ANALYSIS OF ALGORITHMS
HOMEWORK 2

Name: Ali Çağatay Yüksel

ID : 040050220

CRN : 20484

1. Introduction

The aim of this project is to develop and implementation of Huffman Coding algorithm.This method uses the entropy of the pattern to be coded i.e. the less probable the character is the more bits are used to represent it hence resulting in less bits overall than in fixed-length coding with no loss in data.

Huffman Coding algorithm works finding the frequencies of all characters in a string and builds a binary tree to assign a code each of them. Huffman Coding satisfies the uniquely decode-ability property with the help of this tree structure. The algorithm outputs the code map as well as the encoded text since decoder client needs this mapping to reverse the encoding process.

2. Development

In this project, Huffman Coding is implemented using C++ programming language and object-oriented approach. An encoder class is created with the required methods. One can benefit from this class in any program to encode any string using the interface provided.

Coding work is done using Visual Studio 2010 Express Integrated Development Environment on a Windows 7 box. Source code also compiles using GNU C++ compiler as it should compile in any standards-complying C++ compiler. It does not use any non-standard libraries so does not depend on any operating system or platform either.

A Makefile is also provided along with the source code to make the compiling and testing processes easier. Typing "*make*" in the path of the source code simply compiles the program, executes it with the sample input and decodes the output with the given Python program. This results in 5 output files: "*encoded.txt*", the encoded text, "*freq.txt*", the list of frequencies of the characters, "*hcodes.txt*", assigned codes of each character, "*stats.txt*", some info about encoding and "*decoded.txt*", the same of the input, but decoded from the encoded text.