Лема за покачването за безконтекстни езици

Иво Стратев

9 юни 2020 г.

Въведение

Лемата до някъде прилича на тази за регулярни езици, но е по-сложна!

Нека L е безкраен език над азбука Σ.

Ако L е контекстно свободен език, то същесвува $p \in \mathbb{N}_+$, такова че за всяка дума $\alpha \in L$ ако $|\alpha| \geq p$, то същесвуват $u, v, x, y, z \in \Sigma^*$, такива че

$$\begin{aligned} \alpha &= uvxyz \\ \& \, |vxy| &\leq p \\ \& \, vy &\neq \varepsilon \\ \& \, (\forall i \in \mathbb{N})(uv^i xy^i z &\in L) \end{aligned}$$

Контрапозиция на лемата

Нека L е безкраен език над азбука Σ. Ако за всяко $p \in \mathbb{N}_+$ е вярно, че съществува дума $\alpha \in L$, такаква че $|\alpha| \geq p$ и за всеки $u, v, x, y, z \in \Sigma^*$, такива че

$$\begin{aligned} \alpha &= uvxyz \\ \& \, |vxy| &\leq p \\ \& \, vy &\neq \varepsilon \end{aligned}$$

съществува $i \in \mathbb{N}$ такова, че $uv^i xy^i z \notin L$, то L не е контекстно свободен език.

Обща схема на доказателство

Нека е даден безкраен език L над азбука Σ.

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

Избираме $\alpha \in L$, такаква че $|\alpha| \geq p$ и проверяваме, че за всяко едно разбиване на α на пет части $u, v, x, y, z \in \Sigma^*$, такива че $\alpha = uvxyz$ и $|vxy| \leq p$ и $|vy| \geq 1$ можем да намерим $i \in \mathbb{N}$, такова че $uv^ixy^iz \notin L$.

Тогава от лемата за покачването (разрастването) L не е безконтекстен език.

Класика в жанра

Да се докаже, че езикът $L = \{a^nb^nc^n \mid n \in \mathbb{N}\}$ не е контекстно свободен.

Нека $p \in \mathbb{N}_+$. Избираме $\alpha = a^pb^pc^p$. В сила са $\alpha \in L$ и $|\alpha| = 3p \geq p$. Разглеждаме произволно разбиване на 5 части спазващо условията на лемата. Нека $u, v, x, y, z \in \{a, b, c\}^*$ са такива, че $\alpha = uvxyz$, $|vxy| \leq p$ и $|vy| \geq 1$. Възможни са два случая.

Случай 1. $uvxy$ е префикс на a^pb^p

Тогава $\mathcal{N}_a(vy) > 0$ или $\mathcal{N}_b(vy) > 0$.

Избираме $i = 0$ и тогава

$\mathcal{N}_a(uv^ixy^iz) < \mathcal{N}_c(uv^ixy^iz)$ или $\mathcal{N}_b(uv^ixy^iz) < \mathcal{N}_c(uv^ixyz^i)$.

Следователно $uv^ixy^iz \notin L$.

Случай 2. a^p е префикс на u.

Тогава $\mathcal{N}_b(vy) > 0$ или $\mathcal{N}_c(vy) > 0$.

Избираме $i = 0$ и тогава

$\mathcal{N}_b(uv^ixy^iz) < \mathcal{N}_a(uv^ixy^iz)$ или $\mathcal{N}_c(uv^ixy^iz) < \mathcal{N}_a(uv^ixyz^i)$.

Следователно $uv^ixy^iz \notin L$.

Заключение:

Следователно $L = \{a^nb^nc^n \mid n \in \mathbb{N}\}$ не е контекстно свободен.

Забележка:

Най-лесно е да гледаме къде се намира vxy понеже той е непразен. Удобно е или да е сред първите две части a-та и/или b-та или сред вторите b-та и/или c-та.

Случаите не са непресичащи, но обхващат всички възможни понеже $|vxy| \leq p$ и няма как да обвахне пове от два вида букви.

Пример 1.

Да се докаже, че езикът $L = \{b^{n^2} \mid n \in \mathbb{N}\}$ не е контекстно свободен.

Нека $p \in \mathbb{N}_+$. Избираме $\alpha = b^{p^2}$. В сила са $\alpha \in L$ и $|\alpha| = p^2 \geq p$.

Разглеждаме произволно разбиване на 5 части спазващо условията на лемата.

Нека $u, v, x, y, z \in \{a, b, c\}^*$ са такива, че $\alpha = uvxyz$, $|vxy| \leq p$ и $|vy| \geq 1$.

