Лема за разрастване (за регулярни езици).

Лема за разрастване. (Л.Р.) Нека L е регулярен език. Тогава съществува естествено число $p > 0$, т.ч. за всяка дума $w \in L$, за която $|w| \geq p$ съществува разбиване $w = xyz$, изпълняващо:

(а) $|y| > 0$ (т.е. $y \neq \varepsilon$)

(б) $|xy| \leq p$

(в) за всяко $i \geq 0$, $xy^iz = x\underbrace{y \dots y}_{i}z \in L$

Означения. Ако L е рег. език и p е т.ч. за вс. дума w от L с $|w| \geq p$ същ. разбиване $w = xyz$, което изпълнява (а), (б) и (в), то p ще наричаме свидетел за регулярността на L.

Л.Р. е само необходимо условие за регулярността на един език. С други думи, съществуват нерегулярни езици, за които има $p > 0$ т.ч. всяка дума от езика по-дълга от p може да се разбие на 3 части, удовлетворяващи (а), (б) и (в).

Ще използваме Л.Р. в контрапозиция, за да показваме, че даден език не е регулярен. Именно, за да покажем, че езикът L не е регулярен, достатъчно е да покажем, че:

$$\forall p\, \exists w\, \forall x,y,z\, (w = xyz \;\&\; y \neq \varepsilon \;\&\; |xy| \leq p \Rightarrow \exists i\, (xy^i z \notin L))$$

Задачи:

Докажете, че езикът L не е регулярен, където:

0. $L = \{ a^n b^n \mid n \geq 0 \}$

<u>док.</u>: Да допуснем, че L е регулярен език и нека $p > 0$ е свидетел за това. Нека $W = a^p b^p = \underbrace{a \ldots a}_{p}\underbrace{b \ldots b}_{p}$. Тогава $w \in L$ и $|w| = 2p \geq p$.

Следователно, съществува разбиване $w = xyz$, за което: $y \neq \varepsilon$, $|xy| \leq p$ и $\forall i \geq 0\, (xy^i z \in L)$.

Да забележим, че xy е <u>начало</u> на w, което е с дължина не по-голяма от p; но първите p символа на w са <u>a</u>; следов., $xy \in a^*$ (състоят се само от a-та); в частност, $y \in a^*$; понеже $y \neq \varepsilon$, то $y \in a^+$ (y е непразна последователност от a-та).

$$w = a \ldots a\, b \ldots b$$

$$x \quad y \quad z$$

Тогава думата $w_0 = xy^0 z = xz = \underbrace{a \ldots a}_{|x|}\underbrace{a \ldots a}_{|z|_a}\underbrace{b \ldots b}_{p}$

Но $|x| + |y| + |z|_a = |x|_a + |y|_a + |z|_a = |w|_a = p$,

откъдето $|x| + |z|_a = p - |y|$.

Следователно $w_0 = x y^0 z = \underbrace{a^{p-|y|}}_{\text{всички a-та, без тези в y}} b^p$

Според лемата, за вс. $i \geq 0$, $xy^iz \in L$; в частност $xy^0z \in L$, т.е. $a^{p-|y|}b^p \in L$. Тогава

$p - |y| = p$, т.е. $|y| = 0$ – противоречие.

Така L не е регулярен.

<u>Упр.</u> Намерете как изглежда $w_i = xy^iz$ за всяко $i \in \mathbb{N}$. Можете ли да дадете други i, за които се получава противоречие.

---

1. $L = \{ a^n b^m \mid n < m \}$.

<u>док.</u> Да допуснем, че L е рег. и нека $p > 0$ е свидетел за това. Нека $w = a^p b^{p+1}$. Така $w \in L$ и $|w| = 2p+1 \geq p$. Следователно w може да се представи като $w = xyz$ като: $y \neq \varepsilon$, $|xy| \leq p$ и $\forall i \geq 0\ (xy^iz \in L)$.

Понеже xy е начало на w, не по-дълго от p, а първите p символа на w са a-та, то както x, така и y са съставени изцяло от a-та. Според ЛР, думата $w_2 = xy^2z = xyyz$ е от L. Но

$$w_2 = \underbrace{a\ldots a}_{|x|}\ \underbrace{a\ldots a}_{|y|}\ \underbrace{a\ldots a}_{|y|}\ \underbrace{a\ldots a}_{|z|_a}\ b^{p+1} =$$

$$= a^{p+|y|} b^{p+1}$$

Щом $w_2 \in L$, то $p+|y| < p+1$, т.е. $|y| < 1$.
Така $|y| = 0$ — противоречие. След. L не е рег.

3^* $L = \{ 0^{n^2} \mid n \geq 0 \}$

// $L = \{ \varepsilon, 0, 0000, \underbrace{0\ldots0}_{9}, 0^{16}, \ldots \}$

<u>док.</u> Да допуснем, че L ~~не~~ регулярен език и нека $p > 0$ е свидетел за това. Нека $w = 0^{p^2} = \underbrace{0\ldots0}_{p^2}$. Така $w \in L$ и $|w| \geq p$. Според ЛР има разбиване на $w = xyz$, за което: $y \neq \varepsilon$, $|xy| \leq p$ и $\forall i \geq 0\ (xy^iz \in L)$.

