

## 1- Introduction

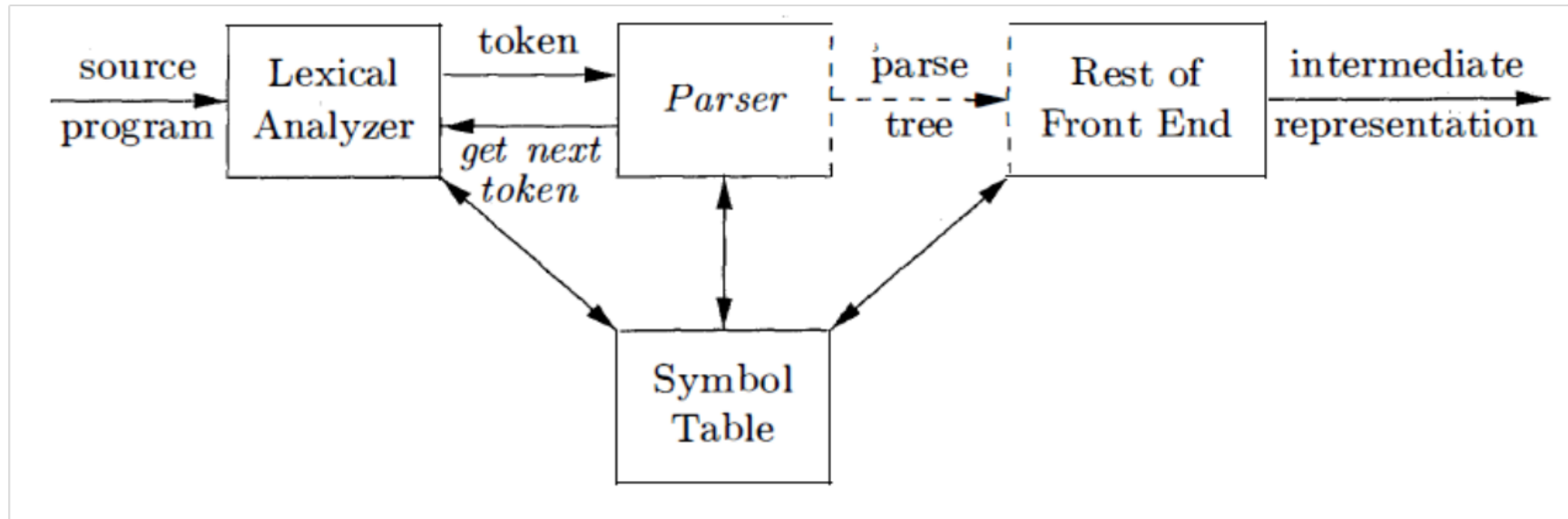
- Si l'analyse lexicographique résout le problème de l'appartenance, des mots du texte source, au langage. L'analyse syntaxique a pour objectif de vérifier si la structure des phrases du texte source appartient au langage en question.
- L'analyse syntaxique ou parsing, consiste à déterminer la syntaxe ou la structure d'un programme.
- La syntaxe d'un langage de programmation est donnée par une grammaire hors-contexte.

## Analyse Syntaxique

---

- L'analyseur syntaxique reçoit une suite d'unités lexicales de la part de l'analyseur lexical et doit vérifier que cette suite peut être engendrée par la grammaire du langage.
- Cela consiste à construire un arbre syntaxique de la suite des unités lexicales formant le programme (le texte en entrée)

## 2- Position de l'analyseur syntaxique



### 3- Grammaire hors-contexte versus expression régulière

Les grammaires sont des notations plus fortes que les expressions régulières. Chaque construction pouvant être décrite par une expression régulière peut être décrite par une grammaire hors-contexte.

Pourquoi utiliser les expressions régulières pour définir le lexique d'un langage de programmation ?

- La séparation de la structurer syntaxique d'un langage en deux modules de tailles faciles à gérer,
- Les règles lexicales d'un langage sont généralement simple, dont leur description ne nécessite pas un mécanisme assez puissant tel que les grammaires,

## Analyse Syntaxique

---

- Les expressions régulières fournissent une notation plus concise et facile à comprendre pour les tokens que les grammaires
- Des analyseurs lexicaux plus efficaces peuvent être construits automatiquement à partir des expressions régulières qu'à partir des grammaires

### Rappel

Une grammaire est dite ambiguë si pour une chaîne du texte source. Il existe au moins deux arbres syntaxiques générés par *les dérivations les plus gauches*.

- **Dérivation la plus à gauche** : est entièrement composée de dérivations en une étape dans lesquelles à chaque fois c'est le non-terminal le plus à

gauche qui est remplacé.  $\alpha \Rightarrow_{lm}^* \beta$

- **Dérivation la plus à droite** : à chaque étape de dérivation c'est le non-terminal le plus à droite qui est remplacé.  $\alpha \xRightarrow{*}_{rm} \beta$

### 4- Méthodes d'Analyse Syntaxique

Il existe deux méthodes d'analyse standards, formelles, et complètement maîtrisées :

- L'analyse descendante (Top-Down Parsing)
- L'analyse ascendante (Bottom-Up Parsing)

## Analyse Syntaxique

---

- **Méthode d'analyse descendante :** Etant donné une grammaire et un texte source, partir de l'axiome et tenter de créer une chaîne identique à la chaîne d'entrée par dérivations successives. L'opération de base est le remplacement d'un symbole non terminal par une de ses parties droites. On dit qu'on procède par génération (dérivation).

