

Langages et Compilation

Analyse descendante prédictive

Introduction

Grammaires $LL(1)$

Une famille de grammaires analysables de façon efficace.

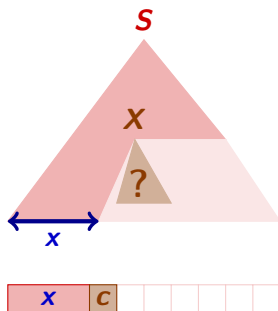
Caractéristiques de l'analyse $LL(1)$

- analyse descendante

Construction de l'arbre de dérivation selon l'ordre préfixé :
on part de la racine (l'axiome) et
on descend vers les feuilles (les
terminaux) en développant le nœud
interne (une variable) le plus à
gauche.

- analyse prédictive

Pour développer le nœud, on choisit
la règle à appliquer en prévisualisant
le symbole courant du mot analysé.



Introduction

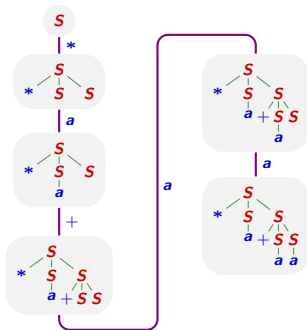
Exemple $S \rightarrow +SS \mid *SS \mid a$

Les chaînes dérivées de **S** commencent toutes par un terminal distinct

~> À partir du terminal courant du mot analysé, on sait déterminer la bonne règle.

	+	*	a
S	$S \rightarrow +SS$	$S \rightarrow *SS$	$S \rightarrow a$

Analyse du mot $*a + aa$



Introduction

Exemple $S \rightarrow aSbS \mid \epsilon$

Le choix devient plus délicat lorsque la grammaire comprend des ϵ -productions. Quel critère doit-on prendre en compte pour choisir entre la règle $S \rightarrow aSbS$ et la règle $S \rightarrow \epsilon$?

On est amené à considérer les terminaux qui peuvent suivre S .

Convention pratique, on introduit un terminal particulier $\$$ pour marquer la fin du mot à analyser.

$\$$ et b (mais pas a) peuvent se retrouver à droite de S .

\leadsto On sait alors déterminer la règle à appliquer au vu du terminal courant.

	$\$$	a	b
S	$S \rightarrow \epsilon$	$S \rightarrow aSbS$	$S \rightarrow \epsilon$

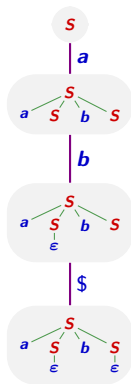
Introduction

Exemple $S \rightarrow aSbS \mid \epsilon$

sa table d'analyse

	\$	a	b
S	$S \rightarrow \epsilon$	$S \rightarrow aSbS$	$S \rightarrow \epsilon$

Analyse du mot $ab\$$



Introduction

Condition pour une analyse $LL(1)$ à chaque étape, pour la variable **X** à développer et le terminal courant **c** en entrée, le choix de la dérivation à appliquer doit être déterministe.

- L** *lecture de l'entrée de la gauche vers la droite (Left to right scanning)*
- L** *construction d'une dérivation gauche (Left derivation)*
- 1** *symbole de l'entrée pour prédire la bonne règle*

L'analyse $LL(1)$ se base sur une table qui indique, pour la variable **X** et le terminal **c**, la règle correcte à appliquer.

Pour construire cette table d'analyse, on détermine au préalable :

1. les variables effaçables
2. les ensembles Premier
3. les ensembles Suivant

Les variables effaçables

Une chaîne $\alpha \in N^*$ est dite **effaçable** si le mot vide se dérive de α : $\alpha \xrightarrow{*} \varepsilon$

Calcul des variables effaçables

On construit par récurrence sur i , l'ensemble \mathbf{Eff}_i des variables \mathbf{A} .

$$\mathbf{Eff}_0 = \{\mathbf{A} \in N : \mathbf{A} \rightarrow \varepsilon \in P\}$$

$$\mathbf{Eff}_{i+1} = \{\mathbf{A} \in N : \mathbf{A} \rightarrow \alpha \in P \text{ et } \alpha \in \mathbf{Eff}_i^*\}$$

