

Compilador Pascal escrito em Python

Universidade Federal de São João Del Rei (UFSJ) - Ciência da Computação

Trabalho prático da disciplina de compiladores - 2021/2

Felipe Henrique Faria\ Fabrício Lopes

Instruções de execução

Em uma máquina ou ambiente virtual com Python3 instalar as dependências listadas no arquivo `requirements.txt`.

```
pip install -r requirements.txt
```

Executar o arquivo `main.py` passando como parâmetro o caminho do arquivo de código para compilação.

```
python main.py <caminho_do_arquivo>
```

Exemplo:

```
python main.py codes/code1.pas
python main.py codes/code2.pas
python main.py codes/code3.pas
...
python main.py codes/code7.pas
```

O resultados da compilação fica disponível na pasta `results`.

Lexemas e tokens

Lexema	Token
program	PROGRAM
begin	BEGIN
end	END
var	VAR
boolean	BOOLEAN
integer	INTEGER
real	REAL
string	STRING
if	IF
then	THEN
else	ELSE
for	FOR
while	WHILE
do	DO
to	TO
read	READ

Lexema	Token
write	WRITE
true	TRUE
false	FALSE
(LBRACKET
)	RBRACKET
{	LBRACE
}	RBRACE
[LCOL
]	RCOL
,	COMMA
;	PCOMMA
:	TWOPOINT
:=	ASSIGN
==	EQUAL
!=	DIFERENT
<	LT
<=	LTE
>	GT
>=	GTE
+	PLUS
-	MINUS
*	MULT
/	DIV

Implementação do analisador léxico

O analisador léxico foi implementado utilizando um autômato com dez estados. O autômato é implementado em uma classe chamada `Scanner` que contém um método `scan`, este método itera por todas as linhas e palavras do arquivo lido alternando de estado conforme as regras definidas. Neste processo também são coletados os erros léxicos que ocorrem quando o automato fica preso em um estado, não tendo opção que satisfaça o caractere lido.

<https://github.com/felipevisu/python-pascal-compiler/blob/d0830eb35a23edfef222b0f4df5c30dad5317baf/scanner.py>

Tabela de simbolos

A tabela de simbolos é implementada utilizando uma classe `Symtable`. Ela ocorre durante o processo de análise sintática. Quando um token do tipo `'ID'` é lido durante a etapa de declaração de variáveis ele é inserido na tabela. O processo de inserção verifica se a variável já foi declarada ou não.

As verificações semânticas podem ser observados no link a baixo:

<https://github.com/felipevisu/python-pascal-compiler/blob/d0830eb35a23edfef222b0f4df5c30dad5317baf/analyzer.py#L50-L88>