

Context free grammars - Gramatici independente de context. Automate stiva

April 2, 2019

$\{a^n b^n\}$ e regulat?

- ▶ Pt a fi limbaj regulat, ar trebui sa existe un **automat finit** care sa-l recunoasca.
- ▶ ar trebui sa tina minte cati a a citit, dar n nu este limitat
- ▶ dupa ce a citit a^m ar trebui sa fie intr-o stare ce specifica o multime de simboluri nonterminale din care sa fie derivate exact b^m . \Rightarrow pt fiecare m ar trebui sa fie o stare distinctica
- ▶ deci automatul ar trebui sa aiba evidenta unui numar nelimitat de posibilitati

acest lucru nu se poate face cu un numar finit de stari

- ▶ dar, nu tot ce pare a avea nevoie de memorie nelimitata, chiar are:
 - ▶ $C = \{w \mid w \text{ are un numar egal de } 0 \text{ si } 1\}$
 - ▶ $D = \{w \mid w \text{ are un numar egal de aparitii } 01 \text{ si } 10\}$

$\{a^n b^n\}$ e regulat?

- ▶ Pt a fi limbaj regulat, ar trebui sa existe un **automat finit** care sa-l recunoasca.
- ▶ ar trebui sa tina minte cati a a citit, dar n nu este limitat
- ▶ dupa ce a citit a^m ar trebui sa fie intr-o stare ce specifica o multime de simboluri nonterminale din care sa fie derivate exact b^m . \Rightarrow pt fiecare m ar trebui sa fie o stare distincta
- ▶ deci automatul ar trebui sa aiba evidenta unui numar nelimitat de posibilitati

acest lucru nu se poate face cu un numar finit de stari

- ▶ dar, nu tot ce pare a avea nevoie de memorie nelimitata, chiar are:
 - ▶ $C = \{w \mid w \text{ are un numar egal de } 0 \text{ si } 1\}$
 - ▶ $D = \{w \mid w \text{ are un numar egal de aparitii } 01 \text{ si } 10\}$

D este limbaj regulat

$D = \{w \mid w \text{ are un numar egal de aparitii } 01 \text{ si } 10 \text{ ca substringuri}\}$

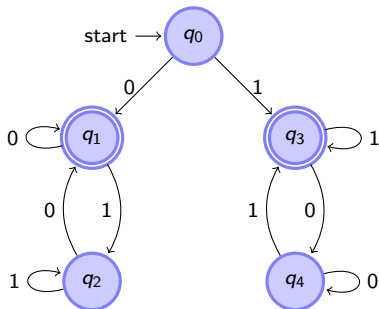
$$D = \{0, 1, \varepsilon$$

w daca incepe cu 0 se termina cu 0

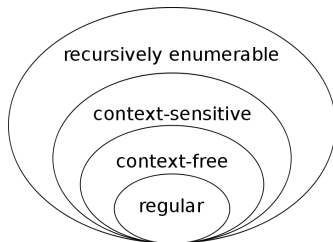
w daca incepe cu 1 se termina cu 1}

? 101, 1010, 0110

$$1 + 0 + \varepsilon + 0(0 + 1)^*0 + 1(0 + 1)^*1$$



Gramatici independente de context. Context-free grammars



- ▶ $G = (T, N, Z, P)$ e independenta de context daca
- ▶ fiecare productie are forma

$$X \rightarrow \chi, X \in N, \chi \in V^*$$

Un limbaj care e definit de o gramatica independenta de context este limbaj independent de context.

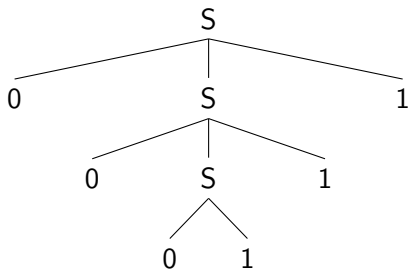
Exista CFL care nu sunt Regular languages

$0^n 1^n$

Fie $G = (\{0, 1\}, \{S\}, \{S \rightarrow 01|0S1\}, S)$

► $S \Rightarrow 0S1 \Rightarrow 00S11 \Rightarrow 000111$

Arbore de parsare pentru 000111:



Limbajul parantezelor

Fie $G = (\{(,)\}, \{S\}, \{S \rightarrow SS|(S)|(())\}, S)$

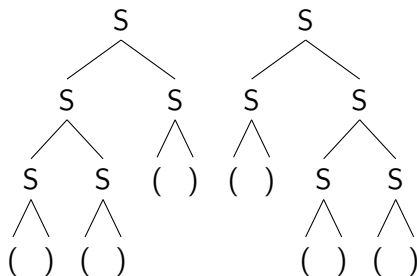
- ▶ left-most derivation: $S \Rightarrow \mathbf{SS} \Rightarrow (\mathbf{S})S \Rightarrow (())S \Rightarrow (())()$
Deci $S \Rightarrow^L (())()$
- ▶ right-most derivation: $S \Rightarrow \mathbf{SS} \Rightarrow \mathbf{S}() \Rightarrow (S)() \Rightarrow (())()$
Deci $S \Rightarrow^R (())()$

Gramatica ambigua - reamintire

O gramatica e ambigua daca exista un string in limbaj care e parsat in doi arbori de derivare.

Pentru $G = (\{(,)\}, \{S\}, \{S \rightarrow SS|(S)|() \}, S)$

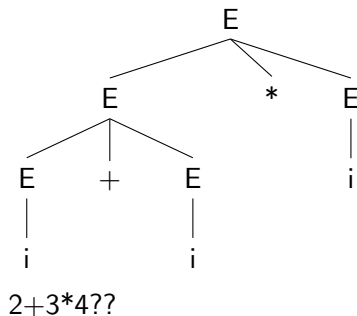
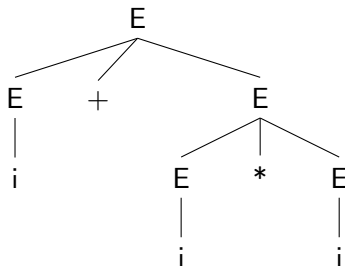
Derivarea $()()()$:



Ambiguitate

Fie $G_4 = (\{+, *, i\}, \{E\}, E, P)$ Doua derivari distincte stanga, doua derivari distincte dreapta

- ▶ $E \rightarrow E + E$
- ▶ $E \rightarrow E * E$
- ▶ $E \rightarrow i$



CFG pentru Engleza

- ▶ $T = \{sleeps, saw, man, woman, telescope, the, with, at\}$
- ▶ $N = \{S, NP, VP, PP, DT, Vi, Vt, NN, IN\}$

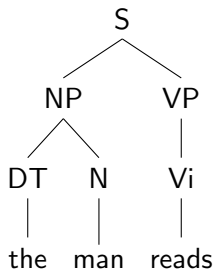
▶ P=	S	→	NP VP	Vi	→	eats
	VP	→	Vi	Vt	→	saw
	VP	→	Vt NP	N	→	man
	VP	→	VP PP	N	→	woman
	NP	→	DT NN	N	→	telescope
	NP	→	NP PP	DT	→	the
	PP	→	IN NP	IN	→	with
				IN	→	at

S = sentence, VP = verb phrase, NP = noun phrase, PP = prepositional phrase, DT = determiner, Vi = intransitive verb, Vt = transitive verb, N = noun, IN = preposition

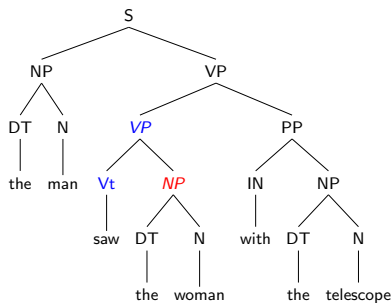
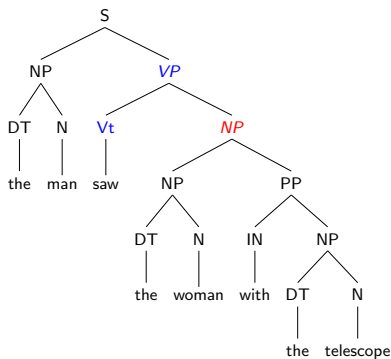
Arbore de derivare

