

# **PROGRAMSI PREVODIOCI**

## **- LL(1) gramatike -**

# LL( $k$ ) gramatike

- Omogućavaju top-down sintaksnu analizu bez vraćanja.
- Ograničenje koja treba da zadovolje smene gramatike:
  - Ukoliko za preslikavanje jednog neterminalnog simbola postoji veći broj smena, na osnovu sledećih  $k$  ulaznih simbola koje treba dobiti jednoznačno je određeno koja će se smena primeniti.

# Prosta LL (1) gramatika

- Prosta **LL(1) gramatika** – Beskonteksna gramatika u kojoj sve smene za isti neterminalni simbol počinju različitim terminalnim simbolima:

$$A \rightarrow a_1\alpha_1 \mid a_2\alpha_2 \mid \dots \mid a_n\alpha_n$$

$$a_i \in V_t, \quad \alpha_i \in V^*, \quad a_i \neq a_j \text{ za } i \neq j$$

# Prosta LL (1) gramatike - Primer

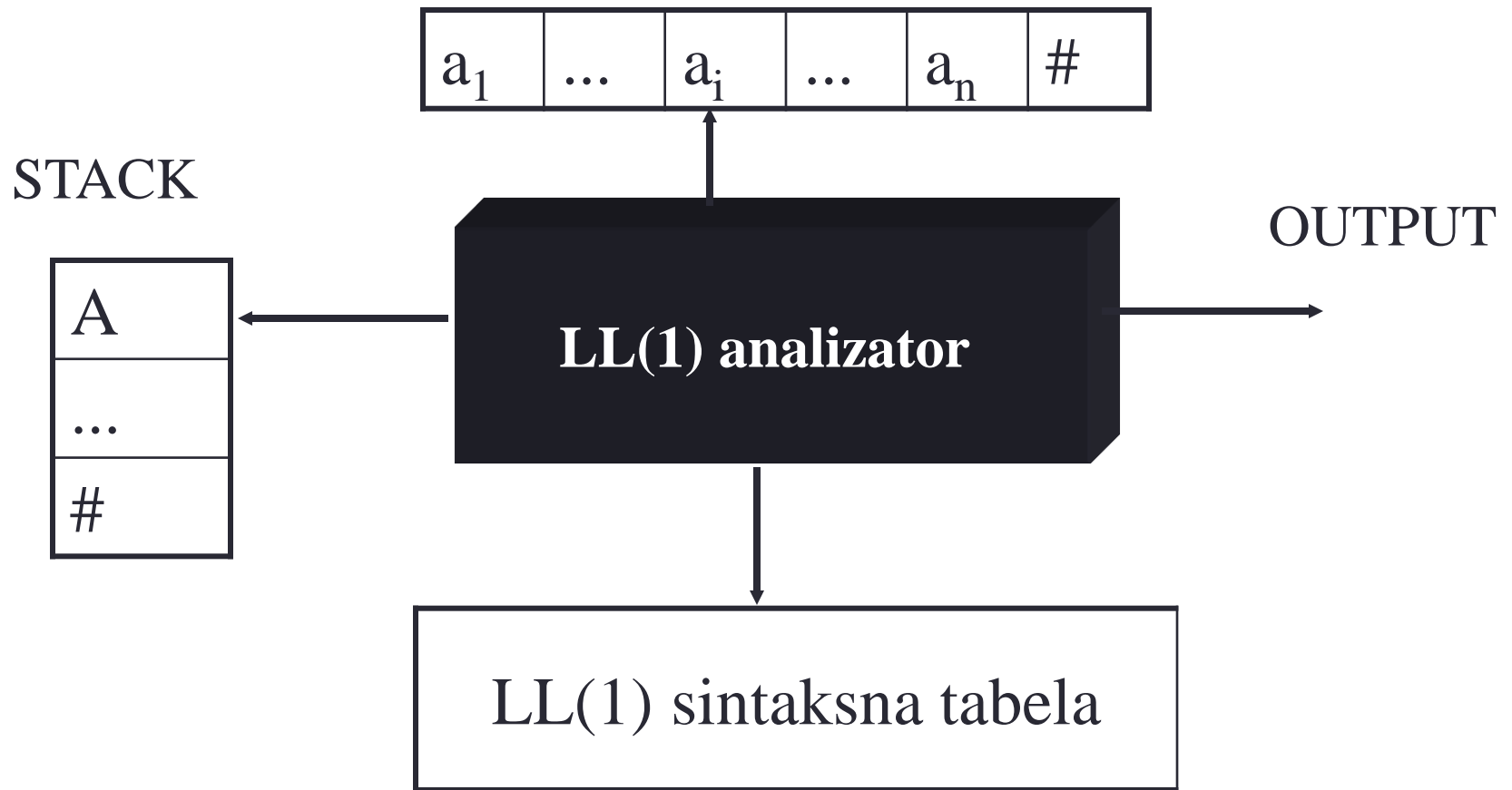
1.  $S \rightarrow aS$
2.  $S \rightarrow bA$
3.  $A \rightarrow d$
4.  $A \rightarrow ccA$