Arrêt lorsque $\mathbf{Eff}_{i+1} = \mathbf{Eff}_i$ (au bout d'au plus $|N|$ étapes)

Exemple

$$\left\{ \begin{array}{ll} S & \rightarrow XY \\ X & \rightarrow aXb \mid \varepsilon \\ Y & \rightarrow cZ \mid Ze \\ Z & \rightarrow dcZ \mid \varepsilon \end{array} \right.$$

$$\mathbf{Eff}_0 = \{X, Z\} = \mathbf{Eff}_1$$

Les ensembles Premier

Soit α une chaîne de $(\Sigma \cup \mathbf{N})^*$. $\mathbf{Premier}(\alpha)$ est l'ensemble des terminaux qui débutent les chaînes dérivées de α :

$$\left\{ \mathbf{a} \in \Sigma : \alpha \xrightarrow{*} \mathbf{a}\beta \text{ où } \beta \in (\Sigma \cup \mathbf{N})^* \right\}$$

Calcul des ensembles Premier

Pour un terminal \mathbf{a} :

$$\mathbf{Premier}(\mathbf{a}\beta) = \{\mathbf{a}\}$$

Pour une variable \mathbf{X} :

$$\mathbf{Premier}(\mathbf{X}) = \bigcup_{\mathbf{X} \rightarrow \beta \in P} \mathbf{Premier}(\beta)$$

$$\mathbf{Premier}(\mathbf{X}\beta) =$$

$$\begin{cases} \mathbf{Premier}(\mathbf{X}) & \text{si } \mathbf{X} \text{ n'est pas effaçable} \\ \mathbf{Premier}(\mathbf{X}) \cup \mathbf{Premier}(\beta) & \text{sinon} \end{cases}$$

Les ensembles Premier

Exemple

$$\left\{ \begin{array}{lcl} S & \rightarrow & XY \\ X & \rightarrow & aXb \mid \varepsilon \\ Y & \rightarrow & cZ \mid Ze \\ Z & \rightarrow & dcZ \mid \varepsilon \end{array} \right.$$

Calcul des ensembles Premier pour les variables :

$$\begin{aligned} \mathbf{Premier(S)} &= \mathbf{Premier(XY)} && S \rightarrow XY \\ &= \mathbf{Premier(X) \cup Premier(Y)} && X \text{ effaçable} \end{aligned}$$

$$\begin{aligned} \mathbf{Premier(X)} &= \mathbf{Premier(aXb)} && X \rightarrow aXb \mid \varepsilon \\ &= \{a\} \end{aligned}$$

$$\begin{aligned} \mathbf{Premier(Y)} &= \mathbf{Premier(cZ) \cup Premier(Ze)} && Y \rightarrow cZ \mid Ze \\ &= \{c\} \cup \mathbf{Premier(Z)} \cup \{e\} && Z \text{ effaçable} \end{aligned}$$

$$\begin{aligned} \mathbf{Premier(Z)} &= \mathbf{Premier(dcZ)} && Z \rightarrow dcZ \mid \varepsilon \\ &= \{d\} \end{aligned}$$

Les ensembles Suivant

Soit \mathbf{X} une variable. $\mathbf{Suivant}(\mathbf{X})$ est l'ensemble des terminaux qui peuvent apparaître après \mathbf{X} dans une dérivation :

$$\left\{ a \in \Sigma : S \xrightarrow{*} \alpha \mathbf{X} a \beta \text{ où } \alpha, \beta \in (\Sigma \cup \mathbf{N})^* \right\}$$

i.e.

$$\mathbf{Suivant}(\mathbf{X}) = \bigcup_{Y \rightarrow \alpha \mathbf{X} \beta \in P} \mathbf{Premier}(\beta \mathbf{Suivant}(\mathbf{Y}))$$

Calcul des ensembles Suivant

Mettre $\$$ dans $\mathbf{Suivant}(\mathbf{S})$ (où \mathbf{S} est l'axiome)

Pour chaque variable \mathbf{X} , examiner chaque production $\mathbf{Y} \rightarrow \alpha \mathbf{X} \beta$ où \mathbf{X} est à droite :

