Gramatici. Automate finite.

Curs 3

Gramatica $G = (N, \Sigma, P, S)$

Un ansamblu $G = (N, \Sigma, P, S)$ unde:

- N este un alfabet de simboluri ***neterminale***;
- Σ este un alfabet de simboluri ***terminale***;
- $P \subseteq (N \cup \Sigma)^*N\,(N \cup \Sigma)^* \times (N \cup \Sigma)^*$, P este o multime finita de ***productii***;
- S∈N, S este simbolul initial;

se numeste ***gramatica***.

Daca (α,β)∈P, atunci notam productia α→β si inseamna ca α se inlocuieste cu β

N∩Σ=Φ

Folosite pentru specificarea limbajelor de programare.

Clasificarea gramaticilor (Chomsky)

Gramatici de tip 1 Gramatici dependente de context

- Gramatici in care pentru orice productie $\alpha \rightarrow \beta \in P$ avem $|\alpha| \leq |\beta|$
- Daca $S \rightarrow \varepsilon \in P$, atunci S nu apare in membrul drept al niciunei productii.

Gramatici de tip 2 Gramatici independente de context

- Gramatici in care productiile sunt de forma $A \rightarrow \alpha, A \in N \; si \; \alpha \in (N \cup \Sigma)^*$

Gramatici de tip 3 Gramatici regulare

- Gramatici in care orice productie are una din formele urmatoare:
  - $A \rightarrow aB \; sau \; A \rightarrow b, unde \; A, B \in N \; si \; a, b, \in \Sigma$
  - Daca $S \rightarrow \varepsilon \in P$, atunci S nu apare in membrul drept al niciunei productii.

Relatia de Derivare

Definita pe $(N \cup \Sigma)^*$

$\Longrightarrow$ *derivare directa*

$\gamma \Longrightarrow \delta \Leftrightarrow \exists \gamma_1, \gamma_2, \alpha, \beta \in (N \cup \Sigma)^* \ astfel\ incat\ \gamma = \gamma_1 \alpha \gamma_2, \delta = \gamma_1 \beta \gamma_2, \quad iar\ (\alpha \longrightarrow \beta) \epsilon P$

$\overset{k}{\Rightarrow}$ *k-derivare*

Succesiune de k derivari directe

$\gamma \overset{k}{\Rightarrow} \delta \Leftrightarrow \exists \alpha_1, \alpha_2, \ldots, \alpha_{k-1} \in (N \cup \Sigma)^* astfel\ incat \gamma \Rightarrow \alpha_1 \Rightarrow \cdots \Rightarrow \alpha_{k-1} \Rightarrow \delta$

$\overset{*}{\Rightarrow} *$ *derivare*

$\gamma \overset{*}{\Rightarrow} \delta, \text{daca } \gamma = \delta\ sau\ \gamma \overset{k}{\Rightarrow} \delta$

$\overset{+}{\Rightarrow} +$ *derivare*

$\gamma \overset{+}{\Rightarrow} \delta\ daca\ \exists\ k > 0\ astfel\ incat\ \gamma \overset{k}{\Rightarrow} \delta$

$\varepsilon -$ productie. Gramatica ε-independenta

- O productie de forma $A \rightarrow \varepsilon$ se numeste ***ε-productie.***

- Gramatica G=(N, Σ, P, S) se numeste ***ε-independenta***:
  - Daca $\varepsilon \notin L(G)$ atunci G nu are ε-productii
  - Daca ε∈L(G) atunci G are o singura productie $S \rightarrow \varepsilon$, iar S nu apare in membrul drept al niciunei productii.

Teorema

∀ G= (N, Σ, P, S), ∃ G' = (N', Σ', P', S) echivalentã, ε-independentã.

Notatii

- A, B, C – neterminale/nonterminale
- S∈N – simbol de start
- a, b, c ∈ Σ -terminale
- $\alpha, \beta, \gamma \epsilon (N \cup \Sigma)^*$- forme propozitionale
- x, y, z ∈Σ^* - cuvinte
- ε – cuvantul vid
- X, Y∈(N∪Σ) – simboluri din gramatica

Forma propozitionala. Limbajul generat de o gramatica.

