Advanced Algorithms for Data Science
Week 3 – Homework

Exercice 1. Dynamic programming

You are in charge of logistics in a company. Each week your company ships some amount of product (in kg). You plan your shipments for n weeks, and you know in advance that at week i you will be shipping s_i kg.

For your shipments, you have a choice between two transport companies. Company A charges you r roubles per kg (i.e. it would cost you $r \cdot s_i$ to ship s_i kg). Company B charges a fixed price c per week but imposes an additional condition: you have to sign for 4 consecutive weeks at a time.

Question: *Propose an algorithm to compute the optimal shipping schedule, that is an optimal sequence of n choices between A and B generating the minimal cost..*

Example: Assume $r = 1$, $c = 10$, and the sequence of s_i is $[11, 9, 9, 12, 12, 13, 12, 9, 9, 11]$. Then the optimal schedule would be to choose A for the first three weeks, then B for the block of next four weeks, and then A for the final three weeks. The total cost would be 98.

Exercice 2. Protein sequence alignment

Question: *Compute a maximal-score alignment of protein sequences* HEAGAWGHEE *and* PAWHEAE *under the scoring matrix* BLOSUM62 *(shown in slide 7 in* `12_sequence-alignment.pdf`*) and the indel penalty $d = 8$.*

You have to provide the full dynamic programming matrix with your answer.

Exercice 3. Hidden Markov Models

Question: *Finish the example of the Forward-Backward algorithm that we didn't finish in class.*

The example is given in slide 44 of `13_HMM.pdf`. The model is defined in slide 35.