#### Analyse syntaxique LR

Alexis Nasr Carlos Ramisch Manon Scholivet Franck Dary

Compilation – L3 Informatique Département Informatique et Interactions Aix Marseille Université

#### Grammaires hors-contexte

Une grammaire hors-contexte est un 4-uplet  $\langle N, \Sigma, P, S \rangle$  où :

- N est un ensemble de symboles non terminaux, appelé l'alphabet non terminal.
- $oldsymbol{\Sigma}$  est un ensemble de symboles terminaux, appelé l'alphabet terminal, tel que N et  $\Sigma$  soient disjoints.
- *P* est un sous ensemble fini de :

$$N \times (N \cup \Sigma)^*$$

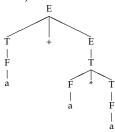
un élément  $(\alpha, \beta)$  de P, que l'on note  $\alpha \to \beta$  est appelé une règle de production ou règle de réécriture.  $\alpha$  est appelé partie gauche de la règle  $\beta$  est appelé partie droite de la règle

■ *S* est un élément de *N* appelé l'axiome de la grammaire.

#### Analyse syntaxique

Etant donné  $m \in \Sigma^*$  et  $G = \langle \Sigma, N, P, A \rangle$ , analyser m consiste à trouver pour m son (et éventuellement ses) arbre de dérivation.

$$\begin{array}{ccc} E & \rightarrow & T+E \mid T \\ T & \rightarrow & F*T \mid F \\ F & \rightarrow & (E) \mid a \end{array}$$



#### Sens d'analyse

#### Analyse descendante

L'arbre de dérivation est construit depuis la racine vers les feuilles

Séquence de dérivations gauches à partir de l'axiome  $E \Rightarrow T + E \Rightarrow F + E \Rightarrow a + E \Rightarrow a + T \Rightarrow a + F * T \Rightarrow a + a * T \Rightarrow a + a * F \Rightarrow a + a * a$ 

#### Analyse ascendante

L'arbre de dérivation est construit des feuilles vers la racine Séquence de dérivation telle que la séquence inverse soit une dérivation droite de *m*.

$$a + a * a \Leftarrow F + a * a \Leftarrow T + a * a \Leftarrow T + F * a \Leftarrow T + F * F \Leftarrow T + F * T \Leftarrow T + E \Leftarrow E$$

#### Utilisation d'une pile

- Pour l'analyse descendante, comme pour l'analyse ascendante on utilise une pile
- Cette dernière permet de stocker les résultats intermédiaires du processus d'analyse

#### Analyse Descendante

- Empiler l'axiome S
- **2** Remplacer *S* par la partie droite d'une règle de la forme  $S \to \alpha$  de telle sorte que le premier symbole x de  $\alpha$  se trouve en sommet de pile.
  - Si x est un terminal alors on le compare avec le caractère se trouvant sous la tête de lecture. S'ils sont égaux alors on dépile.
  - Si x est un non terminal alors on le remplace par la partie droite d'une règle de P de la forme  $x \to \beta$ .

Reconnaissance du mot :

$$a + a * a$$

avec la grammaire:

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$

Ε

Reconnaissance du mot :

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



Reconnaissance du mot :

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



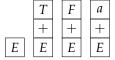
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



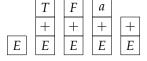
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



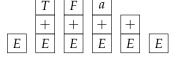
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



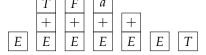
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



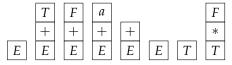
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



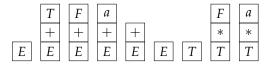
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



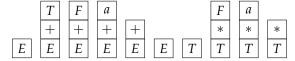
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



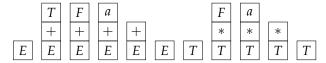
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



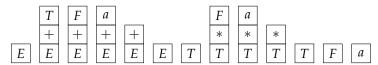
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



#### Non déterminisme

- Lorsqu'un non terminal X doit être remplacé au sommet de la pile, il peut l'être par la partie droite d'une règle de la forme  $X \to \beta$ .
- Plusieurs règles de cette forme peuvent exister dans la grammaire.
- L'algorithme est non déterministe.

# Analyse Ascendante ou analyse par décalage-réduction

- On empile les terminaux au fur et à mesure qu'ils sont lus.
- L'opération qui consiste à empiler un terminal est appelée décalage.
- lorsque les k symboles au sommet de la pile constituent la partie droite d'une production, ils peuvent être dépilés et remplacés par la partie gauche de la production.
- Cette opération s'appelle réduction.
- La séquence de symboles dépilés s'appelle un manche
- Lorsque la pile ne comporte que l'axiome et que tous les symboles de la chaîne d'entrée ont été lus, l'analyse a réussi.

Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$

Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$

Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$

Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$

Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$

Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$

Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$

$$a$$
  $F$   $T$   $T$ 

Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$

Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$

Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$

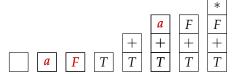
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



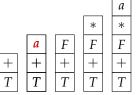
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



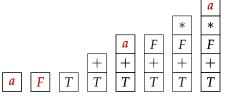
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



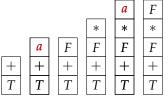
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



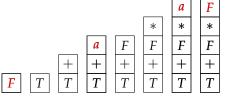
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



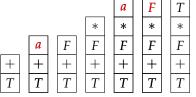
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



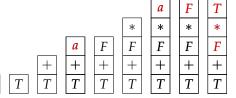
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



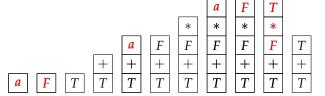
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



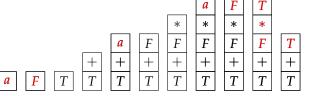
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



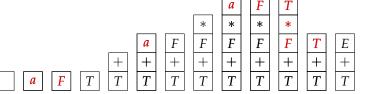
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



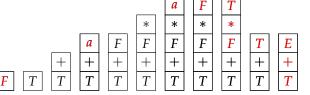
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



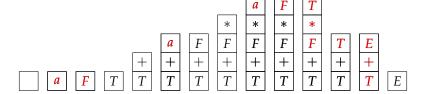
Reconnaissance du mot:

$$a + a * a$$

$$E \to T + E \mid T$$

$$T \to F * T \mid F$$

$$F \to (E) \mid a$$



#### Non déterminisme

- Si les symboles au sommet de la pile constituent la partie droite de deux productions distinctes alors chacune de ces deux règles peut être utilisée pour effectuer une réduction.
- Lorsque les symboles au sommet de la pile constituent la partie droite d'une ou plusieurs productions, on peut réduire tout de suite ou bien continuer à décaler, afin de permettre ultérieurement une réduction plus juste.

# Automate à pile

Un automate à pile est un 6-uplet  $\langle Q, \Sigma, \Gamma, \delta, q_0, F \rangle$ 

- Q est l'ensemble des états
- Σ est l'alphabet d'entrée
- $\blacksquare$   $\Gamma$  est l'alphabet de symboles de pile
- $\bullet$  st la fonction de transition :

$$\delta: Q \times (\Sigma \cup \{\varepsilon\}) \times (\Gamma \cup \{\varepsilon\}) \to \wp(Q \times \Gamma^*)$$

- $q_0 \in Q$  est l'état initial
- $F \subseteq Q$  est l'ensemble des états d'acceptation

## Grammaires hors-contexte ⇔ Automate à pile

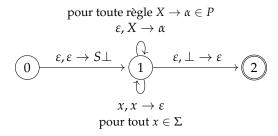
Un langage est hors-contexte si et seulement si il existe un automate à pile qui le reconnaît.

- Si un langage est hors-contexte alors il existe un automate à pile qui le reconnaît.
- Si un langage est reconnu par un automate à pile alors il est hors-contexte.

## Grammaires hors-contexte $\Rightarrow$ Automate gauche

- Soit  $G = \langle N, \Sigma, P, S \rangle$  une grammaire hors-contexte, on construit un automate à pile A qui accepte un mot m s'il existe une dérivation pour m dans G ( $S \stackrel{+}{\Rightarrow} m$ ).
- A est conçu de telle sorte à déterminer une dérivation gauche conduisant de S à m.
- Idée clef : écrire dans la pile de *A* les proto-mots qui constituent la dérivation recherchée.

# Automate gauche correspondant à la grammaire $G = \langle N, \Sigma, P, S \rangle$



#### Construction de l'automate gauche

Automate à pile A correspondant à la grammaire  $G = \langle N, \Sigma, P, S \rangle$ :

$$A = \langle \{0,1,2\}, \Sigma, N \cup \Sigma \cup \{\bot\}, \delta, 0, \{2\} \rangle$$

La fonction de transition  $\delta$  est définie de la façon suivante :

- $\delta(0, \varepsilon, \varepsilon) = \{(1, S\bot)\}$  On empile l'axiome.
- $\delta(1, \varepsilon, X) = \{(1, \alpha) \text{ pour tout } X \to \alpha \in P\}$ Si un symbole non terminal X occupe le sommet de la pile, on le remplace par la partie droite  $\alpha$  d'une règle  $X \to \alpha$ .
- $\delta(1, a, a) = \{(1, \varepsilon) \mid \text{avec } a \in \Sigma\}$ Si le même symbole terminal occupe le sommet de la pile et la case courante de la bande d'entrée, on dépile.
- $\delta(1, \varepsilon, \bot) = \{(2, \varepsilon)\}$ Si le mot en entrée a été reconnu et que la pile ne contient que le symbole de fond de pile, on passe à l'état d'acceptation.

