Design & Analysis of Algorithms Project - 2

Table of Contents:

DESCRIPTION

Longest Common Subsequence **Brute Force approach,** Longest Common Subsequence **Dynamic Programming approach,** Longest Common Subsequence Dynamic Programming approach with b and c arrays algorithms are implemented as part of this project. **Prudhvi Thumma** (Student ID: 1002033401) and **Rohith Kumar Boddu** (Student ID: 1002037081) are members of this team. We both contributed equally towards the implementation of algorithms, drawing comparisons, and documentation. File operations required for the project are mainly read operations. Given LCS1 and LCS2 textfiles consisting of two strings in each line are used to implement LCS_BF.py and LCS_DP.py and LCS_DP_CB.py reads the LCS2 text file to implement the dynamic programming version of LCS-LENGTH(X, Y).

How to run files:

LCS_BF.py, LCS_DP.py, LCS_DP_CB.py are files with code to find the longest common subsequence given in the text files.

Step 1: Install python and set up a runtime environment for python (check system environment variables in the path)

Step 2: Copy the files mentioned above to any directory

Step 3: Open the command prompt, navigate to your directory with files

Step 4: Enter the python "file_name.py" command compile and run the given files. The output will begenerated displaying sorted output and the time taken to complete the given sort operation

Modules to import
- **time**: This module provides various time-related functions.
  - **time()** function returns us the current time based on which I could calculate the difference in time to perform my sort algorithms.

- **read_file** reads content from files, which are written after generating random numbers.

- **time()** gives the current time and is used to calculate the time elapsed to find the longest common subsequence between two strings

- **open()** opens a file with the given mode.

Contributions towards the project:

File read operations: Prudhvi Thumma

Sort algorithms: Team

Report Documentation: Rohith Kumar Boddu

Time complexity comparison: Team

LONGEST COMMON SUBSEQUENCE BRUTE FORCE APPROACH

Run LCS_BF.py file in command prompt as directed, which provides us the output with the time as below

```
E:\Rohith\Spring 2022\Design and Analysis of Algorithms\Project-2>python LCS_BF.py

######################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill
 is 5
Time taken to find length of longest common subsequence is : 0.0468907356262207 seconds
######################################################


######################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1
 is 5
Time taken to find length of longest common subsequence is : 0.08461117744445801 seconds
######################################################


######################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill12
 is 5
Time taken to find length of longest common subsequence is : 0.11593794822692871 seconds
######################################################


######################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill123
 is 5
Time taken to find length of longest common subsequence is : 0.20049309730529785 seconds
######################################################
```

```
########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234
 is 5
Time taken to find length of longest common subsequence is : 0.2539818286895752 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill12345
 is 5
Time taken to find length of longest common subsequence is : 0.34770870208740234 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill123456
 is 5
Time taken to find length of longest common subsequence is : 0.645578145980835 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567
 is 5
Time taken to find length of longest common subsequence is : 0.7031824588775635 seconds
########################################################
```

```
########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567B
 is 5
Time taken to find length of longest common subsequence is : 1.06699538230896 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Bo
 is 6
Time taken to find length of longest common subsequence is : 1.4509756565093994 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Boa
 is 7
Time taken to find length of longest common subsequence is : 1.4875454902648926 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Boar
 is 8
Time taken to find length of longest common subsequence is : 1.3378477096557617 seconds
########################################################
```

```
########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Board
 is 9
Time taken to find length of longest common subsequence is : 0.12591052055358887 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Board1
 is 9
Time taken to find length of longest common subsequence is : 0.47899389266967773 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Board12
 is 9
Time taken to find length of longest common subsequence is : 1.0069780349731445 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Board123
 is 9
Time taken to find length of longest common subsequence is : 1.500481128692627 seconds
########################################################
```

```
########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Board1234
 is 9
Time taken to find length of longest common subsequence is : 3.0752527713775635 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Board12345
 is 9
Time taken to find length of longest common subsequence is : 5.238419532775879 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Board123456
 is 9
Time taken to find length of longest common subsequence is : 8.206242322921753 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Board1234567 is 9
Time taken to find length of longest common subsequence is : 9.06346607208252 seconds
########################################################
```