Prepoznavemo niz: *a a b c c d #*

Dodaje se smena :  $S' \rightarrow S\#$

$S' \rightarrow S\# \rightarrow aS\# \rightarrow aaS\# \rightarrow aabA\# \rightarrow$   
 $aabccA\# \rightarrow aabccd\#$

# LL(1) sintaksni analizator



# LL(1) sintaksna tabela

Definiše akciju koja će se izvršiti u postupku analize zavisno od simbola sa vrha radnog magacina i tekućeg ulaznog simbola.

$$M : \{V \cup \{\#\}\} \times \{V_t \cup \{\#\}\}$$

$$M(A, a) = \begin{cases} pop & \text{ako je } A = a, \quad a \in V_t \\ acc & \text{ako je } A = \# \wedge a = \# \\ (a\alpha, i) & \text{ako je } i - \text{to pravilo: } A \rightarrow a\alpha \\ err & \text{u svim ostalim slucajevima} \end{cases}$$

# LL (1) sintaksna tabela - Primer

1.  $S \rightarrow aS$
2.  $S \rightarrow bA$
3.  $A \rightarrow d$
4.  $A \rightarrow ccA$

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>#</i>
<i>S</i>	( <i>aS</i> ,1)	( <i>bA</i> ,2)			
<i>A</i>			( <i>ccA</i> ,4)	( <i>d</i> ,3)	
<i>a</i>	<i>pop</i>				
<i>b</i>		<i>pop</i>			
<i>c</i>			<i>pop</i>		
<i>d</i>				<i>pop</i>	
<i>#</i>					<i>acc</i>

# Postupak sintaksne analize

U postupku analize pratile trojka  $(z, \alpha, p)$  gde je

$z$  - deo ulaznog niza koji još nije prepoznat

$\alpha$  - sadržaj radnog magacina

$p$  - niz primenjenih pravila

Na početku analize se u radni magacin upisuje granični simbol ( $\#$ ) i startni simbol gramatike, zatim se na osnovu simbola sa vrha radnog magacina i tekućeg ulaznog simbola iz LL(1) sintaksne tabele čita akcija koja će biti izvršena.



# Postupak sintaksne analize

$$(az, A\alpha, p) = \begin{cases} (z, \alpha, p) & \text{ako je } M(A, a) = pop \\ kraj & \text{ako je } M(A, a) = acc \\ (az, \beta\alpha, pi) & \text{ako je } M(A, a) = (\beta, i) \\ greška & \text{ako je } M(A, a) = err \end{cases}$$

.

