~ Seminar 6 ~

Lema de pompare pentru limbajele regulate: **[vezi și CURS 4]**

Fie L un limbaj regulat. Atunci $\exists p \in \mathbb{N}$ (număr natural) astfel încât pentru $\forall \alpha \in L$ cuvânt, cu $|\alpha| \geq p$ și există o descompunere $\alpha = u \cdot v \cdot w$ cu proprietățile:

(1) $|u \cdot v| \leq p$
(2) $|v| \geq 1$
(3) $u \cdot v^i \cdot w \in L,\ \forall i \geq 0.$

Vrem să demonstrăm că un limbaj NU este regulat.
Presupunem prin reducere la absurd că "limbajul este regulat" (predicatul P) și atunci rezultă că "afirmația din lemă este adevărată" (predicatul Q).
Obs: Știm de la logică faptul că $(P \rightarrow Q) \equiv (\neg Q \ \rightarrow \ \neg P)$. Așa că vom nega afirmația lemei $(\neg Q)$ și va rezulta că limbajul nu este regulat $(\neg P)$. Practic negarea constă în interschimbarea cuantificatorilor logici (∃ și ∀) între ei, iar la condiția (3) ∈ devine ∉.

➔ *Schema demonstrației*

Alegem (adică ∃) un cuvânt α din limbajul L dat, care să respecte ipoteza lemei anterioare, adică să aibă lungimea cel puțin p, $\forall p \in \mathbb{N}$, iar descompunerea sa $\alpha = u \cdot v \cdot w$ să respecte condițiile (1) și (2). Apoi alegem convenabil un număr natural *i* pentru care să obținem o contradicție a condiției (3), adică să rezulte că cuvântul $\beta = u \cdot v^i \cdot w \notin L$ și deci presupunerea făcută este falsă.

Lema de pompare pentru limbajele independente de context: **[vezi și CURS 10]**

Fie L un limbaj independent de context.
Atunci $\exists p \in \mathbb{N}$ (număr natural) astfel încât pentru $\forall \alpha \in L$ cuvânt, cu $|\alpha| \geq p$, există o descompunere $\alpha = u \cdot v \cdot w \cdot x \cdot y$ cu proprietățile:

(1) $|v \cdot w \cdot x| \leq p$
(2) $|v \cdot x| \geq 1$
(3) $u \cdot v^i \cdot w \cdot x^i \cdot y \in L,\ \forall i \geq 0.$

➔ *Schema demonstrației*

- (∃) Alegem un cuvânt α din limbajul L astfel încât $|\alpha| \geq p$, $\forall p \in \mathbb{N}$. *Este important ca pentru acel α să NU poată fi construit un automat push-down (sau o gramatică independentă de context), pentru că altfel NU vom putea obține contradicția dorită.*
- Știm că $\alpha = u \cdot v \cdot w \cdot x \cdot y$. Vom presupune proprietățile (1) și (2) îndeplinite și vom găsi o contradicție pentru (3).
- $(\forall u, v, w, x, y)$ Trebuie să analizăm pe rând ***orice*** împărțire posibilă a lui α în cele 5 componente (altfel spus, trebuie să poziționăm vwx în toate modurile posibile în α). Pentru ***fiecare*** caz, trebuie să alegem câte un număr natural *i* (nu neapărat același pentru toate cazurile) astfel încât cuvântul $\beta = u \cdot v^i \cdot w \cdot x^i \cdot y \notin L$, rezultând o contradicție a lemei (a proprietății (3)), deci a presupunerii că limbajul L era independent de context.
  (Atenție, demonstrația este complete doar dacă obținem contradicție pentru fiecare posibilitate de descompunere a lui α, nu doar pe unele cazuri.)

Exemplu: L1 = {a^nb^n, n ≥ 1} = {ab, a^2b^2, a^3b^3, a^4b^4, a^5b^5, ...}

Vrem să demonstrăm că L1 **nu este limbaj regulat**. Observăm că L1 conține cuvinte formate din k litere de ”a” urmate tot de k litere de ”b”. În demonstrație trebuie să obținem un cuvânt β care să nu respecte această proprietate (adică să fie de forma a^*b^*, dar să aibă număr diferit de a-uri și b-uri).

Demonstrație: Presupunem prin reducere la absurd că L1 este limbaj regulat. Atunci $\exists p \in \mathbb{N}$ și putem aplica lema de pompare. *(În continuare negăm afirmația lemei.)*
Alegem cuvântul $\alpha = a^p b^p$, cu $|\alpha| = 2p \geq p, \forall p \in \mathbb{N}$ *(deci lungimea cuvântului respectă ipoteza lemei)*. Conform lemei, cuvântul poate fi scris sub forma $\alpha = u \cdot v \cdot w$.
Din condiția (1), avem $|u \cdot v| \leq p$. Rezultă că cuvântul $u \cdot v$ conține doar litere de ”a” (pentru că $u \cdot v$ este un prefix al primelor p caractere din α).
Atunci notăm $v = a^k$. Din condițiile (1) și (2) avem $1 \leq |v| \leq p$ *(pentru că se poate ca $u = \lambda$, adică $|u| = 0$)*. Deci $1 \leq |a^k| \leq p$, adică $1 \leq k \leq p$ (*).
De asemenea, condiția (3) spune că $u \cdot v^i \cdot w \in L1$, $\forall i \geq 0$. Alegem $i = 0$. Atunci avem cuvântul $\beta = u \cdot v^0 \cdot w = u \cdot w = a^{p-k} b^p \notin L1$ (pentru că în relația (*) avem $k \geq 1$, deci numărul de ”a”-uri este strict mai mic decât numărul de ”b-uri” din cuvântul β). Am obținut o contradicție pentru (3), deci presupunerea făcută este falsă și L1 nu este limbaj regulat.

