05.05.23

$X \overset{\ell}{\triangle} \alpha$ релация: синтактично дърво с корен X, дълбочина ℓ и листа α
за извод

$\mathcal{L}_G(X) = \{\alpha \in \Sigma^* \mid X \overset{*}{\triangle} \alpha\}$ аналог. $\mathcal{L}_A(q)$

X – произволен елемент, $X \in V$, $X \in \Sigma$

$\mathcal{L}^{\ell}_G(X) = \{\alpha \in \Sigma^* \mid X \overset{\leq \ell}{\triangle} \alpha\}$ апроксимации на езика спрямо променлива

$\mathcal{L}^{\ell+1}_G(X) = \bigcup\{\mathcal{L}^{\ell}_G(X_1) \dots \mathcal{L}^{\ell}_G(X_n) \mid X \to X_1 \dots X_n \text{ е правило в } G\}$

$\ell+1$ апроксимация = обединение на ℓ-тите апроксимации

Критерий за доказателство, че един език не е безконтекстен

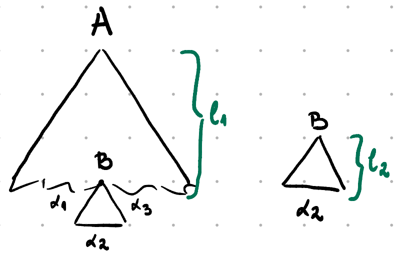
Имаме пътища в дърво, дефинираме лема на покачването за тях.

Търсим достатъчно голяма дума, такава, че да има пътища с повторения на променливите

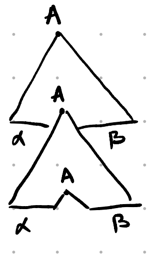
Свойства:

1) Имаме следния извод:

$$\frac{\underset{\text{корен}}{A} \overset{\ell_1}{\triangle} \underset{\text{листа}}{\alpha_1 B \alpha_3} \qquad \underset{\text{корен}}{B} \overset{\ell_2}{\triangle} \underset{\text{листа}}{\alpha_2}}{A \overset{\leq \ell_1+\ell_2}{\triangle} \alpha_1\alpha_2\alpha_3}$$



2) $$\frac{A \overset{\ell}{\triangle} \alpha A \beta \qquad i \in \mathbb{N}}{A \overset{\leq \ell \cdot i}{\triangle} \alpha^i A \beta^i}$$



$A \overset{\leq 2\ell}{\triangle} \alpha^2 A \beta^2$

3) Нека $b = \max\{|\alpha| : A \to \alpha \text{ е правило в } G\}$ (максимална разклоненост на дървото)

- Ако $A \overset{\ell}{\triangle} \alpha$, то $|\alpha| \leq b^{\ell}$
- Ако $\left.\begin{matrix} \alpha \in \mathcal{L}_G(A) \\ |\alpha| > b^{\ell} \end{matrix}\right\} \Rightarrow A \overset{\geq \ell+1}{\triangle} \alpha$

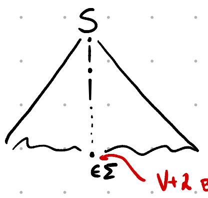
Доказателство:

- $L = \mathcal{L}(G)$ безкраен език, $b = \{|\alpha| : A \to \alpha \text{ е правило в } G\}$,

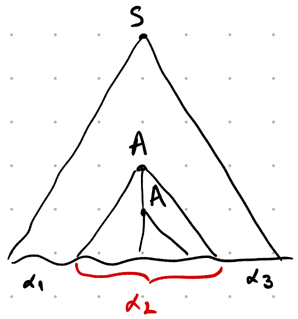
$\alpha \in \mathcal{L}(G)$, като $|\alpha| > b^{|V|}$ (дост. голяма дума > разклон.$^{\text{брой пром.}}$)

• Знаем, че $S \overset{\geq|V|+1}{\Delta} \alpha$. Тогава: имаме път с дължина $|V|+1$

(броя срещани променливи е $|V|+1$)

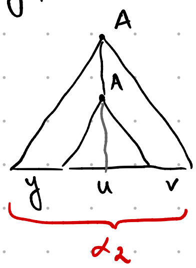


V+2 възела
V+1 ребра } има повторение на променливи

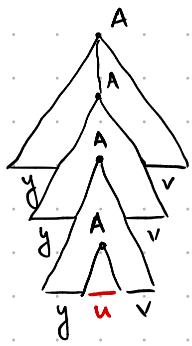


$A \overset{<|V|+1}{\Delta} \alpha_2 \Rightarrow |\alpha_2| \leq b^{|V|+1}$

дърво за α_2



дърво за α_2 след премахване на u



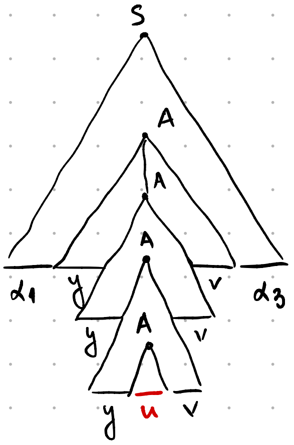
2) $A \overset{*}{\Delta} yAv$

1) $A \overset{*}{\Delta} y^iAv^i$

$A \overset{*}{\Delta} y^iuv^i$

$\forall i \in \mathbb{N}$

$\Rightarrow$ Получаваме следното дърво



$S \overset{*}{\Delta} \alpha_1 A \alpha_3$
$A \overset{*}{\Delta} y^iuv^i$ } $\Rightarrow S \overset{*}{\Delta} \alpha_1 y^i u v^i \alpha_3$

! Може да се появи проблем, ако y и v са ε!