- Si $\beta \neq \varepsilon$, ajouter les éléments de $\mathbf{Premier}(\beta)$ à $\mathbf{Suivant}(\mathbf{X})$
- Si $\beta \xrightarrow{*} \varepsilon$, ajouter les éléments de $\mathbf{Suivant}(\mathbf{Y})$ à $\mathbf{Suivant}(\mathbf{X})$

Les ensembles Suivant

Exemple

$$\left\{ \begin{array}{lcl} S & \rightarrow & XY \\ X & \rightarrow & aXb \mid \varepsilon \\ Y & \rightarrow & cZ \mid Ze \\ Z & \rightarrow & dcZ \mid \varepsilon \end{array} \right.$$

Calcul des ensembles Suivant :

Suivant(*S*) contient \$

Suivant(*X*)

$$\begin{aligned} S \rightarrow XY &\leadsto \text{Premier}(Y \text{ Suivant}(S)) \subseteq \text{Suivant}(X) \\ &\leadsto \text{Premier}(Y) \subseteq \text{Suivant}(X) \end{aligned} \quad (Y \text{ non effaçable})$$

$$X \rightarrow aXb \leadsto b \in \text{Suivant}(X)$$

Suivant(*Y*)

$$S \rightarrow XY \leadsto \text{Suivant}(S) \subseteq \text{Suivant}(Y)$$

Suivant(*Z*)

$$Y \rightarrow cZ \leadsto \text{Suivant}(Y) \subseteq \text{Suivant}(Z)$$

$$Y \rightarrow Ze \leadsto e \in \text{Suivant}(Z)$$

$$Z \rightarrow dcZ \leadsto \text{Suivant}(Z) \subseteq \text{Suivant}(Z)$$

Les variables Effaçables et les ensembles Premier et Suivant

Exemple

$$\begin{cases} S & \rightarrow XY \\ X & \rightarrow aXb \mid \varepsilon \\ Y & \rightarrow cZ \mid Ze \\ Z & \rightarrow dcZ \mid \varepsilon \end{cases}$$

Bilan des calculs :

	Effaçable	Premier	Suivant
S	non	$a \quad c \quad d \quad e$	$\$$
X	oui	a	$b \quad c \quad d \quad e$
Y	non	$c \quad d \quad e$	$\$$
Z	oui	d	$\$ \quad e$

Construction de la table d'analyse

Table à deux dimensions indexée par les variables et les terminaux

Pour toute production $X \rightarrow \alpha$:

- Ajouter $X \rightarrow \alpha$ à l'entrée $T[X, a]$ pour tout terminal a dans **Premier**(α)
- Si α est effaçable, ajouter $X \rightarrow \alpha$ dans la case $T[X, a]$ pour tout terminal a dans **Suivant**(X)

		Effaçable	Premier	Suivant
$\left\{ \begin{array}{l} S \rightarrow XY \\ X \rightarrow aXb \mid \epsilon \\ Y \rightarrow cZ \mid Ze \\ Z \rightarrow dcZ \mid \epsilon \end{array} \right.$	S	non	$a \quad c \quad d \quad e$	$\$$
	X	oui	a	$b \quad c \quad d \quad e$
	Y	non	$c \quad d \quad e$	$\$$
	Z	oui	d	$\$ \quad e$

	\$	a	b	c	d	e
S		$S \rightarrow XY$		$S \rightarrow XY$	$S \rightarrow XY$	$S \rightarrow XY$
X		$X \rightarrow aXb$	$X \rightarrow \epsilon$	$X \rightarrow \epsilon$	$X \rightarrow \epsilon$	$X \rightarrow \epsilon$
Y				$Y \rightarrow cZ$	$Y \rightarrow Ze$	$Y \rightarrow Ze$
Z	$Z \rightarrow \epsilon$				$Z \rightarrow dcZ$	$Z \rightarrow \epsilon$

Grammaire $LL(1)$

Définition

Une grammaire est $LL(1)$ s'il existe au plus une production par entrée dans la table.