LONGEST COMMON SUBSEQUENCE DYNAMIC PROGRAMMING APPROACH

Run LCS_DP.py file in command prompt as directed, which provides us the output with the time below

```
E:\Rohith\Spring 2022\Design and Analysis of Algorithms\Project-2>python LCS_DP.py

########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill
 is 5
Time taken to find length of longest common subsequence is : 8.070000330917537e-05 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1
 is 5
Time taken to find length of longest common subsequence is : 9.689998114481568e-05 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill12
 is 5
Time taken to find length of longest common subsequence is : 0.00024310001754201949 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill123
 is 5
Time taken to find length of longest common subsequence is : 0.00017970000044442713 seconds
########################################################
```

```
########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234
 is 5
Time taken to find length of longest common subsequence is : 0.0002615000121295452 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill12345
 is 5
Time taken to find length of longest common subsequence is : 0.00020570002379827201 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill123456
 is 5
Time taken to find length of longest common subsequence is : 0.00025189999723806977 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567
 is 5
Time taken to find length of longest common subsequence is : 0.00022770001669414341 seconds
########################################################
```

```
########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567B
 is 5
Time taken to find length of longest common subsequence is : 0.00023340000018890947 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Bo
 is 6
Time taken to find length of longest common subsequence is : 0.00030440001864917576 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Boa
 is 7
Time taken to find length of longest common subsequence is : 0.00027740001678466797 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Boar
 is 8
Time taken to find length of longest common subsequence is : 0.00034619998768903315 seconds
########################################################
```

```
########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Board
 is 9
Time taken to find length of longest common subsequence is : 0.00023669999791309237 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Board1
 is 9
Time taken to find length of longest common subsequence is : 0.000363900006050244 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Board12
 is 9
Time taken to find length of longest common subsequence is : 0.0003541999903973192 seconds
########################################################


########################################################
Length of the Longest common Subsequence of Strings BillBoard & BoardBill1234567Board123
 is 9
Time taken to find length of longest common subsequence is : 0.0004796000139322132 seconds
########################################################
```

Time Complexity Comparison

We will analyze differences in theoretical time complexity and achieve experimental time complexity. Let's check the time complexity of individual sort algorithms

Longest Common Subsequence Brute Force Approach Theoretical:

- Time Complexity: $O(n * 2^n)$

Theoretically, the time taken to find the longest common subsequence using the brute force approach will increase by order $n * 2^n$

Longest Common Subsequence Dynamic Programming Approach Theoretical:

- Average Time Complexity: $O(n * m)$

Theoretically, the time taken to find the longest common subsequence using the dynamic programming approach will increase by order n*m
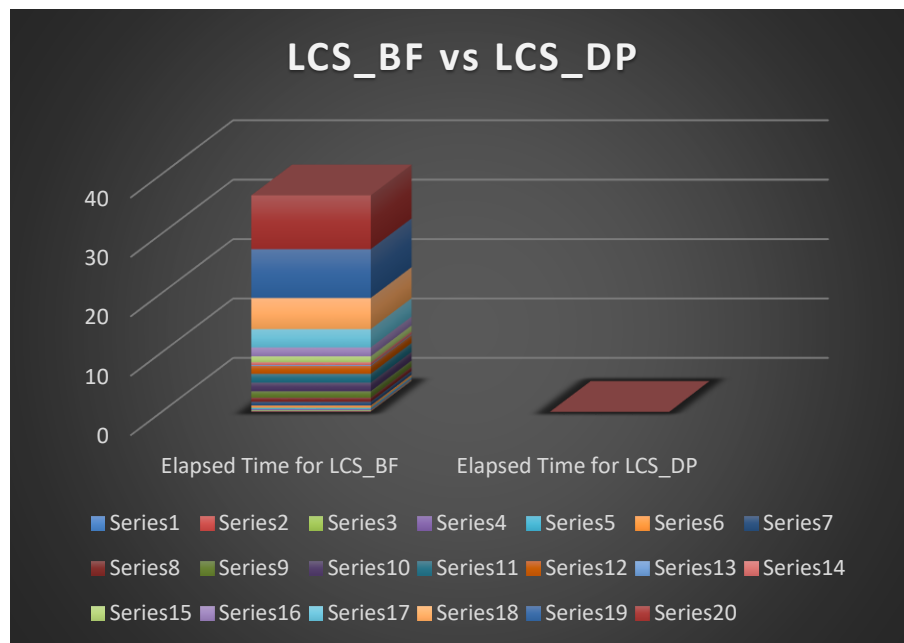
Elapsed Time Comparison between Brute Force and Dynamic Programming Approach:

String 1(X)	String 2(Y)	Length of LCS	Elapsed Time for LCS_BF	Elapsed Time for LCS_DP
BillBoard	BoardBill	5	0.0468907356262207	8.070000330917537e-05
BillBoard	BoardBill1	5	0.08461117744445801	9.689998114481568e-05
BillBoard	BoardBill12	5	0.11593794822692871	0.00024310001754201949
BillBoard	BoardBill123	5	0.20049309730529785	0.00017970000044442713
BillBoard	BoardBill1234	5	0.2539818286895752	0.0002615000121295452
BillBoard	BoardBill12345	5	0.34770870208740234	0.00020570002379827201
BillBoard	BoardBill123456	5	0.645578145980835	0.00025189999723806977
BillBoard	BoardBill1234567	5	0.7031824588775635	0.00022770001669414341
BillBoard	BoardBill1234567B	5	1.06699538230896	0.0002334000018890947
BillBoard	BoardBill1234567Bo	6	1.4509756565093994	0.00030440001864917576

BillBoard	BoardBill1234567Boa	7	1.4875454902648926	0.00027740001678466797
BillBoard	BoardBill1234567Boar	8	1.3378477096557617	0.00034619998768903315
BillBoard	BoardBill1234567Board	9	0.12591052055358887	0.00023669999791309237
BillBoard	BoardBill1234567Board1	9	0.47899389266967773	0.000363900006050244
BillBoard	BoardBill1234567Board12	9	1.0069780349731445	0.0003541999903973192
BillBoard	BoardBill1234567Board123	9	1.500481128692627	0.0004796000139322132
BillBoard	BoardBill1234567Board1234	9	3.0752527713775635	0.00019190000602975488
BillBoard	BoardBill1234567Board12345	9	5.238419532775879	0.00033760000951588154
BillBoard	BoardBill1234567Board123456	9	8.206242322921753	0.00038240000139921904
BillBoard	BoardBill1234567Board1234567	9	9.06346607208252	0.0004285000031813979

In the Brute force approach, all the subsequences are enumerated and tested which are the subsequences of Y and the longest one is picked.
In the Dynamic programming approach, we find the optimal solution using a recursive algorithm.

LCS_BF vs LCS_DP

| | Elapsed Time for LCS_BF | Elapsed Time for LCS_DP |

Legend: Series1, Series2, Series3, Series4, Series5, Series6, Series7, Series8, Series9, Series10, Series11, Series12, Series13, Series14, Series15, Series16, Series17, Series18, Series19, Series20

The above graph depicts the time difference between LCS_BF and LCS_DP algorithms for the given inputs.