Derivarea stanga:

$S \Rightarrow \mathbf{NP} \ VP \Rightarrow \mathbf{DT} \ N \ VP \Rightarrow the \ \mathbf{N} \ VP \Rightarrow the \ man \ \mathbf{VP}$
 $\Rightarrow the \ man \ \mathbf{Vi} \Rightarrow the \ man \ eats$



The man saw the woman with the telescope.



Ce a vazut "the man"?

if then else grammar

► $T = \{if, then, else, E1, E2, S1, S2, S3\}$, $N = \{stmt, expr\}$

► $P =$

stmt \rightarrow if *expr* then *stmt*

stmt \rightarrow if *expr* then *stmt* else *stmt*

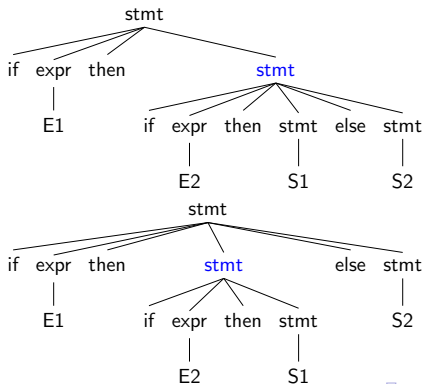
stmt \rightarrow S1 | S2 | S3

expr \rightarrow E1 | E2

if E1	
then	if E2
	then S1
	else S2

sau?

if E1	
then	if E2
	then S1
else S2	



if then else - rezolvare ambiguitate

- ▶ $T = \{if, then, else, E1, E2, S1, S2, S3\}$,
 $N = \{stmt, matched_stmt, unmatched_stmt, expr\}$
- ▶ $P =$

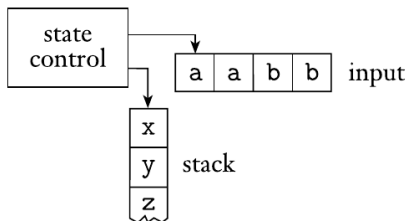
$stmt$	\rightarrow	m_stmt
	$ $	um_stmt
m_stmt	\rightarrow	if $expr$ then m_stmt else m_stmt
um_stmt	\rightarrow	if $expr$ then $stmt$
	$ $	if $expr$ then m_stmt else um_stmt
$stmt$	\rightarrow	$S1 \mid S2 \mid S3$
$expr$	\rightarrow	$E1 \mid E2$

Intre un *then* si un *else* e permis doar *matched_stmt*.
 $m_stmt = matched_stmt$ (if cu ambele *then* si *else*),
 $um_stmt = unmatched_stmt$

Letia and Chifu. 2.3, 2.3.1 Sipser - 2.1,2.2

Automat stiva. Push down automaton

- ▶ Niciun automat finit nu poate fi construit pt a recunoaste $a^n b^n$ sau limbajul parantezelor - structuri imbricate
- ▶ Se creste puterea automatelor finite prin adaugarea unei stive (stack) drept structura aditionala de memorie



- ▶ Daca gramaticile regulate sunt o subclasa a gramaticilor independente de context, de ce se dezvolta metode specifice gramaticilor regulate si nu se aplica pt acestea cele de la gramaticile independente?
- ▶ Datorita complexitatii analizei gramaticilor independente de context: gramaticile regulate sunt mai simplu de analizat

