

Aplicação de algoritmos de análise sintáctica

Diogo Fernandes nº 95000

Jorge Gonçalves nº 96180

ISCTE - IUL, Lisboa, Portugal

1 Introdução

Este trabalho teve como objectivo a construção de uma única gramática livre de contexto e a aplicação do algoritmo Earley. O algoritmo que leva o nome do seu inventor, Jay Earley, é utilizado para analisar *strings* que pertencem a uma determinada gramática livre de contexto e utiliza programação dinâmica. As frases atribuídas ao nosso grupo foram as seguintes:

"Os autores do acórdão revelam um louvável desejo"
"A infalibilidade papal é um dogma"

2 Gramática Livre de Contexto

Com base na análise sintáctica das frases que foram atribuídas, foi construída uma única gramática livre de contexto que aceita ambas as frases como válidas. Abaixo as tabelas que demonstram o conteúdo da gramática construída divididas em regras sintácticas e regras lexicais.

Regras Lexicais	Regras Sintáticas
<i>Art</i> → <i>os</i>	<i>S</i> → <i>NP VP</i>
<i>Art</i> → <i>um</i>	<i>NP</i> → <i>Noun</i>
<i>Art</i> → <i>a</i>	<i>NP</i> → <i>NP PP</i>
<i>Noun</i> → <i>autores</i>	<i>NP</i> → <i>Art Noun</i>
<i>Noun</i> → <i>desejo</i>	<i>NP</i> → <i>Art Adj Noun</i>
<i>Noun</i> → <i>acórdão</i>	<i>NP</i> → <i>Art Noun Adj</i>
<i>Noun</i> → <i>infalibilidade</i>	<i>VP</i> → <i>Verb NP</i>
<i>Noun</i> → <i>dogma</i>	<i>PP</i> → <i>Prep NP</i>
<i>Prep</i> → <i>do</i>	
<i>Verb</i> → <i>revelam</i>	
<i>Verb</i> → <i>é</i>	
<i>Adj</i> → <i>louvável</i>	
<i>Adj</i> → <i>papal</i>	

Table 1. Regras Lexicais e Regras Sintáticas.

3 Aplicação do Algoritmo de Earley

State	Rule	Position	Back Pointer
S0	$ROOT \rightarrow . S$	[0,0]	□
S1	$S \rightarrow . NP VP$	[0,0]	□
S2	$NP \rightarrow . Noun$	[0,0]	□
S3	$NP \rightarrow . NP PP$	[0,0]	□
S4	$NP \rightarrow . Art Noun$	[0,0]	□
S5	$NP \rightarrow . Art Adj Noun$	[0,0]	□
S6	Art $\rightarrow os .$	[0,1]	□
S7	$NP \rightarrow Art. Noun$	[0,1]	[S6]
S8	$NP \rightarrow Art. Adj Noun$	[0,1]	[S6]
S9	Noun $\rightarrow autores .$	[1,2]	□
S10	$NP \rightarrow Art Noun .$	[0,2]	[S6,S9]
S11	$S \rightarrow NP . VP$	[0,2]	[S10]
S12	$NP \rightarrow NP . PP$	[0,2]	[S10]
S13	$VP \rightarrow Verb . NP$	[2,2]	□
S14	$PP \rightarrow Prep . NP$	[2,2]	□
S15	Prep $\rightarrow do .$	[2,3]	□
S16	$PP \rightarrow Prep . NP$	[2,3]	[S15]
S17	$NP \rightarrow . Noun$	[3,3]	□
S18	$NP \rightarrow . NP PP$	[3,3]	□
S19	$NP \rightarrow . Art Noun$	[3,3]	□
S20	$NP \rightarrow . Art Adj Noun$	[3,3]	□
S21	Noun $\rightarrow acórdão .$	[3,4]	□
S22	$NP \rightarrow Noun .$	[3,4]	[S22]
S23	$PP \rightarrow Prep NP .$	[2,4]	[S15,S22]
S24	$NP \rightarrow NP . PP$	[3,4]	[S22]
S25	$NP \rightarrow NP PP .$	[0,4]	[S10,S23]
S26	$PP \rightarrow . Prep NP$	[4,4]	□
S27	$S \rightarrow NP . VP$	[0,4]	[S25]
S28	$NP \rightarrow NP . PP$	[0,4]	[S25]
S29	$VP \rightarrow . Verb NP$	[2,4]	□
S30	Verb $\rightarrow revelam .$	[4,5]	□
S31	$VP \rightarrow Verb . NP$	[4,5]	[S30]
S32	$NP \rightarrow . Noun$	[5,5]	□
S33	$NP \rightarrow . NP PP$	[5,5]	□
S34	$NP \rightarrow . Art Noun$	[5,5]	□
S35	$NP \rightarrow . Art Adj Noun$	[5,5]	□
S36	Art $\rightarrow um .$	[5,6]	□
S37	$NP \rightarrow Art . Noun$	[5,6]	[S36]
S38	$NP \rightarrow Art . Adj Noun$	[5,6]	[S36]
S39	Adj $\rightarrow louvável .$	[6,7]	□
S40	$NP \rightarrow Art Adj. Noun$	[5,7]	[S36, S39]

S41	$Noun \rightarrow \text{desejo} .$	[7,8]	\square
S42	$NP \rightarrow Art \ Adj \ Noun.$	[5,8]	[S36, S39, S41]
S43	$VP \rightarrow Verb \ NP .$	[4,8]	[S30, S42]
S44	$NP \rightarrow NP . \ PP$	[5,8]	[S42]
S45	$S \rightarrow NP \ VP .$	[0,8]	[S25, S43]
S46	$PP \rightarrow . \ Prep \ NP$	[8,8]	\square
S47	$ROOT \rightarrow S .$	[0,8]	[S45]

