Théorie des langages : THL

CM 9

Uli Fahrenberg

EPITA Rennes

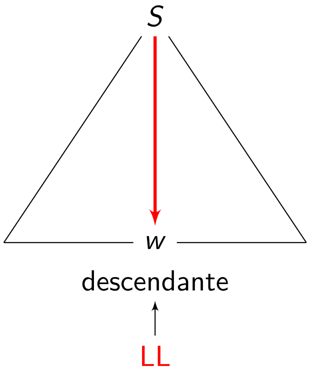
S5 2023

Aperçu

Programme du cours

1. Langages rationnels, automates finis
3. Langages algébriques, grammaires hors-contexte, automates à pile
   - TP 1 : `flex`
   - QCM 1 : langages rationnels
5. Parsage LL
7. Parsage LR, parties 1–3
8. Parsage LR, partie 3,5
9. Introduction `flex` & `bison`
10. TP 3, 4 : `flex` & `bison`

Re : approches parsage



Re : parsage LL(1)

0. entrée : une grammaire hors contexte $G = (N, \Sigma, P, S)$
   - si-dessous, $V = N \cup \Sigma$
   - éliminer récursion à gauche dans G ; factoriser G à gauche
1. calculer NULL
   - $\text{NULL} = \{A \in N \mid A \Rightarrow^* \varepsilon\}$
2. construire la table FIRST
   - $\text{FIRST}(A) = \{a \in \Sigma \mid \exists w \in V^* : A \Rightarrow^* aw\}$
3. construire la table FOLLOW
   - $\text{FOLLOW}(A) = \{a \in \Sigma \mid \exists B \in N, \alpha, \beta \in V^* : B \Rightarrow^* \alpha A a \beta\}$
4. construire la TABLE de parsage :
   1. pour chaque production $X \to w \quad (n)$:
      1. pour chaque $a \in \text{FIRST}(w)$: $\text{TABLE}(X, a)$ += $\{n\}$
      2. si $w \in \text{NULL}$ ou $w = \varepsilon$:
         – pour chaque $a \in \text{FOLLOW}(X)$: $\text{TABLE}(X, a)$ += $\{n\}$

Re : parsage ascendant : the basics

function $\text{BULRP}(\alpha)$
    if $\alpha = S$ **then**
        return True
    for $i \leftarrow 1$ **to** $|\alpha|$ **do**
        for $j \leftarrow i$ **to** $|\alpha|$ **do** ▷ décalage / SHIFT
            for $A \in N$ **do**
                if $A \to \alpha_i \dots \alpha_j$ **then** ▷ réduction / REDUCE
                    return $\text{BULRP}(\alpha_1 \dots \alpha_{i-1} A \alpha_{j+1} \dots \alpha_n)$
    return False

Définition (8.8)

Soit G une grammaire hors-contexte. Une production pointée de G est une paire $(A, \alpha \bullet \beta)$ telle que $A \to \alpha\beta$ est une production de G.

Re : automate de parsage LR(0)

Définition (8.10)

Soit G une grammaire hc et $\mathcal{I}$ un ensemble de productions pointées de G. La clôture de $\mathcal{I}$ est le plus petit ensemble $\mathrm{cl}(\mathcal{I})$ t.q. $\mathcal{I} \subseteq \mathrm{cl}(\mathcal{I})$ et

- si $(A, \alpha \bullet B \beta) \in \mathrm{cl}(\mathcal{I})$ et $B \to \gamma$ est une production de G, alors $(B, \bullet\gamma) \in \mathcal{I}$.

Définition

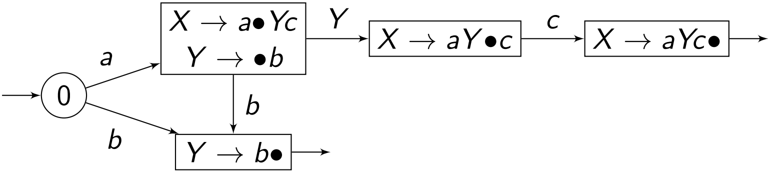
L'automate de parsage LR(0) d'une grammaire hors-contexte G est l'automate fini déterministe (Q, q_0, F, δ) avec

- $Q = \{\mathrm{cl}(\mathcal{I}) \mid \mathcal{I}$ ensemble de productions pointées de $G\}$;
- $q_0 = \mathrm{cl}\big(\{(Z, \bullet S\$)\}\big)$;
- $F = \{q \in Q \mid \exists$ production $X \to w$ de G t.q. $(X, w\bullet) \in q\}$
- et $\delta : Q \times V \to Q$ donnée par

$$\delta(q, \beta) = \mathrm{cl}\big(\{(X, \alpha\beta\bullet\gamma) \mid (X, \alpha\bullet\beta\gamma) \in q\}\big).$$

Re : exemple

$$X \to aYc \qquad (1)$$
$$Y \to b \qquad (2)$$



Re : algorithme de parsage

1. empiler q_0
2. `repeat`
   1. $q \leftarrow$ état en haut de la pile
   2. si $q =$ état final $X \to w\bullet$: REDUCE
      1. dépiler $|w|$ états
      2. $q' \leftarrow$ état en haut de la pile
      3. empiler $\delta(q', X)$ $\longleftarrow$ possible ✗
   3. sinon : SHIFT
      1. $a \leftarrow \text{next}(\text{input})$ $\longleftarrow$ possible ✗
      2. empiler $\delta(q, a)$ $\longleftarrow$ possible ✗
3. `until` $q =$ état final $Z \to S\$\bullet$ (✓) ou échec (✗)

Re : exemple : automate de parsage

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to (S) && (1)\\
&\mid n && (2)
\end{aligned}
$$

$$\longrightarrow \boxed{Z \to \bullet S\$}$$

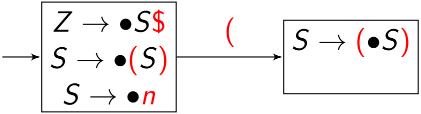
Re : exemple : automate de parsage

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to (S) && (1)\\
&\mid n && (2)
\end{aligned}
$$

$$
\longrightarrow \boxed{\begin{array}{c} Z \to \bullet S\$ \\ S \to \bullet (S) \\ S \to \bullet n \end{array}}
$$

Re : exemple : automate de parsage

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to (S) && (1)\\
&\mid n && (2)
\end{aligned}
$$



