

Sequence Bioinformatics

WS 2022/23

Tutor: Anupam Gautam

1	2	3	4	5	Σ

Marina Dittschar & Clarissa

Auckenthaler

Blatt 3

(Abgabe am 09.11.2022)

Task 1: Counting (2 points)

To run our code, please navigate to the src folder and run the following command:

```
java assignment03/CountEdgesSimpleMixedCycles_Auckenthaler_Dittschar.java 4
```

4 5 where the three integer arguments are the number of letters in three sequences. There are 56 edges and 3184 simple mixed cycles.

Task 2 3: Simple mixed cycles & Objective function

To run our code, run the following command:

```
java assignment03/AlignmentILP_Auckenthaler_Dittschar.java "assignment03/three.fasta"  
"assignment03/output_three_Auckenthaler_Dittschar.lp"
```

where the first argument is the input fasta file and the second argument is the desired output file.

If you omit the output file argument, the constraints will be printed to the terminal.

Task 4: Run the ILP (3 points)

You can see the output of lp_solve in Figure 1.

Variables	MILP ...	MILP ...	MILP ...	MILP ...	MILP ...	MILP ...	re... ▼
	24	26	29	31	34	36	36
X00_10	1	1	1	1	1	1	1
X01_11	1	1	1	1	1	1	1
X02_13	0	0	0	0	0	1	1
X11_20	0	0	0	1	1	1	1
X12_21	0	0	1	1	1	1	1
X13_22	0	0	1	0	1	1	1
X01_20	0	0	0	1	1	1	1
X02_22	1	1	1	1	1	1	1
X03_23	1	1	1	1	1	1	1

Figure 1: Cutout of our result from 'lp_solve' for the 'three.fasta' sequences.

Task 5: Report the alignment (1 point)

All edges that are part of the optimal solution can be seen in Figure 1 as having a value of 1 in the "result" column of the output, contrasting of a value of 0 for all edges that are not part of the final alignment.

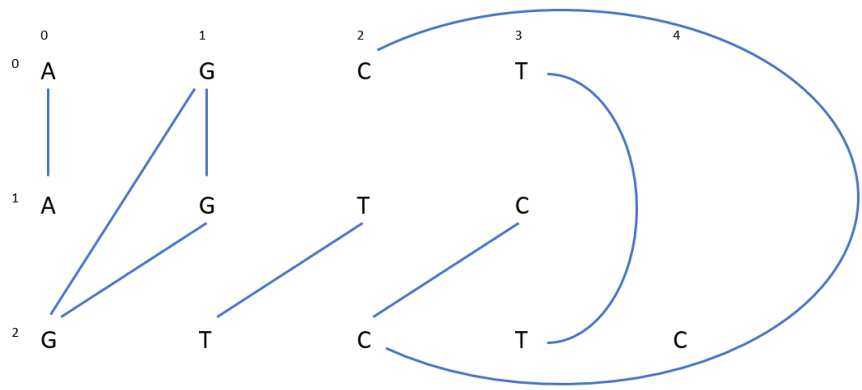


Figure 2: Alignment graph based on our result from task 4.

The alignment that corresponds to Figure 2 looks as follows:

A	G	-	C	T	-
A	G	T	C	-	-
-	G	T	C	T	C