Ankit Agarwal <ankit.agarwal365@gmail.com>

Theory-5, 6, 7: Lexical Analysis

1 message

Compiler Design <bit.compiler.2014@gmail.com>   Tue, Jan 28, 2014 at 3:08 PM
Bcc: ankit.agarwal365@gmail.com

Required Reading
- Specification of Tokens: Sec.3.3, Sec.3.5
- Recognition of Tokens:
    - Basics: Sec.2.6
    - Transition diagrams and algorithms: Sec.3.4
    - Input Buffering: Sec.3.2
- Lexical Errors: Sec.3.1.4

Recommended Reading
- Review of RE-NFA-DFA conversions and NFA, DFA simulations from FLAT course.

Advanced
- Reverse engineering the lex.yy.c file
    - use flex with -Cf --trace --verbose options:
        - trace will give the DFA state transitions
        - verbose will give the summary of the DFA.
        - yy_nxt and yy_accept inside lex.yy.c file contains the
tables of interest
    - use flex without the -Cf option:
        - Optimizing DFAs (Sec.3.9) contains the algorithm
        - start with yy_ec inside lex.yy.c file.
    You can email the writer of flex Vern Paxson (www.icir.org/vern)
at vern@berkeley.edu for any specific questions. He will respond.

- Fast String Matching algorithms: Knuth-Morris-Pratt (KMP): Exercise
3.4.2 onwards.

Arun