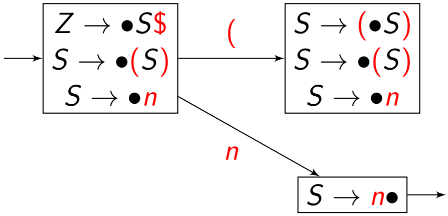
Re : exemple : automate de parsage

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to (S) && (1)\\
&\mid n && (2)
\end{aligned}
$$

$$
\begin{tikzcd}
{} \arrow[r] & \boxed{\begin{array}{c} Z \to \bullet S\$ \\ S \to \bullet (S) \\ S \to \bullet n \end{array}} \arrow[r, "("] & \boxed{\begin{array}{c} S \to (\bullet S) \\ S \to \bullet (S) \\ S \to \bullet n \end{array}}
\end{tikzcd}
$$

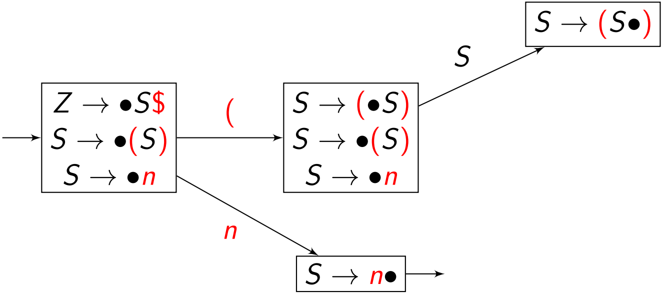
Re : exemple : automate de parsage

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to (S) && (1)\\
&\mid n && (2)
\end{aligned}
$$



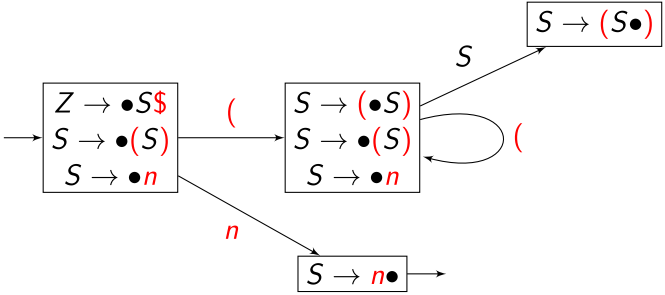
Re : exemple : automate de parsage

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to (S) && (1)\\
&\mid n && (2)
\end{aligned}
$$



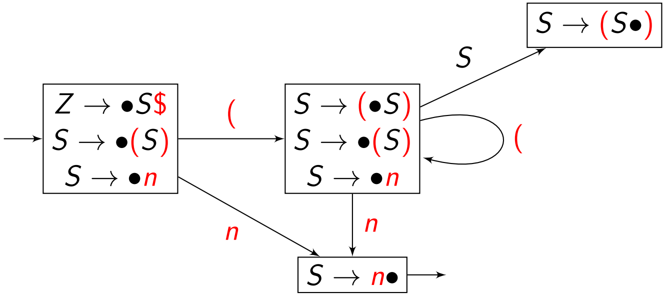
Re : exemple : automate de parsage

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to (S) && (1)\\
&\mid n && (2)
\end{aligned}
$$



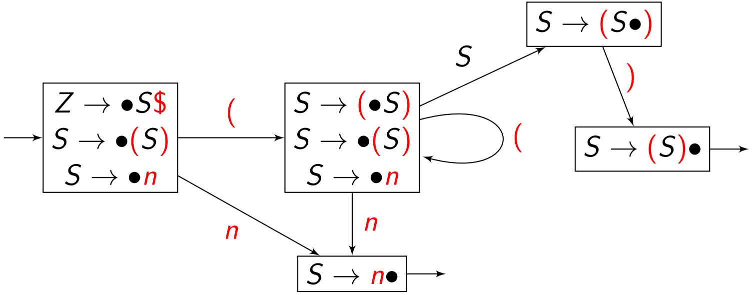
Re : exemple : automate de parsage

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to (S) && (1)\\
&\mid n && (2)
\end{aligned}
$$



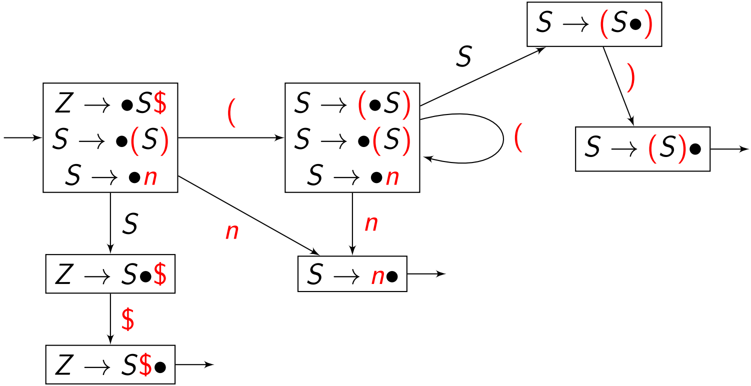
Re : exemple : automate de parsage

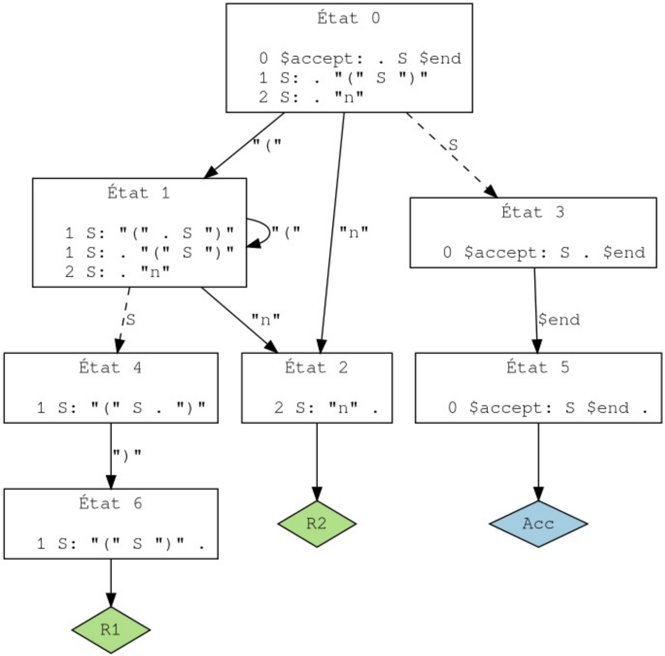
$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to (S) && (1)\\
&\mid n && (2)
\end{aligned}
$$



Re : exemple : automate de parsage

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to (S) && (1)\\
&\mid n && (2)
\end{aligned}
$$

