November 14, 2019

Jean-Phillippe Durand

40052679

Concordia University

TEXT CORRECTOR PROJECT REPORT

COEN 352

# Table of Content

# Problem Description:

The goal of the project is to develop a simple text corrector. The final solution shall focus on typographical and orthograph errors. The user shall provide a word bank using a plain text file (.txt), which will be used by the program as a dictionary.

The input file used as a dictionary will contain an unknown number of words; each word (separated by a whitespace character). To prevent long computing times when analyzing the text to be corrected, the dictionary will be stored into a binary search tree (BST). Due to it’s low time complexity in the order of O(LogN), where N is the number of words stored in the BST, the BST will prevent the program from slowing down noticeably when inputting dictionaries containing a large number of words.

The BST created from the user inputted dictionary will constitute the foundation of the text corrector, as all correction operations will require searching in the dictionary for the validity of the word, or the closest match.

The corrector has three (3) functions. It will correct word separation in the case where the user forgot to add space between two know (to the dictionary) words. It will then correct words containing a mutated character (e.g. d=ctionary will be corrected to dictionary) and warn the user of the correction by providing him with the initial word and its correction. For the last step, the corrector will analyse the frequency of know words in the text and output a list of the words found, sorted by their frequency and then alphabetically.

# Problem Breakdown:

**TEXT CORRECTOR WBS**

# Solution & Design:

# Results and Analysis

# User Manual: