11th Homework

Prof. Helmut Alt

Introduction to Algorithms

Due: June 5, 2017, 10 a.m.

Exercise 1 PATRICIA-trees *6 points*

Describe and analyze efficient algorithms for search, insertion, and deletion in PATRICIA-trees. Suppose that edges have the space efficient labelling as shown in class.

Exercise 2 suffix trees *8 points*

Use suffix trees to design efficient algorithms for the following problems:

(a) Find the longest substrings of a string w that occur more than once.

(b) Find the shortest substrings of w that occur only once.

(c) Find the longest substrings of w which are palindromes, i.e., read the same forward and backward.

Exercise 3 Huffman code *8 points*

(a) The following table gives the frequencies of the letters of the English language (including the blank for separating words) in a particular corpus.

blank	18.3%	r	4.8%	y	1.6%
e	10.2%	d	3.5%	p	1.6%
t	7.7%	l	3.4%	b	1.3%
a	6.8%	c	2.6%	v	0.9%
o	5.9%	u	2.4%	k	0.6%
i	5.8%	m	2.1%	j	0.2%
n	5.5%	w	1.9%	x	0.2%
s	5.1%	f	1.8%	q	0.1%
h	4.9%	g	1.7%	z	0.1%

What is the optimum Huffman encoding of this alphabet? What is the expected number of bits per letter?

(b) Suppose a data file contains a sequence of 8-bit characters such that all 256 characters are about as common: the maximum character frequency is less than twice the minimum character frequency. Prove that Huffman coding in this case is no more efficient than using an ordinary 8-bit fixed-length code.