Résultat : arbre syntaxique dont la racine de l'arbre est l'axiome, et les feuilles sont les unités lexicales

- **Méthode d'analyse ascendante**

Cette méthode travaille de façon inverse de la 1<sup>ère</sup> méthode. Elle consiste à trouver une chaîne de symboles identique à une partie droite et la remplacer par le symbole non-terminal correspondant jusqu'à arriver à l'axiome. Dans ce cas on opère par réduction.



## 5- Analyse descendante

Deux méthodes : L'analyse descendante récursive, et l'analyse descendante prédictive.

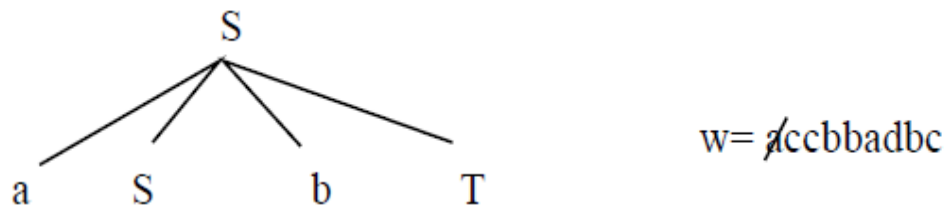
### Exemple

$$S \rightarrow aSbT \mid cT \mid d$$

$$T \rightarrow aT \mid bS \mid c$$

Avec le mot  $w = \text{accbbadbc}$ , on part avec l'arbre contenant le seul sommet  $S$

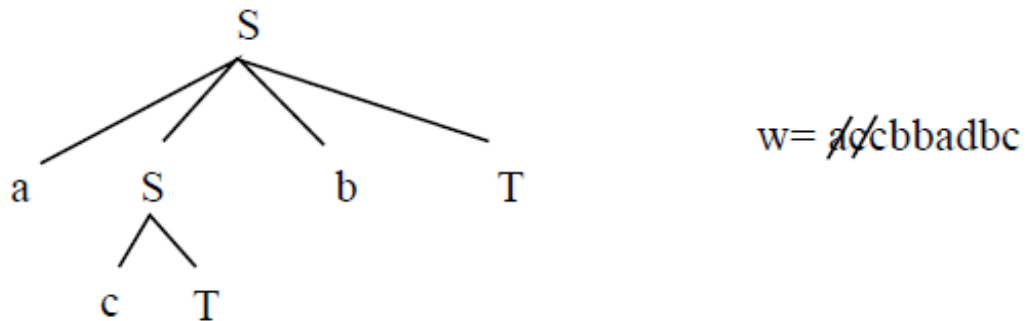
La lecture de la première lettre du mot  $a$  nous permet d'avancer la construction



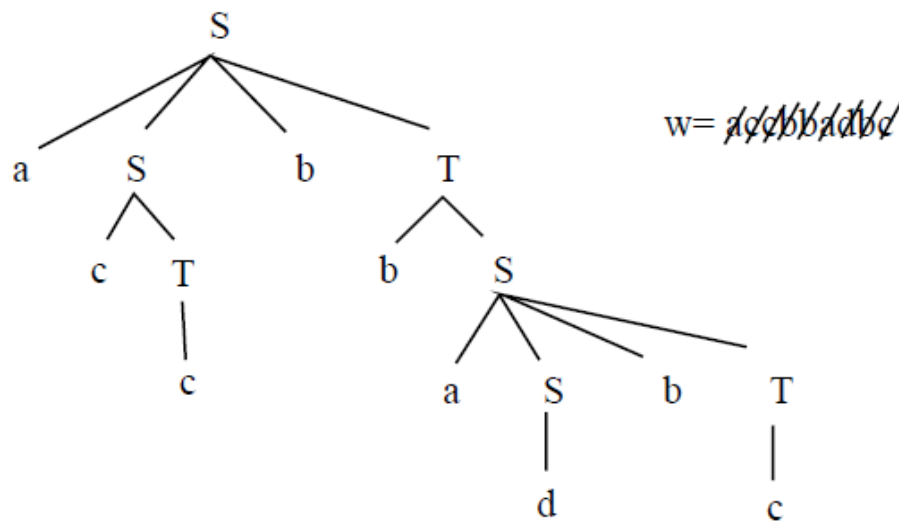
# Analyse Syntaxique

---

Puis la deuxième lettre c nous amène à :



Et ainsi de suite jusqu'à :



Ainsi le mot  $w$  appartient au langage généré par la grammaire.

Sur cet exemple c'est très facile car chaque règle commence par un terminal différent, donc on sait immédiatement quelle règle est utilisée.

