TUTORIAT LFA 4: LEMA DE POPARE PENTRU LIMBAJE REGULATE

RADU COSTACHE, MARIA PREDA

1. Breviar Teoretic

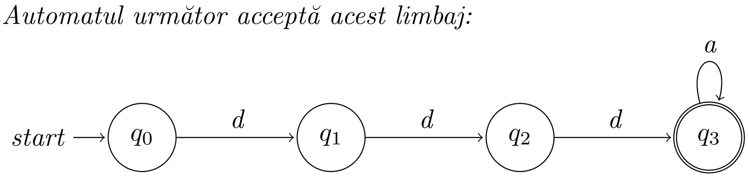
Observaţie 1.1. *Este decidabil dacă un limbaj poate fi sau nu modelat cu ajutorul unui DFA.*

Observaţie 1.2. *Ştim că fiecare limbaj finit poate fi modelat cu ajutorul unui automat (de exemplu, pentru fiecare cuvânt din limbaj putem avea câte o ramură a automatului). Problema apare când dorim săne d*
avm seama dacă se poate construi un automat pentru un limbaj infinit.

Exemplu 1.3. *Un limbaj infinit regulat este:*

$$L = \{d^3a^n | n \in \mathbb{N}\}$$

Automatul următor acceptă acest limbaj:



Exemplu 1.4. *Un limbaj inifinit care nu este regulat este:*

$$L = \{n^p u^p | p \in \mathbb{N}\}$$

Intuitie: *Având un limbaj modelat de un DFA, pentru fiecare cuvânt din limbaj, există o ramura a automatului pe care cuvântul este acceptat (un drum de la starea iniţială la cea finală în care este acceptat). Dacă un cuvânt are mai multe litere decât numarul de stări din automat, atunci este clar că va trece de mai multe ori prin minim o stare (Principiul lui Dirchlet, cunoscut şi sub denumirea de Principiul Cutiei). Aşadar, pe drumul dinspre starea initiala spre cea finala, la un moment dat, fie se va invarti pe o bucla de deasupra unei stari, fie sa va intoarce intr-o stare prin care mai trecuse, continuand dupa pe acelasi drum. Astfel, daca am repeta de mai multe ori acest ciclu in timpul parcurgerii drumului, noul cuvant ar fi in continuare acceptat.*

Teoremă 1.5. ***(Lema de pompare)*** *Fie $M = (Q, \Sigma, \delta, s, F)$ un DFA si p numarul de stari ale lui M. Pentru orice cuvant $x \in L(M), |x| \geq p$ (orice cuvant din limbajul acceptat de automat, cu numarul de simboluri mai mare decat numarul de stari ale lui M), x poate fi descompus in trei "subcuvinte" u, v, w, cu $x = uvw$, astfel incat:*

(1) $|uv| \leq p$

(2) $|v| \geq 1$

(3) $\forall i \geq 0, uv^i w \in L(M)$

Observaţie 1.6. *Orice limbaj regulat indeplineste proprietatile din teorema. Totusi, sa demonstram ca teorema se aplica pentru un limbaj regulat este, in general, mai complicat decat sa aratam prin alte metode.*

Observaţie 1.7. *Asa cum stim de la logica, din primul semestru:*

$$A \to B \iff \neg B \to \neg A$$

Observaţie 1.8. *Un limbaj este regulat daca si numai daca lema de pompare functioneaza. Asadar, daca putem gasi un cuvant dintr-un limbaj pentru care lema nu tine, limbajul nu este regulat. In general, Lema de pompare se foloseste in sens negativ, pentru a demonstra ca limbajele nu sunt regulate.*

Exemplu 1.9. *Este urmatorul limbaj regulat? Demonstrati!*

$$L = \{a^n b^n | n \in \mathbb{N}\}$$

Demonstraţie. Presupunem prin reducere la absurd ca limbajul L este regulat. Daca este regulat, inseamna ca lema de pompare este adevarata pentru el. Fiind un limbaj regulat, poate fi construit un automat finit care sa il accepte (Atentie! Automatul este finit, nu limbajul). Fie $p = \#Q$ (numarul de stari din automat).

Alegem cuvantul $\alpha = a^p b^p$. In mod evident, acest cuvant apartine limbajului si are mai multe litere decat stari in automat: $|\alpha| = 2p > p$. Conform lemei, cuvantul poate fi scris ca $\alpha = uvw$.

Din **1)** avem ca $|uv| \leq p$. Cuvantul ales de noi are p litere a, ceea ce inseamna ca "subcuvintele" u si v contin doar litere a.

Atunci v este de forma a^k, unde $k \leq p$.

Din conditia **3)** avem ca $uv^iw \in L, \forall i \geq 0$. Vom alege $i = 2$ (in acest caz, am fi putut alege orice $i \geq 2$). Vom nota cu β acest nou cuvant $\beta = uv^iw = a^{p+k}b^p$. In mod evident, acest cuvant nu mai apartine limbajului, puterile simbolurilor fiind diferite. Inseamna ca ultimul punct al lemei nu este respectat $\Rightarrow$ limbajul nu este regulat $\Rightarrow$ contradictie.