Нека $k = |vy|$. Тогава $vy = b^k$ и $1 \leq k \leq p$.

Тогава $uv^ixy^iz = uxz.(vy)^i = b^{n^2-k}.(b^k)^i = b^{n^2-k+ik} = b^{n^2+(i-1)k}$.

Нека $i = 2$. Тогава $uv^ixy^iz = b^{n^2+k}$.

Трябва да покажем, че n^2+k не е точен квадрат на естествено число.

В сила са $n^2 < n^2 + 1 \leq n^2 + k$ и

$(n+1)^2 = n^2 + 2n + 1 = n^2 + n + (n+1) \geq n^2 + k + (n+1) > n^2 + k$.

Следователно $n^2 < n^2 + k < (n+1)^2$ и $k \in \mathbb{N}_+$ и значи $n^2 + k$ не е точен квадрат. Следователно $uv^2xy^2z \notin L$.

Следователно от лемата за разрастването следва, че L не е безконтекстен.

Пример 2

Нека $A = \{\omega\omega^{rev} \mid \omega \in \{c, d\}^*\}$ и нека $B = \{\omega.\omega \mid \omega \in \{c, d\}^*\}$.

Да се определи кой от двата езика е контекстно свободен.

Решение. A е контекстно свободен език, защото

$\langle \{S\}, \{c, d\}, S, \{S \to cSc,\ S \to dSd,\ S \to \varepsilon\} \rangle$ го генерира.

За B ще покажем, че е не е контекстно свободен чрез лемата за покачването. Очевидно не е краен език.

Нека $p \in \mathbb{N}_+$. Избираме $\alpha = (a^pb)(a^pb)$. В сила са $\alpha \in L$ и $|\alpha| = 2(p+1) > p$.

Разглеждаме произволно разбиване на 5 части спазващо условията на лемата.

Нека $u, v, x, y, z \in \{a, b, c\}^*$ са такива, че $\alpha = uvxyz$, $|vxy| \leq p$ и $|vy| \geq 1$.

Възможен е следния случай: $p \geq 3$, $u = a^{p-1}$, $v = a$, $x = b$, $y = a$, $z = a^{p-1}b$.

Тогава $(\forall i \in \mathbb{N})(uv^ixy^iz = a^{p-1}a^iba^ia^{p-1}b = (a^{p+i-1}b)(a^{p+i-1}b) \in L)$.

Това означава, че a^pba^pb като дума не става с цел аргумент, че B не е безконтекстен. Трябва да изберем друга дума.

Избираме $\alpha = (a^pb^p)(a^pb^p)$. В сила са $\alpha \in L$ и $|\alpha| = 2(2p) \geq p$.

Разглеждаме произволно разбиване на 5 части спазващо условията.

Нека $u, v, x, y, z \in \{a, b, c\}^*$ са такива, че $\alpha = uvxyz$, $|vxy| \leq p$ и $|vy| \geq 1$. Възможни са пет случаи.

Случай 1. $uvxy = a^k$ за някое $k \in \mathbb{N}$

Тогава $z = a^{p-k}b^pa^pb^p$. Избираме $i = 0$.

Тогава $uv^ixy^iz = a^{k-|vy|}a^{p-k}a^pb^p = a^{p-|vy|}b^p.a^pb^p \notin L$.

Случай 2.

За някои $k, m, t, s \in \mathbb{N}$ $v = a^k$, $x = a^mb^t$, $y = b^s$ и $(|k| > 0 \ \vee \ |s| > 0)$ и $k+s \leq p-m-t \leq p$. Тогава $u = a^{p-k-m}$ и $z = b^{p-t-s}a^pb^p$ и значи избираме $i = 2$. Тогава $uv^ixy^iz = a^{p-k-m}a^{2k}a^mb^tb^{2s}b^{p-t-s}a^pb^p = a^{p+k}b^{p+s}a^pb^p \notin L$.

Случай 3.

За някои $k, t \in \mathbb{N}$ $vxy = b^k$ и $u = a^pb^{p-k-t}$ и $v = b^ta^pb^p$.

Избираме $i = 0$. Тогава $uv^ixy^iz = a^pb^{p-|vy|}a^pb^p \notin L$.

Случай 4.

