COMPSCI 220 Programming Methodology

Regular Expression Derivatives

W12

Overview

We will begin this week with a review of regular expressions and their use with the **grep** tool. We will follow with a brief discussion of how regular expressions are implemented and used in Java/Scala. We will then turn our attention on how we might implement regular expressions using a method known as *regular expression derivatives*. This approach is a straightforward technique for matching an input string to a regular expression using simple rules that manipulate the regular expression. In particular, the derivative of a regular expression is an algebraic manipulation of the regular expression that calculates its *partially matched* tail with respect to a particular character.

E1: $\delta(\varnothing) = \varnothing$

E2: $\delta(\varepsilon) = \varepsilon$

E3: $\delta(c) = \varnothing$

E4: $\delta(re_1\ re_2) = \delta(re_1)\ \delta(re_2)$

E5: $\delta(re_1 | re_2) = \delta(re_1) \mid \delta(re_2)$

E6: $\delta(re^*) = \varepsilon$

D1: $D_c(\varnothing) = \varnothing$

D2: $D_c(\varepsilon) = \varnothing$

D3: $D_c(c) = \varepsilon$

D4: $D_c(c') = \varnothing$ if $c \neq c'$

D5: $D_c(re_1\ re_2) = \delta(re_1)D_c(re_2) \mid D_c(re_1)re_2$

D6: $D_c(re_1 \mid re_2) = D_c(re_1) \mid D_c(re_2)$

D7: $D_c(re^*) = D_c(re)\ re^*$;

Reading, Resources, and Material

- ✓ Regular Expressions, Wikipedia
  https://goo.gl/7g6b0
- ✓ Sculpting Text, Matt Might
  http://goo.gl/m3jT1
- ✓ Scala Regular Expressions, tutorialspoint
  http://goo.gl/4Dwhr
- ✓ Using Pattern Matching with Regex in Scala, Ikai Lan
  http://goo.gl/hc4kTj
- ✓ Implementing Regular Expressions with Derivatives, Matt Might
  http://goo.gl/tIZOqb
- ✓ scala.util.matching.Regex
  http://goo.gl/aU0o8J
- ✓ java.util.regex.Pattern
  http://goo.gl/N7kse