## Construction — Exemple

Grammaire:

$$\langle \{E,T,F\}, \{a,+,*(,)\}, P,E \rangle$$

avec:

$$P = \left\{ \begin{array}{ccc} E & \rightarrow & T + E \mid T, \\ T & \rightarrow & F * T \mid F, \\ F & \rightarrow & (E) \mid a \end{array} \right\}$$

Automate:

$$A_1 = \langle \{0,1,2\}, \{a,+,*,(,)\}, \{a,+,*,(,),E,T,F,\bot\}, \delta, 0, \{2\} \rangle$$

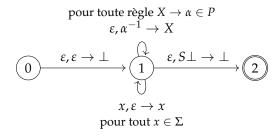
avec:

$$\begin{array}{llll} \delta(0,\varepsilon,\varepsilon) & = & \{(1,E\bot)\} & \delta(1,+,+) & = & \{(1,\varepsilon)\} \\ \delta(1,\varepsilon,E) & = & \{(1,T+E),(1,T)\} & \delta(1,*,*) & = & \{(1,\varepsilon)\} \\ \delta(1,\varepsilon,T) & = & \{(1,F*T),(1,F)\} & \delta(1,(,()) & = & \{(1,\varepsilon)\} \\ \delta(1,\varepsilon,F) & = & \{(1,(E)),(1,a)\} & \delta(1,),) & = & \{(1,\varepsilon)\} \\ \delta(1,\varepsilon,\bot) & = & \{(2,\varepsilon)\} & \delta(1,a,a) & = & \{(1,\varepsilon)\} \end{array}$$

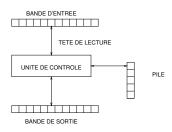
#### Grammaires hors-contexte $\Rightarrow$ Automate droit

- Soit  $G = \langle N, \Sigma, P, S \rangle$  une grammaire hors-contexte, on construit un automate à pile A qui accepte un mot m s'il existe une dérivation pour m dans G ( $S \stackrel{+}{\Rightarrow} m$ ).
- A est conçu de telle sorte à déterminer une réduction droite conduisant de m à S.

# Automate droit correspondant à la grammaire $G = \langle N, \Sigma, P, S \rangle$



# Transducteur à pile



- Un transducteur à pile est un automate à pile qui émet, à chaque déplacement, un suite finie de symboles de sortie.
- Une configuration d'un transducteur à pile est un quadruplet  $(q, w, \alpha, y)$  où y est une séquence de symboles de sortie.

#### Transducteur à pile — définition

Un transducteur à pile est un 8-uplet  $\langle Q, \Sigma, \Gamma, \Delta, \delta, q_0, F \rangle$ 

- Q est l'ensemble des états
- Σ est l'alphabet d'entrée
- **Γ** est l'alphabet de symboles de pile
- lacksquare  $\Delta$  est l'alphabet de sortie
- $\bullet$   $\delta$  est la fonction de transition

$$\delta: Q \times (\Sigma \cup \{\varepsilon\}) \times (\Gamma \cup \{\varepsilon\}) \to \wp(Q \times \Gamma^* \times \Delta^*)$$

- $q_0 \in Q$  est l'état initial
- lacksquare  $F\subseteq Q$  est l'ensemble des états d'acceptation

# Analyseur gauche

1: 
$$E \rightarrow T + E$$
 2:  $E \rightarrow T$   
3:  $T \rightarrow F * T$  4:  $T \rightarrow F$   
5:  $F \rightarrow (E)$  6:  $F \rightarrow a$ 

■ Dérivation gauche de a + a \* a:

$$E \Rightarrow T + E \Rightarrow F + E \Rightarrow a + E \Rightarrow a + T \stackrel{*}{\Rightarrow} a + a * a$$

■ Elle correspond à l'application des règles suivantes : 1,4,6,2,3,6,4,6

# Analyseur gauche

Soit une grammaire hors contexte G dont les règles ont été numérotées de 1 à p. On appelle un analyseur gauche de G, un transducteur à pile non déterministe  $T_G^g$  qui produit pour un mot  $m \in L(G)$ , une dérivation gauche de m. Performances :

- Espace :  $\mathcal{O}(|m|)$
- Temps :  $\mathcal{O}(c^{|m|})$

$$\begin{array}{cccc}
\varepsilon, E & \to & T+E, 1 \\
\varepsilon, E & \to & T, 2 \\
\varepsilon, T & \to & F*T, 3 \\
\varepsilon, T & \to & F*T, 4 \\
\varepsilon, F & \to & (E), 5 \\
\varepsilon, F & \to & a, 6
\end{array}$$

$$\underbrace{\begin{array}{c}
\varepsilon, \varepsilon \to E \bot \\
0
\end{array}}_{x, x \to \varepsilon} \underbrace{\begin{array}{c}
\varepsilon, \bot \to \varepsilon \\
0
\end{array}}_{x, x \to \varepsilon}$$
pour tout  $x \in \{a, +, *, (, )\}$ 

$$(0, a+a*a, \varepsilon, \varepsilon)$$

$$(0, a+a*a, \varepsilon, \varepsilon)$$

$$(1, a+a*a, E\bot, \varepsilon)$$

$$(1, a+a*a, T+E\bot, 1)$$

$$\varepsilon, E \to T+E, 1$$

$$\varepsilon, E \to T, 2$$

$$\varepsilon, T \to F*T, 3$$

$$\varepsilon, T \to F*4$$

$$\varepsilon, F \to (E), 5$$

$$\varepsilon, F \to a, 6$$

$$0 \xrightarrow{\varepsilon, \varepsilon \to E\bot} 1 \xrightarrow{\varepsilon, \bot \to \varepsilon} 2$$

$$x, x \to \varepsilon$$

$$\text{pour tout } x \in \{a, +, *, (,)\}$$

$$(0, a+a*a, & \varepsilon, \varepsilon)$$

$$(1, a+a*a, & E\bot, \varepsilon)$$

$$(1, a+a*a, & T+E\bot, 1$$

$$\varepsilon, E \to T+E, 1$$

$$\varepsilon, E \to T, 2$$

$$\varepsilon, T \to F*T, 3$$

$$\varepsilon, T \to F, 4$$

$$\varepsilon, F \to (E), 5$$

$$\varepsilon, F \to a, 6$$

$$(1, a+a*a, T+E\bot, 1)$$

$$(1, a+a*a, F+E\bot, 14)$$

$$(1, a+a*a, F+E\bot, 14)$$

$$(1, a+a*a, F+E\bot, 14)$$

$$\begin{array}{cccc}
\varepsilon,E & \to & T+E,1 \\
\varepsilon,E & \to & T,2 \\
\varepsilon,T & \to & F*T,3 \\
\varepsilon,T & \to & F,4 \\
\varepsilon,F & \to & (E),5 \\
\varepsilon,F & \to & a,6
\end{array}$$

$$\underbrace{0}_{x,\varepsilon \to E\perp} \underbrace{1}_{x,x \to \varepsilon} \underbrace{1}_{x,x \to \varepsilon} \underbrace{2}_{x,x,\varepsilon \to \varepsilon}$$
pour tout  $x \in \{a,+,*,(,)\}$ 

$$(0, a + a * a, & \varepsilon, \varepsilon)$$
  
 $(1, a + a * a, & E \bot, \varepsilon)$   
 $(1, a + a * a, & T + E \bot, 1)$   
 $(1, a + a * a, & F + E \bot, 14)$   
 $(1, a + a * a, & a + E \bot, 146)$ 

$$(0, a+a*a, \varepsilon, \varepsilon)$$
  
 $(1, a+a*a, E\bot, \varepsilon)$   
 $(1, a+a*a, T+E\bot, 1)$   
 $(1, a+a*a, F+E\bot, 14)$   
 $(1, a+a*a, a+E\bot, 146)$   
 $(1, +a*a, +E\bot, 146)$ 

$$\begin{array}{cccc}
\varepsilon, E & \to & T+E, 1 \\
\varepsilon, E & \to & T, 2 \\
\varepsilon, T & \to & F*T, 3 \\
\varepsilon, T & \to & F*, 4 \\
\varepsilon, F & \to & (E), 5 \\
\varepsilon, F & \to & a, 6
\end{array}$$

$$\underbrace{0} \xrightarrow{\varepsilon, \varepsilon \to E \perp} \underbrace{1} \xrightarrow{\varepsilon, \perp \to \varepsilon} \underbrace{2} \xrightarrow{x, x \to \varepsilon}$$
pour tout  $x \in \{a, +, *, (, )\}$ 

$$\begin{array}{ccc}
\varepsilon, E & \to & T+E, 1 \\
\varepsilon, E & \to & T, 2 \\
\varepsilon, T & \to & F*T, 3 \\
\varepsilon, T & \to & F*4 \\
\varepsilon, F & \to & (E), 5 \\
\varepsilon, F & \to & a, 6
\end{array}$$

$$\begin{array}{c}
\varepsilon, \varepsilon \to E \bot \\
\downarrow \\
0
\end{array}$$

$$\begin{array}{c}
\varepsilon, \varepsilon \to E \bot \\
\downarrow \\
0
\end{array}$$

$$\begin{array}{c}
\varepsilon, x \to \varepsilon \\
\downarrow \\
0
\end{array}$$

$$\begin{array}{c}
\varepsilon, x \to \varepsilon \\
\downarrow \\
0
\end{array}$$

$$\begin{array}{c}
\varepsilon, x \to \varepsilon \\
\downarrow \\
0
\end{array}$$

$$\begin{array}{c}
\varepsilon, x \to \varepsilon \\
\downarrow \\
0
\end{array}$$

$$\begin{array}{c}
\varepsilon, x \to \varepsilon \\
\downarrow \\
0
\end{array}$$
pour tout  $x \in \{a, +, *_{\varepsilon}(x)\}$ 

$$\begin{array}{ccc}
\varepsilon, E & \to & T+E, 1 \\
\varepsilon, E & \to & T, 2 \\
\varepsilon, T & \to & F*T, 3 \\
\varepsilon, T & \to & F, 4 \\
\varepsilon, F & \to & (E), 5 \\
\varepsilon, F & \to & a, 6
\end{array}$$

$$\begin{array}{c}
\varepsilon, \varepsilon \to E \bot & \\
\downarrow & \downarrow \\
0 & & \downarrow \\
\end{array}$$

$$\begin{array}{c}
\varepsilon, \varepsilon \to E \bot & \\
\downarrow & \downarrow \\
\end{array}$$

$$\begin{array}{c}
\varepsilon, \bot \to \varepsilon \\
\downarrow \\
\end{array}$$

$$\begin{array}{c}
\varepsilon, \bot \to \varepsilon \\
\downarrow \\
\end{array}$$

$$\begin{array}{c}
\varepsilon, \bot \to \varepsilon \\
\downarrow \\
\end{array}$$
pour tout  $x \in \{a, +, *, (, )\}$ 

$$\begin{array}{cccc}
\varepsilon,E & \to & T+E,1 \\
\varepsilon,E & \to & T,2 \\
\varepsilon,T & \to & F*T,3 \\
\varepsilon,T & \to & F,4 \\
\varepsilon,F & \to & (E),5 \\
\varepsilon,F & \to & a,6
\end{array}$$

$$\begin{array}{c}
\varepsilon,\varepsilon \to E \perp & \\
0 & & \\
\end{array}$$

$$\begin{array}{c}
\varepsilon,\varepsilon \to E \perp & \\
\varepsilon,x \to \varepsilon & \\
\end{array}$$

$$\begin{array}{c}
\varepsilon, \bot \to \varepsilon \\
\end{array}$$

pour tout  $x \in \{a, +, *, (,)\}$ 

$$\begin{array}{cccc}
\varepsilon, E & \to & T+E, 1 \\
\varepsilon, E & \to & T, 2 \\
\varepsilon, T & \to & F*T, 3 \\
\varepsilon, T & \to & F, 4 \\
\varepsilon, F & \to & (E), 5 \\
\varepsilon, F & \to & a, 6
\end{array}$$

$$\underbrace{c, F \to (E), 5}_{\varepsilon, F} \xrightarrow{\varepsilon} \underbrace{d}_{\varepsilon, F} \xrightarrow{\varepsilon} \underbrace$$

$$\begin{array}{cccc}
\varepsilon, E & \to & T+E, 1 \\
\varepsilon, E & \to & T, 2 \\
\varepsilon, T & \to & F*T, 3 \\
\varepsilon, T & \to & F, 4 \\
\varepsilon, F & \to & (E), 5 \\
\varepsilon, F & \to & a, 6
\end{array}$$

$$\underbrace{c}_{s, \varepsilon} \rightarrow E \perp \qquad \underbrace{1}_{s, \perp} \rightarrow \varepsilon \qquad \underbrace{1}_{s, \perp} \rightarrow \varepsilon \qquad \underbrace{2}_{s, \perp} \rightarrow \varepsilon \qquad \underbrace{2}_{s$$

$$\begin{array}{cccc}
\varepsilon, E & \to & T+E, 1 \\
\varepsilon, E & \to & T, 2 \\
\varepsilon, T & \to & F*T, 3 \\
\varepsilon, T & \to & F, 4 \\
\varepsilon, F & \to & (E), 5 \\
\varepsilon, F & \to & a, 6
\end{array}$$

$$\underbrace{c}_{x, F} \xrightarrow{\varepsilon} \xrightarrow{\varepsilon} \underbrace{c}_{x, F} \xrightarrow{\varepsilon} \underbrace{c}_{x,$$

```
a + a * a,
(1, a+a*a, E\perp, \varepsilon)
(1, a+a*a, T+E\perp,
(1, a+a*a, F+E\perp, 14)
(1, a+a*a, a+E\perp, 146)
(1, +a*a, +E\perp, 146)
               E\perp, 146)
   a*a,
(1, a*a,
              T_{\perp}, 1462)
(1, a*a,
              F*T\perp,
                        14623)
              a * T \perp, 146236)
    a * a,
                 *T_{\perp}, 146236)
    *a,
                T\bot,
                        1462364)
                 F_{\perp}, 14623646)
                   a\perp, 14623646)
                   \perp, 14623646)
```

## Analyseur gauche: Exemple

```
a + a * a,
(1, a+a*a, E\perp, \varepsilon)
(1, a+a*a, T+E\perp,
(1, a+a*a, F+E_{\perp}, 14)
(1, a+a*a, a+E\perp, 146)
(1, +a*a, +E\perp, 146)
                        146)
   a*a,
(1, a*a,
              T_{\perp}, 1462)
(1, a*a,
               F*T\perp,
                         14623)
               a * T \perp, 146236)
    a * a,
                  *T_{\perp}, 146236)
    *a,
                          1462364)
                  F_{\perp}, 14623646)
                    a\perp, 14623646)
                    \perp, 14623646)
                     \varepsilon, 14623646)
```

## Analyseur droit

Réduction droite de a + a \* a:  $a + a * a \rightarrow F + a * a \rightarrow T + a * a \rightarrow T + F * a \rightarrow T + F * F \rightarrow T + F * T \rightarrow T + T \rightarrow T + E \rightarrow E$ 

■ Elle correspond à l'application des règles suivantes : 6,4,6,6,4,3,2,1

## Analyseur droit

Soit une grammaire hors contexte G dont les règles ont été numérotées de 1 à p. On appelle un analyseur droit de G, un transducteur à pile non déterministe  $T_G^d$  qui produit pour un mot  $m \in L(G)$ , une dérivation droite de m à l'envers. Performances :

- Espace :  $\mathcal{O}(|m|)$
- Temps :  $\mathcal{O}(c^{|m|})$

$$\begin{array}{cccc}
\varepsilon,E+T & \to & E,1 \\
\varepsilon,T & \to & E,2 \\
\varepsilon,T*F & \to & T,3 \\
\varepsilon,F & \to & T,4 \\
\varepsilon,(E) & \to & F,5 \\
\varepsilon,a & \to & F,6
\end{array}$$

$$\underbrace{\varepsilon,\varepsilon \to \bot}_{\varepsilon,a} & \underbrace{\downarrow}_{\varepsilon,E\bot \to \varepsilon}_{\varepsilon,a} & \underbrace{\downarrow}_{\varepsilon,E\bot \to \varepsilon}_{\varepsilon,E\bot \to \varepsilon} & \underbrace{\downarrow}_{\varepsilon,E\bot \to \varepsilon}_{\varepsilon,E\bot \to \varepsilon} & \underbrace{\downarrow}_{\varepsilon,E\bot \to \varepsilon}_{\varepsilon,E\bot \to \varepsilon}_{\varepsilon,E\bot \to \varepsilon} & \underbrace{\downarrow}_{\varepsilon,E\bot \to \varepsilon}_{\varepsilon,E\bot \to \varepsilon} & \underbrace{\downarrow}_{\varepsilon,E\bot \to \varepsilon}_{\varepsilon,E\bot \to \varepsilon}_{\varepsilon,E\bot \to \varepsilon} & \underbrace{\downarrow}_{\varepsilon,E\bot \to \varepsilon}_{\varepsilon,E\bot \to \varepsilon}_{\varepsilon,E\bot \to \varepsilon} & \underbrace{\downarrow}_{\varepsilon,E\bot \to \varepsilon}_{\varepsilon,E\bot \to \varepsilon}$$

pour tout  $x \in \{a, +, *, (, )\}$ 

(0, a + a \* a,

 $\varepsilon$ ,  $\varepsilon$ )

$$(0, a + a * a, b)$$

$$(1, a + a * a, b)$$

$$\varepsilon, E + T \rightarrow E, 1$$

$$\varepsilon, T \rightarrow E, 2$$

$$\varepsilon, T * F \rightarrow T, 3$$

$$\varepsilon, F \rightarrow T, 4$$

$$\varepsilon, (E) \rightarrow F, 5$$

$$\varepsilon, a \rightarrow F, 6$$

$$(0, a+a*a, & \varepsilon, \varepsilon) \\ (1, a+a*a, & \bot, \varepsilon) \\ (1, a+a*a, & \bot, \varepsilon) \\ (1, +a*a, & a\bot, 6)$$

$$\varepsilon, E+T \to E, 1 \\ \varepsilon, T \to E, 2 \\ \varepsilon, T*F \to T, 3 \\ \varepsilon, F \to T, 4 \\ \varepsilon, (E) \to F, 5 \\ \varepsilon, a \to F, 6$$

$$0 \xrightarrow{\varepsilon, \varepsilon \to \bot} 1 \xrightarrow{\varepsilon, E\bot \to \varepsilon} 2$$

$$x, \varepsilon \to x, \varepsilon \\ \text{pour tout } x \in \{a, +, *, (, )\}$$

$$\begin{array}{cccc}
\varepsilon, E + T & \to & E, 1 \\
\varepsilon, T & \to & E, 2 \\
\varepsilon, T * F & \to & T, 3 \\
\varepsilon, F & \to & T, 4 \\
\varepsilon, (E) & \to & F, 5 \\
\varepsilon, a & \to & F, 6
\end{array}$$

$$\underbrace{\varepsilon, \varepsilon \to \bot}_{x, \varepsilon} \underbrace{1}_{x, \varepsilon \to x, \varepsilon} \underbrace{1}_{x, \varepsilon \to x, \varepsilon} \underbrace{2}_{x, \varepsilon \to x, \varepsilon}$$
pour tout  $x \in \{a, +, *, (, )\}$ 

$$(0, a + a * a, & \varepsilon, \varepsilon)$$
 $(1, a + a * a, & \bot, \varepsilon)$ 
 $(1, +a * a, & a\bot, 6)$ 
 $(1, +a * a, & F\bot, 64)$ 

$$\begin{array}{cccc}
\varepsilon, E + T & \to & E, 1 \\
\varepsilon, T & \to & E, 2 \\
\varepsilon, T * F & \to & T, 3 \\
\varepsilon, F & \to & T, 4 \\
\varepsilon, (E) & \to & F, 5 \\
\varepsilon, a & \to & F, 6
\end{array}$$

$$\underbrace{\varepsilon, E + T}_{\varepsilon, T} & \to E, 1 \\
\varepsilon, F & \to T, 3 \\
\varepsilon, E & \to F, 5 \\
\varepsilon, a & \to & F, 6$$

$$\underbrace{\varepsilon, E \perp \to \varepsilon}_{\varepsilon, E \perp \to \varepsilon}$$

$$\underbrace{1}_{\varepsilon, E \perp \to \varepsilon}_{\varepsilon, E \perp \to \varepsilon}$$

$$\underbrace{2}_{\varepsilon, E \to x, \varepsilon}_{\varepsilon, E \to x, \varepsilon}$$
pour tout  $x \in \{a, +, *, (,)\}$ 

$$(0, a+a*a, & \varepsilon, \varepsilon)$$
 $(1, a+a*a, & \bot, \varepsilon)$ 
 $(1, +a*a, & a\bot, 6)$ 
 $(1, +a*a, & F\bot, 64)$ 
 $(1, +a*a, & T\bot, 64)$ 

$$\begin{array}{cccc}
\varepsilon, E + T & \to & E, 1 \\
\varepsilon, T & \to & E, 2 \\
\varepsilon, T * F & \to & T, 3 \\
\varepsilon, F & \to & T, 4 \\
\varepsilon, (E) & \to & F, 5 \\
\varepsilon, a & \to & F, 6
\end{array}$$

$$\underbrace{\varepsilon, \varepsilon \to \bot}_{\varepsilon, a} \underbrace{1}_{\varepsilon, E \bot \to \varepsilon} \underbrace{1}_{\varepsilon, E \bot \to \varepsilon}_{\varepsilon, a \to T, \varepsilon}$$

$$\underbrace{1}_{\varepsilon, \varepsilon \to \bot, \varepsilon}_{\varepsilon, \varepsilon \to \bot, \varepsilon} \underbrace{2}_{\varepsilon, \varepsilon \to \bot, \varepsilon}_{\varepsilon, \varepsilon \to \bot, \varepsilon}$$
pour tout  $x \in \{a, +, *, (, )\}$ 

$$\begin{array}{cccc}
\varepsilon, E + T & \rightarrow & E, 1 \\
\varepsilon, T & \rightarrow & E, 2 \\
\varepsilon, T * F & \rightarrow & T, 3 \\
\varepsilon, F & \rightarrow & T, 4 \\
\varepsilon, (E) & \rightarrow & F, 5 \\
\varepsilon, a & \rightarrow & F, 6
\end{array}$$

$$\underbrace{\varepsilon, E + T}_{\varepsilon, T} \xrightarrow{\varepsilon} \underbrace{\varepsilon, 2}_{\varepsilon, 2}$$

$$\underbrace{\varepsilon, T * F}_{\varepsilon, 1} \xrightarrow{\varepsilon} \underbrace{\varepsilon, E \perp \to \varepsilon}_{\varepsilon, 2}$$

$$\underbrace{\varepsilon, E \to \bot}_{\varepsilon, \varepsilon \to \bot} \underbrace{1}_{\varepsilon, E \perp \to \varepsilon}$$

$$\underbrace{\varepsilon, E \to \bot}_{\varepsilon, \varepsilon \to \bot} \underbrace{1}_{\varepsilon, \varepsilon \to \bot}$$

$$\underbrace{\varepsilon, E \to \bot}_{\varepsilon, \varepsilon \to \bot} \underbrace{2}_{\varepsilon, \varepsilon \to \bot}$$

$$\underbrace{\varepsilon, E \to \bot}_{\varepsilon, \varepsilon \to \bot} \underbrace{2}_{\varepsilon, \varepsilon \to \bot}$$

$$\underbrace{\varepsilon, E \to \bot}_{\varepsilon, \varepsilon \to \bot} \underbrace{2}_{\varepsilon, \varepsilon \to \bot}$$

$$\underbrace{\varepsilon, E \to \bot}_{\varepsilon, \varepsilon \to \bot} \underbrace{2}_{\varepsilon, \varepsilon \to \bot}$$

$$\underbrace{\varepsilon, E \to \bot}_{\varepsilon, \varepsilon \to \bot} \underbrace{2}_{\varepsilon, \varepsilon \to \bot}$$

$$\underbrace{\varepsilon, E \to \bot}_{\varepsilon, \varepsilon \to \bot} \underbrace{2}_{\varepsilon, \varepsilon \to \bot}$$

$$\underbrace{\varepsilon, E \to \bot}_{\varepsilon, \varepsilon \to \bot} \underbrace{2}_{\varepsilon, \varepsilon \to \bot}$$

$$\underbrace{\varepsilon, E \to \bot}_{\varepsilon, \varepsilon \to \bot} \underbrace{2}_{\varepsilon, \varepsilon \to \bot}$$

$$\underbrace{\varepsilon, E \to \bot}_{\varepsilon, \varepsilon \to \bot} \underbrace{2}_{\varepsilon, \varepsilon \to \bot}$$

$$\underbrace{\varepsilon, E \to \bot}_{\varepsilon, \varepsilon \to \bot} \underbrace{2}_{\varepsilon, \varepsilon \to \bot}$$

$$\underbrace{\varepsilon, E \to \bot}_{\varepsilon, \varepsilon \to \bot} \underbrace{2}_{\varepsilon, \varepsilon \to \bot}$$

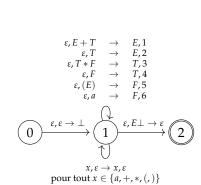
$$\underbrace{\varepsilon, E \to \bot}_{\varepsilon, \varepsilon \to \bot} \underbrace{2}_{\varepsilon, \varepsilon \to \bot}$$

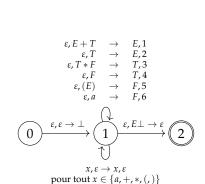
$$\underbrace{\varepsilon, E \to \bot}_{\varepsilon, \varepsilon \to \bot} \underbrace{2}_{\varepsilon, \varepsilon \to \bot}$$

