

Homework 1

1. Write regular expressions for the following languages
 - a. the set of all alphabetic strings
 - b. the set of all lower case alphabetic strings ending in a *b*
 - c. the set of all strings from the alphabet *a, b* such that each *a* is immediately preceded by and immediately followed by a *b*
2. Write regular expressions for the following languages. By “word”, we mean an alphabetic string separated from other words by whitespace, any relevant punctuation, line breaks, and so forth.
 - a. the set of all strings with two consecutive repeated words (e.g., “Humbert Humbert” and “the the” but not “the bug” or “the big bug”)
 - b. all strings that start at the beginning of the line with an integer and that end at the end of the line with a word
 - c. all strings that have both the word grotto and the word raven in them (but not, e.g., words like grottos that merely contain the word grotto)
 - d. write a pattern that places the first word of an English sentence in a register. Deal with punctuation.
3. Edit Distance
 - a. Compute the edit distance (using insertion cost 1, deletion cost 1, substitution cost 2) of “leda” to “deal”. Show your work (using the edit distance grid).
 - b. Figure out whether `drive` is closer to `brief` or to `divers` and what the edit distance is to each.
 - c. Now implement a minimum edit distance algorithm in Java and use your hand-computed results to check your code.