Computational Biology / Bioinformatics
Lab 2 - Report

Group 22

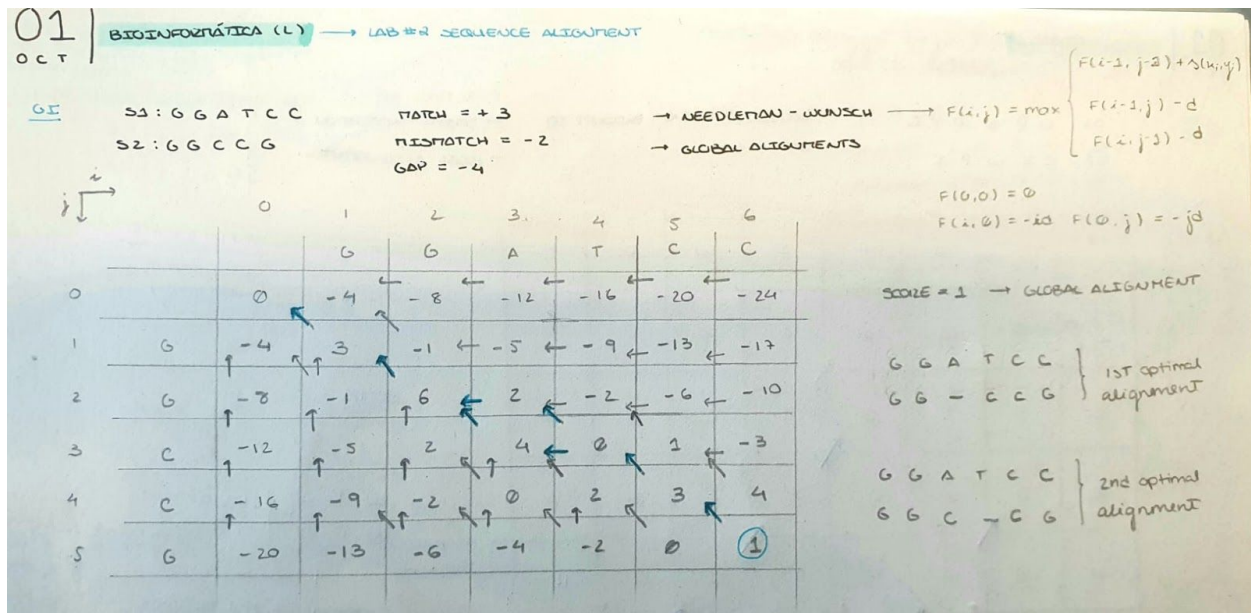**Group Members:**

Alexandra Maroco 86369
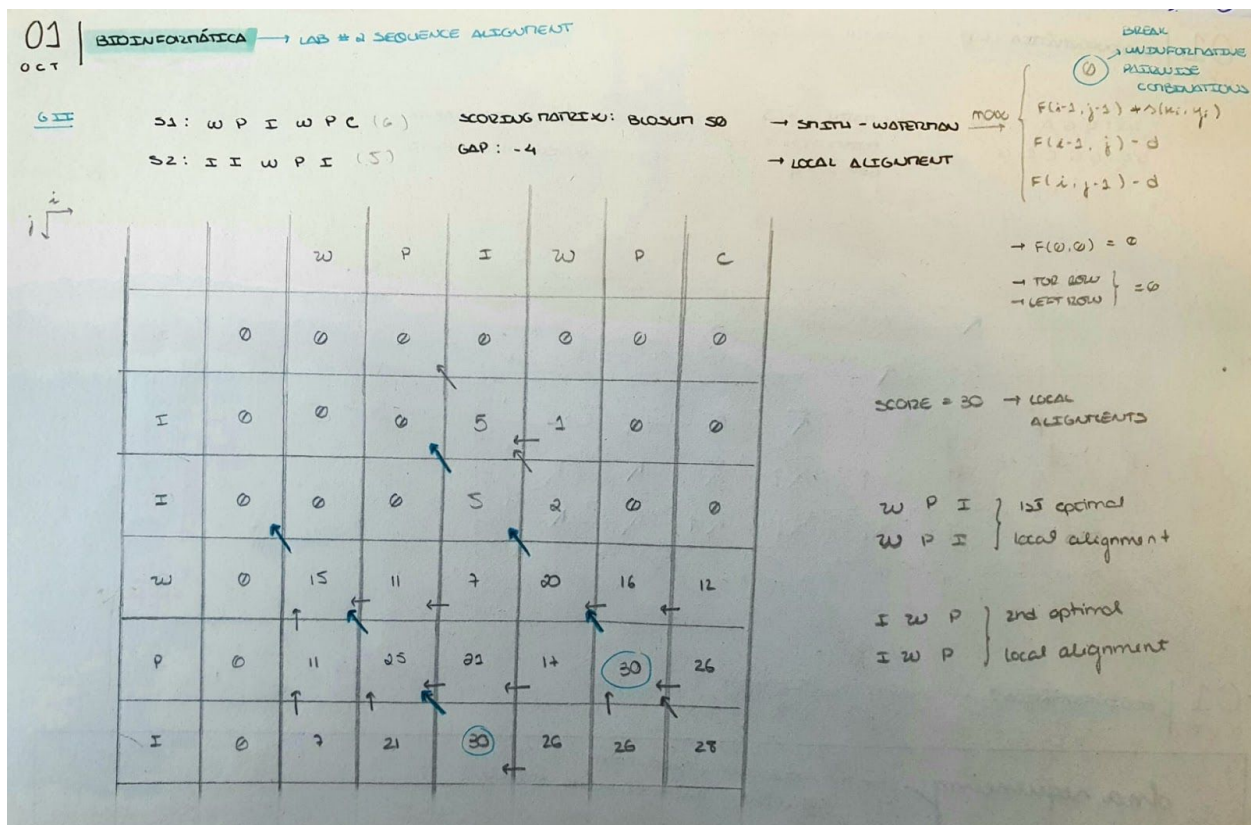André Branco 90013
Joana Alvoeiro 89469
Pedro Nunes 89525