Fie $G=(N, \Sigma, P, S)$. O secventa $x \in (N \cup \Sigma)^*$ astfel incat $\mathbf{S \overset{*}{\Rightarrow} x}$ se numeste **forma propozitionala**.

O forma propozitionala care **nu contine neterminale** se numeste ***propozitie***.

Limbajul generat de o gramatica $G=(N, \Sigma, P, S)$ este:

$$\mathbf{L(G)} = \{w \mid w \in \Sigma^*, S \overset{*}{\Rightarrow} w\}$$

Doua gramatici, G_1 si G_2 sunt *echivalente* daca ele genereaza acelasi limbaj ($L(G_1)=L(G_2)$).

Exemplu

- G=(N, Σ, P, S)
- N={S,A}
- Σ={a, b}
- P: S→AS|a

  A →b

bba ∈L(G)?

S⟹AS ⟹ bS ⟹ bAS ⟹ bbS ⟹ bba

$\exists\ S \overset{*}{\Rightarrow} bba, deci\ w = bba \in L(\mathrm{G})$

ab ∈L(G)?

Provocare

- Gramatica pentru declaratie lista identificatori C++.
- Ex: int a, b, c;
- a,b,c identificatori (notati prin id)

Solutie

- G=(N, Σ, P, Decl)
- N={Decl, Tip, ListaId}
- Σ={",", id, integer, char, double, ";"}
- P: Decl→ Tip ListaId ;

