

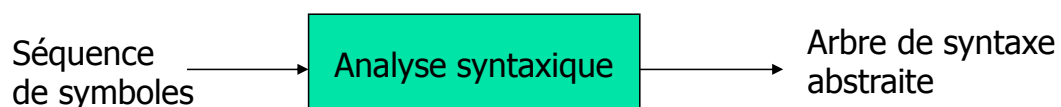


3. Analyse syntaxique

54



3. Analyse syntaxique



- Etant donnés une grammaire et un mot, ce mot appartient-il au langage généré par la grammaire ?
- Représentation sous forme d'arbre syntaxique
- Détection et traitement des erreurs syntaxiques

55

Exemple

- Mot :

var id vir id dpt int ptvir id aff entier ptvir id
aff id mult id add entier ptvir

- Grammaire :

Prog -> Decliste Instliste

Decliste -> Dec | Decliste Dec

Dec -> var Idliste dpt Type ptvir

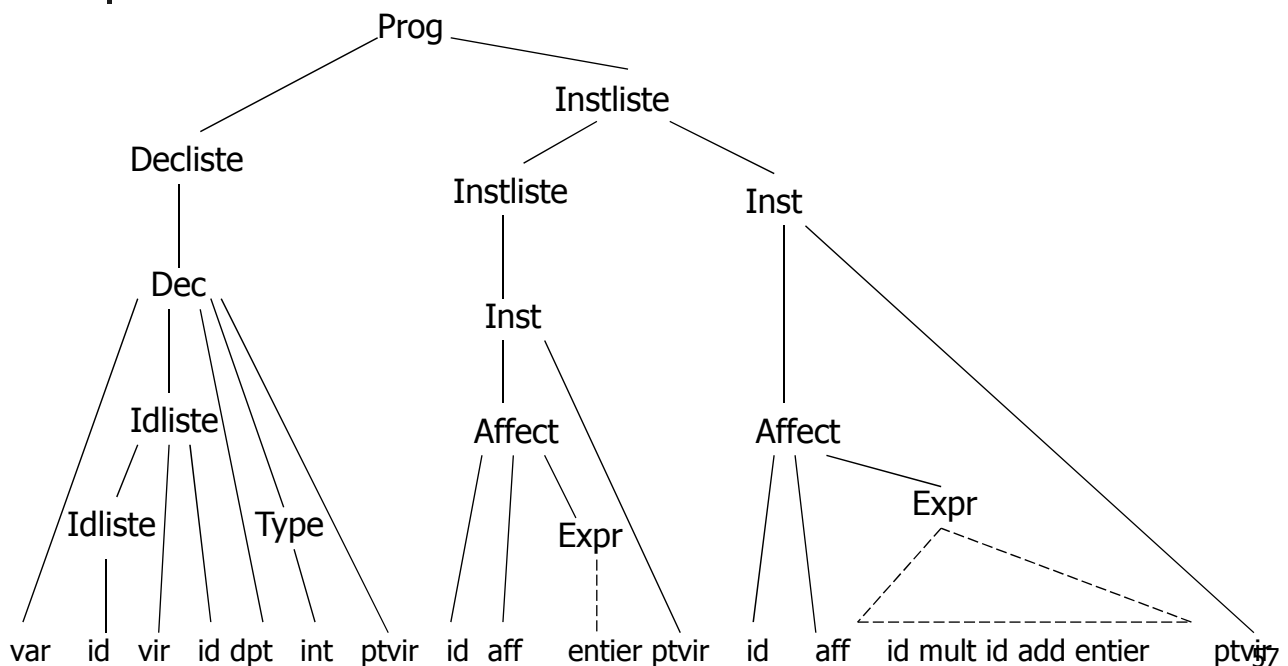
Idliste -> Idliste vir id | id

Instliste -> Instliste Inst | Inst

Inst -> Affect ptvir | Sisinon ptvir | Etc.

56

Exemple suite





Analyse syntaxique : plan

- 3.1. Analyse descendante et analyse ascendante
- 3.2. Analyse syntaxique descendante
 - Ensembles *Premiers* et *Suivants*
 - Table d'analyse
 - Grammaire LL(1)
 - Elimination de la récursivité à gauche et factorisation
- 3.3. Analyse syntaxique ascendante
 - Collection des Items
 - Table d'analyse
 - Grammaire SLR(1)

58



3.1. Analyse descendante et analyse ascendante

- Grammaire G :
 - $S \rightarrow aA \mid bBA$
 - $A \rightarrow aS \mid d$
 - $B \rightarrow c \mid cc$
- Le mot aabcd appartient-il à $L(G)$?

59

Analyse syntaxique descendante

- Grammaire G :
 $S \rightarrow aA \mid bBA$
 $A \rightarrow aS \mid d$
 $B \rightarrow c \mid cc$
- Le mot aabcd appartient-il à $L(G)$?

$S \rightarrow aA \rightarrow aaS \rightarrow aabBA \rightarrow aabcA \rightarrow aabcd$

Ou bien encore :

$S \rightarrow aA \rightarrow aaS \rightarrow aabBA \rightarrow aabBd \rightarrow aabcd$

60

Analyse syntaxique descendante (top-down)

- À partir de l'axiome S , par dérivation, vers le mot en entier
- Choix de la production à utiliser

 ➔ nécessité d'une *table d'analyse* qui indique, étant donné le symbole terminal en entrée, quelle production il faut utiliser
- Construction de l'arbre syntaxique **de la racine vers les feuilles**

61

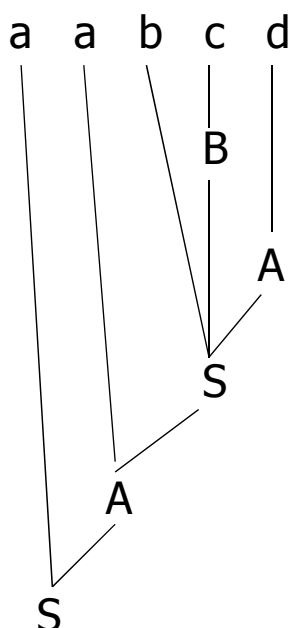
Analyse syntaxique ascendante (bottom-up)

- Grammaire G :
 $S \rightarrow aA \mid bBA$
 $A \rightarrow aS \mid d$
 $B \rightarrow c \mid cc$
- Le mot aabcd appartient-il à $L(G)$?

Construction de l'arbre syntaxique **des feuilles
vers la racine**

62

Analyse syntaxique ascendante



aabcd \leftarrow aabBd

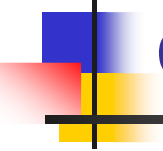
\leftarrow aabBA

\leftarrow aaS

\leftarrow aA

\leftarrow S

63



3.2. Analyse syntaxique descendante (top-down)

64



Parser en analyse descendante

- Pour chaque non terminal X on calcule les ensembles
 - *Premiers*(X) : contient les terminaux qui peuvent **débuter** un mot obtenu par une dérivation à partir de X
 - *Suivants*(X) : contient les terminaux qui peuvent **suivre** X dans une dérivation à partir de l'axiome
- Construction de la table d'analyse (automate à pile)

65



Exemple de Premiers et Suivants

- $S \rightarrow xS \mid AS \mid B$
 $A \rightarrow yA \mid Bx$
 $B \rightarrow zB \mid \varepsilon$
- $S \rightarrow AS \rightarrow yAS$ donc $y \in \text{Premiers}(S)$
- $S \rightarrow AS \rightarrow yAS \rightarrow yAxS$ donc $x \in \text{Suivants}(A)$

66



Calcul des Premiers(X)

Algorithme de construction des ensembles Premiers(X) (pour $X \in T \cup N$)

1. Si X est un non-terminal et $X \rightarrow Y_1Y_2...Y_n$ est une production de la grammaire (avec Y_i symbole terminal ou non terminal) alors
 - ajouter les éléments de $\text{PREMIERS}(Y_1)$ **sauf** ε dans $\text{PREMIERS}(X)$
 - s'il existe j ($j \in \{2, \dots, n\}$) tel que pour tout $i=1, \dots, j-1$ on a $\varepsilon \in \text{PREMIERS}(Y_i)$, alors ajouter les éléments de $\text{PREMIERS}(Y_j)$ **sauf** ε dans $\text{PREMIERS}(X)$
 - si pour tout $i=1, \dots, n$ $\varepsilon \in \text{PREMIERS}(Y_i)$ alors ajouter ε dans $\text{PREMIERS}(X)$
2. Si X est un non-terminal et $X \rightarrow \varepsilon$ est une production, ajouter ε dans $\text{PREMIERS}(X)$
3. Si X est un terminal, $\text{PREMIERS}(X) = \{X\}$.

Recommencer jusqu'à ce qu'on n'ajoute rien de nouveau dans les ensembles PREMIERS .

67

Exemple de calcul des ensembles Premiers

■ $S \rightarrow xS \mid AS \mid B$

$A \rightarrow yA \mid Bx$

$B \rightarrow zB \mid \varepsilon$

■

Premiers	
S	x, y, z, ε
A	y, z, x
B	z, ε

68

Calcul des Premiers(α)

Calcul des PREMIERS(α) pour $\alpha \in (T \cup N)^*$

On a $\alpha = Y_1 Y_2 \dots Y_n$ avec $Y_i \in T$ ou $Y_i \in N$

Si Y_1 est un symbole terminal alors ajouter Y_1 aux PREMIERS(α)

sinon

// Y_1 est un symbole non terminal

ajouter les éléments de PREMIERS(Y_1) **sauf** ε dans PREMIERS(α)

si $\varepsilon \in \text{PREMIERS}(Y_1)$ alors

// faire la même chose avec Y_2

si Y_2 est un symbole terminal alors ajouter Y_2 aux PREMIERS(α)

sinon

// Y_2 est un symbole non terminal

ajouter les éléments de PREMIERS(Y_2) **sauf** ε dans PREMIERS(α)

si $\varepsilon \in \text{PREMIERS}(Y_2)$ alors

-/->

69



Calcul des Premiers(α) - suite

```
// faire la même chose avec  $Y_3$ 
...
// faire la même chose avec  $Y_n$ 
si  $Y_n \in V_T$  alors ajouter  $Y_n$  aux PREMIERS( $\alpha$ )
sinon
    //  $Y_n$  est un symbole non terminal
    ajouter les PREMIERS( $Y_n$ ) sauf  $\epsilon$  aux PREMIERS( $\alpha$ )
    si  $\epsilon \in \text{PREMIERS}(Y_n)$  alors
        ajouter  $\epsilon$  aux PREMIERS( $\alpha$ )
    fin si
fin si
...
fin si
fin si
fin si
fin si
```

70



Calcul des Suivants(X)

1. Ajouter un marqueur de fin de chaîne (symbole \$ par exemple) à SUIVANTS(S) (où S est l'axiome)
2. Pour chaque production $X \rightarrow \alpha Y \beta$ où Y est un non-terminal, alors ajouter le contenu de PREMIERS(β) à SUIVANTS(Y) **sauf** ϵ
3. Pour chaque production $X \rightarrow \alpha Y$, alors ajouter SUIVANTS(X) à SUIVANTS(Y)
4. Pour chaque production $X \rightarrow \alpha Y \beta$ avec $\epsilon \in \text{PREMIERS}(\beta)$, alors ajouter SUIVANTS(X) à SUIVANTS(Y)

Recommencer à partir de l'étape 2 jusqu'à ce qu'on n'ajoute rien de nouveau dans les ensembles SUIVANTS

71

Exemple de calcul des ensembles Suivants

■ $S \rightarrow xS \mid AS \mid B$

$A \rightarrow yA \mid Bx$

$B \rightarrow zB \mid \varepsilon$

■

	Premiers	Suivants
S	x, y, z, ε	\$
A	y, z, x	x, y, z, \$
B	z, ε	\$, x

72

Construction de la table d'analyse

■ tableau : lignes sur N

colonnes sur $T \cup \{\$ \}$

■ Pour chaque production $X \rightarrow \alpha$ faire

1. pour tout $a \in \text{PREMIERS}(\alpha)$ (**et** $a \neq \varepsilon$) rajouter la production $X \rightarrow \alpha$ dans la case $M[X, a]$

2. si $\varepsilon \in \text{PREMIERS}(\alpha)$ alors pour chaque $b \in \text{SUIVANTS}(X)$ ajouter $X \rightarrow \alpha$ dans $M[X, b]$