Group I

GI

S1: G G A T C C MATCH = + 3 → NEEDLEMAN - WUNSCH

S2: G G C C G MISMATCH = - 2 → GLOBAL ALIGNMENTS

GAP = - 4

$$F(i,j) = \max \begin{cases} F(i-1, j-1) + s(x_i, y_i) \\ F(i-1, j) - d \\ F(i, j-1) - d \end{cases}$$

$F(0,0) = 0$

$F(i,0) = -id$ $F(0,j) = -jd$

		0	1	2	3.	4	5	6
			G	G	A	T	C	C
0		0	-4	-8	-12	-16	-20	-24
1	G	-4	3	-1	-5	-9	-13	-17
2	G	-8	-1	6	2	-2	-6	-10
3	C	-12	-5	2	4	0	1	-3
4	C	-16	-9	-2	0	2	3	4
5	G	-20	-13	-6	-4	-2	0	(1)

SCORE = 1 → GLOBAL ALIGNMENT

```
G G A T C C      1ST optimal
G G - C C G      alignment
```

```
G G A T C C      2nd optimal
G G C - C G      alignment
```

Group II

BREAK → UNINFORMATIVE

GII

S1: W P I W P C (6) SCORING MATRIX: BLOSUM 50 → SMITH - WATERMAN

S2: I I W P I (5) GAP: -4 → LOCAL ALIGNMENT

$$\max \begin{cases} F(i-1, j-1) + s(x_i, y_i) \\ F(i-1, j) - d \\ F(i, j-1) - d \end{cases}$$

→ F(0,0) = 0

→ TOP ROW / → LEFT ROW } = 0

PAIRWISE COMBINATIONS

		W	P	I	W	P	C
		0	0	0	0	0	0
I	0	0	0	5	-1	0	0
I	0	0	0	5	2	0	0
W	0	15	11	7	20	16	12
P	0	11	25	21	17	(30)	26
I	0	7	21	(30)	26	26	28

SCORE = 30 → LOCAL ALIGNMENTS

```
W P I      1st optimal
W P I      local alignment
```

```
I W P      2nd optimal
I W P      local alignment
```

Group III

When the input provided to the program is the same as in Question II, our program outputs the following:

```
C:\Users\alexa\Documents\IST\5th\Bioinformática\smith_waterman>python3 smith_waterman.py
Input first sequence:
WPIWPC
Input second sequence:
IIWPI
Input gap penalty:
4

Score Matrix:
[[ 0  0  0  0  0  0]
 [ 0  0  0 15 11  7]
 [ 0  0  0 11 25 21]
 [ 0  5  5  7 21 30]
 [ 0  1  2 20 17 26]
 [ 0  0  0 16 30 26]
 [ 0  0  0 12 26 28]]

The best local alignment score is 30.0

Optimal partial alignments:
Alignment 0:
WPI
WPI

Alignment 1:
IWP
IWP
```