За някои $k, m, t, s \in \mathbb{N}$ $v = b^k$, $x = b^ma^t$, $y = a^s$ и $(|k| > 0 \ \vee \ |s| > 0)$ и $k + s \leq p$. Тогава $u = a^pb^pb^{p-k-m}$ и $z = a^{p-t-s}$ и значи избираме $i = 2$. Тогава $uv^ixy^iz = a^pb^pb^{p-k-m}b^{2k}b^ma^ta^{2s}a^{p-t-s} = a^pb^pb^{p+k}a^{p+s} \notin L$.

Случай 5. $vxyz = b^k$ за някое $k \in \mathbb{N}$

Тогава $u = a^p b^p a^p b^{p-k}$. Избираме $i = 0$.

Тогава $uv^ixy^iz = a^p b^p b^{p-k} b^{k-|vy|} = a^p b^p . a^p b^{p-|vy|} \notin L$.

Заключение.

Следователно от лемата за разрастването следва, че L не е безконтекстен.

Пример 3

Нека $A = \{a^n b^k c^{n+k} \mid n, k \in \mathbb{N}\}$ и нека $B = \{a^n b^k c^{n.k} \mid n, k \in \mathbb{N}\}$.

Да се определи кой от двата езика е контекстно свободен.

Решение. A е контекстно свободен език, защото

$\langle \{S, F\}, \{a, b, c\}, S, \{S \to aSc,\ S \to F,\ F \to bSc,\ F \to \varepsilon\}\rangle$ го генерира.

За B ще покажем, че е не е контекстно свободен чрез лемата за покачването. Очевидно не е краен език.

Нека $p \in \mathbb{N}_+$. Избираме $\alpha = a^p b^p c^{p^2}$. В сила са $\alpha \in L$ и $|\alpha| = 2p + p^2 \geq p$. Разглеждаме произволно разбиване на 5 части спазващо условията на лемата. Нека $u, v, x, y, z \in \{a, b, c\}^*$ са такива, че $\alpha = uvxyz$, $|vxy| \leq p$ и $|vy| \geq 1$. Възможни са няколко случая.

Случай 1. $uvxy = a^k$ за някое $k \in \mathbb{N}$

Тогава $z = a^{p-k} b^p c^{p^2}$. Избираме $i = 0$. Тогава $uv^ixy^iz = a^{k-|vy|} a^{p-k} b^p c^{p^2} = a^{p-|vy|} b^p c^{p^2} \notin L$, защото $(p - |vy|).p < p^2$.

Случай 2.

За някои $k, m, t, s \in \mathbb{N}$ $v = a^k$, $x = a^m b^t$, $y = b^s$ и $(|k| > 0 \ \vee \ |s| > 0)$ и $k + s \leq p$. Тогава $u = a^{p-k-m}$ и $z = b^{p-t-s} c^{p^2}$ и значи избираме $i = 2$. Тогава $uv^ixy^iz = a^{p-k-m} a^{2k} a^m b^t b^{2s} b^{p-t-s} c^{p^2} = a^{p+k} b^{p+s} c^{p^2}$. Така $(p + k)(p + s) = p^2 + (k + s)p + sk \geq p^2 + (k + s)p > p^2 \geq p^2 + p > p^2$.

Следователно $uv^ixy^iz \notin L$.

Случай 3.

За някои $k, m, t, s \in \mathbb{N}$ $v = b^k$, $x = b^m c^t$, $y = c^s$ и $(|k| > 0 \ \vee \ |s| > 0)$ и $k + s \leq p$. Тогава $u = a^p b^{p-k-m}$ и $z = c^{p^2-t-s}$ и значи избираме $i = 2$. Тогава $uv^ixy^iz = a^p b^{p-k-m} b^{2k} b^m c^t c^{2s} c^{p^2-t-s} = a^p b^{p+k} c^{p^2+s}$. Така $p(p+k) = p^2 + kp$.

Ако $k = 0$, то $0 < s \leq p$ и значи $p(p+k) = p^2 < p^2 + s$.

Ако $k > 0$, то $0 \leq s < p$ и значи $p(p+k) = p^2 + kp \geq p^2 + p > p^2 + s$.

Следователно $uv^ixy^iz \notin L$.

Случай 4. $vxyz = c^k$ за някое $k \in \mathbb{N}$

Тогава $u = a^p b^p c^{p^2-k}$. Избираме $i = 0$. Тогава $uv^ixy^iz = a^p b^p c^{p^2-k} b^{k-|vy|} = a^p b^p . c^{p^2-|vy|} \notin L$, защото $p^2 > p^2 - |vy|$.

Заключение.

Следователно от Лемата за разрастването $B = \{a^n b^k c^{n.k} \mid n, k \in \mathbb{N}\}$ не е контекстно свободен.