73

Exemple de construction de la table d'analyse

- $E \rightarrow TE'$
 $E' \rightarrow +TE' \mid -TE' \mid \epsilon$
 $T \rightarrow FT'$
 $T' \rightarrow *FT' \mid /FT' \mid \epsilon$
 $F \rightarrow (E) \mid nb$

	Premiers	Suivants
E	(nb	\$)
E'	+ - ϵ	\$)
T	(nb	+ - \$)
T'	* / ϵ	+ - \$)
F	(nb	* / + - \$)

- table d'analyse :

	nb	(+	-	*	/)	\$
E								
E'								
T								
T'								
F								

74

Exemple de construction de la table d'analyse

- $E \rightarrow TE'$
 $E' \rightarrow +TE' \mid -TE' \mid \epsilon$
 $T \rightarrow FT'$
 $T' \rightarrow *FT' \mid /FT' \mid \epsilon$
 $F \rightarrow (E) \mid nb$

	Premiers	Suivants
E	(nb	\$)
E'	+ - ϵ	\$)
T	(nb	+ - \$)
T'	* / ϵ	+ - \$)
F	(nb	* / + - \$)

- table d'analyse :

	nb	(+	-	*	/)	\$
E	$E \rightarrow TE'$	$E \rightarrow TE'$						
E'			$E' \rightarrow +TE'$	$E' \rightarrow -TE'$			$E' \rightarrow \epsilon$	$E' \rightarrow \epsilon$
T	$T \rightarrow FT'$	$T \rightarrow FT'$						
T'			$T' \rightarrow \epsilon$	$T' \rightarrow \epsilon$	$T' \rightarrow *FT'$	$T' \rightarrow /FT'$	$T' \rightarrow \epsilon$	$T' \rightarrow \epsilon$
F	$F \rightarrow nb$	$F \rightarrow (E)$						

75



Grammaire LL(1)

- **Définition** : si aucune case de la table d'analyse ne contient plus d'une production, alors la grammaire est dite *grammaire LL(1)*.

76



Analyseur syntaxique LL(1)

```
répéter
  Soit X le symbole en sommet de pile
  Soit a la lettre pointée par ps
  si X est un non-terminal alors
    si  $M[X,a]=X \rightarrow Y_1 \dots Y_n$  alors
      enlever X de la pile
      mettre  $Y_n$  puis  $Y_{n-1}$  puis...puis  $Y_1$  dans la pile
      émettre en sortie la production  $X \rightarrow Y_1 \dots Y_n$ 
    sinon (case vide dans la table)
      ERREUR
    fin si
  sinon
    si  $X = \$$  alors
      si  $a = \$$  alors ACCEPTER
      sinon ERREUR
      fin si
    sinon
      si  $X = a$  alors
        enlever X de la pile
        avancer ps
      sinon
        ERREUR
      fin si
    fin si
  fin si
jusqu'à ERREUR ou ACCEPTER
```

77

Exemple d'analyseur syntaxique LL(1)

- table d'analyse LL(1) :

	nb	(+	-	*	/)	\$
E	$E \rightarrow TE'$	$E \rightarrow TE'$						
E'			$E' \rightarrow +TE'$	$E' \rightarrow -TE'$			$E' \rightarrow \epsilon$	$E' \rightarrow \epsilon$
T	$T \rightarrow FT'$	$T \rightarrow FT'$						
T'			$T' \rightarrow \epsilon$	$T' \rightarrow \epsilon$	$T' \rightarrow *FT'$	$T' \rightarrow /FT'$	$T' \rightarrow \epsilon$	$T' \rightarrow \epsilon$
F	$F \rightarrow nb$	$F \rightarrow (E)$						

- Le mot $3+4*(9-1)$ est-il généré par la grammaire ?

