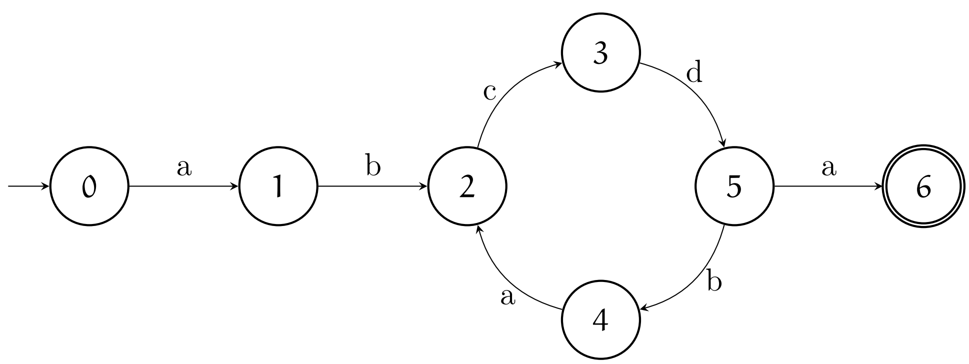
# Лема за покачването за регулярни езици

Иво Стратев

9 юни 2020 г.

## 1 Въведение

Да разгледаме следният автомат




Лесно се съобразява, че езикът на този автомат е безкраен, понеже съществува ориентиран цикъл, в който участва състояние, от което е достижимо финално състояние. Например $5$ е състояние участващо в цикъла $[2, 3, 5, 4, 2]$ и $2$ е достижимо от началното състояние и от $5$ е достижимо $6$, което е финално.

Също така сравнително лесно се съобразява, че ако вземем дума от езика на автомата с дължина по-голяма или равна от $6$, то думата ще е $ab.(cdba)^k.cda$ за някое $k \in \mathbb{N}_+$.

Лесно се съобразява и факта, че ако $k \in \mathbb{N}_+$, то за всяко $i \in \mathbb{N}$ думата $ab.(cdba)^{ik}.cda$ е от езика на автомата. Ако $i = 0$, то все едно не минаваме през цикъла съответстващ на $(cdba)^k$. Ако $i > 1$, то все едно правим допълнителни "обиколки"на цикъла съответстващ на $(cdba)^k$.

## 2 Лема за покачването

Нека $L$ е безкраен език над азбука $\Sigma$. Ако $L$ е регулярен език, то съществува $p \in \mathbb{N}_+$, такова че за всяка дума $\alpha \in L$ ако $|\alpha| \geq p$, то съществуват $x, y, z \in \Sigma^*$, такива че $\alpha = xyz$ и $|xy| \leq p$ и $y \neq \varepsilon$ и за всяко $i \in \mathbb{N}$ думата $xy^iz$ е от езика $L$.

## 3 Контрапозиция на лемата за покачването

Нека $L$ е безкраен език над азбука $\Sigma$. Ако за всяко $p \in \mathbb{N}_+$ същесвува дума $\alpha$ от $L$ с дължина поне $p$ и за всеки $x, y, z \in \Sigma^*$, такива че $\alpha = xyz$ и $|xy| \leq p$ и $y \neq \varepsilon$ същесвува $i \in \mathbb{N}$, такова че думата $xy^iz$ **не** е от езика $L$, то $L$ **не** е регулярен.

## 4 Схема на доказателство, че даден безкраен език не е регулярен

Нека $L$ е безкраен език над азбука $\Sigma$.

Нека $p \in \mathbb{N}_+$ и $p$ е произволно.

Избираме дума $\alpha$ от $L$ с дължина поне $p$.

Показваме, че за всяко разбиване на думата на три части $x, y, z \in \Sigma^*$, такива че $\alpha = xyz$ и $|xy| \leq p$ и $y \neq \varepsilon$ можем да посочим $i \in \mathbb{N}$, такова че $xy^iz$ не е от $L$ (pump-ваме с $i$).

От лемата за покачването заключаваме, че $L$ не е регулярен език.