Proposition

Une grammaire est $LL(1)$ si pour toute paire de productions $A \rightarrow \alpha$ et $A \rightarrow \beta$, on a :

$$\text{Premier}(\alpha \text{ Suivant}(A)) \cap \text{Premier}(\beta \text{ Suivant}(A)) = \emptyset$$

Ainsi une grammaire ne sera pas $LL(1)$ s'il on a :

- soit un **conflit Premier/Premier**, i.e., deux règles $A \rightarrow \alpha \mid \beta$ telles que : $\text{Premier}(\alpha) \cap \text{Premier}(\beta) \neq \emptyset$
- soit un **conflit Premier/Suivant**, i.e., deux règles $A \rightarrow \alpha \mid \beta$ avec $\beta \xrightarrow{*} \epsilon$ telles que : $\text{Premier}(\alpha) \cap \text{Suivant}(A) \neq \emptyset$

Grammaire $LL(1)$

Exemple

$$\left\{ \begin{array}{l} S \rightarrow XY \\ X \rightarrow aXb \mid \varepsilon \\ Y \rightarrow cZ \mid Ze \\ Z \rightarrow dcZ \mid \varepsilon \end{array} \right. \text{ est une grammaire } LL(1)$$

Sa table d'analyse comprend au plus une alternative par case.

→ Le choix de la règle à appliquer se fait de façon déterministe.

	\$	a	b	c	d	e
S		$S \rightarrow XY$		$S \rightarrow XY$	$S \rightarrow XY$	$S \rightarrow XY$
X		$X \rightarrow aXb$	$X \rightarrow \varepsilon$	$X \rightarrow \varepsilon$	$X \rightarrow \varepsilon$	$X \rightarrow \varepsilon$
Y				$Y \rightarrow cZ$	$Y \rightarrow Ze$	$Y \rightarrow Ze$
Z	$Z \rightarrow \varepsilon$				$Z \rightarrow dcZ$	$Z \rightarrow \varepsilon$

Grammaire $LL(1)$

Exemple

$$\left\{ \begin{array}{l} S \rightarrow XY \mid Z \\ X \rightarrow aXb \mid \varepsilon \\ Y \rightarrow bX \\ Z \rightarrow bZc \mid c \end{array} \right. \quad \text{n'est pas une grammaire } LL(1)$$

Il existe un conflit Premier/Premier pour les règles $S \rightarrow XY$ et $S \rightarrow Z$

$\text{Premier}(XY) = \text{Premier}(X) \cup \text{Premier}(Y) = \{a, b\}$ X effaçable

$\text{Premier}(Z) = \{b, c\}$

$$\leadsto \text{Premier}(XY) \cap \text{Premier}(Z) = \{b\} \neq \emptyset$$

	\$	b	\dots
\dots			
S		$S \rightarrow XY$ $S \rightarrow Z$	

Grammaire $LL(1)$

Exemple

$$\begin{cases} S \rightarrow aXb \\ X \rightarrow bX \mid \varepsilon \end{cases} \text{ n'est pas une grammaire } LL(1)$$

Il existe un conflit Premier/Suivant pour les règles $X \rightarrow bX$ et $X \rightarrow \varepsilon$

$$\text{Premier}(bX) = \{b\}$$

$$\text{Suivant}(X) = \{b\}$$

$$\leadsto \text{Premier}(X) \cap \text{Suivant}(X) = \{b\} \neq \emptyset$$

	\$	b	\dots
\dots			
X		$X \rightarrow bX$ $X \rightarrow \varepsilon$	

Analyseur $LL(1)$

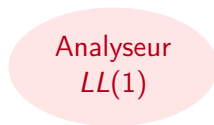
Pour examiner un mot, l'analyseur $LL(1)$ utilise la table d'analyse précédemment construite et une pile.

Initialement, la pile contient le marqueur de fin de mot \$ et l'axiome.

Le mot analysé lu de gauche à droite



La table d'analyse
de la grammaire $LL(1)$



S

L'arbre d'analyse
à construire



La pile

Analyseur $LL(1)$

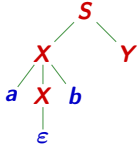
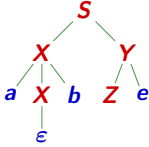
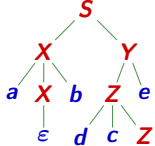
À chaque étape, on examine le terminal courant c du mot analysé et le sommet de la pile (premier symbole non traité de l'arbre en construction).

