Lema de pompare pentru limbaje independente de context

Fie L un limbaj independent de context.

Atunci $\exists p \in N$ (număr natural) astfel încât pentru $\forall$ *cuvant* $\alpha \in L$, cu $|\alpha| \geq p$,

$\exists u, v, w, x, y$ astfel încât cuvântul α poate fi scris ca $\alpha = u \cdot v \cdot w \cdot x \cdot y$ (concatenarea lor)

cu proprietățile:

(1) $|v \cdot w \cdot x| \leq p$

(2) $|v \cdot x| \geq 1$

(3) $u \cdot v^i \cdot w \cdot x^i \cdot y \in L, \forall i \geq 0$

Lema de pompare se folosește negată, pentru a arăta ca un limbaj L nu este independent de context.

Schema generală a demonstrației:

- ($\exists$) Alegem un cuvânt α din limbajul L astfel încât $|\alpha| \geq p$, $\forall p \in N$. *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 un număr natural i (nu neapărat același pentru toate cazurile) astfel încât cuvântul $u \cdot v^i \cdot w \cdot x^i \cdot y \notin L$, rezultând o contradicție a lemei (a proprietății (3)), deci a presupunerii că limbajul L era independent de context.

(2013) Exemplu: $L=\{w\cdot w \mid w\in\{a,b\}^*\}$ 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.

Exemplu: $L = \{w \in \{a\}^*, |w| = nr.prim\}$

Demonstrație: Presupunem că L este limbaj independent de context, rezultă că există acel număr natural p din lemă.

Alegem $\alpha = a^n \in L, n = nr.prim, n \geq p+2$ *(cel mai mic număr natural este 0, dar cel mai mic număr prim este 2)* => $|\alpha| = n \geq p+2 \geq p, \forall p \in N(nr.nat.)$

Avem $\alpha = u \cdot v \cdot w \cdot x \cdot y$ astfel încât $|v \cdot w \cdot x| \leq p$ și $|v \cdot x| \geq 1$.

Fie |vx|=k => vx = a^k.

Din relațiile de mai sus rezultă $1 \leq k = |v \cdot x| \leq |v \cdot w \cdot x| \leq p \leq n-2$ (*)

Alegem $i = n-k \geq 2$ (conform (*)) și după pompare rezultă cuvântul $\beta = u \cdot v^{n-k} \cdot w \cdot x^{n-k} \cdot y$, având lungimea $|\beta| = |u \cdot v^{n-k} \cdot w \cdot x^{n-k} \cdot y| = |u \cdot w \cdot y| + |v^{n-k} \cdot x^{n-k}| = |a^{n-k}| + |vx| * (n-k)$

$= (n-k) + k * (n-k) = \underbrace{(1+k)}_{\geq 2 (cf. *)} * \underbrace{(n-k)}_{\geq 2 (cf. *)}$ => |β| nu poate fi un număr prim *(este produsul a două numere minim 2)*, deci $\beta \notin L$, contradicție cu proprietatea (3) din lemă.