## 5 Пример 1

Нека $L = \{\omega.b^n \mid \omega \in \{a,c\}^* \;\&\; n \in \mathbb{N} \;\&\; |\omega| = n\}$.

Нека $p \in \mathbb{N}_+$ и $p$ е произволно.

Избираме $\alpha = a^p b^p$, която е от $L$ с дължина $2p$, което е повече от $p$.

Избрахме $a^p b^p$, защото от лемата $xy$ е с дължина не повече от $p$, така че $xy$ ще е съставена изцяло от $a$-та и понеже има ясна граница това са $b$-тата и както и да пъпнем надолу ($i = 0$) или нагоре ($i > 1$) pump-натата дума няма да е в езика. Нека все пак формално да изкажем горното наблюдение.

Нека $x, y, z \in \{a, b, c\}^*$ са такива, че $\alpha = xyz$, $|xy| \leq p$ и $y \neq \varepsilon$. Тогава $xy$ е префикс на $a^p$, от където $z = a^{p-|xy|}b^p$. Също $0 < |y| \leq |x| + |y| = |xy| \leq p$. Тогава $xy^iz = a^{|x|}.a^{i|y|}.a^{p-|xy|}.b^p = a^{|x|+i|y|+p-|x|-|y|}.b^p = a^{p+(i-1)|y|}.b^p$ и понеже $a^{p+(i-1)|y|} \in \{a,c\}^*$ търсим такова $i \in \mathbb{N}$, че $p + (i-1)|y| \neq p$. Избираме $i = 0$ понеже $0 < |y| \leq p$ и тогава $p - |y| < p$. Тоест $xy^0z \notin L$.

От лемата за покачването заключваме, че $L$ не е регулярен език.

## 6 Пример 2

Нека $L = \{\alpha.\beta.\alpha \mid \alpha \in \{a,b\}^+ \;\&\; \beta \in \{a,b\}^+\}$.

Нека $p \in \mathbb{N}_+$ и $p$ е произволно.

Избираме $\alpha = a^p b$ и $\beta = b$ и нека $\omega = \alpha.\beta.\alpha$. Тогава $\omega \in L$ и $|\omega| = 2(p+1) + 1 = 2p + 3 > p$.

Нека $x, y, z \in \{a,b\}^*$ са такива, че $\alpha = xyz$, $|xy| \leq p$ и $y \neq \varepsilon$. Тогава $xy$ е префикс на $a^p$, от където $z = a^{p-|xy|}bba^pb$. Също $0 < |y| \leq |x| + |y| = |xy| \leq p$. Тогава $xy^iz = a^{|x|}.a^{i|y|}.a^{p-|xy|}.bba^pb = a^{|x|+i|y|+p-|x|-|y|}.bba^pb = a^{p+(i-1)|y|}.bba^pb$.

Нека изберем $i = 2$. Тогава $xy^2z = a^{p+|y|}.bba^pb$. Искаме да покажем, че $xy^2z \notin L$. Да допусем, че $xy^2z \in L$. Нека тогава $\rho, \sigma \in \{a, b\}^+$ са такива, че $\rho.\sigma.\rho = a^{p+|y|}.bba^pb$.

Възможни са няколко случая спрямо краят на $\rho$ (всеки суфикс на $\rho$ е суфикс на $a^{p+|y|}.bba^pb$).

## 6.1 Случай 1. $\rho = b$

Тогава $xy^2z$ започва с $b$, което не е вярно. Достигаме до Абсурд!

## 6.2 Случай 2. за някое $k \in \{1, 2, \dots, p\}$ е в сила, че $\rho = a^kb$

Тогава $xy^2z$ започва с $a^kb$, което не е вярно понеже първото $b$ е след $p + |y|$ на брой $a$-та и $p + |y| > k$. Достигаме до Абсурд!

## 6.3 Случай 3. за някоя дума $\gamma \in \{a, b\}^*$ е в сила, че $\rho = \gamma.b.a^p.b$

Тогава $xy^2z = (\gamma.b.a^p.b).\sigma.(\gamma.b.a^p.b)$ и значи $\mathcal{N}_b(xy^2z) \geq 4 > 3 = \mathcal{N}_b(xy^2z)$. Това е Абсурд!

