El lema de bombeo de lenguajes regulares

Horst H. von Brand
vonbrand@inf.utfsm.cl

Departamento de Informática
Universidad Técnica Federico Santa María

Contenido

Lenguajes regulares

Las RE, los NFA y los DFA representan los mismos lenguajes, que llamaremos *regulares*. Como tenemos algoritmos que traducen entre las distintas representaciones, para demostraciones podemos suponer que un lenguaje regular nos lo dan en la representación que nos resulte más cómoda.

Quedó abierta la pregunta de si hay lenguajes no regulares.

El lema de bombeo

Teorema

Sea L un lenguaje regular. Entonces existe una constante $N > 0$ (el largo de bombeo) tal que toda palabra $\sigma \in L$, si $|\sigma| \geq N$ puede escribirse $\sigma = \alpha\beta\gamma$, donde $|\alpha\beta| \leq N$ y $\beta \neq \varepsilon$ tal que para todo $k \geq 0$ es $\alpha\beta^k\gamma \in L$.

El lema de bombeo

Demostración.

Sea $M = (Q, \Sigma, \delta, q_0, F)$ un DFA que acepta L. Llamemos $N = |Q|$, y sea $\sigma = a_1 a_2 \ldots a_n \in L$, con $n \geq N$. Consideremos los estados por los que pasa M al procesar $\sigma = a_1 a_2 \ldots a_n$.

$$\begin{array}{|ccccccccc|cccccc}
 & a_1 & & a_2 & & \ldots & & a_N & & a_{N+1} & & \ldots & & a_n & \\
q_0 & & q_1 & & q_2 & \ldots & q_{N-1} & & q_N & & q_{N+1} & \ldots & q_{n-1} & & q_n
\end{array}$$

Alguno de los estados $q_0, \ldots, q_N$ tiene que repetirse, o sea, hay α, β, γ con $\beta \neq \varepsilon$ y $|\alpha\beta| \leq N$ tales que $\delta(q_0, \alpha) = \delta(q_0, \alpha\beta)$ y $\delta(q_0, \alpha\beta\gamma) \in F$. Pero entonces $\delta(q_0, \alpha) = \delta(q_0, \alpha\beta^k)$ y $\delta(q_0, \alpha\beta^k\gamma) \in F$ para todo $k \geq 0$. $\square$

Uso del lema

El principal uso del lema de bombeo es demostrar que ciertos lenguajes no son regulares, por contradicción.

Un lenguaje no regular

Teorema

El lenguaje $L = \{a^k b^k : k \geq 0\}$ *no es regular.*

Un lenguaje no regular

Demostración.
Por contradicción. Supongamos L regular, por lo que cumple el lema de bombeo. Sea N la constante del lema, consideremos $\sigma = a^N b^N \in L$, con lo que $|\sigma| = 2N \geq N$. Sea cual sea la división $\alpha\beta\gamma$ con $|\alpha\beta| \leq N$, α y β están formados solo por a. Repetir $\beta \neq \varepsilon$ dos veces da una palabra con más a que b. Contradicción. $\square$

Esquema del uso del lema de bombeo

Queremos contradecir el lema. Si analizamos su enunciado, hay varios «existe...» (que deberemos demostrar que *ninguno sirve*) y varios «para todo...» (en los que podemos elegir libremente, basta encontrar un contraejemplo). Una forma de esquematizarlo es considerando un juego con un adversario, nosotros estamos a cargo de los $\forall$ (tenemos libertad de elección), el adversario maneja los $\exists$ (debemos conformarnos con lo que nos dé).

Esquema de uso del lema de bombeo

Sansano	Adversario
Propone L	
	Pone N, constante del lema
Elige $\sigma \in L$, $\lvert\sigma\rvert \geq N$	
	Divide $\sigma = \alpha\beta\gamma$, $\lvert\alpha\beta\rvert \leq N$, $\beta \neq \varepsilon$
Elige $k = 0$ o $k \geq 2$ tal que $\alpha\beta^k\gamma \notin L$	
	#&%@$0 sansano, me ganó de nuevo... ¡Ya verá la próxima!