### Exemple 2

$$S \rightarrow cAd$$

$$A \rightarrow ab \mid a$$

$$w = cad$$

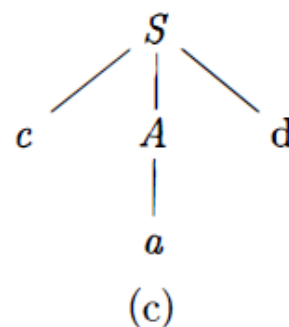
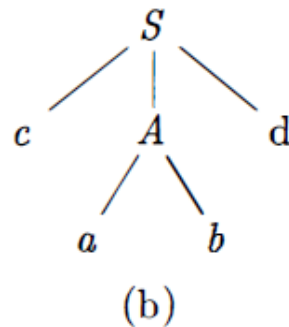
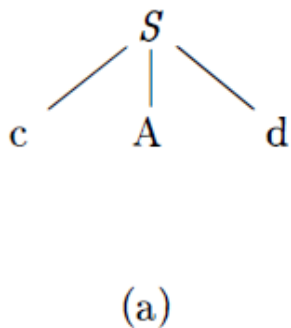
Après la lecture du premier symbole  $c$ , on se trouve avec l'arbre de la Figure (a). on avance le pointeur sur le symbole  $a$ . Le symbole non terminal le plus à gauche est  $A$ . nous remplaçons  $A$  avec la première alternative  $A \rightarrow ab$  (Figure (b)), et nous avons une correspondance pour le deuxième symbole  $c$  de  $w$ , alors

## Analyse Syntaxique

---

nous avançons le pointeur sur  $d$  qui ne correspond pas l'étiquète  $b$  de la feuille suivante. Dans ce cas nous reportons un échec et nous retournons à la deuxième alternative  $A \rightarrow a$  (backtracking). Cette dernière se termine avec succès (Figure (c)).

Lors du retour arrière, nous devons réinitialiser le pointeur d'entrée à la position 2.



### Analyse descendante récursive

Consiste à écrire pour chaque symbole non-terminal  $A$  une procédure récursive  $A()$ . L'exécution commence par la procédure du symbole de départ, qui arrête et annonce succès si la procédure scanne toute la chaîne en entrée.

```
void A() {  
1)    Choose an  $A$ -production,  $A \rightarrow X_1 X_2 \cdots X_k$ ;  
2)    for (  $i = 1$  to  $k$  ) {  
3)        if (  $X_i$  is a nonterminal )  
4)            call procedure  $X_i()$ ;  
5)        else if (  $X_i$  equals the current input symbol  $a$  )  
6)            advance the input to the next symbol;  
7)        else /* an error has occurred */;  
    }  
}
```

## Analyse Syntaxique

---

Le pseudo-code ci-dessus du symbole non-terminal  $A$  est non déterministe. Si  $A \rightarrow A_1 \mid A_2 \mid \dots \mid A_n$ , et si au cours de l'analyse on a dérivé le symbole non-terminal  $A$ , alors le problème qui se pose est celui du choix de la règle à appliquer. Dans ce cas l'analyse récursive nécessite des retours arrière (backtracking) nécessitant des scans répétés de la chaîne en entrée pour chaque règle.

Pour permettre des retours arrière le pseudo-code doit être modifié. Comment ? à la ligne (1) nous devons essayer chacune des alternatives de  $A$  dans un certain ordre. L'échec dans la ligne (7) n'est pas ultime, mais il faut retourner à la ligne (1) pour essayer une autre règle. Pour pouvoir essayer une autre règle, il

faut réinitialiser le pointeur de la chaîne en entrée sur la position dans laquelle il a été lors de ligne (1).

### **Analyse descendante prédictive (déterministe)**

L'analyse descendante prédictive permet toujours de choisir une règle de production unique en se basant sur le prochain symbole de l'entrée et sans effectuer aucun retour en arrière. Elle s'applique à une classe restreinte de grammaires hors-contexte, dite grammaire **LL(k)**.

## Analyse Syntaxique

---

Il ya deux façons pour effectuer une analyse descendante prédictive :

- L'analyse prédictive récursive : implémentée avec des procédures récursives sans retour arrière
- L'analyse prédictive non récursive : dirigée par une table d'analyse

Ces deux types d'analyse nécessitent le calcul de deux ensembles :

- l'ensemble Premier (First), et
- l'ensemble Suivant (Follow).