# Postupak analize - Primer

	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>#</i>
<i>S</i>	$(aS, 1)$	$(bA, 2)$			
<i>A</i>			$(ccA, 4)$	$(d, 3)$	
<i>a</i>	<i>pop</i>				
<i>b</i>		<i>pop</i>			
<i>c</i>			<i>pop</i>		
<i>d</i>				<i>pop</i>	
<i>#</i>					<i>acc</i>

$(aabccd\#, S\#, \varepsilon) \rightarrow$   
 $(aabccd\#, aS\#, 1) \rightarrow$   
 $(abccd\#, S\#, 1) \rightarrow$   
 $(abccd\#, aS\#, 11) \rightarrow$   
 $(bccd\#, S\#, 11) \rightarrow$   
 $(bccd\#, bA\#, 112) \rightarrow$   
 $(ccd\#, A\#, 112) \rightarrow$   
 $(ccd\#, ccA\#, 1124) \rightarrow$   
 $(cd\#, cA\#, 1124) \rightarrow$   
 $(d\#, A\#, 1124) \rightarrow$   
 $(d\#, d\#, 11243) \rightarrow$   
 $(\#, \#, 11243)$

# Leva faktORIZACIJA

$$A \Rightarrow \alpha\beta_1 \mid \alpha\beta_2 \mid \dots \mid \alpha\beta_n$$

$$A \Rightarrow \alpha A'$$

$$A' \Rightarrow \beta_1 \mid \beta_2 \mid \dots \mid \beta_n$$

# Leva faktORIZACIJA - Primer

$$A \Rightarrow \alpha\beta_1 \mid \alpha\beta_2$$

$IfStmt \rightarrow \mathbf{if} \ Expr \ \mathbf{then} \ Stmt \ \mathbf{else} \ Stmt \mid$   
 $\mathbf{if} \ Expr \ \mathbf{then} \ Stmt$

$$A \Rightarrow \alpha A'$$

$$A' \Rightarrow \beta_1 \mid \beta_2$$

$IfStmt \rightarrow \mathbf{if} \ Expr \ \mathbf{then} \ Stmt \ IfStmt'$

$IfStmt' \rightarrow \mathbf{else} \ Stmt \mid \varepsilon$

# Proširenja definicije LL(1) gramatike

1. Šta kada smene ne počinju terminalnim simbolom?
2. Šta kada postoje  $\varepsilon$  smene?

# Funkcija FIRST

$FIRST(\alpha)$  - Skup terminalnih simbola sa kojima počinju reči izvedene iz reči  $\alpha$ .

Ako je  $\alpha \in V^+$

$$FIRST(\alpha) = \{w \mid \alpha \xrightarrow{*} w \dots, w \in V_t\}$$

# Funkcija FIRST - primer

$$S \rightarrow ABc$$

$$A \rightarrow dB$$

$$A \rightarrow aS$$

$$A \rightarrow c$$

$$B \rightarrow AS$$

$$B \rightarrow b$$

$$FIRST(ABc) = \{w \mid A \xrightarrow{*} w\dots, w \in V_t\} = \{a, c, d\}$$

$$FIRST(AS) = \{w \mid ABe \xrightarrow{*} w\dots, w \in V_t\} = \{a, c, d\}$$

# LL(1) gramatike bez $\varepsilon$ pravila

Gramatika bez  $\varepsilon$  pravila je LL(1) gramatika ako su za sva pravila oblika :

$$A \rightarrow \alpha_1 \mid \alpha_2 \mid \dots \mid \alpha_n$$

skupovi

$$FIRST(\alpha_1), FIRST(\alpha_2), \dots, FIRST(\alpha_n)$$

disjunktni po parovima. Odnosno:

$$FIRST(\alpha_i) \cap FIRST(\alpha_j) = \phi, \text{ za } i \neq j.$$



# LL(1) gramatika bez $\varepsilon$ pravila - PRIMER

$$S \rightarrow ABc$$

$$A \rightarrow dB$$

$$A \rightarrow aS$$

$$A \rightarrow c$$

$$B \rightarrow AS$$

$$B \rightarrow b$$

$$FIRST(dB) \cap FIRST(aS) = \{d\} \cap \{a\} = \phi$$

$$FIRST(dB) \cap FIRST(c) = \{d\} \cap \{c\} = \phi$$

$$FIRST(aS) \cap FIRST(c) = \{a\} \cap \{c\} = \phi$$

$$FIRST(AS) \cap FIRST(b) = \{a, c, d\} \cap \{b\} = \phi$$

# LL(1) sintaksna tabela gramatike bez $\varepsilon$ pravila

$$M(A, a) = \begin{cases} pop & \text{ako je } A = a, \quad a \in V_t \\ acc & \text{ako je } A = \# \wedge a \in \#_t \\ (\beta, i) & \text{ako je } i - \text{to pravilo: } A \rightarrow \beta \text{ i } a \in FIRST(\beta) \\ err & \text{u svim ostalim slucajevima} \end{cases}$$

