1 Reading / Source Material

- Scott 2.2
- Cooper and Torczon: Chapter 2

2 Vocabulary

deterministic finite state automata (DFA), nondeterministic finite state automata (NFA), state, transition, \( \epsilon \)-transition, regular expression, \( \epsilon \) (to represent the absence of symbols in REs), alternation/union, concatenation, Kleene closure, configuration of an NFA Thompsons Construction, Subset Construction, DFA Minimization, Kleenes Construction, scanner, \( \epsilon \)-closure, worklist algorithm

3 Objectives: Be able to . . .

1. formalize lexical descriptions using regular expressions
2. determine whether a regular expression recognizes a particular string
3. find a string recognized by a particular regular expression
4. find an NFA equivalent to any given regular expression
5. find a DFA equivalent to any NFA
6. write code implementing a DFA

Outline

1. Regular Languages and Scanners
   (a) role of the Scanner
   (b) finite automata
      i. states \((S)\)
      ii. symbols \((\Sigma)\)
      iii. transitions \((\delta)\)
      iv. final states \((S_F \subseteq S)\)
   (c) regular expressions
      i. alternation: \(a|b\)
      ii. concatenation: \(ab\)
iii. closure $a^*$
iv. grouping (a)

Note: alternation and concatenation are left associative. Highest precedence: *. Lowest precedence: |.

(d) converting regular expressions to NFA’s (Thompson’s Construction)

i. a symbol
   
   ![Diagram of a symbol transition]

ii. concatenation
   
   ![Diagram of concatenation transition]

iii. alternation
   
   ![Diagram of alternation transition]

iv. closure (repetition)
   
   ![Diagram of closure (repetition) transition]

(e) converting NFA’s to DFA’s (Subset Construction)

(f) implementing DFA’s

4 Examples

- moo|meow|baa(baa)*
• Regular expression for a positive integer:

\[(1|2|3|4|5|6|7|8|9)(0|1|2|3|4|5|6|7|8|9)\ast\]

or with syntactic sugar

\[[1 - 9][0 - 9]\ast\]

• Regular expression for Java variable naming convention:

\[(a|b|c|d|f|g|h|i|\ldots|m|n|o|p|q|r|s|t|u|v|w|x|y|z|A|B|C|D|E|F|G|H|I|J|K|L|M|N|O|P|Q|R|S|T|U|V|W|X|Y|Z|0|1|2|3|4|5|6|7|8|9|\$)\]

or with some syntactic sugar

\[[a - z][(a - z)|(A - Z)|(0 - 9)|\$]\ast\]

5 Questions

1. Cooper, Chapter 2, Exercise 5 (p. 80-81)

2. Cooper, Chapter 2, Exercise 7(a) and 8(b) (p. 81)

3. Write pseudo-code for a direct-coded scanner that recognizes the regular expression \[1 - 9][0 - 9]\ast |0.