LONGEST COMMON SUBSEQUENCE DYNAMIC PROGRAMMING APPROACH WITH b AND c ARRAYS

Run LCS_DP_CB.py file in command prompt as directed, which provides us the output with the time below

```
D:\Prudhvi\MSCS\SEM-1\Design Analysis & Algorithms\Project2>python LCS_DP_CB.py
#####################################################
    Y   B   o   a

X   0   0   0   0   0
B   0 \ 1 < 1 < 1 < 1
a   0 ^ 1 ^ 1 \ 2 < 2
t   0 ^ 1 ^ 1 ^ 2 ^ 2
#####################################################
Length of longest common subsequence is : 2
The longest common subsequence of given strings Bat and Boa
 is Ba

Time taken to find length of longest common subsequence is : 0.0018244999992020894 seconds
_____
#####################################################
    Y   B   a   r   t

X   0   0   0   0   0   0
B   0 \ 1 < 1 < 1 < 1 < 1
o   0 ^ 1 ^ 1 ^ 1 ^ 1 ^ 1
g   0 ^ 1 ^ 1 ^ 1 ^ 1 ^ 1
a   0 ^ 1 \ 2 < 2 < 2 < 2
r   0 ^ 1 ^ 2 \ 3 < 3 < 3
t   0 ^ 1 ^ 2 ^ 3 \ 4 < 4
#####################################################
Length of longest common subsequence is : 4
The longest common subsequence of given strings Bogart and Bart
 is Bart

Time taken to find length of longest common subsequence is : 0.0015662000005249865 seconds
```

```
_____
#####################################################
    Y   P   a   p   e   r

X   0   0   0   0   0   0   0
P   0 \ 1 < 1 < 1 < 1 < 1
o   0 ^ 1 ^ 1 ^ 1 ^ 1 ^ 1
p   0 ^ 1 ^ 1 \ 2 < 2 < 2 < 2
p   0 ^ 1 ^ 1 \ 2 ^ 2 ^ 2 ^ 2
e   0 ^ 1 ^ 1 ^ 2 \ 3 < 3 < 3
r   0 ^ 1 ^ 1 ^ 2 ^ 3 \ 4 < 4
#####################################################
Length of longest common subsequence is : 4
The longest common subsequence of given strings Popper and Paper
 is Pper

Time taken to find length of longest common subsequence is : 0.0034866999994846992 seconds
_____
#####################################################
    Y   P   a   r   t   n   e   r
X   0   0   0   0   0   0   0   0
P   0 \ 1 < 1 < 1 < 1 < 1 < 1
a   0 ^ 1 \ 2 < 2 < 2 < 2 < 2 < 2
n   0 ^ 1 ^ 2 ^ 2 ^ 2 \ 3 < 3 < 3
t   0 ^ 1 ^ 2 ^ 2 \ 3 ^ 3 ^ 3 ^ 3
h   0 ^ 1 ^ 2 ^ 2 ^ 3 ^ 3 ^ 3 ^ 3
e   0 ^ 1 ^ 2 ^ 2 ^ 3 ^ 3 \ 4 < 4
r   0 ^ 1 ^ 2 \ 3 ^ 3 ^ 3 ^ 4 \ 5
#####################################################
Length of longest common subsequence is : 5
The longest common subsequence of given strings Panther and Partner is Pater

Time taken to find length of longest common subsequence is : 0.03247670000200742 seconds
_____
```

<u>REFERENCES</u>

[1] https://edutechlearners.com/download/Introduction_to_algorithms-3rd%20Edition.pdf

[2] https://stackoverflow.com/questions/19095796/how-to-print-a-single-backslash

[3] https://docs.python.org/3/library/time.html

HONOR CODE

HONOR CODE

I pledge, on my honor, to uphold UT Arlington's tradition of academic integrity, a tradition that values hard work and honest effort in the pursuit of academic excellence.

I promise that I will Submit only work that I personally Create or that I Contribute to group Collaborations, and I will appropriately reference any work from other Sources. I will follow the highest standards of integrity and uphold the spirit of honor Code.

I will not participate in any form of Cheating/sharing the questions/solutions

PRUDHVI THUMMA
1002033401

ROHITH KUMAR BODDU.
1002037081

28th April 2022