

Lab exercises for CMT304, 17 February 2020

1. Implement a genetic algorithm that develops a regular expression to count the number of occurrences of a given word in a given text. The GA's input should be a word and a text containing this word. The output should be a regular expression that correctly reports the number of occurrences of the word in any text.
2. Implement a genetic algorithm that develops a regular expression for identifying valid e-mail addresses.

Hint 1: Start with a database of labelled instances of e-mail addresses, i.e. valid and non-valid addresses.

Hint 2: Look up the syntax of valid e-mail addresses (e.g. on Wikipedia), and simplify it for your use.