Programming Lab III, WS 2016/2017

Handout 4

Dr. Marc Zimmermann, Jens Dörpinghaus
2016-11-08

Fourth task – Extending the annotator with sequence alignments; perform further unit tests.

Has to be completed finally on 2016-11-22, 11am.

At first you should complete the CollectionReader implementation and the Junit tests for the reader. The next task will be to compute an alignment of the sequences from the FASTA files and to test the pipelet (the AnalysisEngine).

Here's the list of topics you have to complete:

- Compute the pairwise sequence alignments(s) of all sequences within a FASTA file. BioJava has a module called biojava-alignment which helps you. Let the user decide whether to use *Smith-Waterman* or *Needleman-Wunsch*.

- Add the alignment computation to your *annotator*. Write the result into a new **View**, since the documentview must not be modified.

- Add a **parameter** to your annotator which switches between global and local alignment.

- Implement the corresponding unit tests concerning the new parameter and the new view.

- We will checkout your code and run the unit tests – all of them should be **green**.

Please make sure that you commit code that is either self-explaining or has comments to explain your code.

Good luck and have fun!