CS420+500: Advanced Algorithm Design and Analysis

Lectures: March 01 + 03, 2017

Prof. Will Evans *Scribe: ENRIQUE ROSALES*

In this lecture we:

- Discussed an efficient algorithm to find the Longest Increasing Subsequence (LIS);
- And how to solve the Longest Common Subsequence (LCS) using LIS;

Handouts (posted on webpage):

- Group-work exercises (Wednesday March 1)

Reading: ASSIGNED READING(S).

1 Dynamic Programming

1.1 Longest Increasing Subsequence (LIS)

Having $R = \{5, 3, 4, 9, 6, 2, 1, 8\}$ the Longest Increasing Subsequence is: $\{3, 4, 6, 8\}$

$$R = \{5, \textcircled{3}, \textcircled{4}, 9, \textcircled{6}, 2, 1, \textcircled{8}\}$$

Using Longest Common Subsequence (LCS):

1. Sort R $\mathcal{O}(n \log n)$
2. Report LCS $(\underbrace{R}_{n}, \underbrace{\text{sort } (R)}_{n})$ $\mathcal{O}(n^2)$

Faster solution to LIS
When finding LIS of $R\{1...k\}$ what information about $R\{1...k-1\}$ would be useful?

A) LIS $R\{1...k-1\}$

Not enough because $R\{\overbrace{5, 3, 4, 9, 6, 2, 1}^{349 \text{ vs } 346}, 8\}$ ↑

B) Best LIS $R\{1...k-1\}$ (one with smallest last value)

Not enough because $R\{1, 2, 5, 3, 4\}$ ↑

C) Best Increasing Subsequence (BIS) of lengths $\{1, 2, 3...j\}$

Example step by step:

1. $R\{8, 3, 4, 9, 6, 2, 1, 5, 7, 6\}$

   $\uparrow$

   $BIS[1] = \{1\}$
   $BIS[2] = \{3, 4\}$
   $BIS[3] = \{3, 4, 6\}$
   $BIS[4] = \{\emptyset\}$

2. $R\{8, 3, 4, 9, 6, 2, 1, 5, 7, 6\}$

   $\uparrow$

   $BIS[1] = \{1\}$
   $BIS[2] = \{3, 4\}$
   $BIS[3] = \{3, 4, 5\}$
   $BIS[4] = \{\emptyset\}$

3. $R\{8, 3, 4, 9, 6, 2, 1, 5, 7, 6\}$

   $\uparrow$

   $BIS[1] = \{1\}$
   $BIS[2] = \{3, 4\}$
   $BIS[3] = \{3, 4, 5\}$
   $BIS[4] = \{3, 4, 5, 7\}$

This is probably the solution but, how can we compute this in less time?

The solution is to avoid writing everything let's see an example step by step:

1. $R\{8, 3, 4, 9, 6, 2, 1, 5, 7, 6\}$

   $\uparrow$

   $BIS[1] = \{\textcircled{8}\}$
   $BIS[2] = \{\emptyset\}$
   $BIS[3] = \{\emptyset\}$
   $BIS[4] = \{\emptyset\}$

2. $R\{8, 3, 4, 9, 6, 2, 1, 5, 7, 6\}$

   $\uparrow$

   $BIS[1] = \{\cancel{8}, \textcircled{3}\}$
   $BIS[2] = \{\emptyset\}$
   $BIS[3] = \{\emptyset\}$
   $BIS[4] = \{\emptyset\}$

3. $R\{8, 3, 4, 9, 6, 2, 1, 5, 7, 6\}$

   $\uparrow$

   $BIS[1] = \{\cancel{8}, \textcircled{3}\}$
   $\nearrow$
   $BIS[2] = \{\textcircled{4}\}$
   $BIS[3] = \{\emptyset\}$
   $BIS[4] = \{\emptyset\}$

4. $R\{8, 3, 4, 9, 6, 2, 1, 5, 7, 6\}$

   $\uparrow$

   $BIS[1] = \{\cancel{8}, \textcircled{3}\}$
   $\nearrow$
   $BIS[2] = \{\textcircled{4}\}$
   $\uparrow$
   $BIS[3] = \{\textcircled{9}\}$
   $BIS[4] = \{\emptyset\}$