  Tip → integer|char|double

  ListaId → id| id , ListaId

```
int a;
char x, y, z;
double max, min;
```

AUTOMATE FINITE

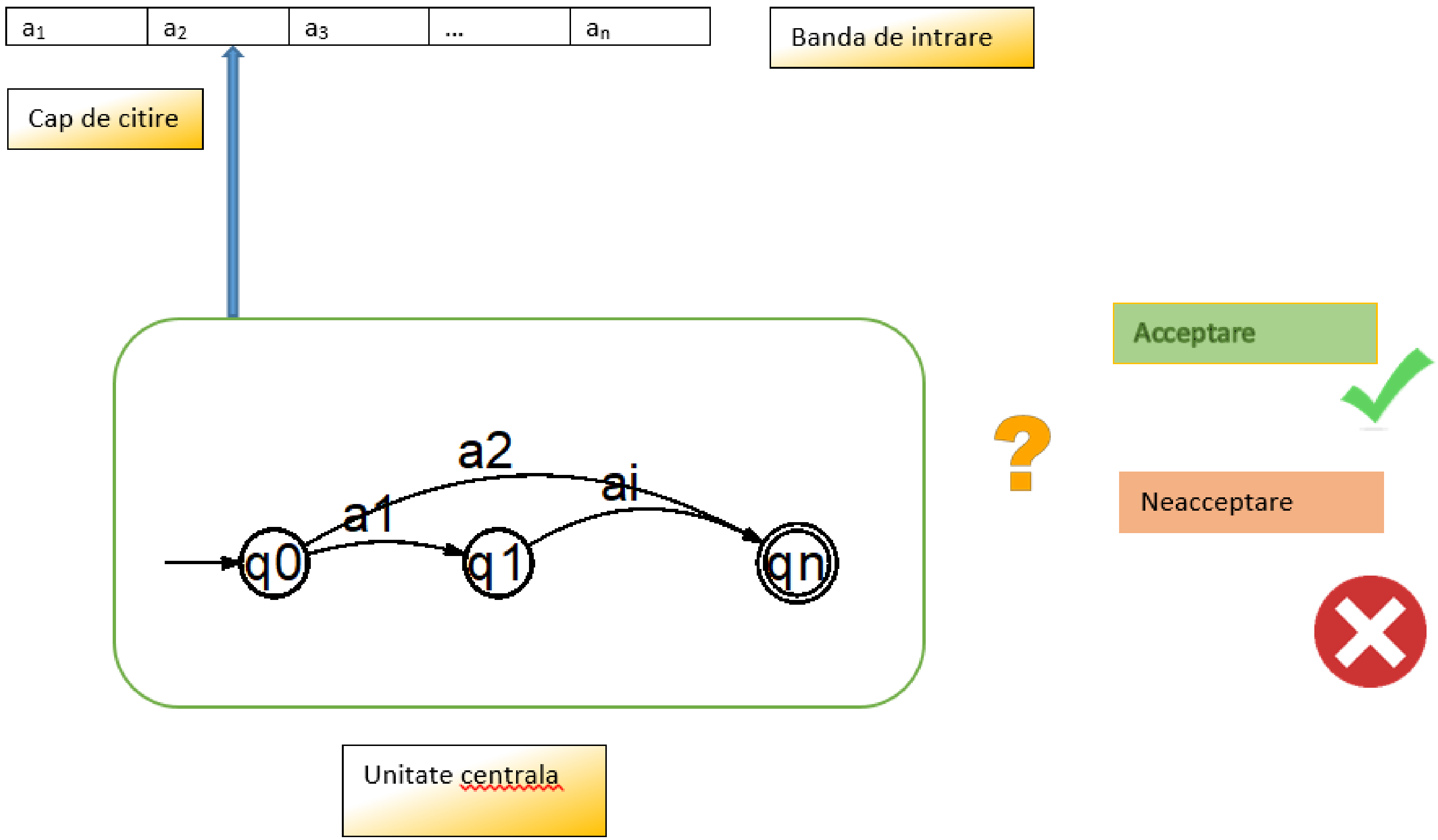
$$M = (Q, \Sigma, \delta, q_0, F)$$

- Un <u>automat finit</u> este un ansamblu

$M = (Q, \Sigma, \delta, q_0, F)$ in care:

  - Q- multime finita si nevida de elemente, numite ***stari***;
  - Σ – multime finita si nevida numita, ***alfabet de intrare***;
  - $\delta: Q \times \Sigma \to \mathcal{P}(Q)$ este o aplicatie, numita ***functie de tranzitie***;
  - $q_0 \in Q$ este ***stare initiala***;
  - $F \subseteq Q$ este o multime nevida numita, ***multimea starilor finale***.

Modelul fizic



Automat finit determinist (AFD). Automat finit nedeterminist (AFN)

Fie M un automat finit. Daca $\forall q \in Q, \forall a \in \Sigma$ $avem$ $|\delta(q, a)| \leq 1$ atunci automatul M se numeste ***automat finit determinist (AFD),*** iar in caz contrar se numeste ***automat finit nedeterminist (AFN).***

Fie M un automat finit. Daca $\forall q \in Q, \forall a \in \Sigma$ $avem$ $|\delta(q, a)| = 1$ atunci automatul M se numeste ***automat finit determinist complet (total) definit.***
In acest caz, avem intotdeauna $\forall q \in Q, \forall a \in \Sigma, \delta(q, a) = \{p\}, p \in Q$

Configuratii si relatii de tranzitie

c = (q, x), q- stare, x –secventa necitita de pe banda de intrare, x $\in\Sigma^*$

Configuratia initiala

- (q_0, w), w secventa

Configuratia finala

- (q_f, ε), $q_f \in F$, ε-secventa vida **(acceptare)**

Relatii intre configuratii

Tranzitie simpla (move) $\vdash$

$$(\boldsymbol{q}, \boldsymbol{ax}) \vdash (\boldsymbol{p}, \boldsymbol{x}), \boldsymbol{p} \in \boldsymbol{\delta}(\boldsymbol{q}, \boldsymbol{a})$$

k-tranzitie $\overset{k}{\vdash}$

Succesiune de k tranzitii directe $c_0 \vdash c_1 \vdash \cdots \vdash c_k$

+-tranzitie $\overset{+}{\vdash}$

$$c \overset{+}{\vdash} c' \ daca\ \exists k > 0\ astfel\ incat\ c \overset{k}{\vdash} c'$$