Cuadrados

Teorema

El lenguaje $L = \{a^{k^2} : k \geq 0\}$ *no es regular.*

Demostración.

Por contradicción. Supongamos L regular, N la constante del lema de bombeo. Sea $\sigma = a^{N^2} \in L$, $|\sigma| = N^2 \geq N$. Sea $\sigma = \alpha\beta\gamma$ con $|\alpha\beta| \leq N$, $\beta \neq \varepsilon$. Llamemos $|\beta| = u$, así $0 < u \leq N$, $|\alpha\beta^2\gamma| = N^2 + u$. Por el lema, para algún u esto es un cuadrado perfecto. Pero:

$$N^2 < N^2 + u \leq N^2 + N < N^2 + 2N + 1 = (N+1)^2$$

Contradicción. $\square$

Números primos

Teorema

El lenguaje $L = \{a^p : \ p \text{ es primo}\}$ *no es regular.*

Demostración.

Por contradicción. Supongamos L regular, sea N la constante de bombeo. Sea $p \geq N$ primo. Consideremos $\sigma = a^p \in L$, $|\sigma| = p \geq N$. El lema asegura que existe una división $\sigma = \alpha\beta\gamma$ con $0 < |\beta| \leq N$ tal que $\alpha\beta^{p+1}\gamma \in L$ (estamos eligiendo $k = p + 1$). Sea $|\beta| = u$. Pero:

$$|\alpha\beta^{p+1}\gamma| = p + pu = p(u+1)$$

que es compuesto. Contradicción. $\square$

Más a que b

Teorema

El lenguaje $L_> = \{\omega \in \{a, b\}^* : \ \omega \text{ tiene más } a \text{ que } b\}$ *no es regular.*

Demostración.

Por contradicción. Sea N la constante del lema de bombeo, consideremos $\sigma = a^{N+1} b^N \in L_>$, $|\sigma| = 2N + 1 \geq N$. El lema asegura que hay una división $\sigma = \alpha\beta\gamma$ con $|\alpha\beta| \leq N$ (son solo a), $\beta \neq \varepsilon$ tal que $\alpha\gamma \in L$ (estamos eligiendo $k = 0$), pero esto tiene a lo más N veces a, contradicción. $\square$

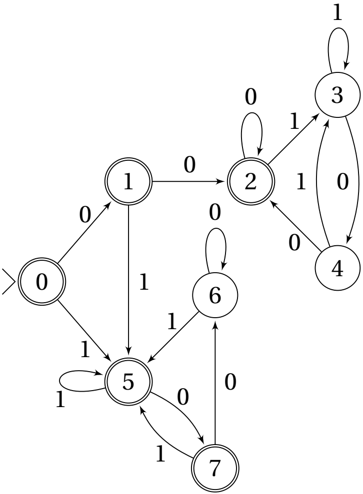
Cuidado

Vimos que los DFA «no pueden contar» (no pueden guardar más que finita información en sus estados). Pero no todo lenguaje descrito exigiendo diferencias o igualdades o con expresiones es no-regular.

Mismo número de veces 001 que 100

El lenguaje de palabras sobre $\{0, 1\}$ que contienen tantas veces 001 como 100 es regular.
Si lo intenta, verá que en una palabra solo puede sobrar o faltar un 100 (pueden haber traslapos, 11001 tiene un 100 y un 001). Lo que debemos verificar es si sobra una de alguna de las dos, cosa que un DFA puede hacer.

Mismo número de veces 001 que 100



Resumen

- El lema de bombeo da una condición que todo lenguaje regular cumple.
- Podemos usarlo por contradicción para demostrar que ciertos lenguajes no son regulares.
- Un resumen del lema de bombeo es que «autómatas finitos no pueden contar»
- Vimos un ejemplo, hay varios más en el apunte. Es un tema bastante popular en https://cs.stackexchange.com