UNIVERSITY OF LIVERPOOL

Department of Computer Science
Dr. Sebastian Wild · Ben Smith

Date: 2022-11-07
Version: 2022-11-09 22:03

Exam Preparation – Transfer Problems for COMP 526 – Efficient Algorithmics, Fall 2022

The final exam will consist of reproduction (bookwork-style) questions, application tasks (similar the tasks in the classtests), and transfer problems.

Since the latter type of tasks is not as widely covered by continuous assessment as the others, this problem sheet gives a few transfer problems from past exams that you can use to practice your understanding and get a feeling for the type of problems on the exam.

The problems below exemplify the style of questions, but do not comprise an exhaustive list of topics; the actual exam may contain questions on any material from Units 0–8 (including tutorials).

Problem 1 (Double String Problem)

Describe an efficient algorithm for the *double string problem*:

Given a text $T[0..n)$, output a substring w that occurs at (at least) 2 different positions in T, i.e., there are $i_1 < i_2$, so that $T[i_1..i_1 + |w|) = T[i_2..i_2 + |w|) = w$.

Output: Such a w.

1. Describe your algorithm (textual description is enough, but be precise).
2. Briefly comment on its running time.

Problem 2 (BWT Pattern Guessing)

Consider the sequence of strings in which the kth string $w_k = \texttt{ab}^k\texttt{ab}^k\texttt{a\$}$ for $k = 1, 2, 3, \ldots$

1. Compute w_1, w_2, w_3, and w_4.
2. Compute the BWT for each of these.
3. Explain the structure of the BWT of w_k for general k, and argue why it looks like this.

Problem 3 (Robustness of Codes)

Consider transmitting a coded text that is produced by the encoding methods below. Now suppose that in this coded text a *single character* is *transmitted incorrectly*, i.e., it is replaced by another character.

Which of the encoding methods listed below can still decode *all but a fixed number* of characters of the source text correctly when such a single-character error occurs?

Here, "can still decode" includes the option of slightly modifying the decoding procedure specifically for this scenario, i.e., you may assume that the decoding method expects a (potential) single-character error, but (obviously) does not know if or where it happened.

Our Encodings:

1. ASCII
2. UTF-8
3. Huffman code
4. run-length encoding
5. Lempel-Ziv-Welch algorithm
6. Move-to-front transform
7. Burrows-Wheeler transform
8. 4+3 Hamming code

Justify your answers.