

Théorie des langages : THL

CM 8

Uli Fahrenberg

EPITA Rennes

S5 2021

Aperçu

Programme du cours

- ① Langages rationnels
- ② Automates finis
- ③ Langages algébriques, grammaires hors-contexte, automates à pile
- ④ Parsage LL
- ⑤ TP 1 : flex
- ⑥ Parsage LR, partie 1
- ⑦ Parsage LR, partie 2
- ⑧ Parsage LR, partie 3
- ⑨ TP 2 : bison
- ⑩ 3 : flex & bison

QCM 3

Re : passage ascendant : the basics

```
function 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  $A \rightarrow \alpha_i \dots \alpha_j$  then                            ▷ réduction / REDUCE  
          if BULRP( $\alpha_1 \dots \alpha_{i-1} A \alpha_{j+1} \dots \alpha_n$ ) then  
            return True  
  return False
```

Re : passage LR(0)

$$Z \rightarrow S\$ \quad (0)$$

$$S \rightarrow S-n \quad (1)$$

$$| n \quad (2)$$

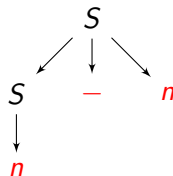
parser $n - n\$$:

entrée	pile	action
$n - n\$$	$\perp 0$	décaler
$-n\$$	$\perp 01$	réduire 2
$-n\$$	$\perp 02$	décaler
$n\$$	$\perp 024$	décaler
$\$$	$\perp 0245$	réduire 1
$\$$	$\perp 02$	décaler
	$\perp 023$	✓

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

$$S \rightarrow n$$

$$S \rightarrow S-n$$



Re : parcours SLR(1)

- ① calculer la table LR(0)
- ② si conflits : **conditionner** l'action par le FOLLOW

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

$S \rightarrow n-S$ (1)

$| n$ (2)

état	action	n	$-$	$\$$	S		état	n	$-$	$\$$	S
0	décaler	2			1		0	d.2			d.1
1	décaler			4			1			d.4	
2	rédu. 2, déc.		3			\Rightarrow	2		d.3	r.2	
3	décaler	2			5		3	d.2			d.5
4	accepter						4		— accepter —		
5	réduire 1						5			r.1	

Re : parcour LR(1)

- conditionner l'action par le **contexte** : les symboles qui peuvent suivre

Exemple :

 $Z \rightarrow S\$ \quad (0)$
 $S \rightarrow L=E \quad (1)$
 $| E \quad (2)$
 $L \rightarrow x \quad (3)$
 $| *E \quad (4)$
 $E \rightarrow L \quad (5)$

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

Exemple

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

Parsage LALR(1) et GLR

Exemple, bis

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

Exemple, bis

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

Parsage LALR(1)

Définition

Deux productions pointées élargies $A \rightarrow \alpha \bullet \beta [a]$ et $A \rightarrow \alpha' \bullet \beta' [b]$ 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

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

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

$Z \rightarrow E\$$ (0)

$E \rightarrow E+E$ (1)

| $E * E$ (2)

| n (3)

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

Résolution de conflits

Exemple :

$Z \rightarrow E\$$ (0)

$E \rightarrow E+E$ (1)

| $E * E$ (2)

| n (3)

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 :

$Z \rightarrow E\$$ (0)

$E \rightarrow E+E$ (1)

| $E * E$ (2)

| n (3)

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 parallèle**
- « parsage parallèle », « parsage Tomita »
- implémenter l'automate (non-déterministe) de parsage sans détermination
- états : productions pointées, **pas de clôture**
- algorithme en temps **exponentiel**, pas linéaire
- optimisation : partager préfixes et suffixes de piles

Résumé du cours

Hiérarchie de Chomsky

type	langages	grammaires	automates
4	finis ⋈	à choix finis ⋈	finis acycliques ⋈
3	réguliers ⋈	régulières ⋈	finis ⋈
2	algébriques ⋈	hors-contexte ⋈	à pile
1	contextuels ⋈	contextuelles ⋈	linéairement bornés ⋈
0	récurivement énumérables	syntagmatiques	de Turing

Hiérarchie de Chomsky

type	langages	grammaires	automates
4	finis ⋈	à choix finis ⋈	finis acycliques ⋈
3	réguliers ⋈	régulières ⋈	finis ⋈
2	algébriques ⋈	hors-contexte ⋈	à pile
1	contextuels ⋈	contextuelles	linéairement bornés
0	récurivement énumérables	syntagmatiques	de Turing

Zoom sur type 3

syntaxe

aut. finis dét. complets

\cap

aut. finis déterministes

\cap

automates finis

\cap

aut. finis à trans. spontanées

expressions rationnelles

grammaires régulières

$\xrightarrow{L(\cdot)}$

sémantique

langages reconnaissables

\parallel

langages reconnaissables

\parallel

langages reconnaissables

\parallel

langages reconnaissables

\parallel

langages rationnelles

\parallel

langages réguliers

Zoom sur type 2

syntaxe

grammaires hc forme Greibach

\cap

grammaires hors-contexte

\cup

grammaires hc forme Chomsky

$\xrightarrow{L(\cdot)}$

automates à pile

\cup

automates à pile sans trans. spont.

\cup

automates à pile déterministes

sémantique

langages algébriques

\parallel

langages algébriques

\parallel

langages algébriques

\parallel

langages algébriques

\parallel

langages algébriques

\cup

langages algébriques déterministes

Zoom sur LR

syntaxe

sémantique

grammaires hors-contexte
 \cup
 grammaires hc non-ambiguës
 \cup
 grammaires hc déterministes
 \cup
 grammaires LR(k)
 \cup
 \dots
 \cup
 grammaires LR(1)
 \cup
 grammaires LALR(1)
 \cup
 grammaires SLR(1)
 \cup
 grammaires LR(0)

$L(\cdot)$
 \longrightarrow

langages algébriques
 \cup
 lang. alg. non-ambiguës
 \cup
 lang. alg. déterministes
 \parallel
 lang. alg. déterministes
 \dots
 \parallel
 lang. alg. déterministes
 \cup
 langages LALR(1)
 \cup
 langages SLR(1)
 \cup
 langages LR(0)

Zoom sur LL

syntaxe

grammaires hors-contexte

⊂

grammaires hc non-ambiguës

⊂

grammaires hc déterministes

⊂

grammaires LL(k)

⊂

...

⊂

grammaires LL(2)

⊂

grammaires LL(1)

$\xrightarrow{L(\cdot)}$

sémantique

langages algébriques

⊂

lang. alg. non-ambiguës

⊂

lang. alg. déterministes

⊂

langages LL(k)

⊂

...

⊂

langages LL(2)

⊂

langages LL(1)

The image features a classic target graphic with concentric circles. The outer rings are a deep red, while the inner rings transition to a lighter, more vibrant red. At the very center is a solid dark blue circle. Overlaid on this target is the text "That's all Folks!" in a white, elegant cursive script. The text is positioned diagonally, starting from the lower left and ending towards the upper right, with the central blue bullseye acting as a focal point for the phrase.

That's all Folks!