**2012004236 정구열**

**Project Name :** Bioinformatics LAB #2 Sequence Alignment

**Environment :**

OS : MAC 0S X High Sierra v. 10.13.6

IDE : Xcode v 9.0

Compiler : Apple LLVM version 10.0.0 (clang-1000.10.44.2)

**Usage :**

* There are 4 command line arguments.

1st - The Absolute Path of Sequence File

2nd - The Absolute Path of Sequence File

3rd - The Absolute Path of Score File

4th - The Absolute Path of ouput File

ex) in my case,

/Users/jeong-guyeol/Developer/BioInfomatics/Lab2\_Alignment/seq1.txt

/Users/jeong-guyeol/Developer/BioInfomatics/Lab2\_Alignment/seq2.txt

/Users/jeong-guyeol/Developer/BioInfomatics/Lab2\_Alignment/score.txt

/Users/jeong-guyeol/Developer/BioInfomatics/Lab2\_Alignment/output.txt

* The output file contains the results of global and local alignment along with various measurements each (matches, mismatches, gaps and score).
* Note. The above algorithm only covers lowercase letters. Uppercase letters will converted to lowercase.

**Result :**

Test case #1

sequence 1 : gcccagtctatgtcagggggcacgagcatgcaca

sequence 2 : gccgccgtcgttttcagcagttatgttcagat

1 for match, -1 for mismatch and -1 for gap

1. Global Alignment

matches : 22

mismatches : 5

gaps : 12

score : 5

gcc-cagtc-tatgtcagggggcacgagcatg--caca-

gccgccgtcgttt-tcag----ca-gtt-atgttcagat

2. Local Alignment

matches : 12

mismatches : 0

gaps : 2

score : 10

cagtctatgt-cag

cagt-tatgttcag

test case #2

sequence 1 : aactgagtca

sequence 2 : accccctgagagttct

3 for match, -1 for mismatch and -3 for gap

1. Global Alignment

matches : 8

mismatches : 2

gaps : 6

score : 4

aac---tgag--t-ca

accccctgagagttct

2. Local Alignment

matches : 5

mismatches : 0

gaps : 0

score : 15

ctgag

ctgag

**Code Explanations :**

1. **Sequence Alignment Class**

* Member Variables
* int \*\*gScore, \*\*lScore;
* int \*\*gBacktrack, \*\*lBacktrack;
* int maxCol, maxRow, maxLScore;
* string v, w;
* **functions**

LCSBackTrack() : Get the matrices of global & local alignment scores and backtrack.

writeResult() : Output the results to a file.

1. **Global Alignment**

* Needleman-Wunsch algorithm

gScore[i][j] = max(gScore[i-1][j] + gap, gScore[i][j-1] + gap, gScore[i-1][j-1] + match or mismatch)

// The score for the global alignment is stored in the gScore matrix.

* The score calculation uses Dynamic Programming.
* When score matrix is updated, matrix for backtrack is also updated.
* if comes from the upside
* gBacktrack[i][j] = 1;
* else if comes from the side
* gBacktrack[i][j] = 2;
* else if comes from the diagonal side
* gBacktrack[i][j] = 3;

//The backtrack matrix stores which direction it comes from.

1. **Local Alignment**

* Smith-Waterman algorithm

lScore[i][j] = max(lScore[i-1][j] + gap, lScore[i][j-1] + gap, lScore[i-1][j-1] + match or mismatch, 0)

// The difference from the global score is that 0 is added**.**