

CSCI 2270

Data Structures and Algorithms

2/24/2014

finish longest common subsequence

Elizabeth White

elizabeth.white@colorado.edu

Office hours: ECCS 128 or ECCS 112

Wed 9:30am-11:30am

Thurs 10:00am-11:00am

Administrivia

- First midterm exam is in grading
- We'll discuss it today in class
- Read Recursion (chapter 4); concentrate on factorial and binary search sections for now
- This week's lab is on LCS alignment
- If you are BORED and you want to go for extra credit, figure out how to extend this to backtrack and display the matching for the 2 input strings, and show it to me next week.

Finished alignment

-	-	B	A	N	A	N	A
-	0	0	0	0	0	0	0
N	0	0	0	1	1	1	1
A	0	0	1	1	2	2	2
N	0	0	1	2	2	3	3
A	0	0	1	2	3	3	4

Backtracking to get the match is harder:

B	A	N	A	N	A
-	-	N	A	N	A

Deleting the array

-	-	B	A	N	A	N	A
-	0	0	0	0	0	0	0
N	0	0	0	1	1	1	1
A	0	0	1	1	2	2	2
N	0	0	1	2	2	3	3
A	0	0	1	2	3	3	4

```
string a = "BANANA";
```

```
string b = "NANA";
```

```
int** lcs_array = new int*[length(b) + 1];
```

```
for (int k = 0; k < length(b) + 1; ++k)
```

```
    lcs_array[k] = new int[length(a) + 1];
```

Deleting the array

-	-	B	A	N	A	N	A
-	0	0	0	0	0	0	0
N	0	0	0	1	1	1	1
A	0	0	1	1	2	2	2
N	0	0	1	2	2	3	3
A	0	0	1	2	3	3	4

```
for (int k = 0; k < length(b) + 1; ++k)
    delete [] lcs_array[k];
delete [] lcs_array;
```

More than one alignment

-	-	B	O	N	O	B	O
-	0	0	0	0	0	0	0
B	0	1	1	1	1	1	1
O	0	1	2	2	2	2	2
B	0	1	2	2	2	3	3
O	0	1	2	2	3	3	4

Backtracing the match is harder (multiple matches can happen)

B	O	N	O	B	O
B	-	-	O	B	O

or

B	O	N	O	B	O
B	O	-	-	B	O

Longest common subsequence

Real sequence data is hundreds to millions of letters long...

These 2D tables can get really big!

What if we had to align 3 strings? Oh, the humanity.