

Documentation to implementation of task No. 1 to the IPP 2018/2019

Name and surname: Juraj Holub
Login: xholub40

1 Proposal of solution

Scheme of project solution is split to three blocks.

First block split input string to tokens representation where one token represents atomic lexeme of input programming language IPPcode19. Token is implemented like data structure with two attributes. First attribute is identification of token, for example keywords or integer constant. Second attribute is value of token specified by his identification, for example integer constant contain specific integer value.

Second block is finite state machine (FSM). FSM input is string of tokens and machine verify lexical and syntactical correctness of input program.

Third and last block is generator of output in eXtensible Markup Language (XML).

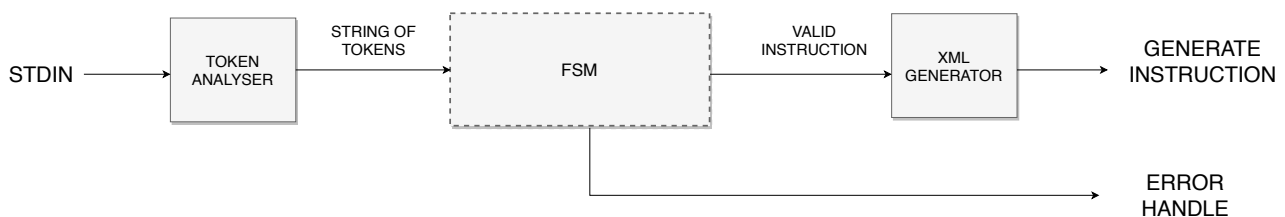


Figure 1: Scheme of implemented parse solution.

Figure 1 shows scheme of implemented algorithm of parsing IPPcode19 and generating XML output.

2 Implementation details

Suggested solution parted to three blocks described in section 1 is implemented in three logical units:

- **Tokens analyser:** Analyse input string from standard input of program char by char and convert them to the token representation. Supported tokens are instruction code, symbol, label, variable, type, white space, new line. Instruction codes are sorted to groups by number and types of arguments. Comments are ignored and token analyser skip them in input program string.
- **Parse:** Implementation of finite state machine. FSM input is actual token received from token analyser. Machine output depends on current state and current token which lead to Mealy FSM. If input is valid then after every successive analyse of one IPPcode19 instruction is this instruction send to XML generator. Incorrect input lead to error state where FSM read rest of tokens and then inform about error.
- **XML generator:** Generate XML representation of IPPcode19. Input of class are one instruction with their arguments. Class implementation uses DOMDocument library. This library is optimized for XML generating and brings advantages like converting XML key words inside of input data to correct representation.