

## Programming Assignment 2

Handed Out: May 15

Due: May 30 (Fri), 11:59 PM

Your second programming assignment asks you to implement the **Longest Common Subsequence** algorithm. Your program should read two text character strings, and it should output (1) the **length** of the LCS, and (2) the common subsequence as a character string.

Recall that the dynamic program to determine the LCS length is the following:

**LCS-Length** ( $X, Y$ )

```

for  $i = 1$  to  $m$  do  $c[i, 0] \leftarrow 0$ ;
for  $j = 1$  to  $n$  do  $c[0, j] \leftarrow 0$ ;
for  $i = 1$  to  $m$  do
  for  $j = 1$  to  $n$  do
    if  $x_i = y_j$  then
       $c[i, j] \leftarrow c[i - 1, j - 1] + 1$ 
       $b[i, j] \leftarrow D$ 
    else if  $c[i - 1, j] \geq c[i, j - 1]$  then
       $c[i, j] \leftarrow c[i - 1, j]$ 
       $b[i, j] \leftarrow U$ 
    else
       $c[i, j] \leftarrow c[i, j - 1]$ 
       $b[i, j] \leftarrow L$ 
return  $b, c$ 

```

The pseudo code for producing the actual subsequence of maximum length is given below. It uses the  $b$  array computed above.

**PRINT-LCS**( $b, X, i, j$ )

```

if  $i = 0$  or  $j = 0$  then return
if  $b[i, j] = D$  then
  PRINT-LCS( $b, X, i - 1, j - 1$ )
  print  $x_i$ 
elseif  $b[i, j] = U$  then
  PRINT-LCS( $b, X, i - 1, j$ )
else PRINT-LCS( $b, X, i, j - 1$ )

```

## Makefile [5 pts]

You will need a makefile once again. Be sure to follow the specification—*this is an easy 5 points!* The default target (the first one in the makefile) should be called `all` and should compile all required executables. The `clean` target should cleanup all output generated by `all`. `clean` should not complain if there's nothing to do and should not require any interaction, so use `rm -f` not just `rm`. There should be targets for each required executable (`lcs` and `lcs-all`).

## LCS [75 pts]

Your first program should be called `lcs`. It will take exactly two command line arguments, each of them an input file. Each input file will contain a sequence of white-space-separated integer values. Your program should read them in and then print out a longest common subsequence of the two input sequences. The sequence should be output as space-separated integer values on a single line.

## All LCS [20 pts]

The second program should be called `lcs-all`. It will behave just like `lcs` except that it will print out not just one of the longest common subsequences for the two input sequences, but *all* of the longest common subsequences. They will of course all have the same number of items in them since otherwise they wouldn't all be longest. You should print them out in the same format as before, but with one subsequence per line.

**Note:** If you wish, you may produce a single program for this assignment that satisfies the requirements of `lcs-all`. In this case, you should make `lcs` be a symbolic link to the `lcs-all` program by running `ln -sf lcs-all lcs` from your makefile when generating the `lcs` target. The `lcs` part of the assignment will be graded based only on the first line of output; the `lcs-all` part will be graded based on all lines of output.