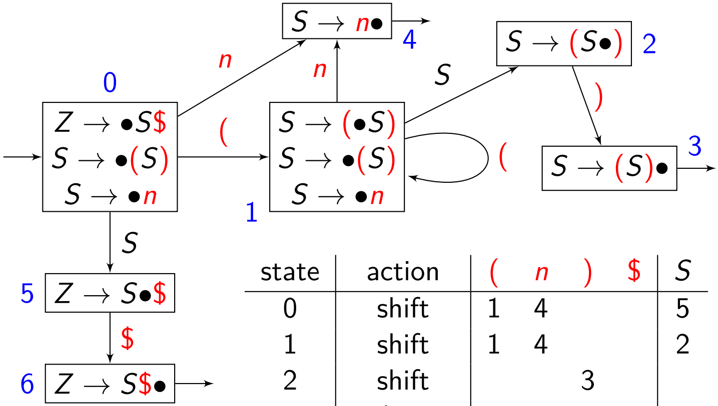




État 0
0 $accept: . S $end
1 S: . "(" S ")"
2 S: . "n"
"("
"n"
S
État 1
1 S: "(" . S ")"
1 S: . "(" S ")"
2 S: . "n"
"("
État 3
0 $accept: S . $end
S
"n"
$end
État 4
1 S: "(" S . ")"
État 2
2 S: "n" .
État 5
0 $accept: S $end .
")"
État 6
1 S: "(" S ")" .
R2
Acc
R1

Re : exemple : table de parsage

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to (S) && (1)\\
&\mid n && (2)
\end{aligned}
$$



state	action	(	n	)	$	S
0	shift	1	4			5
1	shift	1	4			2
2	shift			3		
3	reduce 1					
4	reduce 2					
5	shift				6	
6	accept					

Re : parsage LR(0)

- lire l'entrée de gauche à droite (L)
- approche ascendant
- construire une dérivation droite (R)
- pas de regard avant (0)

Re : parsage SLR(1) : exemple

$$Z \to S\$ \qquad (0)$$
$$S \to n{-}S \qquad (1)$$
$$\mid n \qquad (2)$$

état	productions pointées
0	$Z \to \bullet S\$$, $S \to \bullet n{-}S$, $S \to \bullet n$
1	$Z \to S\bullet\$$
2	$S \to n\bullet{-}S$, $S \to n\bullet$
3	$S \to n{-}\bullet S$, $S \to \bullet n{-}S$, $S \to \bullet n$
4	$Z \to S\$\bullet$
5	$S \to n{-}S\bullet$

état	action	n	$-$	$\$$	S
0	décaler	2			1
1	décaler			4	
2	réduire 2, décaler		3		
3	décaler	2			5
4	accepter				
5	réduire 1				

conflit SHIFT/REDUCE

Re : Simple LR(1)

1. calculer la table LR(0)
2. si conflits : conditionner l'action par le FOLLOW
3. passer du type état $\to$ action $\to$ entrée
   au type état $\to$ entrée $\to$ action

Exemple : $Z \to S\$$ (0)

$S \to n{-}S$ (1) $\qquad S \to n$ (2)

état	action	n	$-$	$\$$	S
0	décaler	2			1
1	décaler			4	
2	réd. 2, déc.		3		
3	décaler	2			5
4	accepter				
5	réduire 1				

$\Longrightarrow$

état	n	$-$	$\$$	S
0	d.2			d.1
1			d.4	
2		d.3	r.2	
3	d.2			d.5
4	— accepter —			
5			r.1	

Parsage LR(1)

Exemple

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to L{=}E && (1)\\
&\mid E && (2)\\
L &\to x && (3)\\
&\mid {*}E && (4)\\
E &\to L && (5)
\end{aligned}
$$

- manipulation des pointeurs

Exemple

$$
\begin{array}{rll}
Z \to & S\$ & (0) \\
S \to & L{=}E & (1) \\
 & \mid E & (2) \\
L \to & x & (3) \\
 & \mid *E & (4) \\
E \to & L & (5)
\end{array}
$$

état	x	$*$	$=$	$\$$	S	L	E
0	d.4	d.5			d.1	d.2	d.3
1				d.6			
2			d.7				
			r.5	r.5			
3				r.2			
4			r.3	r.3			
5	d.4	d.5				d.9	d.8
6			— accepter —				
7	d.4	d.5				d.9	d.10
8			r.4	r.4			
9			r.5	r.5			
10			r.1	r.1			

Exemple

$$
\begin{array}{rcll}
Z & \to & S\$ & (0) \\
S & \to & L{=}E & (1) \\
 & | & E & (2) \\
L & \to & x & (3) \\
 & | & *E & (4) \\
E & \to & L & (5)
\end{array}
$$

état	x	$*$	$=$	$\$$	S	L	E
0	d.4	d.5			d.1	d.2	d.3
1				d.6			
2		conflit	d.7 r.5	r.5			
3				r.2			
4			r.3	r.3			
5	d.4	d.5				d.9	d.8
6			— accepter —				
7	d.4	d.5				d.9	d.10
8			r.4	r.4			
9			r.5	r.5			
10			r.1	r.1			

Exemple, bis

Le problème :

$$
\begin{array}{rll}
Z \to & S\$ & (0) \\
S \to & L{=}E & (1) \\
\mid & E & (2) \\
L \to & x & (3) \\
\mid & *E & (4) \\
E \to & L & (5)
\end{array}
$$

état	productions pointées
0	$Z \to \bullet S\$$

Exemple, bis

Le problème :

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to L{=}E && (1)\\
&\mid E && (2)\\
L &\to x && (3)\\
&\mid *E && (4)\\
E &\to L && (5)
\end{aligned}
$$

état	productions pointées
0	$Z \to \bullet S\$$, $S \to \bullet L{=}E$, $S \to \bullet E$

Exemple, bis

Le problème :

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to L{=}E && (1)\\
&\mid E && (2)\\
L &\to x && (3)\\
&\mid {*}E && (4)\\
E &\to L && (5)
\end{aligned}
$$

état	productions pointées
0	$Z \to \bullet S\$$, $S \to \bullet L{=}E$, $S \to \bullet E$
	$L \to \bullet x$, $L \to \bullet {*}E$, $E \to \bullet L$

Exemple, bis

Le problème :

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to L{=}E && (1)\\
&\mid E && (2)\\
L &\to x && (3)\\
&\mid *E && (4)\\
E &\to L && (5)
\end{aligned}
$$

état	productions pointées
0	$Z \to \bullet S\$$, $S \to \bullet L{=}E$, $S \to \bullet E$
	$L \to \bullet x$, $L \to \bullet * E$, $E \to \bullet L$
1	$Z \to S \bullet \$$
2	$S \to L \bullet {=} E$, $E \to L \bullet$ ✓

Exemple, bis

Le problème :

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to L{=}E && (1)\\
&\mid E && (2)\\
L &\to x && (3)\\
&\mid *E && (4)\\
E &\to L && (5)
\end{aligned}
$$

état	productions pointées
0	$Z \to \bullet S\$$, $S \to \bullet L{=}E$, $S \to \bullet E$
	$L \to \bullet x$, $L \to \bullet * E$, $E \to \bullet L$
1	$Z \to S \bullet \$$
2	$S \to L \bullet {=} E$, $E \to L \bullet$ ✓

- l'état 2 ne doit accepter que si le L est suivi d'un $\$$

Regard en avant

Définition

Soit G une grammaire hors-contexte. Une production pointée élargie de G est un triplet $(A, \alpha\bullet\beta, a)$ telle que $A \to \alpha\beta$ est une production de G et $a \in \Sigma \cup \{\varepsilon\}$.