*-tranzitie $\overset{*}{\vdash}$

$$c \overset{*}{\vdash} c' daca\ \exists k \geq 0\ astfel\ incat\ c \overset{k}{\vdash} c'$$

Limbajul acceptat de un automat

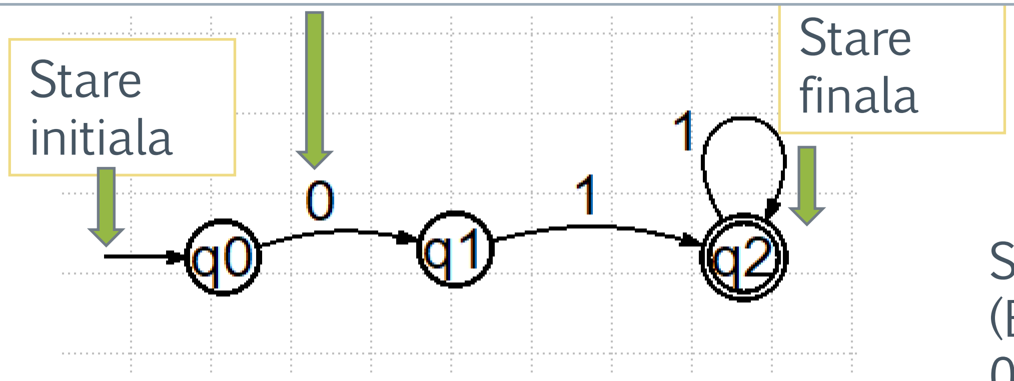
Limbajul acceptat de automatul $M = (Q, \Sigma, \delta, q_0, F)$ este
$T(M) = \{w \in \Sigma^* \mid (q_0, w) \vdash^* (q_f, \varepsilon), q_f \in F\}$.

- Doua automate M_1 si M_2 sunt echivalente daca si numai daca $T(M_1)=T(M_2)$
- **$\varepsilon \in L(M) \Leftrightarrow q_0 \in F$**

$$M = (Q, \Sigma, \delta, q_0, F)$$

Reprezentare

Daca $\delta(q_0,0) \ni q_1$ *atunci vom avea un arc etichetat cu* 0 *intre* q_0 *si* q_1



Graf orientat

Stare finala (Boolean 0/1)

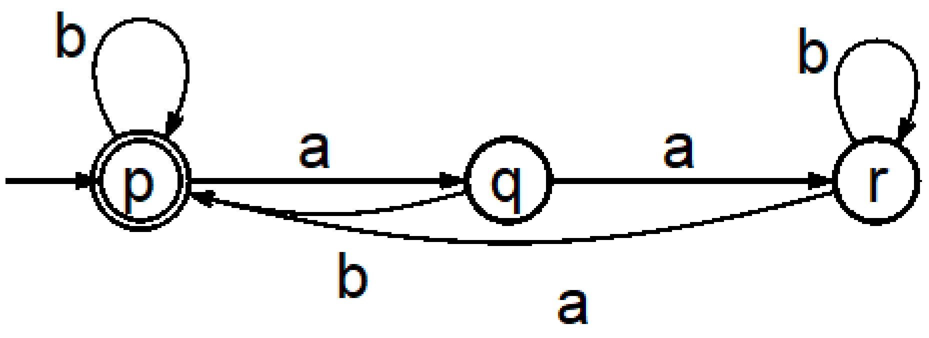
	a_1	…	a_j	…	a_m	
q_0						
…						
q_i	…	…	$\delta(q_i, a_j)$	…	….	z_i
…						
q_n						

Tabelar

Exercitii

Reprezentati ca graf automatul:

Q	Σ	a	b	
p		q	p	1
q		r	p	0
r		p	r	0



Exercitii

Construiti un AFD pentru identificatori.