# LL(1) Sintaksna tabela gramatike bez $\varepsilon$ pravila – Primer

1.  $S \rightarrow ABc$

2.  $A \rightarrow dB$

3.  $A \rightarrow aS$

4.  $A \rightarrow c$

5.  $B \rightarrow AS$

6.  $B \rightarrow b$

	a	b	c	d	#
S	(ABc,1)		(ABc,1)	(ABc,1)	
A	(aS,3)		(c,4)	(dB,2)	
B	(AS,5)	(b,6)	(AS,5)	(AS,5)	
a	<i>pop</i>				
b		<i>pop</i>			
c			<i>pop</i>		
d				<i>pop</i>	
#					<i>acc</i>

# LL(1) gramatike sa $\varepsilon$ pravilima

$\varepsilon$  pravila su pravila oblika:  $A \rightarrow \varepsilon$

Menja se definicija funkcije FIRST:

$$FIRST(\alpha) = \{w \mid \alpha \xrightarrow{*} w\dots, \quad |w| \leq 1, w \in V_t^*\}$$

*FIRST* može da sadži prazan niz.

# Gramatika sa $\varepsilon$ pravilima (Funkcija FIRST) - Primer

$$E \rightarrow TE' \qquad FIRST(TE') = \{ (, a \}$$

$$E' \rightarrow +TE' \qquad FIRST(+TE') = \{ + \}$$

$$E' \rightarrow \varepsilon \qquad FIRST(\varepsilon) = \{ \varepsilon \}$$

$$T \rightarrow FT' \qquad FIRST(FT') = \{ (, a \}$$

$$T' \rightarrow *FT' \qquad FIRST(*FT') = \{ * \}$$

$$T' \rightarrow \varepsilon \qquad FIRST(\varepsilon) = \{ \varepsilon \}$$

$$F \rightarrow (E) \qquad FIRST((E)) = \{ ( \}$$

$$F \rightarrow a \qquad FIRST(a) = \{ a \}$$

# Funkcija FOLLOW

$$FOLLOW(A) = \{w \mid S' \xrightarrow{*} \alpha A \gamma \wedge w \in FIRST(\gamma)\}$$

$S'$  - pomocni startni simbol gramatike ( $S' \rightarrow S\#$ ).

$$FOLLOW(A) = \{w \mid S' \xrightarrow{*} \alpha A w \beta \wedge w \in V_t \cup \{\#\}\}$$