### ➤ Calcul de l'ensemble Premier

Soit  $\alpha \in (N \cup T)^*$  une forme d'une **GHC**  $G = (N, T, P, S)$ . On définit l'ensemble  $\text{Premier}(\alpha)$  (ou  $\text{First}(\alpha)$ ), comme étant l'ensemble des terminaux qui peuvent apparaître au début d'une forme dérivable à partir de  $\alpha$ .  
Formellement :

$$\text{Premier}(\alpha) = \{ a \in T \mid \alpha \Rightarrow^* a\beta, \beta \in (N \cup T)^* \}$$

Algorithme de construction des ensembles Premier :

- 1- Si  $X$  est non-terminal et  $X \rightarrow Y_1 Y_2 \dots Y_n$  est une règle de la grammaire ( $Y_i$  symbole terminal ou non-terminal) alors :
  - 1- ajouter les éléments de  $\text{Premier}(Y_1)$  sauf  $\varepsilon$  dans  $\text{Premier}(X)$ ,
  - 2- s'il existe un  $j$  ( $j \in \{ 2, \dots, n \}$ ) tel que pour tout  $i=1, \dots, j-1$  on a  $\varepsilon \in \text{Premier}(Y_i)$ , alors ajouter les éléments de  $\text{Premier}(Y_j)$  sauf  $\varepsilon$  dans  $\text{Premier}(X)$
  - 3- si pour tout  $i=1, \dots, n$  on a  $\varepsilon \in \text{Premier}(Y_i)$ , alors ajouter  $\varepsilon$  dans  $\text{Premier}(X)$

2- si  $X \rightarrow \varepsilon$  est une règle de la grammaire, alors ajouter  $\varepsilon$  dans  $\text{Premier}(X)$

3- Si  $X$  est terminal, alors  $\text{Premier}(X) = \{X\}$

Recommencer jusqu'à ce qu'on n'ajoute rien de nouveau dans les ensembles  $\text{Premier}$

Il est clair que

$\text{Premier}(X) = \text{Premier}(X_1) \cup \text{Premier}(X_2) \cup \dots \cup \text{Premier}(X_n)$ , pour tout  $X \rightarrow X_1 | X_2 | \dots | X_n$

### Exemple 1 :

$$E \rightarrow TE'$$

$$E' \rightarrow +TE' \mid -TE' \mid \varepsilon$$

$$T \rightarrow FT'$$

$$T' \rightarrow *FT' \mid /FT' \mid \varepsilon$$

$$F \rightarrow (E) \mid nb$$

$$\text{Premier}(E) = \text{Premier}(T) = \{ (, nb \}$$

$$\text{Premier}(E') = \{ +, -, \varepsilon \}$$

$$\text{Premier}(T) = \text{Premier}(F) = \{ (, nb \}$$

$$\text{Premier}(T') = \{ *, /, \varepsilon \}$$

$$\text{Premier}(F) = \{ (, nb \}$$

### Exemple 2 :

$$S \rightarrow ABCe$$

$$A \rightarrow aA \mid \varepsilon$$

$$B \rightarrow bB \mid cB \mid \varepsilon$$

$$C \rightarrow de \mid da \mid dA$$

$$\text{Premier}(S) = \text{Premier}(A) \cup \text{Premier}(B) \cup \text{Premier}(C) = \{ a, b, c, d \}$$

$$\text{Premier}(A) = \{ a, \varepsilon \}$$

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

$$\text{Premier}(C) = \{ d \}$$

### ➤ Calcul de l'ensemble Suivant

Soit  $A$  un symbole non-terminal,  $\text{Suivant}(A)$  est l'ensemble de tous les symboles terminaux  $a$  qui peuvent apparaître immédiatement à droite de  $A$  dans une dérivation  $S \Rightarrow^* \alpha A a \beta$

$$\text{Suivant}(A) = \{a \in T \mid S \Rightarrow^* \alpha A a \beta, \text{ avec } \alpha, \beta \in (N \cup T)^*\}$$

Algorithme de construction des ensembles Suivant :

- 1- ajouter  $\$$  dans  $\text{Suivant}(S)$ , où  $S$  est l'axiome de la grammaire et  $\$$  est le marqueur de fin de chaîne d'entrée
- 2- pour chaque règle  $A \rightarrow \alpha B \beta$  où  $B$  est un non-terminal, alors ajouter les éléments de  $\text{Premier}(\beta)$  dans  $\text{Suivant}(B)$  sauf  $\varepsilon$

## Analyse Syntaxique

---

- 3- pour chaque règle  $A \rightarrow \alpha B$ , ajouter  $\text{Suivant}(A)$  à  $\text{Suivant}(B)$
- 4- pour chaque règle  $A \rightarrow \alpha B \beta$  avec  $\varepsilon \in \text{Premier}(\beta)$ , ajouter  $\text{Suivant}(A)$  à  $\text{Suivant}(B)$
- 5- pour chaque règle  $A \rightarrow Ba\beta$ , avec  $a$  un symbole terminal, alors ajouter  $a$  à  $\text{Suivant}(B)$

## Exemple 1

$$E \rightarrow TE'$$

$$E' \rightarrow +TE' \mid -TE' \mid \varepsilon$$

$$T \rightarrow FT'$$

$$T' \rightarrow *FT' \mid /FT' \mid \varepsilon$$

$$F \rightarrow (E) \mid nb$$

$$\text{Suivant}(E) = \{ \$ , ) \}$$

$$\text{Suivant}(E') = \text{Suivant}(E) = \{ \$ , ) \}$$

$$\text{Suivant}(T) = \text{Premier}(E') = \{ + , - , ) , \$ \}$$

$$\text{Suivant}(T') = \text{Suivant}(E') = \{ + , - , ) , \$ \}$$

$$\text{Suivant}(F) = \text{Premier}(T') = \{ * , / , + , - , \$ , ) \}$$



### Exemple 2 :

$$S \rightarrow aSb \mid cd \mid SAe$$

$$A \rightarrow aAdB \mid \varepsilon$$

$$B \rightarrow bb$$

	Premier	Suivant
S	{ a , c }	{ \$ , b , a , e }
A	{ a , $\varepsilon$ }	{ d , e }
B	{ b }	{ d , e }

### ➤ Construction de la table d'analyse

La table d'analyse est un tableau  $M$  à deux dimensions qui indique pour chaque non-terminal  $A$  et chaque symbole terminal  $a$  ou le symbole  $\$$  la règle de production à appliquer. La table d'analyse permet de choisir de façon explicite la production à appliquer lors de l'analyse.

Algorithme de construction :

- Pour chaque règle de production  $A \rightarrow \alpha$  faire
  - 1- pour chaque  $a \in \text{Premier}(\alpha)$  (et  $a \neq \varepsilon$ ), ajouter la production  $A \rightarrow \alpha$  dans la case  $M[A, a]$ ,

- 2- si  $\varepsilon \in \text{Premier}(\alpha)$ , alors pour chaque  $b \in \text{Suivant}(A)$  ajouter  $A \rightarrow \alpha$  dans la case  $M[A, b]$
- Chaque case  $M[A, a]$  vide est une erreur syntaxique

# Analyse Syntaxique

## Exemple

La table d'analyse de la grammaire de l'exemple 1

	<b>nb</b>	+	-	*	/	(	)	\$
<b>E</b>	$E \rightarrow TE'$					$E \rightarrow TE'$		
<b>E'</b>		$E' \rightarrow +TE'$	$E' \rightarrow -TE'$				$E' \rightarrow \varepsilon$	$E' \rightarrow \varepsilon$
<b>T</b>	$T \rightarrow FT'$					$T \rightarrow FT'$		
<b>T'</b>		$T' \rightarrow \varepsilon$	$T' \rightarrow \varepsilon$	$T' \rightarrow *FT'$	$T' \rightarrow /FT'$		$T' \rightarrow \varepsilon$	$T' \rightarrow \varepsilon$
<b>F</b>	$F \rightarrow nb$					$F \rightarrow (E)$		

### Grammaire LL(1)

On appelle grammaire **LL(1)** une grammaire pour la quelle chaque case de la table d'analyse contient au plus une règle. Dans le cas contraire, la grammaire n'est pas **LL(1)**.

**LL(k)** signifie :

- **L** : lecture de la chaine d'entrée de gauche à droite (scanning the input from **Left** to right)
- **L** : création de l'arbre syntaxique par les dérivations les plus gauches (producing a **Leftmost** derivation).
- **k** : on lit **k** symbole de prévision à la fois du texte à analyser (with **k** symbols of lookahead).

On ne s'intéresse qu'aux grammaires **LL(1)**. Quand  $k > 1$  les algorithmes de mise en œuvre deviennent difficiles.

### Définition formelle d'une grammaire LL(1)

Une grammaire  $G$  est dite **LL(1)** si pour toute règle de production de la forme

$A \rightarrow X_1 \mid X_2 \mid \dots \mid X_n$  les conditions suivantes sont vérifiées:

$\forall i, j :$

- 1-  $\text{Premier}(X_i) \cap \text{Premier}(X_j) = \emptyset$ , pour tout  $i \neq j$

- 2- S'il existe  $X_i$  annulable (c.-à-d.  $X_i \rightarrow^* \varepsilon$ ), alors  
 $\text{Premier}(X_j) \cap \text{Suivant}(A) = \emptyset$ , pour tout  $i \neq j$
- 3- Il existe au plus une règle  $X_i \rightarrow^* \varepsilon$

**Ces conditions entraînent l'unicité de l'existence d'une seule règle dans une case de la table d'analyse prédictive.**

Avantages : cette technique est simple, très efficace et facile à implémenter.

### Condition nécessaire pour qu'une grammaire soit LL(1)

Pour qu'une grammaire soit LL(1), il faut qu'elle soit *non ambiguë*, *non récursive à gauche* et *factorisée à gauche*.

#### a- Récursivité à gauche

Une grammaire est récursive à gauche s'il existe un non-terminal  $A$  et une dérivation de la forme  $A \rightarrow^* A\alpha$  où  $\alpha$  est une chaîne quelconque.

Cas particulier, on dit qu'on a une récursivité à gauche immédiate si la grammaire possède une production de la forme  $A \rightarrow A\alpha$



- *Elimination de la récursivité à gauche immédiate*

Remplacer toute règle de la forme  $A \rightarrow A\alpha \mid \beta$  par

$$A \rightarrow \beta A'$$

$$A' \rightarrow \alpha A' \mid \varepsilon$$

La grammaire obtenue après élimination de la récursivité à gauche génère le même langage que la grammaire initiale

### Exemple 1

$$S \rightarrow ScA \mid B$$

$$A \rightarrow Aa \mid \varepsilon$$

$$B \rightarrow Bb \mid d \mid e$$

Cette grammaire contient plusieurs récursivités à gauche immédiates.

Après élimination de la récursivité immédiate

$$S \rightarrow BS'$$

$$S' \rightarrow cAS' \mid \varepsilon$$

$$A \rightarrow A'$$

$$A' \rightarrow aA' \mid \varepsilon$$

$$B \rightarrow dB' \mid eB'$$

$$B' \rightarrow bB' \mid \varepsilon$$

De façon générale **la récursivité à gauche immédiate** peut être éliminée comme suite :

- 1- Grouper toutes les  $A$ -productions (les règles de productions ayant la même partie droite  $A$ ):  $A \rightarrow A\alpha_1 \mid A\alpha_2 \mid \dots \mid A\alpha_m \mid \beta_1 \mid \beta_2 \mid \dots \mid \beta_n$

où aucun  $\beta_i$  ne commence par  $A$

2- Remplacer les  $A$ -productions par :

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

$$A' \rightarrow \alpha_1 A' \mid \alpha_2 A' \mid \dots \mid \alpha_m A' \mid \varepsilon$$

- *Elimination de la récursivité à gauche*

### Exemple 2

$$S \rightarrow Aa \mid b$$

$$A \rightarrow Ac \mid Sd \mid \varepsilon$$

Le non terminal  $S$  est récursive à gauche car  $S \rightarrow Aa \rightarrow Sda$  mais pas immédiatement récursive à gauche.

## Analyse Syntaxique

---

L'algorithme ci-dessous, systématiquement élimine les récursivité à gauche d'une grammaire ne contenant aucun cycle (c.-à-d. des dérivations de la forme  $A \Rightarrow^* A$ ) ou des  $\varepsilon$ -règles (c.-à-d. règle de la forme  $A \rightarrow \varepsilon$ ).

### Début

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

$A_j \rightarrow \beta_1 \mid \beta_2 \mid \dots \mid \beta_p$  par  $A_i \rightarrow \beta_1 \alpha \mid \beta_2 \alpha \mid \dots \mid \beta_p \alpha$

**fin pour**

    éliminer les récursivités à gauche immédiates des productions  $A_i$

**fin pour**

### Fin

### Exemple 2

On ordonne  $S, A$

$i = 1$  pas de récursivité immédiate dans  $S \rightarrow Aa \mid b$

$i = 2$  et  $j=1$  on obtient  $A \rightarrow Ac \mid Aad \mid bd \mid \varepsilon$

on élimine la récursivité immédiate :

$$A \rightarrow bdA' \mid A'$$

$$A' \rightarrow cA' \mid adA' \mid \varepsilon$$

On obtient la grammaire

$$S \rightarrow Aa \mid b$$

$$A \rightarrow bdA' \mid A'$$

$$A' \rightarrow cA' \mid adA' \mid \varepsilon$$

### **b- Factorisation à gauche**

Nous cherchons à écrire des analyseurs prédictifs. Cela veut dire qu'à tout moment le choix entre productions qui ont le même membre gauche doit pouvoir se faire, sans risque d'erreur, en comparant le symbole courant de la chaîne à analyser avec les symboles susceptibles de commencer les dérivations des membres droits des productions en compétition.

Une grammaire contenant des productions comme  $A \rightarrow \alpha\beta_1 \mid \alpha\beta_2$  viole ce principe.

La factorisation à gauche corrige ce problème comme suit :

Pour chaque non-terminal  $A$  :

- trouver le préfixe le plus long  $\alpha$  ( $\alpha \neq \varepsilon$ ) commun à deux ou plusieurs de ses alternatives.
- Remplacer  $A \rightarrow \alpha\beta_1 | \alpha\beta_2 | \dots | \alpha\beta_n | \lambda_1 | \dots | \lambda_p$  (où  $\lambda_i$  ne commencent pas par  $\alpha$ ) par

$$A \rightarrow \alpha A' | \lambda_1 | \dots | \lambda_p$$

$$A' \rightarrow \beta_1 | \beta_2 | \dots | \beta_n$$

Répéter pour tout non-terminal ayant un préfixe commun à ses alternatives.

