv^ix \in L$, $\forall i \geq 0$.. En particular $uv^2x \in L$. Por lo que, para algún k entero $|uv^2x| = k^2$. Luego,

$$
\begin{aligned}
n^2 &= |a^{n^2}| \\
&= |uvx| \\
&< |uv^2x| \\
&= |u| + |v| + |v| + |x| \\
&= |uvx| + |v| \\
&\leq n^2 + n
\end{aligned}
$$

pues $n \geq |uv| \geq |v|$. Se sigue que

$$
\begin{aligned}
n^2 < |uv^2x| &\leq n^2 + n \\
&< n^2 + 2n + 1 \\
&= (n+1)^2
\end{aligned}
$$

entonces

$$
n^2 < \underbrace{|uv^2x|}_{k^2} < (n+1)^2
$$

esto es, $n^2 < k^2 < (n+1)^2$, i.e., $n < k < n+1$ con k entero: un absurdo.