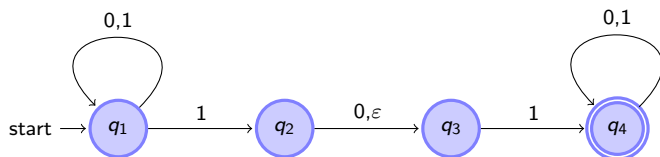
Idee: Push down automata: 00001111

1. citește simboluri de la intrare
2. la fiecare 0 citit, împinge-l pe stivă
3. la fiecare 1 citit, scoate de pe stivă un 0
4. dacă citirea stringului se termină când stivă se golește, acceptă stringul. Dacă stivă devine goală când mai sunt 1 de citit sau s-a terminat sirul și în stivă încă mai sunt 0-uri, respinge stringul

NFA- Sipser 1.2 (optional)

Un automat finit nedeterminist este $(Q, \Sigma, \delta, q_0, F)$, unde:

1. Q este setul de stari
2. Σ un alfabet finit de intrare
3. $\delta : Q \times \Sigma_{\epsilon} \rightarrow P(Q)$ este o functie de tranzitie
4. $q_0 \in Q$ este starea de start
5. $F \subseteq Q$ setul de stari finale



$$\delta(q_1, 1) = \{q_1, q_2\}$$

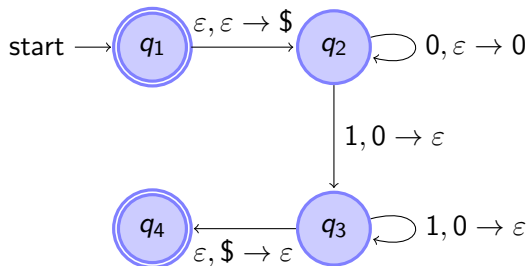
Definitie formală 1 a automatului stivă (Sipser)

Un automat stivă este $(Q, \Sigma, \Gamma, \delta, q_0, F)$, unde Q, Σ, Γ, F sunt seturi finite:

1. Q este setul de stări
2. Σ un alfabet de intrare
3. Γ este alfabetul stivei
4. $\delta : Q \times \Sigma_{\varepsilon} \times \Gamma_{\varepsilon} \rightarrow P(Q \times \Gamma_{\varepsilon})$ este o funcție de tranziție
5. $q_0 \in Q$ este starea de start
6. $F \subseteq Q$ setul de stări finale

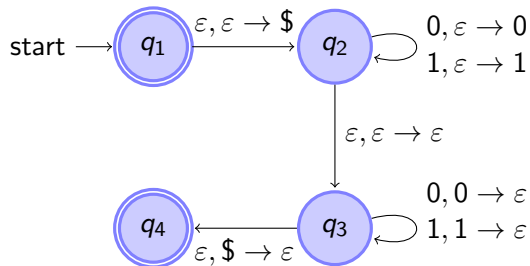
PDA for 0^n1^n

Let $M_1 = (Q, \Sigma, \Gamma, \delta, q_1, F)$



PDA for ?

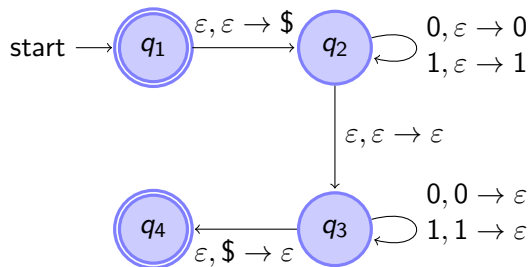
File $M_1 = (Q, \{0, 1\}, \{\$, 0, 1\}, \delta, q_1, \{q_1, q_4\})$