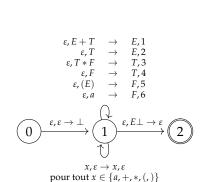
$$\underbrace{\varepsilon, E \to \bot}_{\varepsilon, \varepsilon \to \bot}$$

$$\begin{array}{cccc}
\varepsilon, E + T & \to & E, 1 \\
\varepsilon, T & \to & E, 2 \\
\varepsilon, T * F & \to & T, 3 \\
\varepsilon, F & \to & T, 4 \\
\varepsilon, (E) & \to & F, 5 \\
\varepsilon, a & \to & F, 6
\end{array}$$

$$\underbrace{\varepsilon, E + T}_{\varepsilon, T} \xrightarrow{\varepsilon, E} \xrightarrow{\varepsilon} \underbrace{0}_{\varepsilon, E}$$







$$\begin{array}{cccc}
\varepsilon, E + T & \rightarrow & E, 1 \\
\varepsilon, T & \rightarrow & E, 2 \\
\varepsilon, T * F & \rightarrow & T, 3 \\
\varepsilon, F & \rightarrow & T, 4 \\
\varepsilon, (E) & \rightarrow & F, 5 \\
\varepsilon, a & \rightarrow & F, 6
\end{array}$$

$$\underbrace{\varepsilon, \varepsilon \to \bot}_{\varepsilon, a} \underbrace{1}_{\varepsilon, E \bot \to \varepsilon} \underbrace{2}_{\varepsilon, E \bot \to \varepsilon}_{\varepsilon, a}$$

$$\underbrace{1}_{\varepsilon, E \bot \to \varepsilon}_{\varepsilon, E \bot \to \varepsilon} \underbrace{2}_{\varepsilon, E \bot \to \varepsilon}_{\varepsilon, E \bot \to \varepsilon}$$
pour tout  $x \in \{a, +, *, (, )\}$ 

$$\begin{array}{cccc}
\varepsilon, E + T & \rightarrow & E, 1 \\
\varepsilon, T & \rightarrow & E, 2 \\
\varepsilon, T * F & \rightarrow & T, 3 \\
\varepsilon, F & \rightarrow & T, 4 \\
\varepsilon, (E) & \rightarrow & F, 5 \\
\varepsilon, a & \rightarrow & F, 6
\end{array}$$

$$\underbrace{\varepsilon, \varepsilon \to \bot}_{\varepsilon, a} \underbrace{1}_{\varepsilon, E \bot \to \varepsilon} \underbrace{1}_{\varepsilon, E \bot \to \varepsilon}_{\varepsilon} \underbrace{2}_{\varepsilon, \varepsilon \to \bot}_{\varepsilon, \varepsilon \to \bot} \underbrace{2}_{\varepsilon, \varepsilon \to \bot}_{\varepsilon, \varepsilon \to \bot}_{\varepsilon, \varepsilon \to \bot}$$

$$\underbrace{1}_{\varepsilon, E \bot \to \varepsilon}_{\varepsilon, \varepsilon \to \bot} \underbrace{2}_{\varepsilon, \varepsilon \to \bot}_{\varepsilon, \varepsilon \to \bot}_{\varepsilon$$

)
)
)
4)
4)
4)
46)
46)
46)
466)
4664)
46643)
466432)
4664321)

	$\varepsilon, E + T$ $\varepsilon, T$ $\varepsilon, T * F$ $\varepsilon, F$ $\varepsilon, (E)$	$\begin{array}{c} \rightarrow \\ \rightarrow \\ \rightarrow \\ \rightarrow \\ \rightarrow \\ \rightarrow \end{array}$	T, 4
	ε, α	$\overset{'}{\rightarrow}$	F, 6
0	$\xrightarrow{\varepsilon,\varepsilon\to\bot} \left(\begin{array}{c} \\ \\ \end{array}\right)$	$\frac{1}{1}$	$\xrightarrow{\epsilon} \mathbb{Z}$
	$x, \varepsilon$ pour tout $x \in$	$\rightarrow x, \varepsilon \in \{a, +\}$	+,*,(,)}

(0,	a + a * a,	ε,	$\varepsilon)$
(1,	a + a * a,	⊥,	$\varepsilon)$
(1,	+a*a,	$a\perp$ ,	6)
(1,	+a*a,	$F\bot$ ,	64)
(1,	+a*a,	$T\bot$ ,	64)
(1,	a * a,	$+T\bot$ ,	64)
(1,	*a,	$a+T\perp$ ,	646)
(1,	*a,	$F+T\perp$ ,	646)
(1,	<i>a</i> ,	$*F+T\bot$ ,	646)
(1,	ε,	$a*F+T\bot$ ,	6466)
(1,	ε,	$F * F + T \perp$ ,	64664)
(1,	ε,	$T * F + T \perp$ ,	646643)
(1,	ε,	$T+T\bot$ ,	6466432)
(1,	ε,	$E+T\perp$ ,	64664321)
(1,	ε,	E⊥,	64664321)

$$\begin{array}{cccc}
\varepsilon, E+T & \to & E, 1 \\
\varepsilon, T & \to & E, 2 \\
\varepsilon, T*F & \to & T, 3 \\
\varepsilon, F & \to & T, 4 \\
\varepsilon, (E) & \to & F, 5 \\
\varepsilon, a & \to & F, 6
\end{array}$$

$$\underbrace{\varepsilon, E \to \bot}_{x, \varepsilon \to x, \varepsilon} \underbrace{1}_{x, \varepsilon \to x, \varepsilon} \underbrace{1}_{x, \varepsilon \to x, \varepsilon} \underbrace{2}_{x, \varepsilon \to x, \varepsilon}$$

$$\underbrace{\varepsilon, E \to \bot}_{x, \varepsilon \to x, \varepsilon} \underbrace{1}_{x, \varepsilon \to x, \varepsilon} \underbrace{2}_{x, \varepsilon \to x, \varepsilon}$$

$$\underbrace{\varepsilon, E \to \bot}_{x, \varepsilon \to x, \varepsilon} \underbrace{1}_{x, \varepsilon \to x, \varepsilon} \underbrace{2}_{x, \varepsilon \to x, \varepsilon}$$

$$\underbrace{\varepsilon, E \to \bot}_{x, \varepsilon \to x, \varepsilon} \underbrace{2}_{x, \varepsilon \to x, \varepsilon}$$

$$\underbrace{\varepsilon, E \to \bot}_{x, \varepsilon \to x, \varepsilon} \underbrace{2}_{x, \varepsilon \to x, \varepsilon}$$

$$\underbrace{\varepsilon, E \to \bot}_{x, \varepsilon \to x, \varepsilon} \underbrace{2}_{x, \varepsilon \to x, \varepsilon}$$

$$\underbrace{\varepsilon, E \to \bot}_{x, \varepsilon \to x, \varepsilon} \underbrace{2}_{x, \varepsilon \to x, \varepsilon}$$

$$\underbrace{\varepsilon, E \to \bot}_{x, \varepsilon \to x, \varepsilon} \underbrace{2}_{x, \varepsilon \to x, \varepsilon}$$

$$\underbrace{\varepsilon, E \to \bot}_{x, \varepsilon \to x, \varepsilon} \underbrace{2}_{x, \varepsilon \to x, \varepsilon}$$

$$\underbrace{\varepsilon, E \to \bot}_{x, \varepsilon \to x, \varepsilon} \underbrace{2}_{x, \varepsilon \to x, \varepsilon}$$

```
a + a * a,
a + a * a,
+a*a,
                         64)
+a*a,
                   T\bot,
                         64)
+a*a,
                         64)
a * a,
                 +T\perp,
               a+T\perp
                         646)
*a,
*a,
               F+T\perp,
                         646)
             *F+T\bot,
                         646)
                         6466)
           a*F+T\bot,
           F * F + T \perp,
                         64664)
           T*F+T\perp,
                         646643)
              T+T\perp,
                         6466432)
              E+T\perp,
                         64664321)
                   E⊥,
                         64664321)
                         64664321)
```

#### Analyse déterministe

- L'automate non déterministe n'est pas utilisable
- **Idée générale** : rendre déterministe un analyseur gauche ou un analyseur droit en s'autorisant à regarder les *k* prochains symboles de la bande d'entrée (*lookahead*)
- La prochaine action à effectuer est indiquée par une table d'analyse
- Etant donné la configuration courante de l'automate et les *k* prochains symboles, elle indique l'action à effectuer.
- Si l'analyseur gauche correspondant à la grammaire peut être rendu déterministe dans ces conditions, alors on dit que la grammaire est *LL*(*k*) (Left to right, Leftmost derivation)
- Si l'analyseur droit correspondant à la grammaire peut être rendu déterministe dans ces conditions, alors on dit que la grammaire est LR(k) (Left to right, Rightmost derivation)

## Grammaires LR(k)

- Une grammaire est LR(k) s'il est possible d'effectuer une analyse par décalage-réduction déterministe en s'autorisant à lire les k symboles suivant le symbole courant.
- La grammaire suivante n'est pas LR(1) mais elle est LR(2):

 $3: X \rightarrow a$ 

 $4: Y \rightarrow a$ 

#### Table SLR

La table SLR est une structure de données auxiliaire qui permet d'analyser certaines grammaires de manière déterministe.

état		action								
	0	1	*	+	\$	E	B			
0	d1	d2				4	3			
1	r4	r4	r4	r4	r4					
2 3	r5	r5	r5	r5	r5					
	r3	r3	r3	r3	r3					
4			d5	d6	acc					
5	d1	d2					7			
6	d1	d2					8			
7	r2	r2	r2	r2	r2					
8	r1	r1	r1	r1	r1					

#### Structure et utilisation de la table SLR

La table d'analyse est composée de deux parties :

- lacktriangle une fonction action représentée dans la partie ACTION[i,a]
- lacktriangle et une fonction *transfert*, représentée dans la partie GOTO[l, A]

La fonction action prend comme argument un état i et un terminal a (ou le marqueur \$)

la valeur de ACTION[i, a] peut avoir une des quatre formes suivantes :

- dj, où j est un état. L'analyseur effectue un décalage : il empile j et consomme une unité lexicale
- rj, où j est le numéro de la règle  $A \rightarrow \beta$ . L'analyseur effectue une réduction :
  - il dépile  $|\beta|$  symboles de la pile
  - l'état *l* est maintenant au sommet de la pile
  - il empile l'état m, qui correspond à l'entrée GOTO[l, A]
- acc: l'analyseur accepte l'entrée
- err : l'analyseur signale une erreur (signalé en général par une case vide)

état	I		actio	•		~	ıto.	On lit le mot :
etat			actioi	. L		go	ilo	
	0	1	*	+	\$	E	В	1 + 0\$
0	d1	d2				4	3	1   0φ
1	r4	r4	r4	r4	r4			
2	r5	r5	r5	r5	r5			
3	r3	r3	r3	r3	r3			
2 3 4 5			d5	d6	acc			
5	d1	d2					7	
6	d1	d2					8	
6 7	r2	r2	r2	r2	r2			0
8	r1	r1	r1	r1	r1			
	'							

état	l		actio	1		go	to.	On lit le mot :
Ctat			*		ф			
	0	1	*	+	\$	E	В	1+0\$
0	d1	d2				4	3	, , ,
1	r4	r4	r4	r4	r4			
2	r5	r5	r5	r5	r5			
3	r3	r3	r3	r3	r3			
2 3 4 5			d5	d6	acc			
5	d1	d2					7	
6	d1	d2					8	
7	r2	r2	r2	r2	r2			0
8	r1	r1	r1	r1	r1			
						'		

état			actio	า		go	oto	On lit le mot :
ctat	0	1	*	+	\$	E	В	1 + 0\$
0	d1	d2				4	3	1   Οψ
1	r4	r4	r4	r4	r4			
2 3	r5	r5	r5	r5	r5			
	r3	r3	r3	r3	r3			
4			d5	d6	acc			
5	d1	d2					7	
6	d1	d2					8	2
7	r2	r2	r2	r2	r2			0
8	r1	r1	r1	r1	r1			
								—

4L-L	I			_				On lit le mot :
état		action					oto	
	0	1	*	+	\$	E	B	1 + 0\$
0	d1	d2				4	3	2   04
1	r4	r4	r4	r4	r4			
2	r5	r5	r5	r5	r5			
2 3 4	r3	r3	r3	r3	r3			
			d5	d6	acc			
5	d1	d2					7	
6	d1	d2					8	
7	r2	r2	r2	r2	r2			0
8	r1	r1	r1	r1	r1			

<1 - 1	ı			_				On lit le mot :
état			actio	n		gc	oto	
	0	1	*	+	\$	Ε	B	1 + 0\$
0	d1	d2				4	3	Ι   ΟΨ
1	r4	r4	r4	r4	r4			
2	r5	r5	r5	r5	r5			
3 4	r3	r3	r3	r3	r3			
			d5	d6	acc			
5	d1	d2					7	
6	d1	d2					8	3
7	r2	r2	r2	r2	r2			0
8	r1	r1	r1	r1	r1			

état			actio	n		go	oto	On lit le mot :
	0	1	*	+	\$	Ë	B	1 + 0\$
0	d1	d2				4	3	1   0φ
1	r4	r4	r4	r4	r4			
2	r5	r5	r5	r5	r5			
2 3 4	r3	r3	r3	r3	r3			
4			d5	d6	acc			
5	d1	d2					7	
6	d1	d2					8	
7	r2	r2	r2	r2	r2			0
8	r1	r1	r1	r1	r1			
						'		

état			actio	n		go	oto	On lit le mot :
	0	1	*	+	\$	Ĕ	B	1 + 0\$
0	d1	d2				4	3	1 γ οφ
1	r4	r4	r4	r4	r4			
2	r5	r5	r5	r5	r5			
3	r3	r3	r3	r3	r3			
4			d5	d6	acc			
5	d1	d2					7	
6	d1	d2					8	4
7	r2	r2	r2	r2	r2			0
8	r1	r1	r1	r1	r1			
								I <del></del> I

	ı							On lit le mot :
état		action					oto	0
	0	1	*	+	\$	E	B	1 + 0\$
0	d1	d2				4	3	Σ 1 0φ
1	r4	r4	r4	r4	r4			
2	r5	r5	r5	r5	r5			
2 3 4	r3	r3	r3	r3	r3			
			d5	d6	acc			
5	d1	d2					7	6
6	d1	d2					8	$\mid 4 \mid$
7	r2	r2	r2	r2	r2			0
8	r1	r1	r1	r1	r1			
						1		

	état			actio	ı	goto On lit le m			
		0	1	*	+	\$	Ĕ	B	1 + 0\$
•	0	d1	d2				4	3	Ι   ΟΨ
	1	r4	r4	r4	r4	r4			
	2	r5	r5	r5	r5	r5			
	2 3 4 5	r3	r3	r3	r3	r3			1
	4			d5	d6	acc			
	5	d1	d2					7	6
	6	d1	d2					8	$\mid 4 \mid$
	7	r2	r2	r2	r2	r2			0
	8	r1	r1	r1	r1	r1			

état	action goto							On lit le mot :
etat				.1			ito	
	0	1	*	+	\$	E	B	1 + 0\$
0	d1	d2				4	3	Σ   0φ
1	r4	r4	r4	r4	r4			
2	r5	r5	r5	r5	r5			
2 3 4	r3	r3	r3	r3	r3			
			d5	d6	acc			
5	d1	d2					7	6
6	d1	d2					8	$\mid 4 \mid$
7	r2	r2	r2	r2	r2			0
8	r1	r1	r1	r1	r1			

<i></i>	ı					ı		On lit le mot :
état			actio	n		go	oto	
	0	1	*	+	\$	E	B	1 + 0\$
0	d1	d2				4	3	Σ   0φ
1	r4	r4	r4	r4	r4			
2	r5	r5	r5	r5	r5			
2 3 4	r3	r3	r3	r3	r3			
			d5	d6	acc			8
5	d1	d2					7	6
6	d1	d2					8	$\mid 4 \mid$
7	r2	r2	r2	r2	r2			0
8	r1	r1	r1	r1	r1			
	1					ı		

41-1	ı		( !					On lit le mot :
état			actioi	n		goto		
	0	1	*	+	\$	E	B	1 + 0\$
0	d1	d2				4	3	1   0φ
1	r4	r4	r4	r4	r4			
2	r5	r5	r5	r5	r5			
3	r3	r3	r3	r3	r3			
1 2 3 4 5 6 7			d5	d6	acc			
5	d1	d2					7	
6	d1	d2					8	
7	r2	r2	r2	r2	r2			0
8	r1	r1	r1	r1	r1			
						•		

# Exemple d'utilisation de la table

état	action						oto	On lit le mot :
Ctut	0	1	*	+	\$	E	В	1 + 0\$
0	d1	d2				4	3	ΙΙΟΨ
1	r4	r4	r4	r4	r4			
2	r5	r5	r5	r5	r5			
3	r3	r3	r3	r3	r3			
4			d5	d6	acc			
5	d1	d2					7	
6	d1	d2					8	4
7	r2	r2	r2	r2	r2			0
8	r1	r1	r1	r1	r1			

 $3: E \rightarrow B$ 

- La table SLR est construite à partir d'un automate appelé automate LR(0).
- Le langage reconnu par cet automate est l'ensemble des séquences de symboles qui peuvent apparaître sur la pile d'un analyseur par décalage réduction pour cette grammaire.
- L'automate est utilisé pour construire la table SLR.
- La construction consiste à associer à tout état de l'automate des actions (décalage et réduction) en s'autorisant à regarder le prochain symbole.

## Construction de l'automate LR(0)

- Augmentation de la grammaire
- Construction des ensembles d'items (FERMETURE)
- Construction de la fonction de transition (ALLER\_A)
- Construction des SUIVANT() pour la grammaire

# Augmentation de la grammaire

Soit la grammaire suivante :

$$\begin{array}{ccc} E & \rightarrow & E*B \mid E+B \mid B \\ B & \rightarrow & 0 \mid 1 \end{array}$$

## Augmentation de la grammaire

Soit la grammaire suivante :

$$\begin{array}{ccc} E & \rightarrow & E*B \mid E+B \mid B \\ B & \rightarrow & 0 \mid 1 \end{array}$$

On ajoute un non-terminal de départ  $S \to E$ . La grammaire augmentée est donc :

$$\begin{array}{ccc} S & \rightarrow & E \\ E & \rightarrow & E*B \mid E+B \mid B \\ B & \rightarrow & 0 \mid 1 \end{array}$$

#### **Articles**

- Un article d'une grammaire *G* est une règle de *G* avec un marqueur à une certaine position de la partie droite.
- En partant de la règle  $A \rightarrow XYZ$  on peut créer les quatre articles suivants :
  - $\begin{array}{ccc} A & \rightarrow & \bullet XYZ \\ A & \rightarrow & X \bullet YZ \\ A & \rightarrow & XY \bullet Z \\ A & \rightarrow & XYZ \bullet \end{array}$
- A partir de la règle  $A \to \varepsilon$  on ne peut créer que l'article suivant :  $A \to \bullet$
- Un article indique quelle partie d'une règle a déjà été reconnue à un certain point de l'analyse syntaxique (la partie se trouvant à gauche du point) et ce qui reste à reconnaître (à droite du point).

