The Hirschberg Algorithm: Intuition

X = “t o o k “ Y = “ b o o k”

1. Finding the LCSs of partitions of the input can give us the LCS of the inputs. E.g.

X = t o | o k Y = b o | o k

X_{left}= “t o” and X_{right} = “o k”
Y_{left} = “b o” Y_{right} = “o k”

LCS(left) = “o” and LCS(right) = “ o k” LCS(left)+LCS(right) = “o o k”

2. One string can be partitioned in any way but the other has to be partitioned such that the LCS is not broken. E.g. the following partition does not lead to the LCS

X = t o | o k Y = b o o | k

X_{left}= “t o” and X_{right} = “o k”
Y_{left} = “b o o” Y_{right} = “k”

LCS(left) = “o” and LCS(right) = “ k” LCS(left)+LCS(right) = “o k”

Example

X = AGTACGCA
Y = TATGC
Match = +2, Mismatch = -1, Gap = -2

Step 1: Partition X at the mid point

X_{left}=AGTA and X_{right} = CGCA

Step 2: Compute NW-Score(X_{left}, Y) and NW-Score (reverse(X_{right}), reverse(Y))

NW-Score(X_{left}, Y) = NW-Score(“AGTA” , “TATGC”)
NW-Score (reverse(X_{right}), reverse(Y)) = NW-Score(“ACGC” , “CGTAT”)

		T	A	T	G	C
	0	-2	-4	-6	-8	-10
A	-2	-1	0	-2	-4	-6
G	-4	-3	-2	-1	0	-2
T	-6	-2	-4	0	-2	-1
A	-8	-4	0	-2	-1	-3

NW-Score (reverse(X_{right}), reverse(Y)) = NW-Score(“ACGC” , “CGTAT”)

		C	G	T	A	T
	0	-2	-4	-6	-8	-10
A	-2	-1	-3	-5	-4	-6
C	-4	0	-2	-4	-6	-5
G	-6	-2	2	0	-2	-4
C	-8	-4	0	-1	-1	-3

Step-3 $L(mid) = \arg\max_i L(i)$

-8	-4	0	-2	-1	-3
-8	-4	0	-1	-1	-3

i* = 2

So the partition is: X = AGTA | CGCA and Y = TA | TGC