PDA for $\{ww^R \mid w \in \{0,1\}^*\}$

$w^R = w$ scris invers

Fie $M_1 = (Q, \{0,1\}, \{\$,0,1\}, \delta, q_1, \{q_1, q_4\})$



la fiecare pas, ghiceste daca a ajuns la mijlocul stringului sau nu

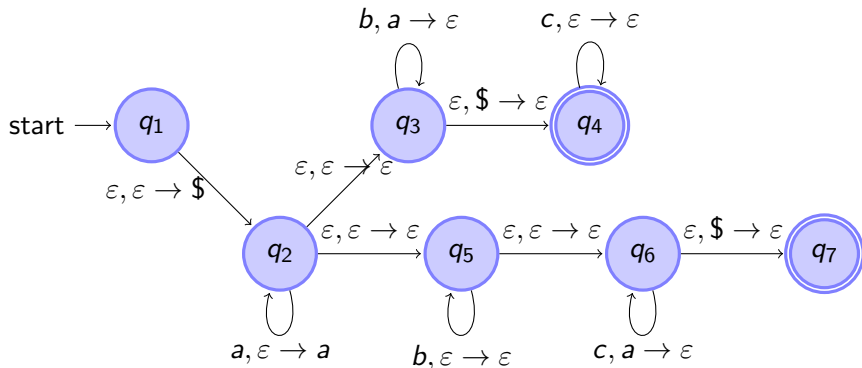
Gramatica palindrom par

- Palindrom: (T, N, P, A) , $P =$

$$\begin{aligned} &\{A \rightarrow 0A0 \mid 1A1 \\ &A \rightarrow \varepsilon\} \end{aligned}$$

PDA for $\{a^i b^j c^k \mid i, j, k \geq 0, i = j \text{ sau } i = k\}$

Fie $M_1 = (Q, \Sigma, \Gamma, \delta, q_1, F)$



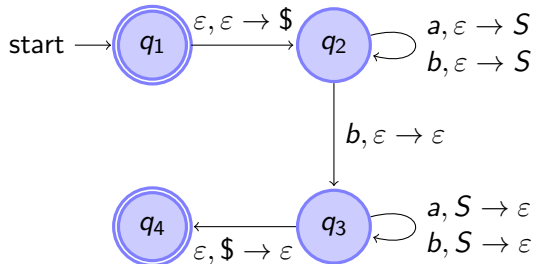
ghiceste daca e acelasi numar de a si b sau a si c

$$\{vbw \mid v, w \in \{a, b\}^*, |v| = |w|\}$$

$$M = (Q, \{a, b\}, \{\$, S\}, \delta, q_1, \{q_4\})$$

$$\{vbw \mid v, w \in \{a, b\}^*, |v| = |w|\}$$

$$M = (Q, \{a, b\}, \{\$, S\}, \delta, q_1, \{q_4\})$$



Exemplu Gramatica independenta de context

- Palindrom: (T, N, P, A) , $P =$

$$\{A \rightarrow 0A0|1A1$$
$$A \rightarrow \varepsilon\}$$

- Acelasi numar de 0 si 1: $(\{0, 1\}, \{A\}, P, A)$, $P =$

$$\{A \rightarrow 0A1A|1A0A$$
$$A \rightarrow \varepsilon\}$$

Automat finit - reamintire

Automat finit (finite automaton, finite state acceptor):

$$A = (T, Q, R, q_0, F)$$

- ▶ Q set nevid - setul starilor interne
- ▶ $(T \cup Q, R)$ sistem de rescriere; $T \cap Q = \emptyset$
- ▶ $q_0 \in Q$ - starea initiala
- ▶ $F \subseteq Q$ - stari finale
- ▶ fiecare element din R are forma $qt \rightarrow q'$, $q, q' \in Q, t \in T$

$$L(A) = \{\tau \in T^* \mid q_0\tau \Rightarrow^* q, q \in F\}$$

Automat stiva - definitie sistem de rescriere

Automat stiva

$$A = (T, Q, R, q_0, F, S, s_0)$$

, unde:

- ▶ Q set nevid - setul starilor interne
- ▶ $(T \cup Q \cup S, R)$ sistem de rescriere; $T \cap Q = \emptyset$
- ▶ $q_0 \in Q$ - starea initiala
- ▶ $s_0 \in S \cup \{\varepsilon\}$ - simboluri stiva, s_0 continutul initial al stivei
- ▶ $F \subseteq Q$ - stari finale
- ▶ fiecare element din R are forma $\sigma q t \tau \rightarrow \sigma' q' \tau$,
 $\sigma, \sigma' \in S^*$, $q, q' \in Q$, $t \in T \cup \varepsilon$, $\tau \in T^*$

Daca automatul e la configuratia $s_1 \dots s_n q \tau$ intr-o derivare, automatul e in starea q , τ este partea necitita din input, s_1, \dots, s_n este continutul pe stiva, s_n in varf.

