Minhala and Grades

Part 1 (20% of final grade). Solve the objective problem for the case where k is always zero.

- Analyze and explain in detail what is the time and space complexity of your algorithm?
- Implement the algorithm in python and exemplify on the dataset how your implementation assists in deciphering the function of an "unknown" COG.

Code (documented notebook) and report (detailed) due date: 15\11\2021

Part 2 (40% of final grade). Solve the objective problem for the more general case (k can assume the value of any integer such that k <= d).

- Implement the algorithm in python and exemplify on the dataset how your implementation assists in deciphering the function of an "unknown" COG. Find at least one example where this extension is informative and helped decipher the function of a gene that was not deciphered in Part 1.
- How did the extension to allow up to k character insertions affect the time and space requirements of your algorithm? Do you see some practical time difference in running time as the value of k increases?

Code (documented notebook) and report (detailed) due date: 13\12\2021

Part 3 (40% of final grade). Creative: propose your own contribution for a special variant of the problem, according to your choice.

For example: Suffix-Trie based search, identify Colinear Gene Blocks that are discriminative in terms of environment or taxonomy, your own suggestions for speeding up or saving memory, add new constraints such as allowing a bounded number of substitutions\deletions, Branch-And-Bound version, reporting k-best results, interesting biological direction of your choice, and more.

Code (documented notebook) and report (detailed) due date: 10\01\2022