El lema de bombeo y los lenguajes no regulares

Elvira Mayordomo

Universidad de Zaragoza

27 de octubre de 2014

Contenido de este tema

- ¿Son todos los lenguajes regulares?
- El lema de bombeo
- Cómo aplicar el lema de bombeo
- Usando las propiedades de clausura

Lenguajes regulares y lenguajes no regulares

- Ya sabemos que los lenguajes regulares tienen propiedades muy buenas (se pueden definir usando autómatas o e.r., se pueden comparar y simplificar, ...)
- Pero ¿cualquier lenguaje es regular?
- Un sospechoso de no serlo es el lenguaje $\{a^n b^n \mid n \in \mathbb{N}\}$
  Intenta pensar en un autómata o e.r. que lo represente
- En este tema vamos a ver una herramienta para distinguir que algunos lenguajes (por ejemplo $\{a^n b^n \mid n \in \mathbb{N}\}$) no son regulares
- Es el lema de bombeo

Lema de bombeo

- **Dado un lenguaje infinito** A
- **si** para todo N
  - existe una palabra w con $|w| \geq N$ y $w \in A$ tal que
  - para cualquier partición en tres trozos $w = xyz$ con $|xy| \leq N$, $|y| \geq 1$ existe un i,

$$xy^iz \notin A$$

- **entonces** A **no es regular**

Intuitivamente

- Dado un lenguaje infinito A
- si existe una palabra $w \in A$ todo lo larga que quiera
- tal que para cualquier partición w no se puede bombear
- entonces A no es regular

Ejemplo, $A = \{a^n b^n \mid n \in \mathbb{N}\}$

- Para todo N existe una palabra en A,
  - $w = a^N b^N \quad |w| = 2N \geq N$
- tal que para cualquier partición de w $\; x, y, z$ con $w = xyz$, $|y| \geq 1$ y $|xy| \leq N$,
  - la partición tiene que ser (con $s = |y| \geq 1$)

$$x = a^r \quad y = a^s, s \geq 1 \quad z = a^{N-r-s} b^N$$

- $\exists i$ con $xy^i z \notin A$
  - para $i = 2$,

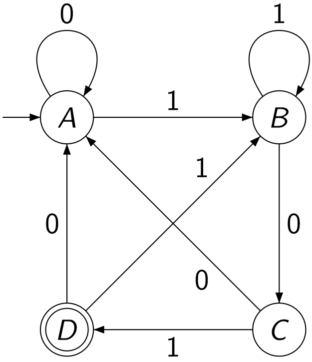
$$xy^2 z = a^r a^s a^s a^{N-r-s} b^N = a^{N+s} b^N$$

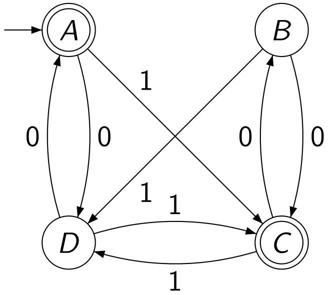
  y $a^{N+s} b^N \notin A$ ya que $s \geq 1$

- luego A no es regular

Justificación del lema de bombeo: Los bucles en los autómatas

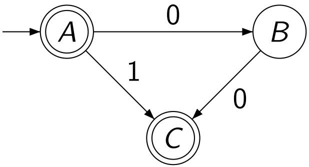
- Si tenemos un autómata vemos que puede haber bucles
- Un par de ejemplos





Los bucles en los autómatas

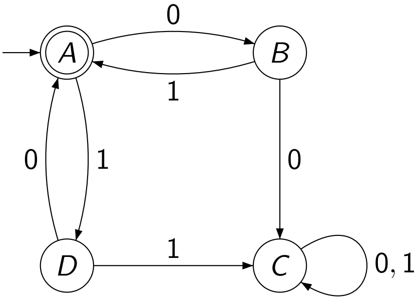
- Sólo cuando el autómata acepta un número finito de palabras podría no haber bucles



- Pero los lenguajes finitos no interesan, son todos regulares
- El lema de bombeo dice que si un lenguaje es regular, las palabras largas se pueden bombear, es decir, se puede alargar la palabra repitiendo un bucle

Los bucles en los autómatas

- Por ejemplo, una palabra aceptada por este autómata es 010110100101 con computación $(A, B, A, B, A, D, A, D, A, B, A, B, A)$



- Podemos repetir el bucle (desde A) B, A con subcadenas 01, por ejemplo obteniendo $01(01)^{100}10100101$ también aceptada

El lema de bombeo, versión fácil

- Dado un autómata de n estados M
- cualquier palabra w con $|w| \geq n$ y w aceptada por M ($w \in L(M)$)
- puede ser bombeada, es decir, existe una partición en tres trozos $w = xyz$, $|y| \geq 1$, tal que para todo i,

$$xy^iz \in L(M)$$

El lema de bombeo, versión fácil

Demostración

- La demostración es sencilla, si $|w| = m \geq n$ los estados por los que paso son
$$(q_0, q_1, \ldots q_m)$$
Aquí hay $m + 1 > n$ estados luego alguno de los n estados posibles está repetido, $q_a = q_b$, $a \neq b$
- El bucle va a ser $y = w_{a+1} \ldots w_b$ que es el fragmento que lleva de q_a a q_b
- $w = xyz$ con $x = w_1 \ldots w_a$, $z = w_{b+1} \ldots w_m$
- Si repetimos y:
$$xy^2z = w_1 \ldots w_a w_{a+1} \ldots w_b w_{a+1} \ldots w_b w_{b+1} \ldots w_m$$
llegamos al mismo estado q_m y aceptamos
- También aceptamos xy^3z, xy^4z, incluso xy^0z

