

Analizador gramatical tabular por deslocamento e redução em Prolog*

Luiz Arthur Pagani (UFPR)

`http://people.ufpr.br/~arthur`
`arthur@ufpr.br`

*Esta apresentação foi preparada no sistema L^AT_EX, através de uma de suas implementações para Linux — o T_EX_E — e de um programa de integração (*IDE*) — o Kile — ambos instalados em computadores funcionando com o sistema operacional Kurumin (Linux), dentro das diretrizes do chamado *software livre*.

1 Introdução

- Gramática como descrição das unidades lingüísticas possíveis × analisador gramatical (*parser*) como procedimento de análise
- Ascendente (*bottom-up*) × Descendente (*top-down*)
- Pilha (*stack*) × Fila (*queue*)
- Algoritmos:
 - Gramática de Cláusula Definida (*Definite Clause Grammar*, DCG)
 - Analisador por Deslocamento e Redução (*Shift-Reduce Parser*)
 - Analisador Tabular (*Chart Parser*)

2 Analisador gramatical por deslocamento e redução

- Fila: expressão a ser analisada
- Pilha: resultados da análise
- Operações:
 - Deslocamento da fila para a pilha
 - Redução da pilha através de alguma regra da gramática

2.1 Exemplo

Gramática:

1. $S \rightarrow SN\ SV$
2. $SV \rightarrow V\ SN$
3. $SN \rightarrow \{\text{Pedro, Maria}\}$
4. $V \rightarrow \{\text{ama}\}$

Fila	Pilha	Operação
Pedro ama Maria		Deslocamento
ama Maria	Pedro	Redução (3)
ama Maria	SN	Deslocamento
Maria	$SN\ ama$	Redução (4)
Maria	$SN\ V$	Deslocamento
	$SN\ V\ Maria$	Redução (3)
	$SN\ V\ SN$	Redução (2)
	$SN\ SV$	Redução (1)
	S	Análise bem sucedida

3 Analisador gramatical tabular

- Tabela (*chart*): armazena informações parciais
 - Nós (*vertices*): posições entre as palavras
 - Arcos (*edges*): ligações entre os nós, rotulados com as operações
- Operações
 - Arco ativo: regra ainda em uso (Ex.: $S \rightarrow SN \bullet SV$)
 - Arco desativado: regra completa (Ex.: $S \rightarrow SN SV \bullet$)

3.1 Exemplo

1	Pedro	2	ama	3	Maria	4
$S \rightarrow \bullet SN SV$	SN	$S \rightarrow \bullet SN SV$	V	$S \rightarrow \bullet SN SV$	SN	
$SV \rightarrow \bullet V SN$		$SV \rightarrow \bullet V SN$		$SV \rightarrow \bullet V SN$		
$S \rightarrow SN \bullet SV$		$SV \rightarrow V \bullet SN$		$S \rightarrow SN \bullet SV$		
		$SV \rightarrow V SN \bullet$				
		$S \rightarrow SN SV \bullet$				

4 Analisador gramatical tabular com deslocamento e ‘expansão’

- Fila ainda com a expressão a ser analisada
- Deslocamento da fila para a tabela (da categoria, e não de uma expressão)
- Tabela no lugar da pilha
- Expansão ao invés de redução

Fila	Tabela	Operação
Pedro ama Maria		Deslocamento
ama Maria	$SN \rightarrow \text{Pedro} \bullet$ (entre 1 e 2)	Expansão
ama Maria	$SN \rightarrow \text{Pedro} \bullet$ (entre 1 e 2) $S \rightarrow SN \bullet SV$ (entre 1 e 2)	Deslocamento
Maria	$SN \rightarrow \text{Pedro} \bullet$ (entre 1 e 2) $S \rightarrow SN \bullet SV$ (entre 1 e 2) $V \rightarrow \text{ama} \bullet$ (entre 2 e 3)	Expansão
Maria	$SN \rightarrow \text{Pedro} \bullet$ (entre 1 e 2) $S \rightarrow SN \bullet SV$ (entre 1 e 2) $V \rightarrow \text{ama} \bullet$ (entre 2 e 3) $SV \rightarrow V \bullet SN$ (entre 2 e 3)	Deslocamento
	$SN \rightarrow \text{Pedro} \bullet$ (entre 1 e 2) $S \rightarrow SN \bullet SV$ (entre 1 e 2) $V \rightarrow \text{ama} \bullet$ (entre 2 e 3) $SV \rightarrow V \bullet SN$ (entre 2 e 3) $SN \rightarrow \text{Maria} \bullet$ (entre 3 e 4)	Expansão
	$SN \rightarrow \text{Pedro} \bullet$ (entre 1 e 2) $S \rightarrow SN \bullet SV$ (entre 1 e 2) $V \rightarrow \text{ama} \bullet$ (entre 2 e 3) $SV \rightarrow V \bullet SN$ (entre 2 e 3) $SN \rightarrow \text{Maria} \bullet$ (entre 3 e 4) $SV \rightarrow V SN \bullet$ (entre 2 e 4) $S \rightarrow SN SV \bullet$ (entre 1 e 4)	Análise bem sucedida

5 Implementação em Prolog

Baseado nos analisadores de Covington [1, ps. 151–195] e Matthews [2, ps. 214–243].

5.1 Representação dos arcos

- $SN \rightarrow \text{Pedro}$ • (entre 1 e 2): $\text{arco}(1, \text{sn} ==> [\text{pedro}]^\circ [], 2)$
- $SV \rightarrow V$ • SN (entre 2 e 3): $\text{arco}(2, \text{sv} ==> [v]^\circ [\text{sn}], 3)$
- $S \rightarrow SN SV$ • (entre 1 e 4): $\text{arco}(1, \text{s} ==> [\text{sn}, \text{sv}]^\circ [], 4)$

5.2 Resultado

```
?- analisa([pedro,ama,maria]).
```

```
arco(1, sn==>[pedro]°[ ], 2)
```

```
arco(1, s==>[ ]°[sn, sv], 1)
```

```
arco(1, s==>[sn]°[sv], 2)
```

```
arco(2, v==>[ama]°[ ], 3)
```

```
arco(2, sv==>[ ]°[v, sn], 2)
```

```
arco(2, sv==>[v]°[sn], 3)
```

```
arco(3, sn==>[maria]°[ ], 4)
```

```
arco(3, s==>[ ]°[sn, sv], 3)
```

```
arco(3, s==>[sn]°[sv], 4)
```

```
arco(2, sv==>[v, sn]°[ ], 4)
```

```
arco(1, s==>[sn, sv]°[ ], 4)
```

Análise bem sucedida.

Categoria da expressão: s

6 Conclusões

- Versão em DCG
- Eficiência: não refaz análises já feitas

“a menina com os cabelos loiros”

- $SN \rightarrow Det\ N$
- $SN \rightarrow Det\ N\ Adj$
- $SN \rightarrow SN\ SP$

6.1 Perspectivas

- Interpretação semântica
- Gramática Categorial

Referências

- [1] Michael A. Covington. *Natural Language Processing for Prolog Programmers*. Prentice Hall, Englewood Cliffs, 1994.
- [2] Clive Matthews. *An Introduction to Natural Language Processing through Prolog*. Longman, London, 1998.

Analizador gramatical tabular por deslocamento e redução em Prolog