

Analísadores Descendentes Tabulares; Cjs First Follow

Marcelo Johann

Conteúdo da aula

Analísadores Descendentes

- Recursivos com Retrocesso
- Recursivos Preditivos
- Conjunto FIRST e Implementação

Analísador Preditivo Tabular

Conjuntos FIRST e FOLLOW

Montagem da Tabela

Implementação

YACC

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 2

Definição: Conjuntos "First"

- Seja α qualquer seqüência de símbolos – terminais ou não terminais

• First(α):

– Definição informal:

- conjunto de todos os terminais que começam qualquer seqüência derivável de α .

– Definição formal:

- Se existe um $t \in T$ e um $\beta \in V^*$ tal que $\alpha \Rightarrow^* t \beta$ então $t \in \text{First}(\alpha)$
- Se $\alpha \Rightarrow^* \epsilon$ então $\epsilon \in \text{First}(\alpha)$

$A \rightarrow B C D$	$\text{First}(A) = \{b, c, d\}$
$B \rightarrow b$	$\text{First}(B) = \{b\}$
$C \rightarrow c$	$\text{First}(C) = \{c\}$
$D \rightarrow d$	$\text{First}(D) = \{d\}$

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 3

Condição para que se possa usar um analisador preditivo

- **Informalmente:** no caso que em os First() dos lados direitos das regras de produção sejam "simpáticos", não terá retrocesso.

- **Formalmente:** para qualquer produção

$$A \rightarrow \alpha_1 | \alpha_2 | \dots | \alpha_n,$$

quer-se:

$$\text{First}(\alpha_1) \cap \text{First}(\alpha_2) \cap \dots \cap \text{First}(\alpha_n) = \emptyset$$

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 4

proc First(α : string of symbols)

```
Repeat {
  Para todas as produções  $\alpha \rightarrow X_1 X_2 X_3 \dots X_n$  do
    if  $X_1 \in T$  then // caso simples onde  $X_1$  é um terminal
      First( $\alpha$ ) := First( $\alpha$ )  $\cup$  { $X_1$ }
    else { // caso menos simples:  $X_1$  é um não-terminal
      First( $\alpha$ ) = First( $\alpha$ )  $\cup$  First( $X_1$ )  $\setminus$  { $\epsilon$ };
      for (i=1 ; i<=n ; i++) {
        if  $\epsilon$  is in First( $X_i$ ) and in First( $X_{i+1}$ ) and in ... First( $X_{i-1}$ )
          First( $\alpha$ ) := First( $\alpha$ )  $\cup$  First( $X_i$ )  $\setminus$  { $\epsilon$ }
      }
    }
  if ( $\alpha \Rightarrow^* \epsilon$ )
    then First( $\alpha$ ) := First( $\alpha$ )  $\cup$  { $\epsilon$ }
} until no change in any First( $\alpha$ )
```

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 5

proc Follow($A \in N$)

```
Follow(S) := {$};
Repeat
  foreach p  $\in$  P do { // Laço sobre as produções
    case p ==  $X \rightarrow \alpha A$  // a produção termina por A
      Follow(A) := Follow(A)  $\cup$  Follow(X);

    case p ==  $X \rightarrow \alpha A \beta$  // a produção NÃO termina por A
      Follow(A) := Follow(A)  $\cup$  First( $\beta$ )  $\setminus$  { $\epsilon$ };
      if  $\epsilon \in \text{First}(\beta)$  then
        Follow(A) := Follow(A)  $\cup$  Follow(X);
      end
  }
} until no change in any Follow()
```

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 6

Observações First/Follow

- Só **terminais** entram em First e Follow.
- O algoritmo de cálculo de $\text{First}(\alpha)$:
 - É trivial quando α é um **terminal** t .
 - varre as **produções** $X \rightarrow t\omega$ quando α é um não-terminal X ;
 - é chato quando o início de uma derivação se α deriva em ϵ .
 - Inclui ϵ apenas quando α pode derivar em ϵ .
- O algoritmo de cálculo de $\text{Follow}(A)$
 - É reservado aos não-terminais A
 - Inclui o $\$$ em alguns casos triviais ($A == \text{start } S$)
 - Varre as produções onde A aparece à direita ($X \rightarrow \omega A \omega'$)
 - É chato quando A aparece no fim (ou logo antes de algo que deriva em ϵ)
 - NUNCA** inclui ϵ