Следователно $xy^2z \notin L$.

От лемата за покачването заключваме, че $L$ не е регулярен език.

# 7 Пример 3

Нека $L = \{\alpha.\gamma \mid \alpha, \gamma \in \{a, b, c\}^* \ \&\ \mathcal{N}_b(\alpha) = 2\mathcal{N}_a(\alpha) \ \&\ \mathcal{N}_c(\gamma) = 3\mathcal{N}_b(\gamma)\}$.

Нека $p \in \mathbb{N}_+$ и $p$ е произволно.

Избираме $\alpha = a^pb^{2p}$ и $\gamma = c^{3p}b^p$ и нека $\omega = \alpha.\gamma = a^pb^{2p}c^{3p}b^p$. Тогава $\omega \in L$ и $|\omega| = p + 2p + 3p + p > 7p > p$.

Нека $x, y, z \in \{a, b, c\}^*$ са такива, че $\alpha = xyz$, $|xy| \leq p$ и $y \neq \varepsilon$. Тогава $xy$ е префикс на $a^p$, от където $z = a^{p-|xy|}b^{2p}c^{3p}b^p$. Също

$0 < |y| \leq |x| + |y| = |xy| \leq p$. Тогава $xy^i z = a^{|x|}.a^{i|y|}.a^{p-|xy|}b^{2p}c^{3p}b^p = a^{|x|+i|y|+p-|x|-|y|}.b^{2p}c^{3p}b^p = a^{p+(i-1)|y|}.b^{2p}c^{3p}b^p$.

Нека изберем $i = 0$. Тогава $xy^0 z = a^{p-|y|}.b^{2p}c^{3p}b^p$. Искаме да покажем, че $xz \notin L$. Да допусем, че $xz \in L$. Нека тогава $\rho, \sigma \in \{a, b, c\}^*$ са такива, че $\rho.\sigma = xz = a^{p-|y|}.b^{2p}c^{3p}b^p$ и $\mathcal{N}_b(\rho) = 2\mathcal{N}_a(\rho)$ и $\mathcal{N}_c(\sigma) = 3\mathcal{N}_b(\sigma)$.

Възможни са два случая спрямо $\rho$.

### 7.1 Случай 1. $\rho = \varepsilon$

Тогава $\sigma = xz = a^{p-|y|}.b^{2p}c^{3p}b^p$. Но $\mathcal{N}_c(\sigma) = 3p$ и $\mathcal{N}_b(\sigma) = 3p$ и $3p < 9p$. Така $xz \notin L$. Достигаме до Абсурд!

### 7.2 Случай 2. $\rho \neq \varepsilon$

Тогава $\mathcal{N}_a(\rho) > 0$ от където $\mathcal{N}_b(\rho) > 0$ Тогава $\rho = a^{p-|y|}.b^{2p-2|y|}$ и $\sigma = b^{2|y|}c^{3p}.b^p$. Така $\mathcal{N}_b(\sigma) = p + 2|y| > p$ и значи $3\mathcal{N}_b(\sigma) > 3p = \mathcal{N}_c(\sigma)$, което е Абсурд!

Следователно $xy^0 z = xz \notin L$.

От лемата за покачването заключваме, че $L$ не е регулярен език.

## 8 Пример 4

Нека $L = \{c^n.a^k.b^s \mid n \in \mathbb{N}_+ \ \& \ s \in \mathbb{N}_+ \ \& \ (\exists l \in \mathbb{N}_+)(k = ls)\}$.

Нека $p \in \mathbb{N}_+$ и $p$ е произволно.

Избираме $\alpha = ca^p b^p$ и $|\omega| = 2p + 1 > p$. Понеже $p = 1.p$, то $\omega \in L$.

Нека $x, y, z \in \{a, b\}^*$ са такива, че $\alpha = xyz$, $|xy| \leq p$ и $y \neq \varepsilon$. Тогава $xy$ е префикс на $ca^{p-1}$. Възможни са два случая.

