Prof. Dr. Daniel Huson Anna Gorska Zentrum für Bioinformatik Fachbereich Informatik Mathematisch-Naturwissenschaftliche Fakultät



Advanced Java for bioinformatics. WS 2015/16 Getting started! Due: 21 October 2015

1 Setting up

- Download and install newest JDK (Java SE Development Kit 8u60).
- Download and install newest Intellij IDEA Java IDE we will be using during the course (https://www.jetbrains.com/idea/).
- If you have not used version control before read about it (even on Wikipedia), and make sure you know basic git functionalities (one of the many online git tutorials: https://www.atlassian.com/git/tutorials/).
- Install git and create a GitHub account.

2 Java reminder (5 points)

Write a command line tool called FastaReader in Java. FastaReader should be able to read multi fasta files and fasta alignment files, containing RNA sequences and print them formatted on the screen. For the exemplary file (exemplary_RNA_aln.fa) the formatting should look like this:

	1 60
AB003409.1/96-167	GGGCCCAU-A-GCUCAGUGGUAGAGUG-C-CUCCU-UUGCAAGGAG-GAU
AB009835.1/1-71	CAUUAGAU-G-ACUGAAAGCAAGUA-C-UGGUC-UCUUAAACCA-UUU
AB013372.1/8-81	GCGCCCGU-A-GCUCAAUUGGAUAGAGCG-U-UUGAC-UACGGAUCAA-AAG
AB013373.1/3754-3825	GCGGAAGU-A-GUUCAGUGGUAGAACA-C-CACCU-UGCCAAGGUG-GGG
	61 119
AB003409.1/96-167	GCCCUG-GGU-UCG-AAUCCCA-G-UGGGUCC-A
AB009835.1/1-71	AAUAGUAAAU-UAG-CACUUAC-U-UCUAAUG-A
AB013372.1/8-81	GUUAGG-GGU-UCG-ACUCCUC-U-CGGGCGC-G
AB013373.1/3754-3825	GU-CGCG-GGU-UCG-AAUCCCG-U-CUUCCGC-U

FastaReader should take the file name from the parameter list. The tool should consist of at least four classes: Nucleotide, Sequence, FastaReader, CommandLine.

Handing in

In order to hand in the assignment - send address of your repository on GitHub to anna.gorska@student.unituebingen.de, I will check out the repository and grade your code.