ECS 222A: Assignment #1

Due on Tusday, January 13, 2015

Daniel Gusfield TR 4:40pm-6:00pm

Wenhao Wu

Contents

Problem 0

In the bit-model every bit-level operation must be counted. For example, to take the OR of two binary strings of length q each, takes q operations, and to use or set an index consisting of q bits, takes q operations. Show in detail that in the bit-model, then the 4-Russians method for bitmatrix multiplication only takes $\mathcal{O}(n^3/(\log n))$ operations.

Answer:

Problem 1

Prove that the edit distance is the same no matter which definition is used.

Answer:

Problem 2

From the mathematical standpoint, an alignment and an edit transcript are equivalent ways to describe a relationship between two strings. An alignment can be easily converted to the equivalent edit transcript and vice-versa. Completely explain and justify the above statement.

Answer:

Problem 3

Prove that any traceback path specifies an optimal edit transcript, and an optimal alignment. In the latter case, explain how the path specifies where the spaces should go in the two strings.

Answer:

Problem 4

Theorem Any path from (n, m) to $(0, 0)$ following pointers established during the computation of $D(i, j)$ specifies an edit transcript with the minimum number of edit operations. Conversely, any optimal edit transcript is specified by such a path. Moreover, since a path describes only one transcript, the correspondence between paths and optimal transcripts is one-one.

Prove this theorem.

Answer:

Problem 5

Since the traceback paths in a dynamic programming table correspond one-to-one with the optimal alignments, the number of distinct co-optimal alignments can be obtained by computing the number of distinct traceback paths. Give an algorithm to compute this number in $O(nm)$ time. Hint: use dynamic programming.

Answer: