

Exemple de limbaj care respectă L.P. \Rightarrow nu este regulat.

$$L_{\text{faxă}} = \{a^m f^m \mid m \geq 0\}.$$

$$L_1 = \underbrace{ac^+ac^+ \dots ac^+bc^+bc^+ \dots bc^+}_{nr. a = nr. b}$$

$len(L.P.) = 3 \Rightarrow$ orice subșir de 3 caract \rightarrow include c^+
 \Rightarrow de elimin c (cf. L.P.) \Rightarrow 2 șiruri a alăturate
 b alăturate $\nrightarrow \notin L_1$

$$L_{\text{faxă}} = \bigcup_{k=0}^{\infty} (a)^k (f)^k$$

$$L_1 = \bigcup_{k=0}^{\infty} (ac^+)^k (bc^+)^k$$

? $L_2 \Rightarrow$ regulat $L_1 \cup L_2$ s. satisf. L.P. dar să nu fie regulat

\downarrow

$$0^n 1^n \cup \{0, 1\}^* \rightarrow \{0, 1\}^*$$

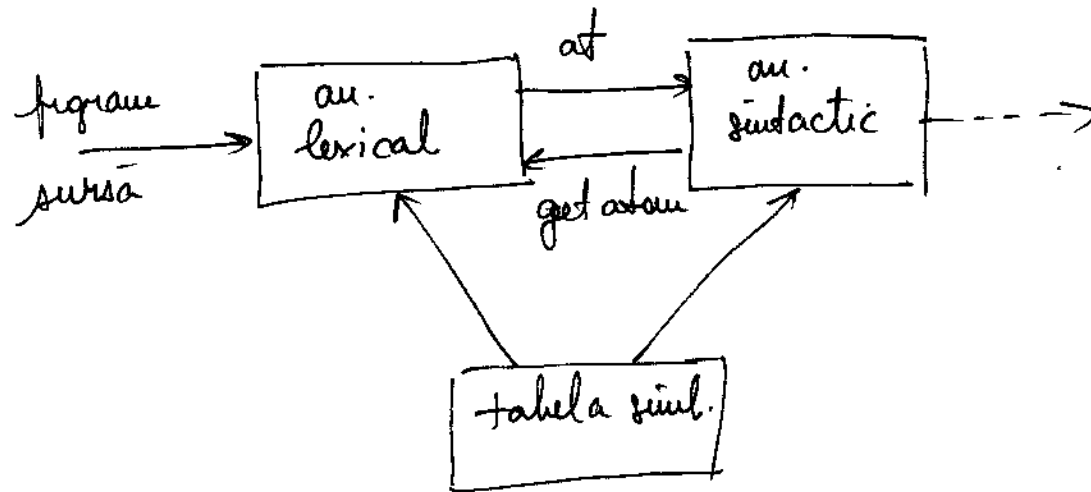
$L_2 \rightarrow$ disjoint L_1

$$L_2 = (a \neq b c)^* \underbrace{(a \neq b)^2}_{\text{}} (a \neq b c)^* \rightarrow \text{regular}$$

$$\text{pt } \bar{c} \in L_1 \supset L_2 \rightarrow \text{disjuncte} \Rightarrow L_1 = \underbrace{L_1 \cup L_2}_{\text{prop. inclusion}} - L_2$$

Analiza lexicală

Am. lexicală → prima etapă a proc. de compilare.



At. lexicali → an. cheie
operatori
identificatori
constante
siruri de literali
simboluri de punctuație

$v = m * 2;$

$\langle id, \uparrow T.S. \text{ pt } v \rangle$

$\langle op - atr \rangle$

$\langle id, \uparrow T.S. \text{ pt } m \rangle$

$\langle op - mult \rangle$

$\langle num, val. 2 \rangle$

$\langle ; \rangle$

Definiții regulate

$\Sigma \rightarrow$ alfabet simboluri de taxă

$d_1 \rightarrow r_1$

$d_2 \rightarrow r_2$

$d_m \rightarrow r_m$

$d_i \rightarrow$ nume distinct

$r_i \rightarrow$ expresie regulată peste $\Sigma \setminus \{d_1, \dots, d_{i-1}\}$.