Да забележим, че $w_2 = xy^2z = xyyz = 0^{p^2+|y|}$
// добавеме толкова 0, колкото има в y.

Тогава: $p^2 < |w_2| = p^2 + |y| \leq p^2 + |xy| \leq$

$\leq p^2 + p < (p^2+p) + (p+1) = (p+1)^2$

Следователно $|w_2|$ е м/у два последователни точни квадрата (на ест. числа) като не е равна на нито един от тях. Следователно $|w_2|$ не е квадрат на естествено число.

Тогава $w_2 \notin L$. Противоречие.

4^*. $L = \{ 1^p \,/\, p$ е просто число $\}$

// p е просто число, ако $p \neq 1$ и единствените

// му делители (в естествените числа) са 1 и p

// $L = \{ 11, 111, 11111, 1^7, \dots \}$

<u>док.</u> Да допуснем, че L е регулярен език и нека $p > 0$ е свидетел за това (това не значи, че p е просто число). Понеже простите числа са безброй много, то има просто число $q \geq p$. Нека $w = 1^q = \underbrace{1 \dots 1}_{q}$.

Тогава $w \in L$ (защото q е просто) и $|w| = q \geq p$.

Нека $w = xyz$ е разбиване, за което е в сила

$|y| > 0$, $|xy| \leq p$, $\forall i \geq 0\ (xy^i z \in L)$.

Да забележим, че за вс. $i \geq 0$, $w_i = xy^i z =$

$= x \underbrace{y \ldots y}_{i} z = 1^{q + (i-1)|y|}$. Понеже $w_i \in L$,

то за вс. $i \geq 0$, $q + (i-1)|y|$ е просто.

Но при $i = q+1$, $q + (q+1-1)|y| = q + q \cdot |y| =$

$= q(|y|+1)$. Понеже q е просто, то

$q \geq 2$; $|y| > 0$, откъдето $|y|+1 \geq 2$.

След. $|w_{q+1}| = q(|y|+1)$ е произведение на

две ест. числа ≥ 2 и не може да бъде просто.

Противоречие. Така L не е регулярен.

5^*. $L = \{ a^n b^m \mid n \neq m \}$.

<u>док.</u>: Да допуснем, че L е регулярен език.

Тогава и $\overline{L} = \{a,b\}^* \setminus L$ също е регулярен

Забележете, че $\overline{L} = \{ a^n b^m \mid n = m \} \cup \{ w \in \{a,b\}^* \mid$
w не е от вида $a^n b^m \}$.

Понеже $a^* b^*$ е регулярен език, то сечението му с $\overline{L}$ (- регулярен) също ще бъде регулярен.

Но $a^*b^* \cap \overline{L} = \{a^n b^m \mid n = m\} =$

$= \{a^n b^n \mid n \geq 0\}$ - рег.

Това противоречи със заключението, направено в зад. 1. След. L не е регулярен

Упр. Док, че $\overline{\{a^n b^n \mid n \in \mathbb{N}\}}$ не е регулярен.

Релация на Myhill-Nerode *

Опр. Нека L е език (не задължително регулярен) над Σ. Над Σ^* определяме релацията $\sim_L$ на еквивалентност като:

$$x \sim_L y \Leftrightarrow \forall w \, (xw \in L \Leftrightarrow yw \in L)$$

// тук $x, y, w \in \Sigma^*$.

$\sim_L$ ще наричаме релация на Майхил-Нероуд за L. Използвайки я можем да дадем необходимо и достатъчно условие един език да бъде краен:

<u>Теорема. (Майхил, Нероуд.)</u> L е регулярен език над Σ, точно тогава, когато $\Sigma^*/\sim_L$ е крайно.

<u>Примери:</u>

① Док., че $L = \{a^n b^n \mid n \geq 0\}$ не е регулярен.

<u>док.</u> Нека $x = a^n$, $y = a^m$, $n < m$.

Тогава $x \cdot b^n \in L$, но $y \cdot b^n \notin L$. Следователно $\neg(x \sim_L y)$ (защото $\exists w\,(xw \in L \;\&\; yw \notin L)$), т.е. $[x]_{\sim_L} \neq [y]_{\sim_L}$. Така за вс. $n \geq 0$, $[a^n]_{\sim_L}$ е отделен клас. В частност, $\Sigma^*/\sim_L$ не е крайно — има поне изброимо много елементи.

② Док. че $L = \{0^{n^2} \mid n \geq 0\}$ не е регулярен.

<u>док.</u> Нека $x = 0^{n^2}$, $y = 0^{m^2}$, $n < m$.

Тогава $x \cdot 0^{2n+1} \in L$, $y \cdot 0^{2n+1} \notin L$, откъдето $[x] \neq [y]$. Следователно, $\Sigma^*/\sim_L$ има поне изброимо много класове, т.е. не е крайно.