Group IV

GIV

S1: A A C G T C   MATCH = +2   → MULTIPLE SEQUENCE ALIGNMENT   match score for u and v

S2: A G C G E C   MISMATCH = -1   → SUM OF PAIRS SCORE MODEL   ⟶   -d

S3: C C C G T   GAP = -3   0

S4: A C A T

① Align S1 & S2

		A	G	C	G	C	C
	0	-3	-6	-9	-12	-15	-18
A	-3	2	-1	-4	-7	-10	-13
A	-6	-1	1	-2	-5	-8	-11
C	-9	-4	-2	3	0	-3	-6
G	-12	-7	-2	0	5	2	-1
T	-15	-10	-5	-3	2	4	1
C	-18	-13	-8	-3	-1	4	⑥

$$F(i,j) = \max \begin{cases} F(i-1,j-1) + s(x_i, y_i) \\ F(i-1,j) - d \\ F(i,j-1) - d \end{cases}$$

$F(0,0) = 0$

$F(i,0) = -id$

$F(0,j) = -jd$

→ A G C G C C

   A A C G T C

② Align S1 & S3

		C	C	C	G	T
	0	-3	-6	-9	-12	-15
A	-3	-1	-4	-7	-10	-13
A	-6	-4	-2	-5	-8	-11
C	-9	-4	-2	0	-3	-6
G	-10	-7	-5	-3	2	-1
T	-15	-10	-8	-6	-1	4
C	-17	-13	-8	-6	-4	(1)

③ Align S1 & S4

		A	C	A	T
	0	-3	-6	-9	-12
A	-3	2	-1	-4	-7
A	-6	-1	1	1	-2
C	-9	-4	-1	0	0
G	-12	-7	-2	0	-1
T	-15	-10	-5	-3	2
C	-18	-13	-8	-6	(-1)

④ Align S2 & S3

		C	C	C	G	T
	0	-3	-6	-9	-12	-15
A	-3	-1	-4	-7	-10	-13
G	-6	-4	-2	-5	-5	-8
C	-9	-4	-2	0	-3	-6
G	-12	-7	-5	-3	2	-1
C	-15	-10	-5	-3	-1	1
C	-17	-13	-8	-3	-4	(-2)

⑤ Align S2 & S4

		A	C	A	T
	0	-3	-6	-9	-12
A	-3	2	-1	-4	-7
G	-6	-1	1	-2	-5
C	-9	-4	1	0	-3
G	-12	-7	-2	0	-1
C	-15	-10	-5	-3	-1
C	-17	-13	-8	-6	(-4)

(6) Align S3 & S4

		A	C	A	T
	0 ←	-3 ←	-6 ←	-9 ←	-12
C	-3 ↑	-1 ↑↖	-1 ↖	-4 ←	-7 ←
C	-6 ↑	-4 ↖↑	1 ↖↑	-2 ←	-5 ←
C	-9 ↑	-7 ↖↑	-2 ↑	0 ↖↑	-3 ←
G	-10 ↑	-10 ↖↑	-5 ↑	-3 ↖↑	-1 ↖
T	-15	-13	-8	-6	(-1)

→ MATRIX:

	S1	S2	S3	S4
S1	0	6	1	-1
S2		0	-2	-4
S3			0	-1
S4				0

$$w(x,y) = \begin{cases} 2 & \text{if } x = y \\ -3 & \text{if } x = - \lor y = - \\ -1 & \text{otherwise} \end{cases}$$

→ THE BEST SCORING ALIGNMENT IS S1 & S2.
AS SUCH, WE SELECT S1 S2 & S3 S4.