#### La fonction FERMETURE

La fonction FERMETURE(I) prend une liste d'articles et, pour tout article  $A \to \alpha \bullet B\gamma$  dans FERMETURE(I), ajoute  $B \to \bullet \beta$  pour toute règle  $B \to \beta$ .

$$\begin{array}{lll} \operatorname{FERMETURE}(S \to \bullet E) & = & \{S \to \bullet E, \\ E \to \bullet E * B, \\ E \to \bullet E + B, \\ E \to \bullet B, \\ B \to \bullet 0, \\ B \to \bullet 1\} \end{array}$$

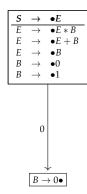
La fonction FERMETURE permet de définir les états de l'automate LR(0)

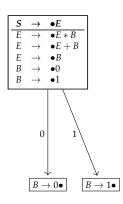
## La fonction ALLER\_A(I, X)

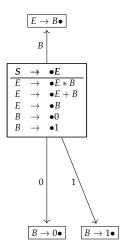
- Si I est un ensemble d'articles et X est un symbole, alors Aller\_A(I, X) est la FERMETURE de l'ensemble de tous les articles  $A \to \alpha X \bullet \beta$  tels que  $A \to \alpha \bullet X \beta$  est dans I
- La fonction Aller\_A(I, X) est utilisée pour définir les transitions de l'automate LR(0) d'une grammaire.
- Exemple :  $I_4 = \{S \to E \bullet, E \to E \bullet *B, E \to E \bullet +B\}$ ALLER\_A( $I_4, *$ ) = FERMETURE  $\{E \to E * \bullet B\}$ =  $\{E \to E * \bullet B, B \to \bullet 0, B \to \bullet 1\}$

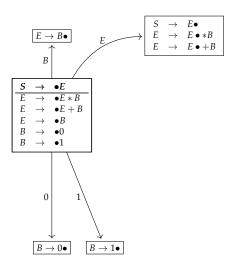


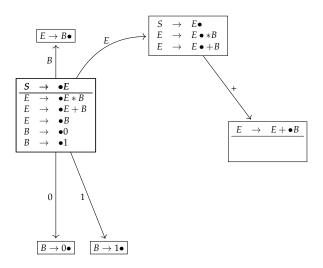
```
\begin{array}{cccc} S & \rightarrow & \bullet E \\ \hline E & \rightarrow & \bullet E * B \\ E & \rightarrow & \bullet E + B \\ E & \rightarrow & \bullet B \\ B & \rightarrow & \bullet 0 \\ B & \rightarrow & \bullet 1 \\ \end{array}
```

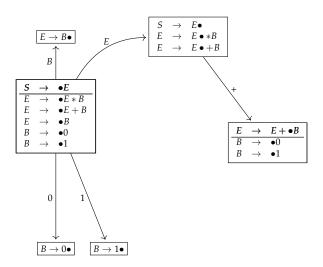


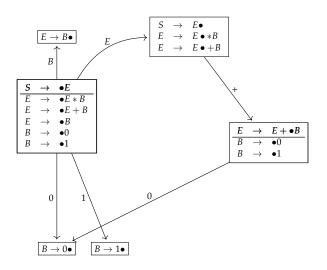


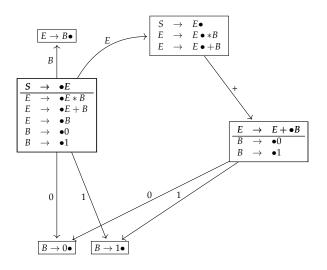


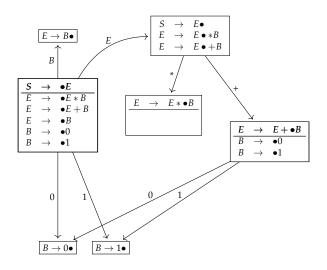


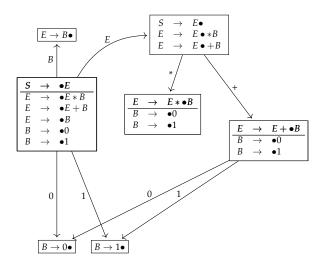


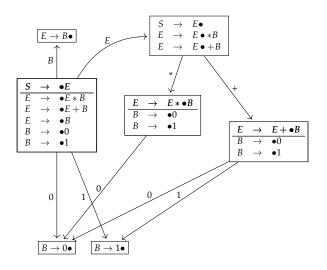


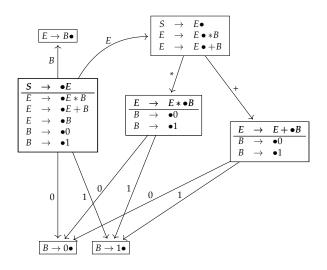


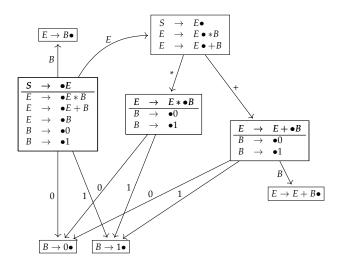


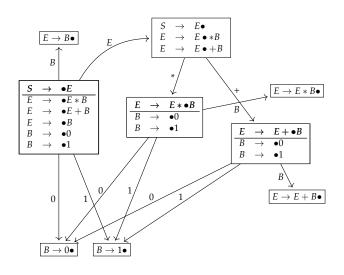


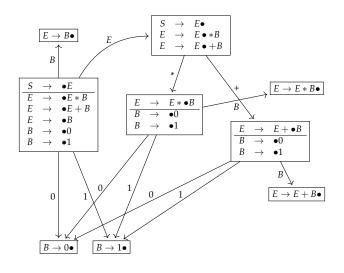


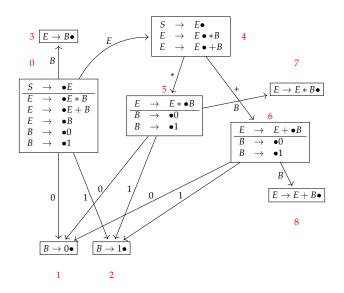




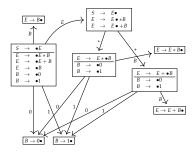






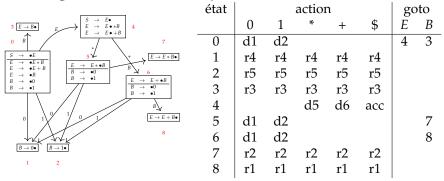


# De l'automate LR(0) à la table SLR



### De l'automate LR(0) à la table SLR

 $d = d\acute{e}$ calage,  $r = r\acute{e}$ duction, acc = acceptation



#### Construction de la table SLR

Entrée : Une grammaire augmentée G'

- Construire  $C = \{I_0, I_1, ..., I_n\}$  la collection d'ensemble d'articles LR(0) pour G'
- 2 L'état i est construit à partir de  $I_i$ . Les actions d'analyse syntaxique pour l'état i sont déterminées comme suit :
  - I Si  $(A \to \alpha \bullet a\beta \in I_i)$  et  $(a \in \Sigma)$  et  $(ALLER\_A(I_i, a) = I_j)$  alors ACTION[i, a] = dj.
  - **2** Si  $(A \to \alpha \bullet \in I_i)$  et  $(A \neq S')$  et  $(a \in SUIVANT(A))$  alors ACTION[i, a] = rj où j est le numéro de la règle  $A \to \alpha$
  - 3 Si S' → S• ∈  $I_i$ , alors ACTION[i, \$] = acc
- 3 Les transitions de transfert pour l'état i sont construites pour tout non terminal A à l'aide des règles suivantes : si  $ALLER_A(I_i, A) = I_j$  alors GOTO[i, A] = j
- Toutes les entrées non remplies sont positionnées à err
- **5** L'état initial de l'analyseur est celui construit à partir de l'ensemble d'articles contenant S' → •S

### SUIVANT(X)

- Permet de savoir quels symboles terminaux peuvent suivre le symbole *X* dans les proto-mots de la grammaire.
- Pour calculer SUIVANT(*X*), il faut connaître les symboles terminaux qui peuvent commencer les proto-mots dérivant d'un symbole *Y* qui peut suivre *X*.
- Ces symboles sont déterminés par la fonction PREMIER(*Y*).

#### **PREMIER**

Si  $\alpha$  est un proto-mot de G, PREMIER $(\alpha)$  est l'ensemble des terminaux qui commencent les chaînes se dérivant de  $\alpha$  :

$$PREMIER(\alpha) = \{a \in \Sigma \mid \alpha \stackrel{*}{\Rightarrow} au\}$$

Si  $\alpha \stackrel{*}{\Rightarrow} ε$  alors ε appartient aussi à PREMIER $(\alpha)$ .

