The Big Baby

Pascal Subset Compiler in C

Todd Sharpe

5/8/2011

Summary

This compiler is meant for use on a subset of Pascal. It is written entirely in C and makes use of both Lex and Yacc. This compiler outputs AT&T syntax Intel assembly for use with the GNU Assembler.

Abilities

The entire Appendix A grammar has been implemented along with a few additional features. This compiler allows for infinitely nested scopes – functions/procedures can be declared with infinite depth, as well as appropriate variables in those scopes (see test infinite.p) In addition, array indexing operations can be performed as input to functions, as well as on the right hand side of another array index operation (see fib.p). Lastly, the compiler supports output of integers with either write({expression}) or PRINT {expression} syntax.

Limitations

The largest limitation of the compiler is that it only implements a subset of Pascal, and therefore no function pointers, set manipulation, or string concatenation. The compiler also doesn't allow for multidimensional arrays.

References

An Introduction to Compiling Techniques – JP Bennett

A Compact Guide to Lex and Yacc – Thomas Neimann

Intel x86 Function-call Conventions - Steve Friedl

(http://www.unixwiz.net/techtips/win32-callconv-asm.html)

Intel Assembler Code Table – Roger Jegerlehner