Limbaj acceptat

Daca automatul e la configuratia $s_1...s_nq\tau$ intr-o derivare, automatul e in starea q , τ este partea necitita din input, s_1, \dots, s_n este continutul pe stiva, s_n in varf.

$$L(A) = \{\tau | s_0q_0\tau\# \Rightarrow^* q\#, q \in F, \tau \in T^*\}$$

$0^n 1^n$

$M_1 = (\{0, 1\}, \{q_2, q_3\}, R, q_2, \{q_3\}, \{0, 1\}, \varepsilon), R = \{$

1. $\varepsilon q_2 0 \rightarrow 0 q_2$

2. $0 q_2 1 \rightarrow \varepsilon q_3$

3. $0 q_3 1 \rightarrow \varepsilon q_3\}$

$??\varepsilon q_2 0011 \Rightarrow^* q_3$

Pe stiva pot fi alte simboluri decat cele din alfabetul de intrare.

CFG - PDA

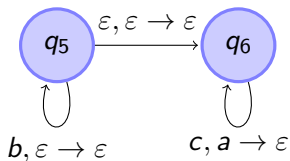
Pentru fiecare gramatica independenta de context G exista un automat stiva A a.i. $L(A)=L(G)$.

Care din urmatoarele afirmatii sunt adevarate pt gramatica

$G = (\{(,)\}, \{S\}, \{S \rightarrow SS|(S)|()\}, S)$

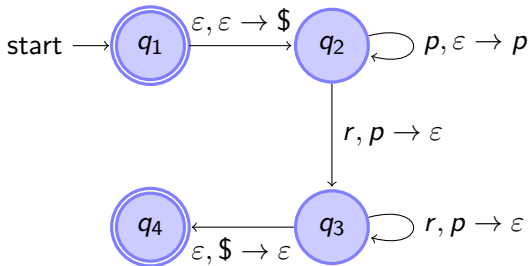
1. $S \Rightarrow^L ()S$
2. $S \Rightarrow^R ()S$
3. $S \Rightarrow^* (())(S)$
4. $()$ nu e generat de gramatica

Pentru fragmentul dintr-un automat stiva A , care afirmatii sunt adevarate?



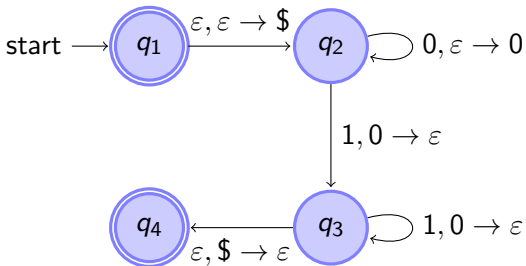
1. daca la intrare este c , starea este q_6 , in varful stivei este a ; A va ramane in q_6
2. din q_5 in q_6 A trece fara a citi niciun simbol de la intrare si fara a modifica stiva
3. tranzitia din q_5 in q_6 are loc doar daca se poate scoate din stiva a si la intrare nu conteaza
4. daca la intrare este b , starea este q_5 ; automatul va ramane cu siguranta in q_5

Pentru automatul din figura, ce este posibil sa avem in stiva?



1. oricat de multi de p , dar doar p
2. oricat de multi de r , dar doar r
3. orice combinatie de p si r
4. secvente de pr

Pentru automatul din figura, daca in stiva la un moment dat avem 0, care afirmatie e adevarata



1. daca A este in q3 si la intrare este 0, ramane in q3
2. stringul acceptat incepe cu 0 si se termina cu 1
3. va fi acceptat stringul care incepe cu 1 si se termin cu 0
4. daca A este in q2 si la intrare este 0, trece in q3