## Exemple

```
\begin{array}{l} A \to BC | a \\ B \to b | \varepsilon \\ C \to c | \varepsilon \end{array}
\begin{array}{l} A \Rightarrow a \\ A \Rightarrow BC \Rightarrow bC \\ A \Rightarrow BC \Rightarrow C \Rightarrow c \\ A \Rightarrow BC \Rightarrow C \Rightarrow \varepsilon \end{array} \qquad \begin{array}{l} a \in \text{PREMIER}(A) \\ b \in \text{PREMIER}(A) \\ c \in \text{PREMIER}(A) \\ c \in \text{PREMIER}(A) \\ c \in \text{PREMIER}(A) \end{array}
```

# Exemple

```
A \to BC|a
B \to b|\varepsilon
C \rightarrow c | \varepsilon
 A \Rightarrow a

A \Rightarrow BC \Rightarrow bC a \in PREMIER(A)

b \in PREMIER(A)
  A \Rightarrow BC \Rightarrow C \Rightarrow c \mid c \in PREMIER(A)
  A \Rightarrow BC \Rightarrow C \Rightarrow \varepsilon \mid \varepsilon \in PREMIER(A)
plus généralement :
PREMIER(B) \subseteq PREMIER(A)
si \varepsilon \in \text{PREMIER}(B) alors \text{PREMIER}(C) \subseteq \text{PREMIER}(A)
si \varepsilon \in \text{PREMIER}(B) et \varepsilon \in \text{PREMIER}(C) alors \varepsilon \in \text{PREMIER}(A)
```

### PREMIER(X)

Pour calculer PREMIER(X) avec  $X \in N \cup \Sigma$ , on applique les règles suivantes jusqu'à ce qu'aucun terminal ni  $\varepsilon$  ne puisse être ajouté aux ensembles PREMIER.

- 1 Si  $X \in \Sigma$  (X terminal), PREMIER(X) = {X}.
- 2 Si  $X \to \varepsilon \in$  productions de la grammaire, on ajoute  $\varepsilon$  à PREMIER(X).
- 3 Si  $X \in N$  (X non terminal) et  $X \to Y_1 \dots Y_k \in P$ , mettre a dans PREMIER(X) s'il existe i tel que a est dans PREMIER( $Y_i$ ) et que  $\varepsilon$ est dans tous les PREMIER( $Y_1$ ) ... PREMIER( $Y_{i-1}$ ). Si  $\varepsilon \in \text{PREMIER}(Y_i) \forall j$ ,  $1 \le j \le k$ , on ajoute  $\varepsilon$  à PREMIER(X).

### PREMIER( $X_1 \dots X_n$ )

On calcule PREMIER $(X_1 ... X_n)$  de la façon suivante :

- **I** Ajouter à PREMIER( $X_1 ... X_n$ ) tous les symboles de PREMIER( $X_1$ ) différents de ε.
- 2 Si  $\varepsilon \in \text{PREMIER}(X_1)$ , ajouter également les symboles de PREMIER $(X_2)$  différents de  $\varepsilon$ . Si  $\varepsilon \in \text{PREMIER}(X_2)$ , ajouter également les symboles de PREMIER $(X_3)$  différents de  $\varepsilon$ , etc.
- **3** Finalement, si ε appartient à PREMIER $(X_j)$  pour tous les j = 1, 2, ..., n, on ajoute ε à PREMIER $(X_1, ..., X_n)$ .

### SUIVANT(X)

Si  $X \in N$ , SUIVANT(X) est l'ensemble des symboles  $a \in \Sigma$  qui peuvent apparaître immédiatement à droite de X dans un proto-mot :

$$SUIVANT(X) = \{ a \in \Sigma \mid S \stackrel{*}{\Rightarrow} \alpha X a \beta \}$$

Si X peut être le symbole le plus à droite d'un proto-mot alors  $\bot$  est dans SUIVANT(X).

# Exemple

```
S \rightarrow Aa

A \rightarrow BC

C \rightarrow c | \varepsilon

S \Rightarrow Aa \Rightarrow BCa \Rightarrow Bca \mid c \in SUIVANT(B)

S \Rightarrow Aa \Rightarrow BCa \Rightarrow Ba \mid a \in SUIVANT(B)
```

# Exemple

```
S \rightarrow Aa

A \rightarrow BC

C \rightarrow c | \varepsilon

S \Rightarrow Aa \Rightarrow BCa \Rightarrow Bca \mid c \in \text{SUIVANT}(B)

S \Rightarrow Aa \Rightarrow BCa \Rightarrow Ba \mid a \in \text{SUIVANT}(B)

plus généralement :

PREMIER(C) \subseteq \text{SUIVANT}(B)

si \varepsilon \in \text{PREMIER}(C) alors SUIVANT(A) \subseteq \text{SUIVANT}(B)
```

### SUIVANT(X)

Pour calculer  $\operatorname{SUIVANT}(X)$  pour tous symbole non terminal X, on applique les règles suivantes jusqu'à ce qu'aucun symbole terminal ne puisse être ajouté aux ensembles  $\operatorname{SUIVANT}$ :

- **1** Mettre  $\perp$  dans SUIVANT(S).
- **2** si  $X \to \alpha B\beta$ , le contenu de PREMIER( $\beta$ ), excepté  $\varepsilon$ , est ajouté à SUIVANT(B).
- 3 s'il existe une règle  $X \to \alpha B$  ou une règle  $X \to \alpha B \beta$  telle que  $\varepsilon \in \text{PREMIER}(\beta)$  (c'est à dire  $\beta \stackrel{*}{\Rightarrow} \varepsilon$ ), les éléments de SUIVANT(X) sont ajoutés à SUIVANT(B).

# Exemple

Soit la grammaire  $G = \langle \{E, E', T, T', F\}, \{a, +, *, (,), a\}, P, E \rangle$  non récursive à gauche où P est composé des règles suivantes :

$$\begin{array}{ll} 1 \ E \rightarrow TE' & 2 \ E' \rightarrow +TE' \\ 3 \ E' \rightarrow \varepsilon & 4 \ T \rightarrow FT' \\ 5 \ T' \rightarrow *FT' & 6 \ T' \rightarrow \varepsilon \\ 7 \ F \rightarrow (E) & 8 \ F \rightarrow a \end{array}$$

#### Alors:

```
\begin{aligned} & \text{PREMIER}(E) = \text{PREMIER}(T) = \text{PREMIER}(F) = \{(,a\} \\ & \text{PREMIER}(E') = \{+, \epsilon\} \\ & \text{PREMIER}(T') = \{*, \epsilon\} \\ & \text{SUIVANT}(E) = \{), \bot\} \\ & \text{SUIVANT}(E') = \text{SUIVANT}(E) = \{), \bot\} \\ & \text{SUIVANT}(T) = \{\text{PREMIER}(E') - \{\epsilon\}\} \cup \text{SUIVANT}(E) = \{+, ), \bot\} \\ & \text{SUIVANT}(T') = \text{SUIVANT}(T) = \{+, ), \bot\} \\ & \text{SUIVANT}(F) = \{\text{PREMIER}(T') - \{\epsilon\}\} \cup \text{SUIVANT}(T) = \{+, *, ), \bot\} \end{aligned}
```

#### **Conflits**

- Si, à l'issue de la construction de la table, une case possède plusieurs actions, alors la grammaire n'est pas SLR(1).
- Si une case possède deux réductions différentes, on dit qu'il y a un conflit réduction/réduction.
- Si une case possède un décalage et une réduction, on qu'il y a conflit décalage/réduction.
- Dans ce cas, soit la grammaire est ambiguë, soit il faut augmenter le regard en avant (la valeur de *k*).
- La majorité des langages de programmation admettent une grammaire qui est SLR(1). En particulier le langage L.

#### Conflit de réduction / réduction

$$\begin{array}{ccccc}
1: & S & \to & X \\
2: & S & \to & Y \\
3: & X & \to & a \\
4: & Y & \to & a
\end{array}$$

- On ne sait pas s'il faut réduire par 3 ou 4.
- La grammaire est ambiguë!

## Conflit de réduction / réduction (2)

- Si on sait que *a* est suivi de *b* (ou de *c*), il n'y a pas conflit!
- La grammaire est LR(1)

# Conflit de décalage / réduction

- Après un décalage d'un *a*, on ne sait pas s'il faut décaler ou réduire.
- La grammaire est ambiguë.

# Conflit de décalage / réduction

- Si on sait que a est suivi de b (ou de c), il n'y a pas conflit!
- La grammaire est LR(1).

## Grammaires LR(2)

#### Sources

- Michael Sipser,
   Introduction to the Theory of Computation.
   PWS Publishing Company, 1997.
- John Hopcroft, Rajeev Motwani, Jeffrey Ullman, Introduction to Automata Theory, Languages and Computation.
   2ème édition, Pearson Education International, 2001.
- John Aho, Jeffrey Ullman, *The Theory of Parsing*, *Translation and Compiling*, *Vol I* : *Parsing*.

  Prentice-Hall, 1972