Scanner/Lexical Analysis - DUE 01/29/14 (11:59pm)

Objective

Objective 1 of this milestone is to produce the scanner and symbol table for the project.

Objective 2 formalize the data structure used for the tokens.

Objective 3 to give you experience using the state-transition diagram as a design tool.

Objective 4 is to give you experience in designing a program using a categorical style design.

Objective 5 is to give you practice in designing test files based on the formal definitions.

Professional Methods and Values

Design and testing. Use of formal method.

Assignment

The purpose of this milestone is to develop a working scanner from the lexical definitions. Scanners are supported by over 50 years of development; these techniques were used by engineers in the 1960s. You are not allowed to regex libraries to help your determine a lexeme!!! Even if you do not have to import a library for this feature, such as in Python, you are still not to use a built-in regex feature!!!
Help is provided that explains finite state machines.

Performance Objectives

1. Develop a formal definition of the scanner algorithm from the lexical structure for our Itty Bitty Teaching Language (IBTL).
2. Develop a scanner algorithm from a formal definition.
3. Develop a data structure for the tokens of the language.
4. Develop a test driver for the developed program .
5. Test the resulting program for correctness based on the formal definition.

This is the proper time to start the symbol table development. Implement the lookup routine as a hash table.

Milestone Report

Your milestone report will include written answers to the finite state machine assignment mentioned above. Turn in all design documentation as a pdf.

Additional Information

The lexical scanner rules are the same as *C* whenever a specfic rule is not stated.

Examples

1. *stdout* is a keyword in IBTL but not in *C*, so IBTL overrides.
2. Direct information on the format of integers has not been given, so *C* rules are used; the same is true for *floats* and *strings*.