ex:

identificatori în Pascal \rightarrow sir de litere și cifre care începe cu o literă

litera $\rightarrow A|B|C \dots |z|a|t| \dots |z$

cifra $\rightarrow 0|1| \dots |9$

id $\rightarrow \text{litera} (\text{litera} | \text{cifra})^*$

ex:

const. numerică fără semn în Pascal: 5350, 39.34, 6.33E4

cifra $\rightarrow 0|1| \dots |9$

cifre $\rightarrow \text{cifra} \text{ cifra}^*$

p-fracl $\rightarrow \cdot \text{cifre} |e$

exponent $\rightarrow (E(+|-|e) \text{ cifre}) |e$

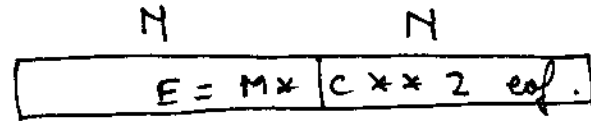
numar $\rightarrow \text{cifre} \text{ p-fracl} \text{ exponent}$

Convenții de notatie

1° etor postfix + \rightarrow

2° $n? \equiv n|e$

Input buffer



$N = 1024 \text{ sau } 4096.$

initial \rightarrow ci 2 ptr indică începutul lexemei

Deplasare de fwd-ptr \Rightarrow potrivire \Rightarrow pos. fwd-ptr pe 1 noul caracter
din dr. lexemei vte $\hat{=}$ aduc leg-ptr în pos. fwd-ptr.