- noté $A \to \alpha\bullet\beta\ [a]$
- on a achevé α dans la production $A \to \alpha\beta$;
- il nous reste à trouver β ;
- la production n'est valable que si A est suivi par a dans l'entrée
- donc $a = \varepsilon$ (pas de contraint) ou $a \in \mathrm{FOLLOW}(A)$

Clôture

Définition

Soit G une grammaire hors-contexte et $\mathcal{I}$ un ensemble de productions pointées élargies de G. La clôture de $\mathcal{I}$ est le plus petit ensemble $\mathrm{cl}(\mathcal{I})$ tel que $\mathcal{I} \subseteq \mathrm{cl}(\mathcal{I})$ et

- si $(A, \alpha \bullet B\beta, a) \in \mathrm{cl}(\mathcal{I})$, $B \to \gamma$ est une production de G et $b \in \mathrm{FIRST}(\beta)$, alors $(B, \bullet\gamma, b) \in \mathrm{cl}(\mathcal{I})$;
- si $(A, \alpha \bullet B, a) \in \mathrm{cl}(\mathcal{I})$ et $B \to \gamma$ est une production de G, alors $(B, \bullet\gamma, a) \in \mathrm{cl}(\mathcal{I})$.

Automate LR(1)

Définition

L'automate de parsage LR(1) d'une grammaire hors-contexte G est l'automate fini déterministe (Q, q_0, F, δ) avec

- $Q = \{\mathrm{cl}(\mathcal{I}) \mid \mathcal{I}$ ensemble de prod. pointées élargies de $G\}$;
- $q_0 = \mathrm{cl}\big(\{(Z, \bullet S\$, \varepsilon)\}\big)$;
- $F = \{q \in Q \mid \exists$ production $X \to w$ de G et $a \in \Sigma \cup \{\varepsilon\}$ tels que $(X, w\bullet, a) \in q\}$
- et $\delta : Q \times V \to Q$ donnée par

$$\delta(q, \beta) = \mathrm{cl}\big(\{(X, \alpha\beta\bullet\gamma, a) \mid (X, \alpha\bullet\beta\gamma, a) \in q\}\big).$$

Exemple, ter

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to L{=}E && (1)\\
&\mid E && (2)\\
L &\to x && (3)\\
&\mid *E && (4)\\
E &\to L && (5)
\end{aligned}
$$

état	productions pointées élargies
0	$Z \to \bullet S\$\ [\varepsilon]$

Exemple, ter

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to L{=}E && (1)\\
&\mid E && (2)\\
L &\to x && (3)\\
&\mid *E && (4)\\
E &\to L && (5)
\end{aligned}
$$

état	productions pointées élargies
0	$Z \to \bullet S\$ \ [\varepsilon]$
	$S \to \bullet L{=}E\ [\$],\ S \to \bullet E\ [\$]$

Exemple, ter

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to L{=}E && (1)\\
&\mid E && (2)\\
L &\to x && (3)\\
&\mid *E && (4)\\
E &\to L && (5)
\end{aligned}
$$

état	productions pointées élargies
0	$Z \to \bullet S\$ \; [\varepsilon]$
	$S \to \bullet L{=}E \; [\$]$, $S \to \bullet E \; [\$]$
	$L \to \bullet x \; [=]$, $L \to \bullet {*}E \; [=]$

Exemple, ter

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to L{=}E && (1)\\
&\mid E && (2)\\
L &\to x && (3)\\
&\mid *E && (4)\\
E &\to L && (5)
\end{aligned}
$$

état	productions pointées élargies
0	$Z \to \bullet S\$\ [\varepsilon]$
	$S \to \bullet L{=}E\ [\$]$, $S \to \bullet E\ [\$]$
	$L \to \bullet x\ [=]$, $L \to \bullet *E\ [=]$
	$E \to \bullet L\ [\$]$

Exemple, ter

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to L{=}E && (1)\\
&\mid E && (2)\\
L &\to x && (3)\\
&\mid *E && (4)\\
E &\to L && (5)
\end{aligned}
$$

état	productions pointées élargies
0	$Z \to \bullet S\$ \ [\varepsilon]$
	$S \to \bullet L{=}E \ [\$],\ S \to \bullet E \ [\$]$
	$L \to \bullet x \ [=],\ L \to \bullet *E \ [=]$
	$E \to \bullet L \ [\$]$
	$L \to \bullet x \ [\$],\ L \to \bullet *E \ [\$]$

Exemple, ter

$Z \to S\$$	(0)
$S \to L{=}E$	(1)
$\mid E$	(2)
$L \to x$	(3)
$\mid {*}E$	(4)
$E \to L$	(5)

état	productions pointées élargies
0	$Z \to \bullet S\$ \ [\varepsilon]$
	$S \to \bullet L{=}E\ [\$]$, $S \to \bullet E\ [\$]$
	$L \to \bullet x\ [=]$, $L \to \bullet {*}E\ [=]$
	$E \to \bullet L\ [\$]$
	$L \to \bullet x\ [\$]$, $L \to \bullet {*}E\ [\$]$
1	$Z \to S\bullet\$ \ [\varepsilon]$
2	$S \to L\bullet{=}E\ [\$]$, $E \to L\bullet\ [\$\checkmark]$

- l'état 2 n'accepte que dans un contexte $

Exemple, complet

$$
\begin{array}{lll lll}
Z \to S\$ & & (0) & L \to x & & (3) \\
S \to L{=}E & & (1) & & \mid *E & (4) \\
& \mid E & (2) & E \to L & & (5)
\end{array}
$$

