"Algorithms and Data Structures for Computational Biology"
Exercises – Module 3 – 7/24/2017

1. A string *P* is a *supersequence* of a string *Q* if *Q* is a *subsequence* of *P*. The *Shortest Common Supersequence* (*SCS*) of two strings *P* and *T* is the shortest supersequence of both strings. For instance, the *SCM* of "A B" and "B C" is "A B C", of length 3, while there are two *SCM* of "D A B" and "D C B", namely "D A C B" and "D C A B", of length 4. Given two strings *S* and *T,* whose lengths are *m* and *n,* respectively, one wants to find the length of an *SCS*. Solve the problem by a *dynamic programming* algorithm, defining first the recurrence relations giving the optimal sub-structure property, and then writing its corresponding pseudo-code and analyzing its complexity.

2. Consider the string "p a p a j a". Write (by hand) its corresponding:

- Suffix trie;
- Suffix tree;
- Suffix array;
- Burrows-Wheeler transform;
- LF mapping.