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 7

Exemplo First/Follow

$S \rightarrow XYZ$
 $X \rightarrow aXb \mid \epsilon$
 $Y \rightarrow cYZcX \mid d$
 $Z \rightarrow eZYe \mid f$

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 8

Exemplo First/Follow

$S \rightarrow XYZ$
 $X \rightarrow aXb \mid \epsilon$
 $Y \rightarrow cYZcX \mid d$
 $Z \rightarrow eZYe \mid f$

$\text{First}(X) = \{a, \epsilon\}$ $\text{Follow}(X) = \{c, d, b, e, f\}$
 $\text{First}(Y) = \{c, d\}$ $\text{Follow}(Y) = \{e, f\}$
 $\text{First}(Z) = \{e, f\}$ $\text{Follow}(Z) = \{\$, c, d\}$
 $\text{First}(S) = \{a, c, d\}$ $\text{Follow}(S) = \{\$\}$

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 9

Gramática LL(1)

Condições necessárias:

- sem ambigüidade
- sem recursão a esquerda

Uma gramática G é LL(1) se e somente se

Para qualquer produção do tipo: $A \rightarrow \alpha \mid \beta \Rightarrow^* t$

- $\text{First}(\alpha) \cap \text{First}(\beta) = \emptyset$
- $\alpha \Rightarrow^* \epsilon$ implies $\neg(\beta \Rightarrow^* \epsilon)$
- $\alpha \Rightarrow^* \epsilon$ implies $\text{First}(\beta) \cap \text{Follow}(A) = \emptyset$

LL(1) = leitura **L**eft -> right + derivação mais a esquerda
(**L**eft) + uso de **1** token lookahead.

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 10

Análise top-down com tabela preditiva

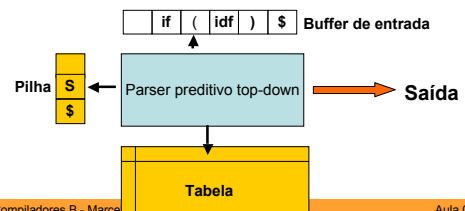
- Os dois métodos apresentados até agora para fazer análise descendente (top-down) usam recursividade.
 - Cada não-terminal tem um procedimento associado;
 - Há chamadas com ou sem retrocesso.
 - Para gramáticas LL1 não tem retrocesso.
- Chamadas recursivas usam uma pilha implícita
 - A pilha das chamadas!
 - Sobrecusto!
- Idéia: de-recursificar o procedimento:
 - Usa-se uma pilha para armazenar os não-terminais encontrados;
 - Usa-se uma **tabela** para orientar as derivações.
 - Ver o uso de tabelas para derecurisificar um algoritmo em programação dinâmica.

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 11

Reconhecedor preditivo com Pilha

- Tem um buffer de símbolos em entrada;
 - $\$$ marca seu fim.
- Tem um fluxo de saída;
- Usa uma pilha cujo fundo é marcado por $\$$
 - Inicializada com S (Start)
- Usa uma **tabela sintática preditiva** M



INF01033 - Compiladores B - Marce

Aula 08 : Slide 12

Funcionamento do parser

- Seja X o símbolo no topo da pilha
- Seja a o símbolo de entrada (terminal/token de lookahead) a analisar
- Etapas:
 1. Se $X == \$$ e $a == \$$: para e reconheceu uma sentença.
 2. Se $X == a$ e $a != \$$: desempilha X e avança de um símbolo na entrada.
 3. Se X é um não-terminal:
 - Consulta a tabela $M(X, a)$
 - Se for vazia: ERRO
 - Se contém $X \rightarrow UVW$, então substitui na pilha X por UVW (U no topo).

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 13

Como construir a tabela?

- (Re-escrever a gramática para satisfazer condições de LL(1) – isso é um pre-requisito!)
- Calcular os conjuntos First e Follow
- Para cada produção $A \rightarrow \alpha$ na gramática:
 1. Para cada terminal $a \in \text{First}(\alpha)$
 - incluir a produção $A \rightarrow \alpha$ em $M[A, a]$
 2. Se $\epsilon \in \text{First}(\alpha)$
 - incluir a produção $A \rightarrow \alpha$ em $M[A, b]$ para cada b em $\text{Follow}(A)$
 3. Se $\epsilon \in \text{First}(\alpha)$ e $\$ \in \text{Follow}(A)$
 - incluir a produção $A \rightarrow \alpha$ em $M[A, \$]$
- Todas as entradas não definidas são erros

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 14

A tabela preditiva $M(X, t)$

- Tabela Bi-dimensional:
 - Dimensão 1: Não-terminal X
 - Dimensão 2: Caractere da entrada (terminal) t
 - A entrada (X, t) contém a regra de produção a aplicar

	a	b	c	\$
S				
A				
B				

$S \rightarrow cAa$ $\text{First}(A) = \{b, c, \epsilon\}$ $\text{Follow}(S) = \{\$\}$
 $A \rightarrow cB \mid B$ $\text{First}(B) = \{b, \epsilon\}$ $\text{Follow}(A) = \{a\}$
 $B \rightarrow bcB \mid \epsilon$ $\text{First}(S) = \{c\}$ $\text{Follow}(B) = \{a\}$

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 15

Exemplo de tabela $M(X, t)$

- Tabela Bi-dimensional:
 - Dimensão 1: Não-terminal X
 - Dimensão 2: Caractere da entrada (terminal) t
 - A entrada (X, t) contém a regra de produção a aplicar

	a	b	c	\$
S			$S \rightarrow cAa$	
A				
B				

$S \rightarrow cAa$ $\text{First}(A) = \{b, c, \epsilon\}$ $\text{Follow}(S) = \{\$\}$
 $A \rightarrow cB \mid B$ $\text{First}(B) = \{b, \epsilon\}$ $\text{Follow}(A) = \{a\}$
 $B \rightarrow bcB \mid \epsilon$ $\text{First}(S) = \{c\}$ $\text{Follow}(B) = \{a\}$

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 16

Exemplo de tabela $M(X, t)$

- Tabela Bi-dimensional:
 - Dimensão 1: Não-terminal X
 - Dimensão 2: Caractere da entrada (terminal) t
 - A entrada (X, t) contém a regra de produção a aplicar

	a	b	c	\$
S			$S \rightarrow cAa$	
A			$A \rightarrow cB$	
B				

$S \rightarrow cAa$ $\text{First}(A) = \{b, c, \epsilon\}$ $\text{Follow}(S) = \{\$\}$
 $A \rightarrow cB \mid B$ $\text{First}(B) = \{b, \epsilon\}$ $\text{Follow}(A) = \{a\}$
 $B \rightarrow bcB \mid \epsilon$ $\text{First}(S) = \{c\}$ $\text{Follow}(B) = \{a\}$

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 17

Exemplo de tabela $M(X, t)$

- Tabela Bi-dimensional:
 - Dimensão 1: Não-terminal X
 - Dimensão 2: Caractere da entrada (terminal) t
 - A entrada (X, t) contém a regra de produção a aplicar

	a	b	c	\$
S			$S \rightarrow cAa$	
A		$A \rightarrow B$	$A \rightarrow cB$	
B				

$S \rightarrow cAa$ $\text{First}(A) = \{b, c, \epsilon\}$ $\text{Follow}(S) = \{\$\}$
 $A \rightarrow cB \mid B$ $\text{First}(B) = \{b, \epsilon\}$ $\text{Follow}(A) = \{a\}$
 $B \rightarrow bcB \mid \epsilon$ $\text{First}(S) = \{c\}$ $\text{Follow}(B) = \{a\}$

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 18

Exemplo de tabela M(X, t)

- Tabela Bi-dimensional:
 - Dimensão 1: Não-terminal X
 - Dimensão 2: Caractere da entrada (terminal) t
 - A entrada (X,t) contém a regra de produção a aplicar

	a	b	c	\$
S			S → cAa	
A	A → B	A → B	A → cB	
B				

$S \rightarrow cAa$ $First(A) = \{b, c, \epsilon\}$ $Follow(S) = \{\$, \}$
 $A \rightarrow cB \mid B$ $First(B) = \{b, \epsilon\}$ $Follow(A) = \{a\}$
 $B \rightarrow bcB \mid \epsilon$ $First(S) = \{c\}$ $Follow(B) = \{a\}$

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 19

Exemplo de tabela M(X, t)

- Tabela Bi-dimensional:
 - Dimensão 1: Não-terminal X
 - Dimensão 2: Caractere da entrada (terminal) t
 - A entrada (X,t) contém a regra de produção a aplicar

	a	b	c	\$
S			S → cAa	
A	A → B	A → B	A → cB	
B		B → bcB		

$S \rightarrow cAa$ $First(A) = \{b, c, \epsilon\}$ $Follow(S) = \{\$, \}$
 $A \rightarrow cB \mid B$ $First(B) = \{b, \epsilon\}$ $Follow(A) = \{a\}$
 $B \rightarrow bcB \mid \epsilon$ $First(S) = \{c\}$ $Follow(B) = \{a\}$

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 20

Exemplo de tabela M(X, t)

- Tabela Bi-dimensional:
 - Dimensão 1: Não-terminal X
 - Dimensão 2: Caractere da entrada (terminal) t
 - A entrada (X,t) contém a regra de produção a aplicar

	a	b	c	\$
S			S → cAa	
A	A → B	A → B	A → cB	
B	B → ε	B → bcB		

$S \rightarrow cAa$ $First(A) = \{b, c, \epsilon\}$ $Follow(S) = \{\$, \}$
 $A \rightarrow cB \mid B$ $First(B) = \{b, \epsilon\}$ $Follow(A) = \{a\}$
 $B \rightarrow bcB \mid \epsilon$ $First(S) = \{c\}$ $Follow(B) = \{a\}$

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 21

Exemplo de tabela M(X, t)

- Tabela Bi-dimensional:
 - Dimensão 1: Não-terminal X
 - Dimensão 2: Caractere da entrada (terminal) t
 - A entrada (X,t) contém a regra de produção a aplicar

	a	b	c	\$
S	ERRO	ERRO	S → cAa	ERRO
A	A → B	A → B	A → cB	ERRO
B	B → ε	B → bcB	ERRO	ERRO

$S \rightarrow cAa$ $First(A) = \{b, c, \epsilon\}$ $Follow(S) = \{\$, \}$
 $A \rightarrow cB \mid B$ $First(B) = \{b, \epsilon\}$ $Follow(A) = \{a\}$
 $B \rightarrow bcB \mid \epsilon$ $First(S) = \{c\}$ $Follow(B) = \{a\}$

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 22

Usando a tabela

- String: "cbca"

Pilha	Entrada	Ação
S\$	cbca\$	S->cAa
cAa\$	cbca\$	casar c
Aa\$	bca\$	A->B
Ba\$	bca\$	B->bcB
bcBa\$	bca\$	casar b
cBa\$	ca\$	casar c
Ba\$	a\$	B->ε
a\$	a\$	casar a
\$	\$	casar \$, sucesso

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 23

Mais um exemplo...

$E \rightarrow TE'$
$E' \rightarrow +TE' \mid \epsilon$
$T \rightarrow FT'$
$T' \rightarrow *FT' \mid \epsilon$
$F \rightarrow (E) \mid id$

Símbolo	First	Follow
E	{(, id}	{\$,)}
E'	{+, ε}	{\$,)}
T	{(, id}	{+, \$,)}
T'	{*, ε}	{+, \$,)}
F	{(, id}	{*, +, \$,)}

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 24

Exemplo LL(1)

$S \rightarrow XYZ$

$X \rightarrow aXb \mid \epsilon$

$Y \rightarrow cYZcX \mid d$

$Z \rightarrow eZYe \mid f$

Construir Tabela
Analisar abcdfc

$\text{First}(X) = \{a, \epsilon\}$ $\text{Follow}(X) = \{c, d, b, e, f\}$

$\text{First}(Y) = \{c, d\}$ $\text{Follow}(Y) = \{e, f\}$

$\text{First}(Z) = \{e, f\}$ $\text{Follow}(Z) = \{\$, c, d\}$

$\text{First}(S) = \{a, c, d\}$ $\text{Follow}(S) = \{\$\}$

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 25

Observação sobre a Tabela

- A tabela indica se há ambigüidade!
 - Mais de uma regra numa entrada!
- Soluções?
 - Tornar a gramática LL(1)
 - Eliminar ambigüidade, recursividade...
 - Usar uma heurística para desempatar as regras
 - Qual?
 - Usar outros algoritmos do que os top-down!
- Exemplo total: if... Then... Else:
 - $S \rightarrow iEtS \mid iEtSeS \mid a$
 - $E \rightarrow b$

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 26

Retornando ao Gerador Yacc

Makefile: continuando...

Produções entre %% e %%

%token

%type, %union e yy.lval

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 27

Leituras e Tarefas sugeridas

Repetir os experimentos com yacc
feitos pelo professor com etapa1

Implementar uma outra gramática com
o código de analisador tabular

INF01033 - Compiladores B - Marcelo Johann - 2010/1

Aula 08 : Slide 28