Тогава не генерираме нови думи

Трябва да осигурим, че поне едно от тях не е ε.

Ако y и v са ε, можем да премахнем тяхното поддърво и да получим минимално дърво.

Pumping Lemma

Нека L е безкраен безконтекстен език.

$\exists$ конст. $p \geq 1$:

$\forall \alpha \in L: |\alpha| > p$:

$\exists$ разбиване на α на 5 части:

$\alpha = xyuvw: |yuv| \leq p, |yv| \geq 1$

$(\forall i \in \mathbb{N})[xy^iuv^iw \in L]$

За да докажем, че един език е безконтекстен, използваме лемата в контрапозиция

Пример: $L = \{a^n b^n c^n \mid n \in \mathbb{N}\}$. Да се докаже, че L не е безконтекстен

$(\forall)$ произволно $p \geq 1$:

$(\exists)$ $\alpha \in L$, $|\alpha| \geq p$.

Нека $\alpha = a^p b^p c^p$

$(\forall)$ произв. разбиване на α като $\alpha = xyuvw$, където:

$|yv| \geq 1$ и $|yuv| \leq p$

- Ако y или v съдържа 2 разл. букви, то $xy^2uv^2w \notin L$.
- Ако y или v съдържа 1 буква: аналогично

$(\exists)$ за $i = 2$: $xy^2uv^2w \notin L$.

Задача: Дадено е $L = \{\alpha\alpha \mid \alpha \in \{a, b\}^*\}$

Докажете, че L не е регулярен

Нека $\alpha = a^p b a^p b \in L$

$|xy| \leq p$ $|y| \geq 1$ $xy^i z \notin L$ $\forall i \neq 1$

=> не е регулярен

Докажете, че L не е безконтекстен

$\alpha = \underset{\uparrow y}{a^p} b \underset{\uparrow v}{a^p} b \in L$

$\alpha = xyuvw$

Ако $y = a^k = v$, то

$xy^iuv^iw \in L$ $\forall i \in \mathbb{N}$

=> не работи за тази α

$\alpha = a^p b^p a^p b^p$ $\alpha = xyuvw$

y и v са в съседни групи букви

Примерен случай: $\alpha = a^p b^p a^p b^p$ (y — край на b^p и начало на a^p, v — в a^p)

$y = ba$ $v = a^k$

Ако $i = 0$, то $xy^0uv^0w \notin L$

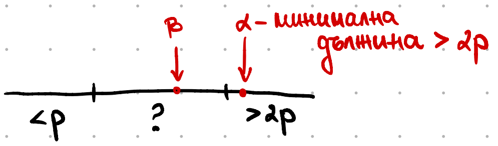
=> L не е безконтекстен

$L_1 = \{a^n b^n c^k \mid n, k \in \mathbb{N}\}$
$L_2 = \{a^k b^n c^n \mid n, k \in \mathbb{N}\}$ — безконтекстни

Сечението на безконтекстни езици не винаги дава безконтекстен език

$L_1 \cap L_2 = \{a^n b^n c^n \mid n \in \mathbb{N}\}$ - не е безконтекстен

Тв: Нека L е безконтекстен. L е безкраен $\iff \exists \alpha \in L : p \leq |\alpha| < 2p$



$\alpha = xyuvw$ $\beta = xy^0uv^0w$ $|\beta| < |\alpha|$

Как мога да изчистя граматиката от променливи, които не допринасят за разпознаване на думи от езика ѝ? (Опростяване)

Стъпка 1: филтриране ненужни променливи

вход: G

изход: G'

като $\mathcal{L}(G)=\mathcal{L}(G')$ $\quad \underbrace{\forall A\in V' : \mathcal{L}_{G'}(A)\neq\emptyset}_{\text{принос за разпозн.}}$

Търсим множество $Gen=\{A\in V \mid \underbrace{\mathcal{L}_G(A)\neq\emptyset}_{A \overset{*}{\triangleright} \alpha \text{ за някое } \alpha\in\Sigma^*}\}$

Дефинираме $Gen[l]=\{A\in V \mid \mathcal{L}^{l}_G(A)\neq\emptyset\}$

$Gen[0]=\emptyset$ (дърво с височина 0 не може да генерира думи)

$Gen[l+1]=\{A\in V \mid \exists \alpha\in(\Sigma\cup Gen[l])^* \mid A\to\alpha \text{ е правило в } G\}$

Генерираме дървото от листата към корена с цел осигуряване, че поддърветата могат да генерират думи също. Спираме когато достигнем стъпка, в която $Gen[l+1]=Gen[l]$ т.е. не можем да добавим нищо.

Стъпка 2: филтриране недостижими променливи