El lema de bombeo, versión útil

- Dado un lenguaje infinito A
- si A es regular
- entonces existe un N tal que
  - para cualquier palabra w con $|w| \geq N$ y $w \in A$
  - existe una partición en tres trozos $w = xyz$ con $|y| \geq 1$, $|xy| \leq N$ tal que para todo i,
$$xy^i z \in A$$

Sólo he cambiado "lenguaje aceptado por un autómata" por "lenguaje regular" y $|xy| \leq n$ (porque para $y = w_{a+1} \dots w_b$, $b \leq n$ en la demostración anterior)

Regulares, no regulares

- Hemos visto que todos los lenguajes regulares se pueden buclear o bombear
- Lo que nos interesa es el converso, si un lenguaje no se puede bombear entonces no es regular

El lema de bombeo, versión muy útil

Lema de bombeo

- Dado un lenguaje infinito A
- si para todo N
  - existe una palabra w con $|w| \geq N$ y $w \in A$ tal que
  - para cualquier partición en tres trozos $w = xyz$ con $|xy| \leq N$, $|y| \geq 1$ existe un i,

$$xy^iz \notin A$$

- entonces A no es regular

Recordad, $\alpha \Rightarrow \beta$ es lo mismo que $\neg\beta \Rightarrow \neg\alpha$

El lema de bombeo, telegráfico

Lema de bombeo

- Dado un lenguaje infinito A
- si $\forall N\ \exists w$ con $w \in A$, $|w| \geq N$
  - tal que $\forall x, y, z$ con $w = xyz$, $|y| \geq 1$ y $|xy| \leq N$
  - $\exists i$ con $xy^iz \notin A$
- entonces A no es regular

Intuitivamente

- Dado un lenguaje infinito A
- si existe una palabra $w \in A$ todo lo larga que quiera
- tal que para cualquier partición w no se puede bombear
- entonces A no es regular

Ejemplo, $A = \{uu \mid u \in \{0,1\}^*\}$

- Para todo N existe una palabra en A,
  - $w = 0^N 1^N 0^N 1^N \quad |w| = 4N \geq N$
- tal que para cualquier partición de w x, y, z con $w = xyz$, $|y| \geq 1$ y $|xy| \leq N$,
  - la partición tiene que ser (con $s = |y| \geq 1$)

$$x = 0^r \quad y = 0^s, s \geq 1 \quad z = 0^{N-r-s} 1^N 0^N 1^N$$

- $\exists i$ con $xy^i z \notin A$
  - para $i = 2$,

$$xy^2 z = 0^r 0^s 0^s 0^{N-r-s} 1^N 0^N 1^N = 0^{N+s} 1^N 0^N 1^N$$

y $0^{N+s} 1^N 0^N 1^N \notin A$ ya que $s \geq 1$

- luego A no es regular

Resumen lema de bombeo

- Para demostrar que A no es regular
- Para cada N elegir $w \in A$ con $|w| \geq N$
- Ver cómo son todas las particiones de w que cumplen $w = xyz$, $|y| \geq 1$ y $|xy| \leq N$
  Hay que elegir w para que las particiones sean fáciles
- Para cada partición, encontrar i con $xy^iz \notin A$

Otras formas de ver que un lenguaje no es regular

Usando las propiedades de clausura:

- Si A y B son regulares entonces $A \cup B$ es regular.
  Si $A \cup B$ no es regular y B es regular entonces A no es regular.
- Si A y B son regulares entonces $A \cdot B$ es regular.
- Si A es regular entonces A^* es regular.
- Si A es regular entonces A^c es regular.
  Si A^c no es regular entonces A no es regular.
- Si A es regular entonces A^R es regular.
  Si A^R no es regular entonces A no es regular.
- Si A y B son regulares entonces $A \cap B$ es regular.
  Si $A \cap B$ no es regular y B es regular entonces A no es regular.

Ejemplo

$A = \{w \mid |w|_a = |w|_b\}$
Palabras con el mismo número de as que de bs

- $A \cap a^* b^* = \{a^n b^n \mid n \in \mathbb{N}\}$
- Hemos visto que $\{a^n b^n \mid n \in \mathbb{N}\}$ no es regular
- Sabemos que $a^* b^*$ es regular

- Luego A no es regular

Ejemplo

$A = \{w \mid |w|_a \neq |w|_b\}$
Palabras con distinto número de as que de bs

- $A^c = \{w \mid |w|_a = |w|_b\}$
- Hemos visto que $\{w \mid |w|_a = |w|_b\}$ no es regular
- Luego A no es regular

Bibliografía

- Kelley, sección 2.9.
- Sipser (2a edición), sección 1.4.