# Funkcija FOLLOW - Primer

$$E \rightarrow TE'$$

$$\text{FOLLOW}(E') = ?$$

$$E' \rightarrow +TE'$$

$$E_1 \rightarrow E\#$$

$$E' \rightarrow \varepsilon$$

$$\rightarrow TE'\#$$

$$\# \in \text{FOLLOW}(E')$$

$$T \rightarrow FT'$$

$$\rightarrow FT'E'\#$$

$$T' \rightarrow *FT'$$

$$\rightarrow (E)T'E'\#$$

$$\rightarrow (TE')T'E'\# \quad ) \in \text{FOLLOW}(E')$$

$$T' \rightarrow \varepsilon$$

$$\text{FOLLOW}(E') = \{\#, , )\}$$

$$F \rightarrow (E)$$

$$F \rightarrow a$$

# Funkcija FOLLOW - Primer

$$E \rightarrow TE'$$

$$\text{FOLOW}(T') = ?$$

$$E' \rightarrow +TE'$$

$$E_1 \rightarrow E\#$$

$$\rightarrow TE'\#$$

$$E' \rightarrow \varepsilon$$

$$\rightarrow FT'E'\#$$

$$T \rightarrow FT'$$

$$(1) \rightarrow FT'+TE'\# \quad + \in \text{FOLOW}(T')$$

$$T' \rightarrow *FT'$$

$$(2) \rightarrow FT'\# \quad \# \in \text{FOLOW}(T')$$

$$T' \rightarrow \varepsilon$$

$$\rightarrow (E)T'\#$$

$$F \rightarrow (E)$$

$$\rightarrow (TE')T'\#$$

$$F \rightarrow a$$

$$\rightarrow (FT'E')T'\#$$

$$\rightarrow (FT')T'\# \quad ) \in \text{FOLOW}(T')$$

$$\text{FOLOW}(T') = \{+, \#, )\}$$



# Odredjivanje FOLLOW funkcije

- FOLLOW funkcija startnog simbola gramatike sadrži granični simbol #.
- Za odredjivanje FOLLOW( $X$ ) posmatraju se desne strane smena u kojima se  $X$  pojavljuje.
  - $Z \rightarrow \alpha X x\beta \wedge x \in V_T \Rightarrow x \in \text{FOLLOW}(X)$
  - $Z \rightarrow \alpha XY\beta \wedge Y \in V_N \Rightarrow \text{FIRST}(Y) \subset \text{FOLLOW}(X)$   
(Napomena: ukoliko postoji izvodjenje  $Y \Rightarrow \varepsilon$ , simbol  $Y$  se preskače i gleda se nastavak smene)
  - $Z \rightarrow \alpha X \Rightarrow \text{FOLLOW}(Z) \subset \text{FOLLOW}(X)$

# Funkcija FOLLOW - Primer

$$E \rightarrow TE'$$

$$E' \rightarrow +TE'$$

$$E' \rightarrow \varepsilon$$

$$T \rightarrow FT'$$

$$T' \rightarrow *FT'$$

$$T' \rightarrow \varepsilon$$

$$F \rightarrow (E)$$

$$F \rightarrow a$$

$$\text{FOLLOW}(E') = ?$$

$$E \rightarrow TE' \Rightarrow \text{FOLLOW}(E) \subset \text{FOLLOW}(E')$$

$$E \text{ je startni simbol} \Rightarrow \# \in \text{FOLLOW}(E')$$

$$F \rightarrow (E) \Rightarrow ) \in \text{FOLLOW}(E')$$

$$\text{FOLLOW}(E') = \text{FOLLOW}(E) = \{\#, )\}$$

# Funkcija FOLLOW - Primer

$$E \rightarrow TE'$$

$$E' \rightarrow +TE'$$

$$E' \rightarrow \varepsilon$$

$$T \rightarrow FT'$$

$$T' \rightarrow *FT'$$

$$T' \rightarrow \varepsilon$$

$$F \rightarrow (E)$$

$$F \rightarrow a$$

$$FOLLOW(T') = ?$$

$$T \rightarrow FT' \Rightarrow FOLLOW(T) \subset FOLLOW(T')$$

$$E' \rightarrow +TE' \Rightarrow FIRST(E') \subset FOLLOW(T)$$

$$+ \in FIRST(E') \Rightarrow + \in FOLLOW(T)$$

$$\varepsilon \in FIRST(E') \Rightarrow FOLLOW(E') \subset FOLLOW(T)$$

$$FOLLOW(T') = FOLLOW(T') = \{+, \#, )\}$$

# LL(1) gramatike sa $\varepsilon$ pravilima

Gramatika  $G$  je LL(1) gramatika ako i samo ako za svaki par pravila za isti neterminal simbol  $A$ ,

$$A \rightarrow \alpha \quad i \quad A \rightarrow \beta$$

$$FIRST(\alpha \circ FOLLOW(A)) \cap FIRST(\beta \circ FOLLOW(A)) = \phi$$

◦ znak za konkatenciju

Za sva pravila  $A \rightarrow \alpha_1 \mid \alpha_2 \mid \dots \mid \alpha_n$

1.  $FIRST(\alpha_i) \cap FIRST(\alpha_j) = \phi$ , za  $i \neq j$

2. Ako  $\alpha_i \xrightarrow{*} \varepsilon$  tada mora da važi:

$$FIRST(\alpha_j) \cap FOLLOW(A) = \phi \quad \text{za svako } j \neq i.$$

# LL(1) gramatika sa $\varepsilon$ pravilima - Primer

1.  $E \rightarrow TE'$

2.  $E' \rightarrow +TE'$

$$\begin{aligned} & FIRST(+TE' \circ FOLLOW(E')) \cap \\ & FIRST(\varepsilon \circ FOLLOW(E')) = \{+\} \cap \{), \#\} = \phi \end{aligned}$$