5. $R\{8, 3, 4, 9, 6, 2, 1, 5, 7, 6\}$

$BIS[1] = \{\cancel{8}, \textcircled{3}\}$

$BIS[2] = \{\textcircled{4}\}$

$BIS[3] = \{\cancel{9}, \textcircled{6}\}$

$BIS[4] = \{\emptyset\}$

6. $R\{8, 3, 4, 9, 6, 2, 1, 5, 7, 6\}$

$BIS[1] = \{\cancel{8}, \textcircled{3}, 2\}$

$BIS[2] = \{\textcircled{4}\}$

$BIS[3] = \{\cancel{9}, \textcircled{6}\}$

$BIS[4] = \{\emptyset\}$

7. $R\{8, 3, 4, 9, 6, 2, 1, 5, 7, 6\}$

$BIS[1] = \{\cancel{8}, \textcircled{3}, \cancel{2}, 1\}$

$BIS[2] = \{\textcircled{4}\}$

$BIS[3] = \{\cancel{9}, \textcircled{6}\}$

$BIS[4] = \{\emptyset\}$

8. $R\{8, 3, 4, 9, 6, 2, 1, 5, 7, 6\}$

$BIS[1] = \{\cancel{8}, \textcircled{3}, \cancel{2}, 1\}$

$BIS[2] = \{\textcircled{4}\}$

$BIS[3] = \{\cancel{9}, \cancel{6}, \textcircled{5}\}$

$BIS[4] = \{\emptyset\}$

9. $R\{8, 3, 4, 9, 6, 2, 1, 5, 7, 6\}$

$BIS[1] = \{\cancel{8}, \textcircled{3}, \cancel{2}, 1\}$

$BIS[2] = \{\textcircled{4}\}$

$BIS[3] = \{\cancel{9}, \cancel{6}, \textcircled{5}\}$

$BIS[4] = \{\textcircled{7}\}$

10. $R\{8, 3, 4, 9, 6, 2, 1, 5, 7, 6\}$

$BIS[1] = \overbrace{\{\cancel{8}, \textcircled{3}, \cancel{2}, 1\}}^{\text{decreasing order}}$

$BIS[2] = \{\textcircled{4}\}$

$BIS[3] = \{\cancel{9}, \cancel{6}, \textcircled{5}\}$

$BIS[4] = \{\cancel{7}, \textcircled{6}\}$

To add the next number in R perform binary search on the last values in the BISs and put the next number in sorted order, pointing to the end value of the previous BIS (black arrows) having a running time of $\mathcal{O}(n \log n)$

1.2 Use LIS to solve LCS

To find a sequence of index pairs of matches so that sequence increases in both coordinates.

Example:

$$\{A, \textcircled{B}, \textcircled{C}, B, \textcircled{A}, \textcircled{C}, \textcircled{C}, B\}$$
$$\{\textcircled{B}, \textcircled{C}, D, \textcircled{A}, B, \textcircled{C}, \textcircled{C}\}$$

We can see the relation between the sequence in table 1

Table 1: LCS pairs

		1	2	3	4	5	6	7	8
		A	B	C	B	A	C	C	B
1	B		X		X				X
2	C			X			X	X	
3	D								
4	A	X				X			
5	B		X		X				X
6	C			X			X	X	
7	C			X			X	X	

In this case, we can see that the coordinates of the LCS pairs have increasing order in both coordinates.

$$\{\overbrace{(1,2),(2,3),(4,5),(6,6),(7,7)}^{\text{increasing order}}\}$$

It is easy to obtain the matching pairs, and to list them by its row and column coordinates:

$$\begin{aligned} \text{Rows } &= \{1,1,1,2,2,2...\} \\ \text{Columns } &= \{2,4,8,3,6,7...\} \end{aligned}$$

The trick is to invert the column numbers that have the same value in row.

$$\begin{aligned} \text{Rows } &= \{1,1,1,2,2,2...\} \\ \text{Columns } &= \{\underbrace{8,4,2}_{\text{inverted}},\underbrace{7,6,3}_{\text{inverted}}...\} \end{aligned}$$

List coordinates of matching characters in order by row index and, within one row, in reverse order by column number, then run LIS on the sequence of columns order.

Another way to see this problem is to determine a value of 1 to vertical and horizontal movements, and a value of 0 to diagonal movements like in a chessboard, having this, the problem can be seen as the shortest path between the upper left corner and the lower right corner as seen in table 2.

Table 2: *LCS* as Shortest Path

		1	2	3	4	5	6	7	8
		A	B	C	B	A	C	C	B
1	B		X		X				X
2	C			X			X	X	
3	D								
4	A	X				X			
5	B		X		X				X
6	C			X			X	X	
7	C			X			X	X	