Assignment 2: Tokenizing Pascal

Assigned: Friday, 9 September 2011
Due: 11:00 p.m. Wednesday, 14 September 2011

Summary: In this assignment, you will write regular expressions for the tokens of Pascal.

Purposes: To get you started with your lexer. To help you think more deeply about tokens.

Expected Time: Three to Five Hours.

Collaboration: You may work in groups of up to size three. You may consult with anyone you’d like on this assignment, but you should clearly document that consultation.

Submitting: Submit a tarball containing your solution on Pioneerweb.

Warning: So that this exercise is a learning assignment for everyone, I may spend class time publicly critiquing your work.

Assignment

Reread (or rescan) the Pascal report and identify the important tokens in the language. Write a regular expression for each token.

For example, if you decide that there should be a BEGIN token, you might write

BEGIN: \[Bb\][Ee][Gg][Ii][Nn]\n
Minimally, you should support:

- integer constants,
- real constants,
- string constants,
- identifiers,
- the various forms of punctuation (multiple token classes), and
- the various keywords.

Submitting

Create a tarball of your files. The tarball should include your regular expressions (probably just one file) and any notes you make.
If possible, ensure that your tarball expands to the directory `user1-user2-user3.HW2`.

Submit your tarball on Pioneerweb.

**Important Evaluation Criteria**

I will evaluate your assignment primarily on its *correctness*. In particular, I will consider whether your program achieves its desired goals.

I will also evaluate your assignment on its *style*. I will consider how you have decomposed the problem, the names of your procedures, your indentation and commenting, and other issues that suggest that you have not only thought about how to code the problem, but how to code it well, and in a way readable to others.