Project #4: String Matching with Dynamic Programming

CPSC 335 – ALGORITHM ENGINEERING

LIREN YIN

CALIFORNIA STATE UNIVERSITY, FULLERTON

Abstract:

This project is a follow on to Project 3. The Local String Alignment algorithm was implemented that solve the same problem. This is a dynamic programming algorithm that extends the one wrote in project 3, but it uses penalties and performs a local alignment instead of a global alignment.

Experimental Data

------------------- Dynamic Programming ------------------
String to Match = PIEPCMGA
a. The name of the protein matched:
Best Protein Description: sp|Q12517|DCP1_YEAST mRNA-decapping enzyme subunit 1
OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=DCP1 PE=1 SV=1

b. The score of the match:
Best Score: 30

c. The two strings representing the alignment of the match:
PIEPCMGA
PLE*CMGV

String to Match = TQGASNIGE
a. The name of the protein matched:
Best Protein Description: sp|Q08213|NGL1_YEAST RNA exonuclease NGL1
OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=NGL1 PE=1 SV=1

b. The score of the match:
Best Score: 36

c. The two strings representing the alignment of the match:
TQGASN*IGE
TDGASNVIGE

String to Match = ALAKLIRYGG
a. The name of the protein matched:
Best Protein Description: sp|P15108|HSC82_YEAST ATP-dependent molecular chaperone
HSC82 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=HSC82 PE=1 SV=4

b. The score of the match:
Best Score: 35

c. The two strings representing the alignment of the match:
ALAKLIRYGG
ALAKLLRYNS

String to Match = CSNPNLSDFGR
a. The name of the protein matched:
Best Protein Description: sp|Q03088|SVL3_YEAST Styryl dye vacuolar localization protein 3
OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=SVL3 PE=1 SV=1

b. The score of the match:
Best Score: 37

c. The two strings representing the alignment of the match:
CSNPNLSDFGR
PSSINMSDFGR

String to Match = MYPEPTIDE
a. The name of the protein matched:
Best Protein Description: sp|P25558|BUD3_YEAST Bud site selection protein 3
OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) GN=BUD3 PE=1 SV=2

b. The score of the match:
Best Score: 35

c. The two strings representing the alignment of the match:
MYPEPTIDE
LFP*PTIDE

Reference:

[1] Wortman , K., & Bernstein, L. (n.d.). CPSC 335 Project 4.docx. Retrieved November 30, 2017, from https://docs.google.com/document/d/1bX7Vs1zP9EKXe9U18-evi6oDTs4glPIee-Vqneg8BlM/edit