□

2. Exercitii

(1) *Verificati daca urmatoarele limbaje sunt regulate. Daca da, construiti automatul, altfel utilizati lema de pompare pentru a demonstra ca nu sunt:*

(a) $L_1 = \{a^k b^{3l} a^l | k \geq 1, l \geq 0\}$

Soluţie:
Limbajul nu este regulat si vom demonstra acest lucru ca in exemplul de la partea de teorie. Presupunem prin reducere la absurd ca limbaj este regulat, deci lema de pompare se aplica pentru el.
Fie p numarul de stari din automat. Fiind un limbaj infinit, exista cuvinte care au mai mult de p litere. Vom alege cuvantul $\alpha = ab^{3p}a^p \in L_1$, cu $|\alpha| = 4p + 1 \geq p, \forall p \in \mathbb{N}$. Conform lemei, cuvantul poate fi scris sub forma $\alpha = uvw$.
*Din conditia **1)** a lemei, avem ca $|u \cdot v| \leq p$, ceea ce inseamna ca "subcuvantul" uv poate avea doar litere "a" sau "b".*
***Cazul 1:** Avem v de forma a (ceea ce inseamna ca u este cuvantul vid). Conform punctului **3)** al lemei, orice cuvant $\beta = uv^iw$ trebuie sa apartina in continuarea limbajului. Alegem $i = 0$, deci $\beta = b^{3p}a^p$, dar $k \geq 1 \Rightarrow$ contradictie.*
***Cazul 2:** Avem v de forma $a^n b^m$, $0 \leq n \leq 1, 1 \leq m \leq p - 1$, deoarece $1 \leq |m + n| \leq p$.*
*Din conditia **3)**, stim ca orice cuvant $\beta = uv^iw$, cu $i \geq 0$ apartine limbajului si alegem $i = 0$:*

$$\beta = uv^iw = uw = a^{1-n}b^{3p-m}a^p \iff 3p - m = 3p \iff m = 0 \Rightarrow \textit{Contradictie}$$

Pentru ca niciunul dintre cazuri nu satisface lema de pompare, inseamna ca ea nu se aplica, deci presupunerea facuta este falsa, asadar limbajul nu este regulat.

(b) $L_2 = \{0^{2k}1^{3k}0^{5k'} | k, k' \geq 2\}$

Soluţie:
Limbajul nu este regulat. Intuitiv, ne putem da seama de acest fapt deoarece exista doua puteri care nu sunt independente.
Presupunem prin reducere la absurd ca limbajul este regulat $\Rightarrow$ lema de pompare este adevarata pentru el.
Fie p numarul de stari din automat. Fiind un limbaj infinit, exista cuvinte care au mai mult de p litere. Vom alege cuvantul $\alpha = 0^{2p}1^{3p}0^{5p} \in L_2$, cu $|\alpha| = 10p \geq p, \forall p \in \mathbb{N}$. Conform lemei, cuvantul poate fi scris sub forma $\alpha = uvw$.
*Din conditia **1)** a lemei, avem ca $|u \cdot v| \leq p$, ceea ce inseamna ca "subcuvantul" uv poate avea doar simboluri "0", deci v este de forma 0^n, unde $1 \leq n \leq p$.*
*Conform punctului **3)** din lema de pompare, orice cuvant $\beta = uv^iw \in L_2$. Alegem $i = 2$ si obtinem:*

$$\beta = 0^{2p+n}1^{3p}0^{5p}$$

Dar stim ca numarul de "1" este de $\frac{3}{2}$ *ori mai mare decat numarul de "0"-uri de la inceputul cuvantului, deci:*

$$\frac{3(2p+n)}{2} = 3p \iff 3p + \frac{3n}{2} = 3p, \textit{ unde } n \geq 1 \iff \frac{3n}{2} \geq \frac{3}{2} \Rightarrow \textit{Imposibil}$$

$\Rightarrow$ *Lema de pompare nu este adevarata pentru orice cuvant din limbaj, deci presupunerea facuta este falsa* $\Rightarrow$ *limbajul nu este regulat.*

(c) $L_3 = \{0^{k-3}1^{2l+3} | k \geq 3l\}$

Soluţie:
Limbajul nu este regulat. Vom presupun prin reducere la absurd ca limbajul este regulat, deci lema de pompare este adevarata.
Fie p numarul de stari din automat. Fiind un limbaj infinit, exista cuvinte care au mai mult de p litere. Vom alege cuvantul $\alpha = 0^{3p-3}1^{2p+3} \in L_3$*, cu* $|\alpha| = 5p \geq p, \forall p \in \mathbb{N}$*. Conform lemei, cuvantul poate fi scris sub forma* $\alpha = uvw$*.*
Din conditia ***1)*** *a lemei, avem ca* $|u \cdot v| \leq p$*, ceea ce inseamna ca "subcuvantul" uv poate avea doar simboluri "0", deci v este de forma* 0^n*, unde* $1 \leq n \leq p$*.*
Conform punctului ***3)*** *din lema de pompare, orice cuvant* $\beta = uv^iw \in L_3$*. Alegem* $i = 0$ *si obtinem:*

$$\beta = 0^{3p-n-3}1^{2p+3}$$

Pentru ca este din limbaj, inseamna ca $3p - n \geq 3p$*, dar stim ca* $n \geq 1 \Rightarrow$ *Contradictie*
Inseamna ca limbajul nu este, de fapt, regulat.

(d) $L_4 = \{wa^kw | w \in \{a, b, c\}^*, k \geq 0\}$

Soluţie:
Presupunem ca limbajul este regulat, deci lema de pompare este adevarata.
Fie p numarul de stari din automat. Fiind un limbaj infinit, exista cuvinte care au mai mult de p litere. Vom alege cuvantul $\alpha = b^pa^pb^p \in L_4$*, cu* $|\alpha| = 3p \geq p, \forall p \in \mathbb{N}$*. Conform lemei, cuvantul poate fi scris sub forma* $\alpha = uvw$*.*
Din conditia ***1)*** *a lemei, avem ca* $|u \cdot v| \leq p$*, ceea ce inseamna ca "subcuvantul" uv poate avea doar simboluri "b", deci v este de forma* b^n*, unde* $1 \leq n \leq p$*.*
Conform punctului ***3)*** *din lema de pompare, orice cuvant* $\beta = uv^iw \in L_4$*. Alegem* $i = 2$ *si obtinem:*

$$\beta = b^{p+n}a^pb^p$$

Singurele letere a sunt cele p, deci w in acest caz este b^{p+n}*, dar observam ca dupa terminarea literelor "a" din centru, avem* b^p*, deci:*

$$b^{p+n} = b^p, \textit{ dar } n \geq 1 \Rightarrow \textit{Contradictie}$$

Presupunerea facuta a fost falsa, deci limbajul nu este regulat.

(e) $L_5 = \{a^{k-1}b^{2k+3} | k \geq 5\}$

Soluţie:
Limbajul nu este regulat, deoarece puterile la care apar cele doua simboluri depind ambele de k.
Presupunem prin reducere la absurd ca limbajul este regulat. Atunci $\exists p \in \mathbb{N}$ *numarul de stari din automatul care accepta limbajul si se aplica lema de pompare.*
Alegem $\alpha = a^{5p-1}b^{10p+3} \in L_5$*, cu* $|\alpha| = 15p + 2 \geq p$ *(deci lungimea cuvantului respecta ipoteza lemei). Conform lemei, cuvantul poate fi scris sub forma* $\alpha = uvw$*.*

*Din conditia **1)** avem* $|uv| \leq p$*, ceea ce inseamna ca* v *este poate contine doar simboluri* a*, adica este de forma* $v = a^n$*, unde* $1 \leq n \leq p$*.*
*De asemenea, conditia **3)** spune ca* $uv^iw \in L_5, \forall i \geq 0$*. Alegem* $i = 0$ *si avem cuvantul* $\beta = uv^0w = uw = a^{5p-n-1}b^{10p+3} \in L_5$*. In acest caz, uitandu-ne la puterea lui* b*, observam ca am avea* $k = 5p$*.*
Inseamna ca $5p - n - 1 = 5p - 1$*, deci avem* $n = 0$*, dar* $n \geq 1$*.*
Am ajuns la o contradictie cu presupunerea facuta, deci aceasta era falsa.
Inseamna ca limbajul nu este regulat.

(f) $L_6 = \{a^n | n\ prim\}$

Soluţie:
Limbajul nu este regulat.
Presupunem prin reducere la absurd ca limbajul este regulat. Atunci $\exists p \in \mathbb{N}$ *numarul de stari din automatul care accepta limbajul si se aplica lema de pompare.*
Alegem $\alpha = a^t \in L_6$*, cu* $|\alpha| = t \geq p$ *(deci lungimea cuvantului respecta ipoteza lemei). Conform lemei, cuvantul poate fi scris sub forma* $\alpha = uvw$*.*
*Din conditia **1)** avem* $|uv| \leq p$*, ceea ce inseamna ca* v *este poate contine doar simboluri* a*, adica este de forma* $v = a^k$*, unde* $1 \leq k \leq p$*.*
*De asemenea, conditia **3)** spune ca* $uv^iw \in L_6, \forall i \geq 0$*. Alegem* $i = t + 1$ *si avem cuvantul* $\beta = uv^{t+1}w = a^{t+kt} \in L_6$*.*
Observam ca $t + tk$ *este divizibil cu* t*, deci nu este un numar prim. Contradictie.*
Inseamna ca limbajul nu este regulat.