LaTeX
UFO Assignment 3
Optimization

Camilla Staunstrup Asger Lundblad Rúni Niclasen

April 2021

Contents

1 Introduction **2**

2 Initial program **2**

2.1 Current performance 2

2.2 Program Bottlenecks 2

2.3 Hypothesis 2

3 Our optimized program **3**

3.1 Solution 3

3.2 Performance 3

Notes

- Remember box plots & hardware details (requested by Martin).
- Run both programs and collect timings for later calculations (Task 1, bullet point 2)
- Use the profiler as hinted in the assignment
- This could probably be done even faster in Python
- Follow the links on the course page. There are some hints on microbenchmarking.

1 Introduction

We decided to use the supplied program, *letterfrequencies*, for the optimization task.

The program source consists of the foundation trilogy by Isaac Asimov (roughly 25.000 lines of text) and a single Java file. The file reads the text and returns the frequencies of the letters within the text.

2 Initial program

The program is purposefully written in an inefficient way. It has two helper methods and at its core it presents a list of the letter frequencies between a-zA-Z.

Below you can find a sample of the output of the program:

Letter	Frequencies
E	118313
T	87190
A	76011
O	74907
N	67908
...	...
K	6630
X	1457
Q	1017
J	806
Z	674

Table 1: The 5 most and least frequent letters in the dataset

2.1 Current performance

Documentation of the current performance . . .

2.2 Program Bottlenecks

Explanation of bottleneck(s) . . .

2.3 Hypothesis

A hypothesis of what causes the problem . . .

3 Our optimized program

A changed program with better performance ...

3.1 Solution

(How did we do it) ...

3.2 Performance

Documentation of the new performance ...