78

Exemple d'analyseur syntaxique LL(1)

Pile	Entrée	Sortie
E\$	3+4*(9-1)\$	E -> TE'
TE'\$	3+4*(9-1)\$	T -> FT'
FT'E'\$	3+4*(9-1)\$	F -> nb
nbT'E'\$	3+4*(9-1)\$	
T'E'\$	+4*(9-1)\$	T' -> ϵ
E'\$	+4*(9-1)\$	E' -> +TE'
+TE'\$	+4*(9-1)\$	
TE'\$	4*(9-1)\$	T -> FT'
FT'E'\$	4*(9-1)\$	F -> nb
nbT'E'\$	4*(9-1)\$	
T'E'\$	*(9-1)\$	T' -> *FT'
*FT'E'\$	*(9-1)\$	
FT'E'\$	(9-1)\$	F -> (E)
(E)T'E'\$	(9-1)\$	

79

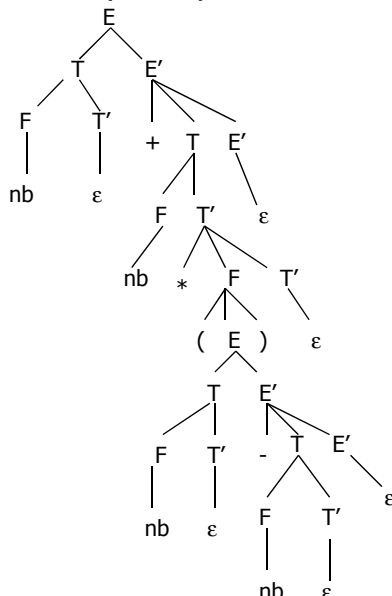
Exemple d'analyseur syntaxique LL(1) - suite

Pile	Entrée	Sortie
E)T'E'\$	9-1)\$	E -> TE'
TE')T'E'\$	9-1)\$	T -> FT'
FT'E')T'E'\$	9-1)\$	F -> nb
nbT'E')T'E'\$	9-1)\$	
T'E')T'E'\$	-1)\$	T' -> ε
E')T'E'\$	-1)\$	E' -> -TE'
-TE')T'E'\$	-1)\$	
TE')T'E'\$	1)\$	T -> FT'
FT'E')T'E'\$	1)\$	F -> nb
nbT'E')T'E'\$	1)\$	
T'E')T'E'\$)\$	T' -> ε
E')T'E'\$)\$	E' -> ε
)T'E'\$)\$	
T'E'\$	\$	T' -> ε
E'\$	\$	E' -> ε
\$	\$	ACCEPTÉ

80

Exemple d'analyseur syntaxique LL(1) - suite

Et en même temps on peut construire l'arbre syntaxique :



81



Analyse LL(k)

- LL(1) car
 - analyse de gauche (**L**eft) à droite
 - dérivations gauches (**L**eft)
 - **1** symbole regardé en avance
- Certaines grammaires ne sont pas LL(1) mais sont LL(k) avec $k > 1$

82



Quelques propriétés des grammaires LL(1)

- **Théorème** : toute grammaire LL(1) est factorisée à gauche
- En effet : s'il existe deux productions
$$A \rightarrow \alpha\beta_1 \text{ et } A \rightarrow \alpha\beta_2$$
elles seront toutes les deux dans chaque case $M[A, a]$ où a est dans $\text{Premiers}(\alpha)$
- **Attention** : la réciproque n'est pas nécessairement vraie

83

Quelques propriétés des grammaires LL(1)

- **Théorème** : toute grammaire LL(1) est non récursive à gauche
- En effet : s'il existe les productions
$$X \rightarrow A\alpha \mid \beta_1 \text{ et } A \rightarrow X\beta_2$$
les cases $M[X, x]$ avec x dans $\text{Premiers}(\beta_1)$ contiendront les deux productions
$$X \rightarrow A\alpha \text{ et } X \rightarrow \beta_1$$
- **Attention** : la réciproque n'est pas nécessairement vraie

84

Quelques propriétés des grammaires LL(1)

- **Conséquence** : avant de chercher à construire une table LL(1), il faut supprimer la récursivité à gauche (s'il y en a), puis factoriser à gauche (si besoin)
- **Attention** cependant : ce n'est pas parce qu'une grammaire est non récursive et factorisée qu'elle sera LL(1)

85

Élimination de la récursivité immédiate à gauche

- Remplacer toute règle de la forme $A \rightarrow A\alpha|\beta$ par :

$$\begin{cases} A \rightarrow \beta A' \\ A' \rightarrow \alpha A' | \varepsilon \end{cases}$$