\Rightarrow buff circular

Recunoașterea atomilor lexicali

instr \rightarrow if expr then instr

! if expr then instr else instr

! e

expr \rightarrow termen oprel termen

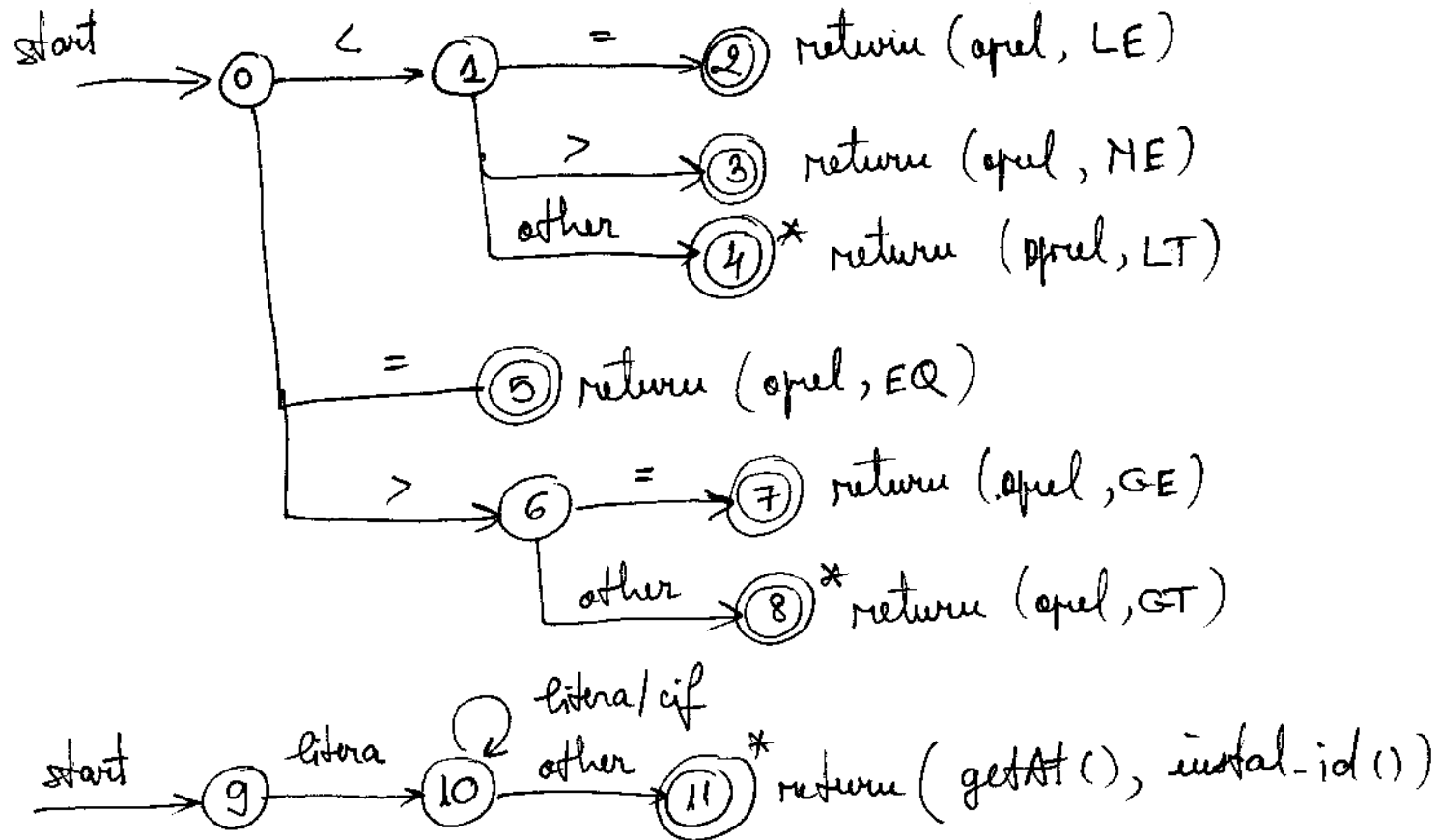
! termen

termen \rightarrow id | num

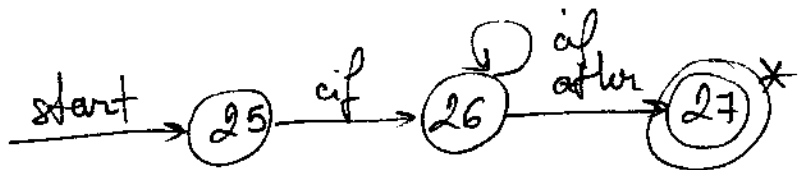
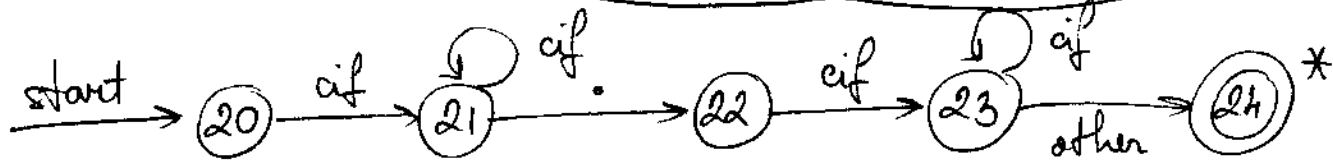
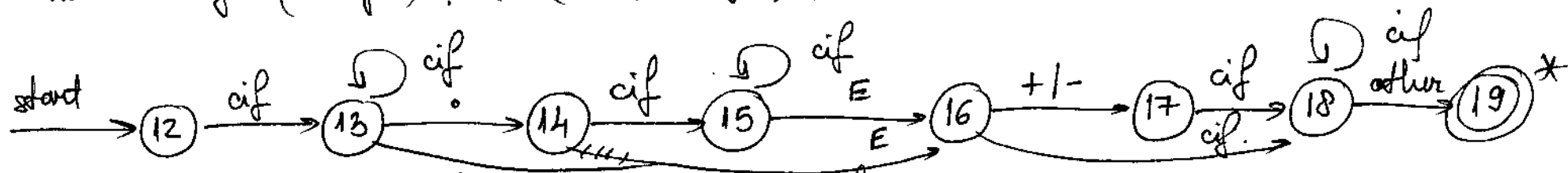
if \rightarrow if
 then \rightarrow then
 else \rightarrow else
 oprel \rightarrow < | <= | = | <> | > | >=
 id \rightarrow litera (litera | cifra)*
 num \rightarrow cifra⁺ (. cifra⁺)? (E (+|-|)? cifra⁺)?
 delim \rightarrow flanks | tab | newline
 ws \rightarrow delim⁺

Expresi regulate	At lexical	Atribut
ws	—	—
if	if	—
then	then	—
else	else	—
id	id	↑ T.S.
num	NUM	↑ T.S.
<	oprel	LT
<=	opul	LE

Diagramme de transition



num \rightarrow cif+ (. cif+)? (E(+|-)? cif+)?



san