### Exemple

$stmt \rightarrow \text{if } expr \text{ then } stmt \text{ else } stmt \mid \text{if } expr \text{ then } stmt$

$stmt \rightarrow \underbrace{\text{if } expr \text{ then } stmt}_{\alpha} \underbrace{\text{elses } stmt}_{\beta} \mid \underbrace{\text{if } expr \text{ then } stmt}_{\alpha}$

$stmt \rightarrow \text{if } expr \text{ then } stmt A'$

$A' \rightarrow \text{else } stmt \mid \varepsilon$



## Analyse Syntaxique

---

Cette grammaire n'est pas LL(1)

	Premier	Suivant
Stmt	{if}	{\$,else}
A'	{else, $\varepsilon$ }	{\$,else}

	If	else	\$
Stmt	$stmt \rightarrow \mathbf{if\ expr\ then\ stmt\ A'}$		$A' \rightarrow \varepsilon$
A'		$A' \rightarrow \mathbf{else\ stmt}$ $A' \rightarrow \varepsilon$	

## Analyse Syntaxique

---

Avec la grammaire factorisée, pour la chaîne d'entrée : **if  $a$  then if  $b$  then  $c$  else  $d$**  il y a deux arbres de dérivation. La grammaire factorisée est **ambiguë**.

### *Règle pour lever l'ambiguïté :*

Dans la dérivation d'un non-terminal, une " $\varepsilon$ -règle ne peut être choisie que lorsque aucune autre production n'est applicable.

Dans cet exemple, cela donne : si la chaîne d'entrée commence par **else** alors on doit nécessairement choisir la première règle.

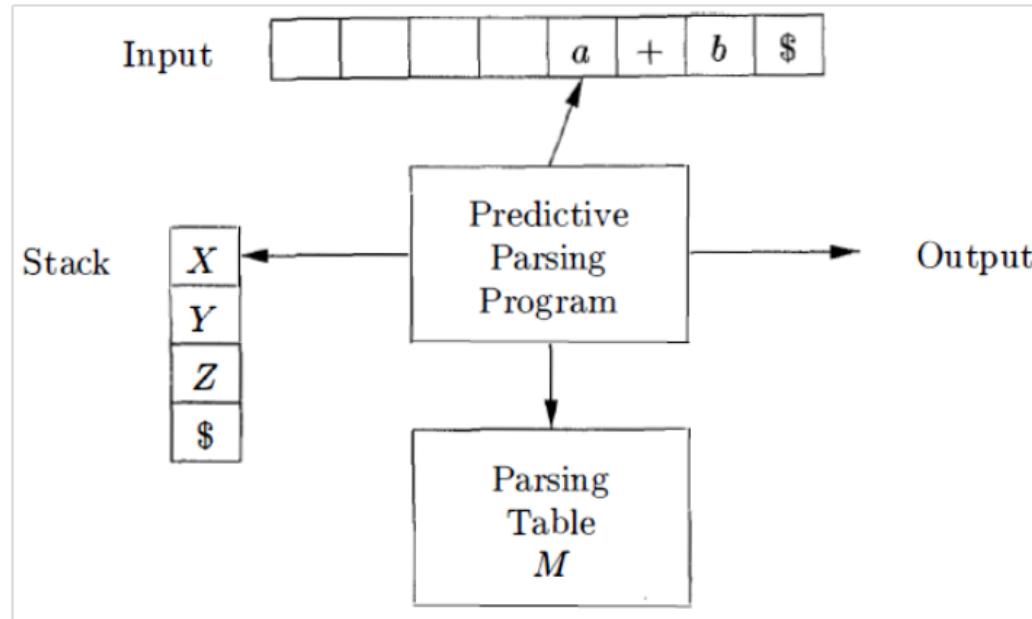
### Algorithme d'analyseur LL(1) dirigé par la table d'analyse

Comme le montre la figure, un analyseur syntaxique prédictif dirigé par une table d'analyse est composé des éléments suivants :

- Un **tampon d'entrée (Input)**: contient la chaîne à analyser, suivie du symbole \$.
- Une **pile (Stack)**: contient une séquence de symboles (terminaux ou non-terminaux), avec un symbole \$ qui marque le **fond de la pile**. Au départ, la pile contient l'**axiome** de la grammaire au sommet et le symbole \$ au **fond**.
- Une **table d'analyse LL(1) (Parsing Table)**.
- Un **flot de sortie (Output) (l'arbre syntaxique)**

# Analyse Syntaxique

---



## Algorithme

**Entrée :** une chaîne  $w$  terminée par  $\$$ ,  $ps$  est le pointeur sur le premier symbole de  $w$  et la table d'analyse  $M$  de la grammaire

**Sorite :** si  $w$  est dans  $L(G)$ , arbre d'analyse de  $w$ , sinon erreur

# Analyse Syntaxique

---

## Répéter

$X$  : reçoit le symbole en sommet de pile ;  $a$  : est le symbole pointé par  $ps$

**Si** ( $X$  est un non-terminal) **alors**

**Si** ( $M[X,a] = X \rightarrow Y_1Y_2...Y_n$ ) **alors**

        dépiler  $X$  ; empiler  $Y_n$  suivie  $Y_{n-1}$  jusqu'à  $Y_1$  ; emmètre la règle  $X \rightarrow Y_1Y_2...Y_n$  en sortie