état	x	$*$	$=$	$\$$	S	L	E
0	d.4	d.5			d.1	d.2	d.3

état	productions pointées élargies
0	$Z \to \bullet S\$\ [\varepsilon]$, $S \to \bullet L{=}E\ [\$]$ $S \to \bullet E\ [\$]$, $L \to \bullet x\ [=]$ $L \to \bullet *E\ [=]$, $E \to \bullet L\ [\$]$ $L \to \bullet x\ [\$]$, $L \to \bullet *E\ [\$]$

Exemple, complet

$$Z \to S\$ \quad (0) \qquad L \to x \quad (3)$$
$$S \to L{=}E \quad (1) \qquad \mid *E \quad (4)$$
$$\mid E \quad (2) \qquad E \to L \quad (5)$$

état	x	$*$	$=$	$\$$	S	L	E
0	d.4	d.5			d.1	d.2	d.3

état	productions pointées élargies
0	$Z \to \bullet S\$ \ [\varepsilon]$, $S \to \bullet L{=}E \ [\$]$ $S \to \bullet E \ [\$]$, $L \to \bullet x \ [=]$ $L \to \bullet{*}E \ [=]$, $E \to \bullet L \ [\$]$ $L \to \bullet x \ [\$]$, $L \to \bullet{*}E \ [\$]$
1	$Z \to S\bullet\$ \ [\varepsilon]$
2	$S \to L\bullet{=}E \ [\$]$, $E \to L\bullet \ [\$✓]$

Exemple, complet

$$
\begin{array}{lll}
Z \to S\$ & (0) \\
S \to L{=}E & (1) \\
\quad\mid E & (2)
\end{array}
\qquad
\begin{array}{lll}
L \to x & (3) \\
\quad\mid *E & (4) \\
E \to L & (5)
\end{array}
$$

état	x	$*$	$=$	$\$$	S	L	E
0	d.4	d.5			d.1	d.2	d.3

état	productions pointées élargies
0	$Z \to \bullet S\$ \ [\varepsilon]$, $S \to \bullet L{=}E \ [\$]$
	$S \to \bullet E \ [\$]$, $L \to \bullet x \ [=]$
	$L \to \bullet{*}E \ [=]$, $E \to \bullet L \ [\$]$
	$L \to \bullet x \ [\$]$, $L \to \bullet{*}E \ [\$]$
1	$Z \to S\bullet\$ \ [\varepsilon]$
2	$S \to L\bullet{=}E \ [\$]$, $E \to L\bullet \ [\$✓]$
3	$S \to E\bullet \ [\$✓]$
4	$L \to x\bullet \ [=✓]$, $L \to x\bullet \ [\$✓]$

Exemple, complet

$$Z \to S\$ \quad (0) \qquad L \to x \quad (3)$$
$$S \to L{=}E \quad (1) \qquad \mid *E \quad (4)$$
$$\mid E \quad (2) \qquad E \to L \quad (5)$$

état	x	$*$	$=$	$\$$	S	L	E
0	d.4	d.5			d.1	d.2	d.3

état	productions pointées élargies
0	$Z \to \bullet S\$\ [\varepsilon]$, $S \to \bullet L{=}E\ [\$]$ $S \to \bullet E\ [\$]$, $L \to \bullet x\ [=]$ $L \to \bullet * E\ [=]$, $E \to \bullet L\ [\$]$ $L \to \bullet x\ [\$]$, $L \to \bullet * E\ [\$]$
1	$Z \to S \bullet \$\ [\varepsilon]$
2	$S \to L \bullet {=} E\ [\$]$, $E \to L \bullet\ [\$✓]$
3	$S \to E \bullet\ [\$✓]$
4	$L \to x \bullet\ [=✓]$, $L \to x \bullet\ [\$✓]$
5	$L \to * \bullet E\ [=]$, $L \to * \bullet E\ [\$]$

Exemple, complet

$$
\begin{array}{lll lll}
Z \to S\$ & & (0) & L \to x & & (3) \\
S \to L{=}E & & (1) & & \mid *E & (4) \\
& \mid E & (2) & E \to L & & (5)
\end{array}
$$

état	x	$*$	$=$	$\$$	S	L	E
0	d.4	d.5			d.1	d.2	d.3

état	productions pointées élargies
0	$Z \to \bullet S\$ \ [\varepsilon]$, $S \to \bullet L{=}E \ [\$]$
	$S \to \bullet E \ [\$]$, $L \to \bullet x \ [=]$
	$L \to \bullet{*}E \ [=]$, $E \to \bullet L \ [\$]$
	$L \to \bullet x \ [\$]$, $L \to \bullet{*}E \ [\$]$
1	$Z \to S\bullet\$ \ [\varepsilon]$
2	$S \to L\bullet{=}E \ [\$]$, $E \to L\bullet \ [\$\checkmark]$
3	$S \to E\bullet \ [\$\checkmark]$
4	$L \to x\bullet \ [=\checkmark]$, $L \to x\bullet \ [\$\checkmark]$
5	$L \to *\bullet E \ [=]$, $L \to *\bullet E \ [\$]$
	$E \to \bullet L \ [=]$, $L \to \bullet x \ [=]$
	$L \to \bullet{*}E \ [=]$, $E \to \bullet L \ [\$]$
	$L \to \bullet x \ [\$]$, $L \to \bullet{*}E \ [\$]$

Exemple, complet

$$Z \to S\$ \quad (0) \qquad L \to x \quad (3)$$
$$S \to L{=}E \quad (1) \qquad \mid {*}E \quad (4)$$
$$\mid E \quad (2) \qquad E \to L \quad (5)$$

état	x	$*$	$=$	$\$$	S	L	E
0	d.4	d.5			d.1	d.2	d.3
1				d.6			

état	productions pointées élargies
0	$Z \to \bullet S\$ \ [\varepsilon]$, $S \to \bullet L{=}E \ [\$]$
	$S \to \bullet E \ [\$]$, $L \to \bullet x \ [=]$
	$L \to \bullet{*}E \ [=]$, $E \to \bullet L \ [\$]$
	$L \to \bullet x \ [\$]$, $L \to \bullet{*}E \ [\$]$
1	$Z \to S\bullet\$ \ [\varepsilon]$
2	$S \to L\bullet{=}E \ [\$]$, $E \to L\bullet \ [\$\checkmark]$
3	$S \to E\bullet \ [\$\checkmark]$
4	$L \to x\bullet \ [=\checkmark]$, $L \to x\bullet \ [\$\checkmark]$
5	$L \to {*}\bullet E \ [=]$, $L \to {*}\bullet E \ [\$]$
	$E \to \bullet L \ [=]$, $L \to \bullet x \ [=]$
	$L \to \bullet{*}E \ [=]$, $E \to \bullet L \ [\$]$
	$L \to \bullet x \ [\$]$, $L \to \bullet{*}E \ [\$]$
6	$Z \to S\$\bullet \ [\varepsilon\checkmark]$

