Hirschberg

Hint: Many test values are taken from project Algorithms
for Bioninformatics of Alexander Mattheis or the lectures.

Test 1 (used Needleman-Waterman-Implementation for calculation)

Input

Sequence A: AATCG
Sequence B: AACG

Deletion: 2
Insertion: 2
Match: -1
Mismatch: 1

Output

1. $Trace(A_1, A_2, T_3, C_4, G_5 | A_1, A_2, C_3, G_4)$:

		A	**A**	**C**	**G**
	0	2	4	6	8
A	2	-1	1	3	5
A	4	1	-2	0	2
T	6	3	0	-1	1
C	8	5	2	-1	0
G	10	7	4	1	-2

		G	**C**	**A**	**A**
	0	2	4	6	8
G	2	-1	1	3	5
C	4	1	-2	0	2
T	6	3	0	-1	1
A	8	5	2	-1	-2
A	10	7	4	1	-2

	0	**1**	**j=2**	**3**	**4**
i=3	6	3	0	-1	1
	1	-1	0	3	6
∑	7	2	0	2	7

		A	**A**	**C**	**G**
A					
A					
T					
C					
G					

1.1. $Trace(A_1, A_2|A_1, A_2)$:

		A	A
	0	2	4
A	2	-1	1
A	4	1	-2

		A	A
	0	2	4
A	2	-1	1
A	4	1	-2

	0	j=1	2
i=1	2	-1	1
	1	-1	2
∑	3	-2	3

		A	A
A			
A			

1.1.1. $Trace(|A_1)$:

```
return length(A) <= 1
→ select most left
```

	j=0	A
i=0		

1.1.**2.** $Trace(A_2|A_1, A_2)$:

```
return length(A) <= 1
→ select most right (because right node)
```

	A	A j=2
A i=2		

1.2. $Trace(C_4, G_5|A_2, C_3, G_4)$:

		A	C	G
	0	2	4	6
C	2	1	1	3
G	4	3	2	0

		G	C	A
	0	2	4	6
G	2	-1	1	3
C	4	1	-2	0

	0	2	j=3	4
i=4	2	1	1	3
	3	1	-1	2
∑	5	2	0	5

	A	C	G
C			
G			

1.2.1. $Trace(|A_2, C_3)$:
`return length(A) <= 1`

1.2.**2.** $Trace(C_4|C_3, G_4)$:
`return length(A) <= 1`
`→ select most right (because right node)`

	C	G j=4
G i=5		

so:

		A	A	C	G
A					
A					
T					
C					
G					

Test 2 (used Needleman-Waterman-Implementation for calculation)

Input

```
Sequence A:       TCG
Sequence B:       TCTAG

Deletion:          2
Insertion:         2
Match:            -1
Mismatch:          1
```

Output

1. $Trace(T_1, C_2, G_3 | T_1, C_2, T_3, A_3, G_4)$:

		T	C	T	A	G
	0	2	4	6	8	10
T	2	-1	1	3	5	7
C	4	1	-2	0	2	4
G	6	3	0	-1	1	1

		G	A	T	C	T
	0	2	4	6	8	10
G	2	-1	1	3	5	7
C	4	1	0	2	2	4
T	6	3	2	-1	1	1

$$i = \left\lceil \frac{\text{length(A)}}{2} \right\rceil = \left\lceil \frac{3}{2} \right\rceil = 2$$

	0	1	j=2	3	4	5
i=2	4	1	-2	0	2	4
	4	2	2	0	1	4
∑	8	3	0	0	3	8

		T	C	T	A	G
T						
C						
G						

1.1. $Trace(T_1 | T_1, C_2)$:
`return length(A) <= 1`
`→ select most left`

1.2. $Trace(G_1 | C_2, T_3, A_3, G_4)$:
`return length(A) <= 1`
`→ select most right (because right node)`

so:

		T	C	T	A	G
T						
C						
G						

logic completion: (maybe wrong and other minima has to be stored)
(not possible)

		T	C	T	A	G
T						
C				X	X	X
G						

(possible)

		T	C	T	A	G
T						
C				X	X	
G						