Exemplu: $\boldsymbol{L2 = \{a^{2^n} | n \geq 0\} = \{a^1, a^2, a^4, a^8, a^{16}, a^{32}, a^{64}, \ldots\}}$

[de adaugat…]

EXERCITII:

EX_1: Demonstrați că următoarele limbaje NU sunt regulate.

$L3 = \{a^m b^n | m > n \geq 0\}$

$L4 = \{w \cdot w^R | w \in \{a, b\}^*\}$

$L5 = \{w \cdot w \mid w \in \{a, b\}^*\}$

$L6 = \{a^{n^2} | n \geq 1\} = \{a^1, a^4, a^9, a^{16}, a^{25}, a^{36}, \ldots\}$

$L7 = \{a^m b^{3n} a^n | m \geq 1, n \geq 0\}$

Exemplu: $L1 = \{w \cdot w \mid w \in \{a,b\}^*\}$ nu este limbaj independent de context.

Demonstrație: Presupunem că L este limbaj independent de context, rezultă că există acel număr natural p din lemă *(numit lungimea de pompare)*.
Alegem $\alpha = a^p b^p a^p b^p \in L \Rightarrow |\alpha| = 4p \geq p, \forall p \in N(nr.nat.) \Rightarrow \alpha = \underbrace{a...a}_{p}\underbrace{b...b}_{p}\underbrace{a...a}_{p}\underbrace{b...b}_{p}$

Avem $\alpha = u \cdot v \cdot w \cdot x \cdot y$ astfel încât $|v \cdot w \cdot x| \leq p$ și $|v \cdot x| \geq 1$
=> $1 \leq |v \cdot x| \leq p$ *(pentru că w are voie să fie inclusiv cuvîntul vid).*

Caz I: Dacă vwx este în prima jumătate a lui α, atunci alegem i=2 și rezultă după pompare cuvântul $\beta = u \cdot v^2 \cdot w \cdot x^2 \cdot y$ => |β|=|α|+|vx| => |β| este mai mare decât |α| cu maxim p litere (și minim o literă) => jumătatea cuvântului se mută spre stânga cu maxim p/2 poziții (și minim 1), deci nu va mai fi între un "b" și un "a", ci va fi între doi de "b" *(indiferent dacă pompăm doar a-uri, doar b-uri, sau amândouă în prima jumătate a cuvântului)*. Rezultă că cuvântul β începe cu litera "a", dar prima literă din a doua lui jumătate este "b", deci cuvântul $\beta \notin L$ *(pentru că nu este de forma ww)*, contradicție cu proprietatea (3) din lemă.

Caz II: Dacă vwx este în a doua jumătate a lui α, atunci alegem i=2 și rezultă după pompare cuvântul $\beta = u \cdot v^2 \cdot w \cdot x^2 \cdot y$ => |β|=|α|+|vx| => |β| este mai mare decât |α| cu maxim p litere (și minim o literă) => jumătatea cuvântului se mută spre dreapta cu maxim p/2 poziții (și minim 1), deci nu va mai fi între un "b" și un "a", ci va fi între doi de "a" *(indiferent dacă pompăm doar a-uri, doar b-uri, sau amândouă în a doua jumătate a cuvântului)*. Rezultă că cuvântul β se termină cu litera "b", dar ultima literă din prima lui jumătate este "a", deci cuvântul $\beta \notin L$ *(pentru că nu este de forma ww)*, contradicție cu proprietatea (3) din lemă.

Caz III: Dacă vwx intersectează mijlocul cuvântului α (vwx conține cel puțin una din cele 2 litere din mijloc), atunci alegem i=0 și rezultă după pompare cuvântul $\beta = u \cdot v^0 \cdot w \cdot x^0 \cdot y = u \cdot w \cdot y$ care este de forma $a^p b^{p-s} a^{p-r} b^p$, cu $1 \leq s + r \leq p$. Rezultă că cuvântul β are mai puțini de "b" în prima parte decât la final sau are mai puțini de "a" în a doua parte decât la început (sau ambele), deci nu este de forma ww => $\beta \notin L$, contradicție cu proprietatea (3) din lemă.

[de adaugat…] **Varianta 2** (demonstratie mai pe scurt).

Exemplu: $\boldsymbol{L2} = \cdots$
[de adaugat…]

EX_2: Demonstrați că următoarele limbaje NU sunt independente de context.
$L3 = \{a^n b^n c^n | n \geq 1\}$
$L4 = \{a^n b^m c^n d^m | n \geq 1, m \geq 1\}$
$L5 = \{w \cdot w \cdot w \mid w \in \{a,b\}^*\}$
$L6 = \{a^n b^m c^r | n > m > r \geq 1\}$
$L7 = \{a^{n^2} | n \geq 1\} = \{a^1, a^4, a^9, a^{16}, a^{25}, a^{36}, \ldots\}$
$L8 = \{a^{2^n} | n \geq 0\} = \{a^1, a^2, a^4, a^8, a^{16}, a^{32}, a^{64}, \ldots\}$