- Remarque : $L(A) = \beta\alpha^n$

86

Élimination de la récursivité à gauche, cas général

Ordonner les non-terminaux A_1, A_2, \dots, A_n

Pour $i=1$ à n faire

Pour $j=1$ à $i-1$ faire

 remplacer chaque production de la forme $A_i \rightarrow A_j\alpha$ où

$A_j \rightarrow \beta_1 | \dots | \beta_p$ par :

$A_i \rightarrow \beta_1\alpha | \dots | \beta_p\alpha$

Fin pour

 éliminer les récursivités à gauche immédiates des productions A_i

Fin pour

87

Exemple d'élimination de la récursivité à gauche

- $S \rightarrow AA \mid y$
 $A \rightarrow BzS \mid xy$
 $B \rightarrow SA \mid z$
- On ordonne : S, A, B
- $i=1$: supprimer la récursivité immédiate de S : rien à faire
- $i=2$:
 - $j=1$: remplacer les S au début des productions de A : rien à faire
 - Supprimer la récursivité immédiate de A : rien à faire
- $i=3$:
 - $j=1$: remplacer les S qui se trouvent au début des productions de B
 $B \rightarrow AAA \mid yA \mid z$
 - $j=2$: remplacer les A qui se trouvent au début des productions de B
 $B \rightarrow BzSAA \mid xyAA \mid yA \mid z$
 - Supprimer la récursivité immédiate de B :
 $B \rightarrow xyAAB' \mid yAB' \mid zB'$
 $B' \rightarrow zSAAB' \mid \varepsilon$

88

Cas particulier : grammaire non propre

- Lorsque l'on élimine la récursivité à gauche d'une grammaire non propre, elle ne sera toujours pas LL(1)
- **Conséquence** : rendre la grammaire propre avant d'éliminer la récursivité à gauche

89



Rendre une grammaire propre

- S'il existe une production $A \rightarrow \varepsilon$, on la supprime et pour chaque A apparaissant en partie droite d'une production, on rajoute une production dans laquelle le A est remplacé par ε .
- Si ε appartient à $L(G)$, on crée l'axiome S' avec $S' \rightarrow S \mid \varepsilon$ (où S est l'ancien axiome)

90



Exemple : rendre une grammaire propre

- $S \rightarrow AxB \mid Sy$
| $A \rightarrow xAyA \mid \varepsilon$
| $B \rightarrow ByA \mid x$

devient :

- $S \rightarrow AxB \mid xB \mid Sy$
| $A \rightarrow xAyA \mid xAy \mid xyA \mid xy$
| $B \rightarrow ByA \mid By \mid x$

91

Factorisation à gauche

Pour chaque non-terminal A

trouver le plus long préfixe α commun à deux de ses alternatives ou plus

Si $\alpha \neq \varepsilon$, remplacer $A \rightarrow \alpha\beta_1 | \dots | \alpha\beta_n | \gamma_1 | \dots | \gamma_p$ (où les γ_i ne commencent pas par α) par

$$\begin{cases} A \rightarrow \alpha A' | \gamma_1 | \dots | \gamma_p \\ A' \rightarrow \beta_1 | \dots | \beta_n \end{cases}$$

Fin si

Recommencer jusqu'à ne plus en trouver

92

Exemple de factorisation à gauche

$S \rightarrow AxB \mid Ay \mid xy$
 $A \rightarrow xAy \mid xA \mid Sx$
 $B \rightarrow xByA \mid xy \mid xBS$

devient :

$S \rightarrow AS' \mid xy$
 $S' \rightarrow xB \mid y$
 $A \rightarrow xAA' \mid Sx$
 $A' \rightarrow y \mid \varepsilon$
 $B \rightarrow xB''$
 $B' \rightarrow yA \mid S$
 $B'' \rightarrow BB' \mid y$

93

Conclusion sur l'analyse descendante

- Type de grammaire : LL(k)
- Faire transformations sur la grammaire si nécessaire
- Calcul des ensembles Premiers_k et Suivants_k
- Construction de la table d'analyse LL(k)

94

Remarque

Soit la grammaire G :
 $S \rightarrow aSb \mid bSa \mid \varepsilon$

Il n'existe pas de k tel que G soit LL(k)

95