**Sinon** (*case vide dans la table d'analyse*)

**ERREUR**

**finsi**

**Sinon**

**Si**  $X = \$$  **alors**

**Si**  $a = \$$  **alors** **ACCEPTER**

**Sion** **ERREUR**

**finsi**

**Sinon**

**Si**  $X = a$  **alors**      dépiler  $X$  ; avancer  $ps$

**Sinon** **ERREUR**

**finsi**

**finsi**

**finsi**

**Jusqu'à ACCEPTER ou ERREUR**

## Exemple

Soit la grammaire  $G$

$$E \rightarrow E+T$$

$$T \rightarrow T*F \mid F$$

$$F \rightarrow (E) \mid id$$

La grammaire  $G'$  équivalent à  $G$  après élimination des récursivités à gauche

$$E \rightarrow TE'$$

$$E' \rightarrow +TE' \mid \varepsilon$$

$$T \rightarrow FT'$$

$$T' \rightarrow *FT' \mid \varepsilon$$

$$F \rightarrow (E) \mid nb$$

- *Calcul des ensembles Premier et Suivant :*

	<b>Premier</b>	<b>Suivant</b>
<b><i>E</i></b>	$\{ (, \text{nb} \}$	$\{ ), \$ \}$
<b><i>E'</i></b>	$\{ + \}$	$\{ ), \$ \}$
<b><i>T</i></b>	$\{ (, \text{nb} \}$	$\{ +, ), \$ \}$
<b><i>T'</i></b>	$\{ * \}$	$\{ +, ), \$ \}$
<b><i>F</i></b>	$\{ (, \text{nb} \}$	$\{ +, *, ), \$ \}$

## Analyse Syntaxique

---

- *Construction de la table d'analyse*

	+	*	(	)	nb	\$
E			$E \rightarrow TE'$		$E \rightarrow TE'$	
E'	$E' \rightarrow +TE'$			$E' \rightarrow \varepsilon$		$E' \rightarrow \varepsilon$
T			$T \rightarrow FT'$		$T \rightarrow FT'$	
T'	$T' \rightarrow \varepsilon$	$T' \rightarrow *FT'$		$T' \rightarrow \varepsilon$		$T' \rightarrow \varepsilon$
F			$F \rightarrow (E)$		$F \rightarrow nb$	



## Analyse Syntaxique

Analysons la chaîne  $w = \text{nb}+\text{nb}*\text{nb}$

Pile	Entrée	Sortie
\$E	<b>nb+nb*nb\$</b>	E->TE'
\$E'T	<b>nb+nb*nb\$</b>	T->FT'
\$E'T'F	<b>nb+nb*nb\$</b>	F->nb
\$E'T'nb	<b>nb+nb*nb\$</b>	
\$E'T'	<b>+nb*nb\$</b>	T'-> $\epsilon$
\$E'	<b>+nb*nb\$</b>	E'->+TE'
\$E'T+	<b>+nb*nb\$</b>	
\$E'T	<b>nb*nb\$</b>	T->FT'
\$E'T'F	<b>nb*nb\$</b>	F->nb
\$E'T'nb	<b>nb*nb\$</b>	
\$E'T'	<b>*nb\$</b>	T'->*FT'

## Analyse Syntaxique

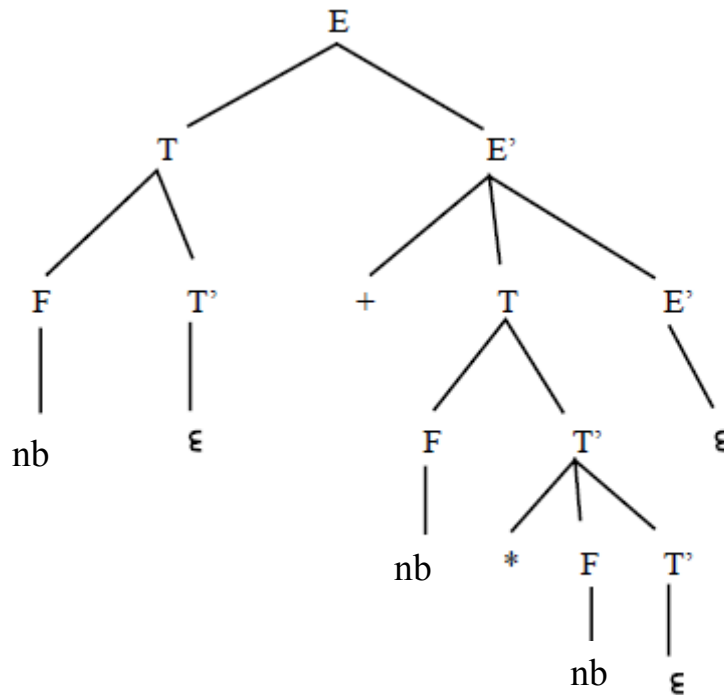
---

\$E'T'F*	*nb\$	
\$E'T'F	nb\$	F->nb
\$E'T'nb	nb\$	
\$E'T'	\$	T'-> $\varepsilon$
\$E'	\$	E'-> $\varepsilon$
\$	\$	ACCEPTE

## Analyse Syntaxique

---

L'arbre syntaxique de la chaîne  $w = \text{nb} + \text{nb} * \text{nb}$



# Analyse Syntaxique

---

## Analyse Ascendante