→ ALIGNMENT FOR S1 & S2:   A G C G C C
                            A A C G T C

→ ALIGNMENT FOR S3 & S4:   C C C G T
(one of the possible
optimal global alignments) A - C A T

→ align A G C G C C with C C C G T
    A A C G T C   A - C A T

$w(--,CA) = w(-,C) + w(-,A) +$
$+ w(-,C) + w(-,A) =$
$= -3 -3 -3 -3 = -12$

$w(--,C-) = w(-,C) + w(-,-) +$
$+ w(-,C) + w(-,-) =$
$= -3 +2 -3 +2 = -2$

$w(AA,CA) = w(A,C) + w(A,A) +$
$+ w(A,C) + w(A,A) =$
$= -1 +2 -1 +2 = 2$

$w(AA,C-) = w(A,C) + w(A,-) +$
$+ w(A,C) + w(A,-) =$
$= -1 -3 -1 -3 = -8$

$w(CC,CA) = 2w(C,C) + 2w(C,A) = 2(2) + 2(-1) = 2$

$w(CC,C-) = 2w(C,C) + 2w(C,-) = 2(2) + 2(-3) = -2$

$w(CC,CC) = 4w(C,C) = 4(2) = 8$

$w(CC,GA) = 2w(C,G) + 2w(C,A) = 2(-1) + 2(-1) = -4$

$w(CC,TT) = 4w(C,T) = 4(-1) = -4$

$w(GG,CA) = 2w(G,C) + 2w(G,A) = 2(-1) + 2(-1) = -4$

$w(GG,C-) = 2w(G,C) + 2w(G,-) = 2(-1) + 2(-3) = -8$

$w(GG,CC) = 4w(G,C) = 4(-1) = -4$

$w(GG,GA) = 2w(G,G) + 2w(G,A) = 2(2) + 2(-1) = 2$

$w(AG,TT) = 2w(A,T) + 2w(G,T) = 2(-1) + 2(-1) = -4$

		CA	C-	CC	GA	TT
		0	-12	-14	-26	-38	-50
AA	-12	2	-10	-18	-24	-36
AG	-24	-10	-6	-14	-16	-27
CC	-36	-22	-8	2	-10	-20
GG	-48	-34	-20	-10	4	-8
TC	-60	-46	-12	-8	-8	6
CC	-72	-58	-14	-4	-12	(-6)

$w(AA,CC) = 4w(A,C) = 4(-1) = -4$

$w(AA,GA) = 2w(A,G) + 2w(A,A)$
$= 2(-1) + 2(2) = 2$

$w(AA,TT) = 4w(A,T) = 4(-1) = -4$

$w(AG,CA) = w(A,C) + w(A,A)$
$+ w(G,C) + w(G,A) =$
$= -1 +2 -1 -1 = -1$

$w(AG,C-) = w(A,C) + w(A,-) +$
$+ w(G,C) + w(G,-) =$
$= -1 -3 -1 -3 = -8$

$w(AG,CC) = 2w(A,C) + 2w(G,C)$
$= 2(-1) + 2(-1) = -4$

$w(AG,GA) = w(A,G) + w(A,A) +$
$+ w(G,G) + w(G,A) =$
$= -1 +2 +2 -1 = 2$

→ TRACING BACK THE MATRIX, WE OBTAIN THE FOLLOWING ALIGNMENT:

$$\hslash(u,v) = \begin{cases} \text{match score for } u \text{ and } v, \\ \quad \text{if both are residues} \\ -d \text{ if } u \text{ or } v \text{ is a gap} \\ 0 \text{ if both } u \text{ and } v \text{ are gaps} \end{cases}$$

(1)	(2)	(3)	(4)	(5)	(6)
A	A	C	G	T	C
A	G	C	G	C	C
C	C	C	G	T	—
A	—	C	A	T	—

SP-SCORE :

SCORE = S(1) + S(2) + S(3) + S(4) + S(5) + S(6)

S(1): A - A PAIRS : 3
A - C PAIRS : 3
SCORE = 3×2 + 3×(-1) = 3

S(2): A - G PAIRS : 1
A - C PAIRS : 1
A - GAP PAIRS : 1
G - C PAIRS : 1
G - GAP PAIRS : 1
C - GAP PAIRS : 1
SCORE = -1 -1 -3 -1 -3 -3 = -12

S(3): C - C PAIRS : 6
SCORE = 6(2) = 12

S(4): G - G PAIRS : 3
G - A PAIRS : 3
SCORE = 3(2) + 3(-1) = 3

S(5): T - T PAIRS : 3
T - C PAIRS : 3
SCORE = 3(2) + 3(-1) = 3

S(6): C - C PAIRS : 1
C - GAP PAIRS : 4
GAP - GAP PAIRS : 1
SCORE = 1(2) + 4(-3) + 0 = -10

TOTAL
SCORE = S(1) + S(2) + S(3) + S(4) + S(5) + S(6)

= 3 + (-12) + 12 + 3 + 3 + (-10)

= -1