Exemple, complet

$$Z \to S\$ \quad (0) \qquad L \to x \quad (3)$$
$$S \to L{=}E \quad (1) \qquad \mid *E \quad (4)$$
$$\mid E \quad (2) \qquad E \to L \quad (5)$$

état	x	$*$	$=$	$\$$	S	L	E
0	d.4	d.5			d.1	d.2	d.3
1				d.6			
2			d.7	r.5			

état	productions pointées élargies
0	$Z \to \bullet S\$ \; [\varepsilon]$, $S \to \bullet L{=}E \; [\$]$ $S \to \bullet E \; [\$]$, $L \to \bullet x \; [=]$ $L \to \bullet *E \; [=]$, $E \to \bullet L \; [\$]$ $L \to \bullet x \; [\$]$, $L \to \bullet *E \; [\$]$
1	$Z \to S \bullet \$ \; [\varepsilon]$
2	$S \to L \bullet {=}E \; [\$]$, $E \to L \bullet \; [\$✓]$
3	$S \to E \bullet \; [\$✓]$
4	$L \to x \bullet \; [=✓]$, $L \to x \bullet \; [\$✓]$
5	$L \to * \bullet E \; [=]$, $L \to * \bullet E \; [\$]$ $E \to \bullet L \; [=]$, $L \to \bullet x \; [=]$ $L \to \bullet *E \; [=]$, $E \to \bullet L \; [\$]$ $L \to \bullet x \; [\$]$, $L \to \bullet *E \; [\$]$
6	$Z \to S\$ \bullet \; [\varepsilon✓]$
7	$S \to L{=}\bullet E \; [\$]$

Exemple, complet

$$
\begin{aligned}
Z &\to S\$ && (0) & L &\to x && (3)\\
S &\to L{=}E && (1) & &\mid {*}E && (4)\\
&\mid E && (2) & E &\to L && (5)
\end{aligned}
$$

état	x	$*$	$=$	$\$$	S	L	E
0	d.4	d.5			d.1	d.2	d.3
1				d.6			
2			d.7	r.5			

état	productions pointées élargies
0	$Z \to \bullet S\$\ [\varepsilon]$, $S \to \bullet L{=}E\ [\$]$
	$S \to \bullet E\ [\$]$, $L \to \bullet x\ [=]$
	$L \to \bullet {*}E\ [=]$, $E \to \bullet L\ [\$]$
	$L \to \bullet x\ [\$]$, $L \to \bullet {*}E\ [\$]$
1	$Z \to S\bullet\$\ [\varepsilon]$
2	$S \to L\bullet{=}E\ [\$]$, $E \to L\bullet\ [\$\checkmark]$
3	$S \to E\bullet\ [\$\checkmark]$
4	$L \to x\bullet\ [=\checkmark]$, $L \to x\bullet\ [\$\checkmark]$
5	$L \to {*}\bullet E\ [=]$, $L \to {*}\bullet E\ [\$]$
	$E \to \bullet L\ [=]$, $L \to \bullet x\ [=]$
	$L \to \bullet {*}E\ [=]$, $E \to \bullet L\ [\$]$
	$L \to \bullet x\ [\$]$, $L \to \bullet {*}E\ [\$]$
6	$Z \to S\$\bullet\ [\varepsilon\checkmark]$
7	$S \to L{=}\bullet E\ [\$]$, $E \to \bullet L\ [\$]$
	$L \to \bullet x\ [\$]$, $L \to \bullet {*}E\ [\$]$

Exemple, complet

$$
\begin{array}{rcll@{\qquad}rcll}
Z & \to & S\$ & (0) & L & \to & x & (3)\\
S & \to & L{=}E & (1) & & | & *E & (4)\\
 & | & E & (2) & E & \to & L & (5)
\end{array}
$$

état	x	$*$	$=$	$\$$	S	L	E
0	d.4	d.5			d.1	d.2	d.3
1				d.6			
2			d.7	r.5			
3				r.2			
4			r.3	r.3			

état	productions pointées élargies
0	$Z \to \bullet S\$\ [\varepsilon]$, $S \to \bullet L{=}E\ [\$]$
	$S \to \bullet E\ [\$]$, $L \to \bullet x\ [=]$
	$L \to \bullet * E\ [=]$, $E \to \bullet L\ [\$]$
	$L \to \bullet x\ [\$]$, $L \to \bullet * E\ [\$]$
1	$Z \to S \bullet \$\ [\varepsilon]$
2	$S \to L \bullet {=} E\ [\$]$, $E \to L \bullet\ [\$\checkmark]$
3	$S \to E \bullet\ [\$\checkmark]$
4	$L \to x \bullet\ [=\checkmark]$, $L \to x \bullet\ [\$\checkmark]$
5	$L \to * \bullet E\ [=]$, $L \to * \bullet E\ [\$]$
	$E \to \bullet L\ [=]$, $L \to \bullet x\ [=]$
	$L \to \bullet * E\ [=]$, $E \to \bullet L\ [\$]$
	$L \to \bullet x\ [\$]$, $L \to \bullet * E\ [\$]$
6	$Z \to S\$ \bullet\ [\varepsilon\checkmark]$
7	$S \to L{=}\bullet E\ [\$]$, $E \to \bullet L\ [\$]$
	$L \to \bullet x\ [\$]$, $L \to \bullet * E\ [\$]$

Exemple, complet

$$
\begin{array}{llll}
Z \to S\$ & (0) & L \to x & (3) \\
S \to L{=}E & (1) & \quad \mid *E & (4) \\
\quad \mid E & (2) & E \to L & (5)
\end{array}
$$

état	x	$*$	$=$	$	S	L	E
0	d.4	d.5			d.1	d.2	d.3
1				d.6			
2			d.7	r.5			
3				r.2			
4			r.3	r.3			
5	d.4	d.5				d.9	d.8

