

Cpt S 450 Homework #6

Please print your name!

All problems are hard.

1. LCS algorithm gives a way to decide how similar two given strings are. However, sometimes, we have to filter away some common subsequences that are in some pattern. Here is a problem for you to solve. Given two strings α and β , let γ to be a longest word satisfying all of the following conditions:

- γ is a subsequence of α ;
- γ is a subsequence of β ;
- γ does not contain *abb*.

Design an algorithm that finds such a γ for any given α and β . Also, analyze its complexity.

2. For any words α, β , we use $d(\alpha, \beta)$ to denote the length of an $\text{LCS}(\alpha, \beta)$. Let L_1 and L_2 be two given regular languages. Design an algorithm that compute the number D with

$$D = \max_{\alpha \in L_1, \beta \in L_2} d(\alpha, \beta).$$

(Warning: L_1 and L_2 could be infinite. Sometimes, D is infinite.)

3. Locality sensitive hash is a way to assign a number (called hash value) to an object so that when I try to find an object x from a set of objects that is “similar” to a given object a , I only need to compute the hash value of a and compare with hash values of the objects in the set and find those objects with almost the same hash value as a ’s. Is there any locality sensitive hash scheme for strings?