- Soit le sommet de la pile est un terminal a :
 - si $a \neq c$,
l'analyse s'arrête et retourne ÉCHEC
 - si $a = c = \$$,
l'analyse s'arrête et retourne SUCCÈS
 - si $a = c \neq \$$,
 a est dépilé et on avance dans la lecture du mot analysé
- Soit le sommet de la pile est une variable X :
 - si l'entrée $T[X, c]$ est vide,
l'analyse s'arrête et retourne ÉCHEC
 - si l'entrée $T[X, c]$ contient une règle $X \rightarrow \alpha$,
 X est dépilé et α est empilé en partant de la droite
(par exemple, si $X \rightarrow YzT$, T est empilé, puis z , puis Y).

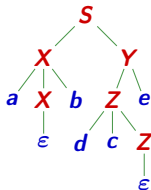
Analyseur $LL(1)$

	\$	a	b	c	d	e
S		$S \rightarrow XY$		$S \rightarrow XY$	$S \rightarrow XY$	$S \rightarrow XY$
X		$X \rightarrow aXb$	$X \rightarrow \varepsilon$	$X \rightarrow \varepsilon$	$X \rightarrow \varepsilon$	$X \rightarrow \varepsilon$
Y				$Y \rightarrow cZ$	$Y \rightarrow Ze$	$Y \rightarrow Ze$
Z	$Z \rightarrow \varepsilon$				$Z \rightarrow dcZ$	$Z \rightarrow \varepsilon$

Mot analysé	Pile	
<i>abdce</i> \$	<i>S</i> \$	$S \rightarrow XY$
<i>abdce</i> \$	<i>XY</i> \$	$X \rightarrow aXb$
<i>abdce</i> \$	<i>aXbY</i> \$	assortiment

Mot analysé	Pile	
$a\textcolor{blue}{b}dce\$$	$\textcolor{red}{X}b\textcolor{red}{Y}\$$	$\textcolor{brown}{X} \rightarrow \epsilon$ 
$a\textcolor{blue}{b}dce\$$ $ab\textcolor{blue}{d}ce\$$	$\textcolor{blue}{b}Y\$$ $\textcolor{red}{Y}\$$	$\textcolor{brown}{assortiment}$ $\textcolor{brown}{Y} \rightarrow \textcolor{blue}{Z}e$ 
$ab\textcolor{blue}{d}ce\$$	$\textcolor{red}{Z}e\$$	$\textcolor{brown}{Z} \rightarrow \textcolor{blue}{d}c\textcolor{red}{Z}$ 
$ab\textcolor{blue}{d}ce\$$	$\textcolor{blue}{d}c\textcolor{red}{Z}e\$$	$\textcolor{brown}{assortiment}$

Mot analysé	Pile	
<i>abd</i> ce \$	c Z <i>e</i> \$	assortiment
<i>abd</i> ce \$	Z <i>e</i> \$	$Z \rightarrow \epsilon$
<i>abd</i> ce \$	e \$	assortiment
<i>abd</i> ce \$	\$	SUCCES

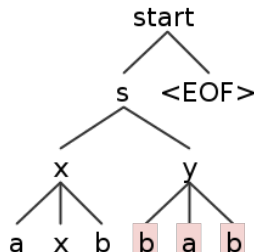


Le coût de l'analyse du mot est linéaire en la taille de l'arbre.

Analyseur $LL(1)$

	\$	a	b	c	d	e
S		$S \rightarrow XY$		$S \rightarrow XY$	$S \rightarrow XY$	$S \rightarrow XY$
X		$X \rightarrow aXb$	$X \rightarrow \epsilon$	$X \rightarrow \epsilon$	$X \rightarrow \epsilon$	$X \rightarrow \epsilon$
Y				$Y \rightarrow cZ$	$Y \rightarrow Ze$	$Y \rightarrow Ze$
Z	$Z \rightarrow \epsilon$				$Z \rightarrow dcZ$	$Z \rightarrow \epsilon$

Mot analysé	Pile	
$abbab\$$	$S\$$	$S \rightarrow XY$
$abbab\$$	$XY\$$	$X \rightarrow aXb$
$abbab\$$	$aXbY\$$	assortiment
$aabbab\$$	$XbY\$$	$X \rightarrow \epsilon$
$aabbab\$$	$bY\$$	assortiment
$ababb\$$	$Y\$$	ECHEC



Rendre une grammaire $LL(1)$

Proposition

Une grammaire ne peut pas être $LL(1)$ si elle est :

- soit ambiguë,
- soit récursive gauche,
- soit n'est pas factorisée à gauche.

On peut modifier la grammaire pour tenter de la rendre $LL(1)$ mais le résultat n'est pas garanti.

Proposition

Il existe des grammaires non ambiguës, non récursives gauche et factorisées à gauche qui ne sont pas $LL(1)$.

Rendre une grammaire $LL(1)$

Par élimination de la récursivité gauche

Exemple

$$\left\{ \begin{array}{l} E \rightarrow E + T \mid T \\ T \rightarrow T * F \mid F \\ F \rightarrow (E) \mid nb \end{array} \right. \quad \text{n'est pas } LL(1) \text{ car récursive gauche}$$

Il existe un conflit Premier/Premier pour les règles $E \rightarrow E + T$ et $E \rightarrow T$ à cause de la récursivité gauche de E

et un conflit Premier/Premier pour les règles $T \rightarrow T * F$ et $T \rightarrow F$ à cause de la récursivité gauche de T

	\$	(<i>nb</i>)	+	*
E		$E \rightarrow E + T$ $E \rightarrow T$	$E \rightarrow E + T$ $E \rightarrow T$			
T		$T \rightarrow T * F$ $T \rightarrow F$	$T \rightarrow T * F$ $T \rightarrow F$			
F		$F \rightarrow (E)$	$F \rightarrow nb$			

Rendre une grammaire **LL(1)**

Par élimination de la récursivité gauche

Supprimer la récursivité gauche rend cette grammaire **LL(1)**

$$\left\{ \begin{array}{l} E \rightarrow E + T \mid T \\ T \rightarrow T * F \mid F \\ F \rightarrow (E) \mid nb \end{array} \right. \longrightarrow \left\{ \begin{array}{l} E \rightarrow TY \\ Y \rightarrow +TY \mid \varepsilon \\ T \rightarrow FZ \\ Z \rightarrow *FZ \mid \varepsilon \\ F \rightarrow (E) \mid nb \end{array} \right.$$

Rendre une grammaire $LL(1)$

Par élimination de la récursivité gauche

$$\left\{ \begin{array}{l} E \rightarrow TY \\ Y \rightarrow +TY \mid \varepsilon \\ T \rightarrow FZ \\ Z \rightarrow *FZ \mid \varepsilon \\ F \rightarrow (E) \mid nb \end{array} \right.$$

	Effaçable	Premier	Suivant
E	non	(nb	\$)
Y	oui	+	\$)
T	non	(nb	\$) +
Z	oui	*	\$) +
F	non	(nb	* \$) +

	\$	(nb)	+	*
E		$E \rightarrow TY$	$E \rightarrow TY$			
Y	$Y \rightarrow \varepsilon$			$Y \rightarrow \varepsilon$	$Y \rightarrow +TY$	
T		$T \rightarrow FZ$	$T \rightarrow FZ$			
Z	$Z \rightarrow \varepsilon$			$Z \rightarrow \varepsilon$	$Z \rightarrow \varepsilon$	$Z \rightarrow *FZ$
F		$F \rightarrow (E)$	$F \rightarrow nb$			

Rendre une grammaire $LL(1)$

Par substitution et factorisation

Exemple

$$\left\{ \begin{array}{l} E \rightarrow TR \\ T \rightarrow id \mid (E) \mid A \\ R \rightarrow +E \mid \varepsilon \\ A \rightarrow id[E] \end{array} \right. \quad \text{n'est pas } LL(1)$$

Il existe un conflit Premier/Premier pour les règles $T \rightarrow id$ et $T \rightarrow A$

	\$	id	\dots
\dots			
T		$T \rightarrow id$ $T \rightarrow A$	

Rendre une grammaire $LL(1)$

Par substitution et factorisation

On effectue une substitution avant de factoriser à gauche.

$$\begin{array}{l|l} E & \rightarrow TR \\ T & \rightarrow id \mid (E) \mid A \\ R & \rightarrow +E \mid \varepsilon \\ A & \rightarrow id[E] \end{array}$$

- Substitution

$$\begin{array}{l|l} E & \rightarrow TR \\ T & \rightarrow id \mid (E) \mid id[E] \\ R & \rightarrow +E \mid \varepsilon \end{array}$$

- Factorisation à gauche

$$\begin{array}{l|l} E & \rightarrow TR \\ T & \rightarrow id X \mid (E) \\ X & \rightarrow [E] \mid \varepsilon \\ R & \rightarrow +E \mid \varepsilon \end{array}$$

Rendre une grammaire $LL(1)$

Par substitution et factorisation

On obtient une grammaire $LL(1)$

$\begin{cases} E \rightarrow TR \\ T \rightarrow id X \mid (E) \\ X \rightarrow [E] \mid \varepsilon \\ R \rightarrow +E \mid \varepsilon \end{cases}$		Effaçable	Premier	Suivant
E	non	id ($\$$)]	
T	non	id ($\$$)]	+
X	oui	[$\$$)]	+
R	oui	+	$\$$)]	

	\$	+	id	()	[]
E			$E \rightarrow TR$	$E \rightarrow TR$			
T			$T \rightarrow id X$	$T \rightarrow (E)$			
X	$X \rightarrow \varepsilon$	$X \rightarrow \varepsilon$		$X \rightarrow \varepsilon$		$X \rightarrow [E]$	$X \rightarrow \varepsilon$
R	$R \rightarrow \varepsilon$	$R \rightarrow +E$			$R \rightarrow \varepsilon$		$R \rightarrow \varepsilon$

Rendre une grammaire $LL(1)$

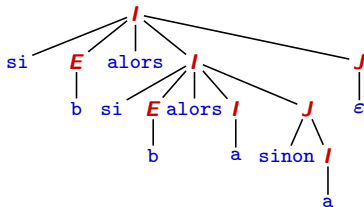
ou rendre l'analyse déterministe

L'ambiguïté engendre un conflit Premier/Suivant pour les règles

$J \rightarrow \text{sinon } I$ et $J \rightarrow \epsilon$

	\$	<i>sinon</i>	...
...			
<i>J</i>	$J \rightarrow \epsilon$	$J \rightarrow \text{sinon } I$ $J \rightarrow \epsilon$...

Avec la convention usuelle d'associer le sinon avec le si le plus proche, on peut rendre l'analyseur déterministe et le forcer à produire l'arbre voulu en privilégiant la règle $J \rightarrow \text{sinon } I$ au détriment de la règle $J \rightarrow \epsilon$.



Analyse $LL(k)$

Généralisation de l'analyse $LL(1)$ avec prévisualisation non pas juste d'un symbole mais d'un nombre fixé k de symboles. Une grammaire est $LL(k)$ si l'analyseur peut choisir de façon déterministe la règle à appliquer en examinant les k symboles courants de l'entrée.

On note :

- $w|_k = \begin{cases} w & \text{si } w \text{ est de longueur au plus } k \\ \text{le préfixe de longueur } k \text{ de } w & \text{sinon} \end{cases}$
- $\text{Premier}_k(\alpha) = \{w|_k : \alpha \xrightarrow{*} w\}$
- $\text{Suivant}_k(A) = \{w : \exists \beta, \gamma \text{ t.q. } S \rightarrow \beta A \gamma \text{ et } w \in \text{Premier}_k(\gamma)\}$

Proposition

Une grammaire est $LL(k)$ si pour toute paire de productions $A \rightarrow \alpha$ et $A \rightarrow \beta$, on a :

$$\text{Premier}_k(\alpha \text{ Suivant}_k(A)) \cap \text{Premier}_k(\beta \text{ Suivant}_k(A)) = \emptyset$$