état	productions pointées élargies
0	$Z \to \bullet S\$\ [\varepsilon]$, $S \to \bullet L{=}E\ [\$]$ $S \to \bullet E\ [\$]$, $L \to \bullet x\ [=]$ $L \to \bullet * E\ [=]$, $E \to \bullet L\ [\$]$ $L \to \bullet x\ [\$]$, $L \to \bullet * E\ [\$]$
1	$Z \to S \bullet \$\ [\varepsilon]$
2	$S \to L \bullet {=} E\ [\$]$, $E \to L \bullet\ [\$\checkmark]$
3	$S \to E \bullet\ [\$\checkmark]$
4	$L \to x \bullet\ [=\checkmark]$, $L \to x \bullet\ [\$\checkmark]$
5	$L \to * \bullet E\ [=]$, $L \to * \bullet E\ [\$]$ $E \to \bullet L\ [=]$, $L \to \bullet x\ [=]$ $L \to \bullet * E\ [=]$, $E \to \bullet L\ [\$]$ $L \to \bullet x\ [\$]$, $L \to \bullet * E\ [\$]$
6	$Z \to S\$ \bullet\ [\varepsilon\checkmark]$
7	$S \to L{=}\bullet E\ [\$]$, $E \to \bullet L\ [\$]$ $L \to \bullet x\ [\$]$, $L \to \bullet * E\ [\$]$
8	$L \to * E \bullet\ [=\checkmark]$, $L \to * E \bullet\ [\$\checkmark$
9	$E \to L \bullet\ [=\checkmark]$, $E \to L \bullet\ [\$\checkmark]$

Exemple, complet

$Z \to S\$$	(0)		$L \to x$	(3)
$S \to L{=}E$	(1)		$\mid *E$	(4)
$\mid E$	(2)		$E \to L$	(5)

état	x	$*$	$=$	$\$$	S	L	E
0	d.4	d.5			d.1	d.2	d.3
1				d.6			
2			d.7	r.5			
3				r.2			
4			r.3	r.3			
5	d.4	d.5				d.9	d.8
6				— accepter —			
7	d.12	d.13				d.11	d.10
8			r.4				
9			r.5				
10				r.1			
11				r.5			
12				r.3			
13	d.12	d.13				d.11	d.14
14				r.4			

état	productions pointées élargies
0	$Z \to \bullet S\$\ [\varepsilon]$, $S \to \bullet L{=}E\ [\$]$ $S \to \bullet E\ [\$]$, $L \to \bullet x\ [=]$ $L \to \bullet {*}E\ [=]$, $E \to \bullet L\ [\$]$ $L \to \bullet x\ [\$]$, $L \to \bullet {*}E\ [\$]$
1	$Z \to S \bullet \$\ [\varepsilon]$
2	$S \to L \bullet {=}E\ [\$]$, $E \to L \bullet\ [\$\checkmark]$
3	$S \to E \bullet\ [\$\checkmark]$
4	$L \to x \bullet\ [=\checkmark]$, $L \to x \bullet\ [\$\checkmark]$
5	$L \to {*}\bullet E\ [=]$, $L \to {*}\bullet E\ [\$]$ $E \to \bullet L\ [=]$, $L \to \bullet x\ [=]$ $L \to \bullet {*}E\ [=]$, $E \to \bullet L\ [\$]$ $L \to \bullet x\ [\$]$, $L \to \bullet {*}E\ [\$]$
6	$Z \to S\$ \bullet\ [\varepsilon\checkmark]$
7	$S \to L{=}\bullet E\ [\$]$, $E \to \bullet L\ [\$]$ $L \to \bullet x\ [\$]$, $L \to \bullet {*}E\ [\$]$
8	$L \to {*}E \bullet\ [=\checkmark]$, $L \to {*}E \bullet\ [\$\checkmark$
9	$E \to L \bullet\ [=\checkmark]$, $E \to L \bullet\ [\$\checkmark]$
10	$S \to L{=}E \bullet\ [\$\checkmark]$
11	$E \to L \bullet\ [\$\checkmark]$
12	$L \to x \bullet\ [\$\checkmark]$
13	$L \to {*}\bullet E\ [\$]$, $E \to \bullet L\ [\$]$ $L \to \bullet x\ [\$]$, $L \to \bullet {*}E\ [\$]$
14	$L \to {*}E \bullet\ [\$\checkmark]$

Parsage LALR(1) et GLR

Exemple, bis

$$Z \to S\$ \quad (0)$$
$$S \to L{=}E \quad (1)$$
$$\mid E \quad (2)$$
$$L \to x \quad (3)$$
$$\mid {*}E \quad (4)$$
$$E \to L \quad (5)$$

état	productions pointées élargies
0	$Z \to \bullet S\$ \; [\varepsilon]$, $S \to \bullet L{=}E \; [\$]$, $S \to \bullet E \; [\$]$, $L \to \bullet x \; [=]$
	$L \to \bullet {*}E \; [=]$, $E \to \bullet L \; [\$]$, $L \to \bullet x \; [\$]$, $L \to \bullet {*}E \; [\$]$
1	$Z \to S \bullet \$ \; [\varepsilon]$
2	$S \to L \bullet {=}E \; [\$]$, $E \to L \bullet \; [\$\checkmark]$
3	$S \to E \bullet \; [\$\checkmark]$
4	$L \to x \bullet \; [=\checkmark]$, $L \to x \bullet \; [\$\checkmark]$
5	$L \to {*}\bullet E \; [=]$, $L \to {*}\bullet E \; [\$]$, $E \to \bullet L \; [=]$, $L \to \bullet x \; [=]$
	$L \to \bullet {*}E \; [=]$, $E \to \bullet L \; [\$]$, $L \to \bullet x \; [\$]$, $L \to \bullet {*}E \; [\$]$
6	$Z \to S\$ \bullet \; [\varepsilon\checkmark]$
7	$S \to L{=}\bullet E \; [\$]$, $E \to \bullet L \; [\$]$, $L \to \bullet x \; [\$]$, $L \to \bullet {*}E \; [\$]$
8	$L \to {*}E \bullet \; [=\checkmark]$, $L \to {*}E \bullet \; [\$\checkmark]$
9	$E \to L \bullet \; [=\checkmark]$, $E \to L \bullet \; [\$\checkmark]$
10	$S \to L{=}E \bullet \; [\$\checkmark]$
11	$E \to L \bullet \; [\$\checkmark]$
12	$L \to x \bullet \; [\$\checkmark]$
13	$L \to {*}\bullet E \; [\$]$, $E \to \bullet L \; [\$]$, $L \to \bullet x \; [\$]$, $L \to \bullet {*}E \; [\$]$
14	$L \to {*}E \bullet \; [\$\checkmark]$

Exemple, bis

$$
\begin{aligned}
Z &\to S\$ && (0)\\
S &\to L{=}E && (1)\\
&\mid E && (2)\\
L &\to x && (3)\\
&\mid *E && (4)\\
E &\to L && (5)
\end{aligned}
$$