4 Frase Não Analisável Com a Gramática Definida

A frase não analisável com a gramática definida que seleccionamos foi a seguinte:

A	infalibilidade	do	acórdão	é	louvável
Art	Noun	Prep	Noun	Verb	Adj

Depois de seleccionada a frase foi aplicado o algoritmo de Earley sobre a mesma, afim de demonstrar que a fra

State	Rule	Position	Back Pointer
S0	$ROOT \rightarrow . \ S$	[0,0]	\square
S1	$S \rightarrow . \ NP \ VP$	[0,0]	\square
S2	$NP \rightarrow . \ Noun$	[0,0]	\square
S3	$NP \rightarrow . \ NP \ PP$	[0,0]	\square
S4	$NP \rightarrow . \ Art \ Noun$	[0,0]	\square
S5	$NP \rightarrow . \ Art \ Adj \ Noun$	[0,0]	\square
S6	Art $\rightarrow a .$	[0,1]	\square
S7	$NP \rightarrow Art. \ Noun$	[0,1]	[S6]
S8	$NP \rightarrow Art. \ Adj \ Noun$	[0,1]	[S6]
S9	Noun $\rightarrow \text{infalibilidade} .$	[1,2]	\square
S10	$NP \rightarrow Art \ Noun .$	[0,2]	[S6,S9]
S11	$S \rightarrow NP . \ VP$	[0,2]	[S10]
S12	$NP \rightarrow NP . \ PP$	[0,2]	[S10]
S13	$VP \rightarrow Verb . \ NP$	[2,2]	\square
S14	$PP \rightarrow Prep . \ NP$	[2,2]	\square
S15	Prep $\rightarrow \text{do} .$	[2,3]	\square
S16	$PP \rightarrow Prep . \ NP$	[2,3]	[S15]
S17	$NP \rightarrow . \ Noun$	[3,3]	\square
S18	$NP \rightarrow . \ NP \ PP$	[3,3]	\square
S19	$NP \rightarrow . \ Art \ Noun$	[3,3]	\square
S20	$NP \rightarrow . \ Art \ Adj \ Noun$	[3,3]	\square
S21	Noun $\rightarrow \text{acórdão} .$	[3,4]	\square
S22	$NP \rightarrow Noun .$	[3,4]	[S22]
S23	$PP \rightarrow Prep \ NP .$	[2,4]	[S15,S22]
S24	$NP \rightarrow NP . \ PP$	[3,4]	[S22]
S25	$NP \rightarrow NP \ PP .$	[0,4]	[S10,S23]
S26	$PP \rightarrow . \ Prep \ NP$	[4,4]	\square
S27	$S \rightarrow NP . \ VP$	[0,4]	[S25]

S28	$NP \rightarrow NP . PP$	[0,4]	[S25]
S29	$VP \rightarrow . Verb NP$	[2,4]	[]
S30	$Verb \rightarrow \acute{e} .$	[4,5]	[]
S31	$VP \rightarrow Verb . NP$	[4,5]	[S30]
S32	$NP \rightarrow . Noun$	[5,5]	[]
S33	$NP \rightarrow . NP PP$	[5,5]	[]
S34	$NP \rightarrow . Art Noun$	[5,5]	[]
S35	$NP \rightarrow . Art Adj Noun$	[5,5]	[]

Como podemos verificar acima não existe na nossa gramática nenhuma regra sintática para lidar com um adjetivo a seguir a um verbo.

5 Implementação do Algoritmo Earley

Para complementar a apresentação dos passos, foi tomada a decisão de fazer uma implementação do algoritmo.

Usou-se a linguagem de programação Python seguindo uma abordagem orientada a objetos que, na nossa opinião, alinha perfeitamente com a ideologia estrutural do algoritmo.

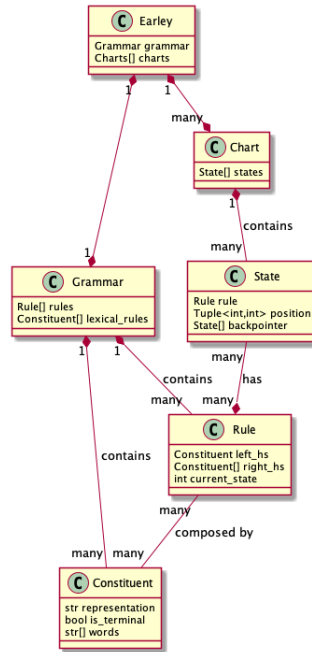


Fig. 1. Class Diagram

Acima temos o diagrama de classes que mostra as várias classes definidas assim como a base da sua interação, que atributos é que foram definidos e como é que permitem montar esta estrutura robusta.

Esta implementação foi um desafio que testou o nosso conhecimento do comportamento do algoritmo para que fosse possível representar toda a informação necessária para o correto funcionamento do mesmo.

Conseguimos então o nosso objetivo inicial que era, através da implementação, detetar quaisquer erros na definição inicial dos passos ao fazermos algumas correções aos indicadores dos backpoints em alguns estados.

6 Conclusão

Neste trabalho foi possível colocar em prática os conhecimentos adquiridos em aula relativamente a aplicação de algoritmos de análise sintáctica, mais especificadamente o algoritmo Earley, e gramática livre de contexto. No nosso entendimento ambos os elementos contribuíram de igual forma nas atividades deste projeto.