Analyse $LL(k)$

Exemple

$$\begin{cases} S \rightarrow abX \mid \varepsilon \\ X \rightarrow Saa \mid b \end{cases} \quad \text{n'est pas une grammaire } LL(1)$$

	Effaçable	Premier	Suivant
S	oui	a	$\$ a$
X	non	$a \ b$	$\$ a$

Conflit Premier/Suivant : $\text{Premier}(abX) \cap \text{Suivant}(S) \neq \emptyset$

	$\$$	a	b
S	$S \rightarrow \varepsilon$	$S \rightarrow abX$ $S \rightarrow \varepsilon$	
X		$X \rightarrow Saa$	$X \rightarrow b$

Analyse $LL(k)$

$$\begin{cases} S \rightarrow abX \mid \varepsilon \\ X \rightarrow Saa \mid b \end{cases} \text{ est une grammaire } LL(2)$$

	Effaçable	Premier ₂	Suivant ₂
S	oui	ab	\$ aa
X	non	$aa \quad ab \quad b$	\$ aa

En prévisualisant deux lettres, il n'y a plus de conflit :

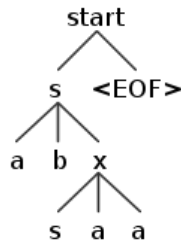
$$\text{Premier}_2(abX) \cap \text{Suivant}_2(S) = \emptyset$$

	\$	aa	ab	b
S	$S \rightarrow \varepsilon$	$S \rightarrow \varepsilon$	$S \rightarrow abX$	
X		$X \rightarrow Saa$	$X \rightarrow Saa$	$X \rightarrow b$

Analyse $LL(k)$

	\$	aa	ab	b
S	$S \rightarrow \epsilon$	$S \rightarrow \epsilon$	$S \rightarrow abX$	
X		$X \rightarrow Saa$	$X \rightarrow Saa$	$X \rightarrow b$

Mot analysé	Pile	
ab aa\$	S \$	$S \rightarrow abX$
ab aa\$	abX \$	assortiment
a ba a\$	bX \$	assortiment
ab aa \$	X \$	$X \rightarrow Saa$
ab aa \$	Saa \$	$S \rightarrow \epsilon$
ab aa \$	aa \$	assortiment
aba a \$	a \$	assortiment
abaa\$	\$	SUCCES



Analyse $LL(k)$

Fait

Il existe des grammaires ni ambiguës, ni récursives gauche qui ne sont $LL(k)$ pour aucun k .

Exemple

$$\begin{cases} S \rightarrow A \mid B \\ A \rightarrow aAb \mid c \\ B \rightarrow aBbb \mid d \end{cases}$$

$a^k \in \text{Premier}_k(A) \cap \text{Premier}_k(B)$ pour tout k

\leadsto Conflit Premier/Premier pour tout k

Exemple

$$\{ S \rightarrow aSb \mid bSa \mid \varepsilon \}$$

$\text{Premier}_k(aSb) \cap \text{Suivant}_k(S) = \{a\}\{a, b\}^{k-2}\{b\}$ pour tout k

\leadsto Conflit Premier/Suivant pour tout k

Analyse syntaxique avec ANTLR

ANTLR4 met en œuvre une analyse descendante prédictive qui a les caractéristiques suivantes

- factorisation à gauche de la grammaire automatique
- suppression automatique des récursivités gauches immédiates (mais pas des récursivités gauches indirectes, e.g. $A \rightarrow B\alpha, B \rightarrow A\beta$)
- résolution de certaines ambiguïtés
 - en jouant sur l'ordre des productions pour lever celles liées aux priorités des opérateurs
 - pour les ambiguïtés dues à l'associativité, par défaut l'association s'effectue à gauche ou de façon explicite à droite, e.g. l'exponentielle, $expr: expr \wedge \langle assoc = right \rangle expr$
- basé sur une stratégie $LL(k)$
- enrichi d'un mécanisme additionnel qui permet de traiter plus que les grammaires $LL(k)$ mais qui induit alors un surcoût en temps