état	productions pointées élargies
0	$Z \to \bullet S\$\ [\varepsilon]$, $S \to \bullet L{=}E\ [\$]$, $S \to \bullet E\ [\$]$, $L \to \bullet x\ [=]$
	$L \to \bullet * E\ [=]$, $E \to \bullet L\ [\$]$, $L \to \bullet x\ [\$]$, $L \to \bullet * E\ [\$]$
1	$Z \to S \bullet \$\ [\varepsilon]$
2	$S \to L \bullet {=} E\ [\$]$, $E \to L \bullet\ [\$ \checkmark]$
3	$S \to E \bullet\ [\$ \checkmark]$
4	$L \to x \bullet\ [= \checkmark]$, $L \to x \bullet\ [\$ \checkmark]$
5	$L \to * \bullet E\ [=]$, $L \to * \bullet E\ [\$]$, $E \to \bullet L\ [=]$, $L \to \bullet x\ [=]$
	$L \to \bullet * E\ [=]$, $E \to \bullet L\ [\$]$, $L \to \bullet x\ [\$]$, $L \to \bullet * E\ [\$]$
6	$Z \to S\$ \bullet\ [\varepsilon \checkmark]$
7	$S \to L{=}\bullet E\ [\$]$, $E \to \bullet L\ [\$]$, $L \to \bullet x\ [\$]$, $L \to \bullet * E\ [\$]$
8	$L \to * E \bullet\ [= \checkmark]$, $L \to * E \bullet\ [\$ \checkmark]$
9	$E \to L \bullet\ [= \checkmark]$, $E \to L \bullet\ [\$ \checkmark]$
10	$S \to L{=}E \bullet\ [\$ \checkmark]$
11	$E \to L \bullet\ [\$ \checkmark]$
12	$L \to x \bullet\ [\$ \checkmark]$
13	$L \to * \bullet E\ [\$]$, $E \to \bullet L\ [\$]$, $L \to \bullet x\ [\$]$, $L \to \bullet * E\ [\$]$
14	$L \to * E \bullet\ [\$ \checkmark]$

Parsage LALR(1)

Définition

Deux productions pointées élargies $A \to \alpha \bullet \beta$ $[a]$ et $A \to \alpha' \bullet \beta'$ $[b]$ sont équivalent LALR(1) si $\alpha = \alpha'$ et $\beta = \beta'$.

- les items sont identiques, mais les contextes peuvent être différents

Définition

L'automate LALR(1) d'une grammaire hors-contexte G est le quotient de l'automate LR(1) de G sous équivalence LALR(1).

Exemple, ter

$$
\begin{array}{rcll}
Z & \to & S\$ & (0) \\
S & \to & L=E & (1) \\
  & | & E & (2) \\
L & \to & x & (3) \\
  & | & *E & (4) \\
E & \to & L & (5)
\end{array}
$$

état	x	$*$	$=$	$\$$	S	L	E
0	d.4	d.5			d.1	d.2	d.3
1				d.6			
2			d.7	r.5			
3				r.2			
4			r.3	r.3			
5	d.4	d.5				d.9	d.8
6			— accepter —				
7	d.12	d.13				d.11	d.10
8			r.4				
9			r.5				
10				r.1			
11				r.5			
12				r.3			
13	d.12	d.13				d.11	d.14
14				r.4			

Exemple, ter

$$
\begin{array}{lll}
Z \to S\$ & & (0) \\
S \to L{=}E & & (1) \\
\quad \mid E & & (2) \\
L \to x & & (3) \\
\quad \mid *E & & (4) \\
E \to L & & (5)
\end{array}
$$

état	x	$*$	$=$	$\$$	S	L	E
0	d.4	d.5			d.1	d.2	d.3
1				d.6			
2			d.7	r.5			
3				r.2			
4			r.3	r.3			
5	d.4	d.5				d.9	d.8
6			— accepter —				
7	d.12	d.13				d.11	d.10
8			r.4	r.4			
9			r.5	r.5			
10				r.1			

Résolution de conflits

Exemple :

$$
\begin{aligned}
Z &\to E\$ && (0)\\
E &\to E+E && (1)\\
&\mid E*E && (2)\\
&\mid n && (3)
\end{aligned}
$$

état	+	*	n	$	E
0			d.2		g.1
1	d.4	d.5		d.3	
2	r.3	r.3		r.3	
3	— accepter —				
4			d.2		g.6
5			d.2		g.7
6	d.4 r.1	d.5 r.1		r.1	
7	d.4 r.2	d.5 r.2		r.2	

- une grammaire ambiguë
- donc pas LR(k) pour n'importe quel k

Résolution de conflits

Exemple :

$$\begin{aligned} Z &\rightarrow E\$ && (0) \\ E &\rightarrow E+E && (1) \\ &\mid E*E && (2) \\ &\mid n && (3) \end{aligned}$$

état	+	*	n	\$	E
0			d.2		g.1
1	d.4	d.5		d.3	
2	r.3	r.3		r.3	
3	— accepter —				
4			d.2		g.6
5			d.2		g.7
6	d.4	d.5			
	r.1	r.1		r.1	
7	d.4	d.5			
	r.2	r.2		r.2	

- une grammaire ambiguë
- donc pas LR(k) pour n'importe quel k
- associativité : d.4 $\Rightarrow$ $n+(n+n)$; r.1 $\Rightarrow$ $(n+n)+n$
- priorité : d.5 $\Rightarrow$ $n*(n+n)$; r.1 $\Rightarrow$ $(n*n)+n$

Résolution de conflits

Exemple :

$$
\begin{aligned}
Z &\to E\$ && (0)\\
E &\to E+E && (1)\\
&\mid E*E && (2)\\
&\mid n && (3)
\end{aligned}
$$

état	+	*	n	$	E
0			d.2		g.1
1	d.4	d.5		d.3	
2	r.3	r.3		r.3	
3	— accepter —				
4			d.2		g.6
5			d.2		g.7
6	d.4	d.5			
	r.1	r.1		r.1	
7	d.4	d.5			
	r.2	r.2		r.2	

- une grammaire ambiguë
- donc pas LR(k) pour n'importe quel k
- associativité : d.4 $\Rightarrow n+(n+n)$; r.1 $\Rightarrow (n+n)+n$
- priorité : d.5 $\Rightarrow n*(n+n)$; r.1 $\Rightarrow (n*n)+n$
- solution : règles de priorité
- ici : r.1 $>$ d.4, r.2 $>$ d.5, r.2 $>$ d.4, d.5 $>$ r.1 $\Leftarrow$!

Parsage LR généralisé

- *embrace non-determinism!*
- parsage GLR : en cas de conflit, suivre tous les chemins en parallel
- « parsage parallel », « parsage Tomita »
- implémenter l'automate (non-déterministe) de parsage sans déterminisation
- états : productions pointées, pas de clôture
- algorithme en temps exponentiel, pas linéaire
- optimisation : partager préfixes et suffixes de piles


That's all Folks!