3.  $E' \rightarrow \varepsilon$

4.  $T \rightarrow FT'$

5.  $T' \rightarrow *FT'$

$$\begin{aligned} & FIRST(*FT' \circ FOLLOW(T')) \cap \\ & FIRST(\varepsilon \circ FOLLOW(T')) = \{*\} \cap \{+, ), \#\} = \phi \end{aligned}$$

6.  $T' \rightarrow \varepsilon$

7.  $F \rightarrow (E)$

$$FIRST((E) \circ FOLLOW(F)) \cap$$

8.  $F \rightarrow a$

$$FIRST(a \circ FOLLOW(F)) = \{( \} \cap \{a\} = \phi$$

# LL(1) sintaksna tabela gramatike sa $\varepsilon$ pravilima

$$M(A, a) = \begin{cases} pop & \text{ako je } A = a, \quad a \in V_t \\ acc & \text{ako je } A = \# \wedge a \in \#_t \\ (\alpha, i) & \text{ako je } i - \text{to pravilo } A \rightarrow \alpha \text{ i } a \in FIRST(\alpha) \\ & \text{ili je } i - \text{to pravilo } A \rightarrow \alpha \quad \varepsilon \in FIRST(\alpha) \text{ i} \\ & a \in FOLLOW(A), \text{ za } A \in V_n \\ err & \text{u svim ostalim slucajevima} \end{cases}$$

# LL(1) sintaksna tabela gramatike sa $\varepsilon$ pravilima - Primer

1.  $E \rightarrow TE'$
2.  $E' \rightarrow +TE'$
3.  $E' \rightarrow \varepsilon$
4.  $T \rightarrow FT'$
5.  $T' \rightarrow *FT'$
6.  $T' \rightarrow \varepsilon$
7.  $F \rightarrow (E)$
8.  $F \rightarrow a$

	+	*	(	)	a	#
E			(TE',1)		(TE',1)	
E'	(+TE',2)			( $\varepsilon$ ,3)		( $\varepsilon$ ,3)
T			(FT',4)		(FT',4)	
T'	( $\varepsilon$ ,5)	(*FT',5)	( $\varepsilon$ ,6)			( $\varepsilon$ ,6)
F			((E),7)		(a,8)	
+	<i>pop</i>					
*		<i>pop</i>				
(			<i>pop</i>			
)				<i>pop</i>		
a					<i>pop</i>	
#						<i>acc</i>

# Primer sintaksne analize

	+	*	(	)	a	#
E			(TE',1)		(TE',1)	
E'	(+TE',2)			(ε,3)		(ε,3)
T			(FT',4)		(FT',4)	
T'	(ε,5)	(*FT',5)	(ε,6)			(ε,6)
F			((E),7)		(a,8)	
+	<i>pop</i>					
*		<i>pop</i>				
(			<i>pop</i>			
)				<i>pop</i>		
a					<i>pop</i>	
#						<i>acc</i>

$(a * a\#, E\#, \varepsilon) \rightarrow$

$(a * a\#, TE'\#, 1) \rightarrow$

$(a * a\#, FT' E'\#, 14) \rightarrow$

$(a * a\#, aT' E'\#, 148) \rightarrow$

$(*a\#, T' E'\#, 148) \rightarrow$

$(*a\#, *FT' E'\#, 1485) \rightarrow$

$(a\#, FT' E'\#, 1485) \rightarrow$

$(a\#, aT' E'\#, 14858) \rightarrow$

$(\#, T' E'\#, 14857) \rightarrow$

$(\#, E'\#, 148576) \rightarrow$

$(\#, \#, 1485763)$

Reč koja se prepoznaje:  $a * (a + a)\#$