**Случай 1.** $x = \varepsilon$. Тогава $\alpha = xyz = yz = ca^p b^p$ и понеже $|y| = |xy| \leq p$, то $z = a^{p-|y|+1}b^p$. Тогава $xy^i z = y^i z = (ca^{|y|-1})^i a^{p-|y|+1}b^p$.

Избираме $i = 0$ тогава $xy^iz = z = a^{p-|y|+1}b^p$, защото тогава $\mathcal{N}_c(xy^iz) = 0 \notin \mathbb{N}_+$ и $xy^iz \notin L$.

**Случай 2.** $x \neq \varepsilon$. Тогава $z = a^{p-|xy|+1}b^p$ и $0 < |y| < p$, понеже $|xy| \leq p$. Тогава $xy^iz = ca^{|x|-1+i|y|+p-|xy|+1}b^p = ca^{p+(i-1)|y|}b^p$. Търсим такова $i \in \mathbb{N}$, че $xy^iz \notin L$. За целта пресмятаме

$$\frac{p + (i-1)|y|}{p} = 1 + \frac{(i-1)|y|}{p}$$

Ако $i = 2$ получаваме $1 + \dfrac{|y|}{p} \notin \mathbb{N}$. Ако $i = p$ получаваме $1 + |y| - \dfrac{|y|}{p} \notin \mathbb{N}$. Тоест не зависимо какво изберем $i = 2$ или $i = p$ думата $xy^iz$ не е в $L$.

От лемата за покачването заключваме, че $L$ не е регулярен език.

# 9 Пример 5.

Нека $L = \{\alpha\#\beta \mid \alpha \in \{0,1\}^+ \ \& \ \beta \in \{0,1\}^+ \ \& \ \alpha_{(2)} + \beta_{(2)} = 3\alpha_{(2)}\}$, където (бинарна дума)$_{(2)}$ е числото с двойчен запис бинарна дума. Тоест

$$\begin{aligned} 0_{(2)} &= 0_{\mathbb{N}} \\ 1_{(2)} &= 1_{\mathbb{N}} \\ (\omega.0)_{(2)} &= 2_{\mathbb{N}} * \omega_{(2)} \\ (\omega.1)_{(2)} &= 2_{\mathbb{N}} * \omega_{(2)} + 1_{\mathbb{N}} \end{aligned}$$

Например $1000_{(2)} = 8$ и $100_{(2)} = 4$ и значи $1000\#100 \in L$, защото $8 + 4 = 12 = 3.4$. Условието на $L$ същност се свежда до решенията на уравнението $2x + x = 3x$, които са безкрайно много.

Нека $p \in \mathbb{N}_+$ и $p$ е произволно.

Избираме $\omega = 100^p\#10^p$, която е в езикът, защото $100^p_{(2)} = 2_{\mathbb{N}} * 10^p_{(2)}$. Дължината на $\omega$ е $2 + p + 1 + 1 + p$, което е поне $p$.

Нека $x, y, z \in \{0,1\}^*$ и са такива, че $\omega = x.y.z$ и $|x.y| \leq p$ и $|y| \geq 1$.

Възможни са два случая.

**Случай 1.** $x = \varepsilon$. Тогава $z = 0^{p+1-|y|+1}\#10^p = 0^{p+2-|y|}\#10^p$. В този случай избираме $i = 0$, защото $x.y^0.z = \varepsilon.\varepsilon.z = 0^{p+2-|y|}\#10^p$ и $0^{p+2-|y|}_{(2)} = 0_{\mathbb{N}}$ и $10^p_{(2)} \neq 0_{\mathbb{N}}$ и значи $x.y^0.z \notin L$.

**Случай 2.** $x \neq \varepsilon$. Тогава $z = 0^{p+1-|xy|+1}\#10^p = 0^{p+2-|xy|}\#10^p$. В този случай избираме $i = 2$, защото $x.y^2.z = 100^{p+|y|}\#10^p$ и

$$100^{p+|y|}_{(2)} = 2^{|y|+1}_{\mathbb{N}} * 10^p_{(2)} > 2_{\mathbb{N}} * 10^p_{(2)}$$

и значи $x.y^2.z \notin L$.