

Pascal to MIPs Compiler

Joseph Miller

January 18, 2018

1 Overview

This program will be able to compile Pascal code into MIPs assembly code. The compiler will be made divided into 4 major sections:

- Scanner
- Parser
- Semantic Analysis
- Code Generation

2 Design

2.1 Scanner

This program will be able to take in a file with Pascal code and output each token in the order that it was written. Each of the following keywords and symbols will be identified as tokens, as well as variable names (IDs).

- Keywords: and, array, begin, div, do, else, end, function, if, integer, mod, not, of, or, procedure, program, real, then, var, while
- Symbols: ;, ,, ., :, [,], (,), +, -, =, <>, <, <=, >, >=, *, /, :=

This project is made of 3 Java/Class files:

- **Scanner:** the DFA scanner made by JFlex to analyze the Pascal code and identify token types.
- **Token:** the token object creator, this file also takes in what kind of token it is and adds a more specific distinct type by using a look-up table (switch function / if-else function). Each Token object contains the lexeme and the type.
- **TokenType:** creates the specifications for a the TokenType attribute of the Token attribute

```
Token: PROGRAM, Program
Token: ID, Lesson1
Token: ERROR, _
Token: ID, Program3
Token: SEMI, ;
Token: VAR, Var
Token: ID, Num1
Token: COMMA, ,
Token: ID, Num2
Token: COMMA, ,
Token: ID, Sum
Token: COLON, :
Token: INTEGER, Integer
Token: SEMI, ;
Token: BEGIN, Begin
Token: COMMENT, {no semicolon}
Token: POSEXP, 12686E20938

The List of Token Arrays:
[Token: PROGRAM, Program, Token: ID, Lesson1, Token: ERROR, _, Token: ID, Program3, Token: SEMI, ;
```

Figure 1: Token output to console

Illegal symbols and sequences not accepted by our simple pascal compiler will have the TokenType: ERROR. A visual of the tokens created can be seen in Figure 1, one of which produces an error.

2.2 Parser

Not written

2.3 Semantic Analysis

Not written

2.4 Code Generation

Not written

3 Change Log

– 18 January 2017 - Miller - Created