**Доклад по проекта**

**Investigation of Gene sequences**

**За куса ООП с Java**

**2020/2021г.**

*Изготвил: Виктор Василев, 2 курс Информатика ФМИ,*

*ф.н. 45676*

1.Съдържание на проекта.

Проекта е съставен от три основни части.

В основата са двата класа Sequence.java и SequenceAnalyzer.java. Sequence.java представлява обект съдържащ името на секвенцията (ID), както и основното текстово съдържание на файла (неформатирано).

SequenceAnalyzer.java е „по-важният“ обект, който има за цел да изследва подадена нуклеотидна секвенция (Sequence). Този клас извършва значителната част от операциите върху секвенцията описани в условието. Той разполага със секвенцията и нейното форматирано текстово съдържание като член-данни. Също са включени изброими типове за основните нуклеотидни бази (A, T, G, C, U) и типовете биологични секвенции (DNA, RNA, Protein).

Съдържанието е в папката BiologicalSequence.

Вторият компонент е SOAP уеб услугата, чиято цел е да осигурява достъпа до импровизирана база данни (файлова директория) като чете текстовото съдържание на файл и го подава на клиента.

Съдържа единствен изпълним файл FileReaderWS.java който публикува локален GlassFish сървър. Намира се в папката FileReaderSOAP.

В папката WebClientFx се намира третият основен компонент на проекта, която представлява JavaFx приловжение с графичен интерфейс, което консумира SOAP услугата на сървъра и изпълнява визуално нагледно операциите върху биологичните секвенции. Съдържа стандартен контролер (Controller.java) където е написана логиката на графичния интерфейс и управлението на събития, sample.fxml файл и изпълним файл Main.java който стартира приложението.

2. Кратко описание на имплементацията.

Класът SequenceAnalyzer е силно зависим от член данната си от тип Sequence. Той може да работи ефективно само когато му е зададена валидна секвенция от тип Sequence. Задачата му е да анализира нейните особености. Използвани са регулярни изрази за разпознаване на под-редици и Stream API за обработка на текст. Изключения (Exceptions) възникват само при невалидно ID на sequence тъй като то трябва да бъде валидно име на файл. Задълженията за проверката за съществуването и валидността на самия файл са прехвърлени на ползвателя. Изброимият тип SequenceType съдържа две допълнителни константи – INVALID – когато секвенцията не спада към нито един от трите типа DNA, RNA и Protein, и QUESTIONABLE – когато типа на секвенцията все още не е определена от analyzer обекта.

SOAP web услугата разполага с единствен уеб метод, който прочита и връща текстовото съдържание на файла като String обект.

Пътят към директорията и разширението на самия файл са изведени като константи. Изключения се хвърлят при невалидни стойности, когато файлът не може да бъде разпознат или правилно прочетен или когато не е с правилното разширение (.txt). Клиентът е длъжен да обработи изключенията по свои съображения взимайки предвид че при възникване на такова методът може да не върне резултат.

Клиентското приложение достъпва и се свързва към уеб услугата чрез WSDL файла и адреса на услугата през UDDI. Достъпва Sequence, SequenceAnalyzer и изброимите типове чрез jar файл отворен като библиотека. Използва основни компоненти в графичния интерфейс – бутони, текстови полета и други. Обработва изключения и използва странични процесорни нишки за зареждането на текст от файлове за да не „задръсти“ основната нишка на JavaFX графичния интерфейс.

Тъй като уеб метода връща String обект чиито метод length връща стойност от тип int с максимална стойност 2,147,483,647 не се препоръчва зареждането на файлове достигащи размери от 2GB и повече.

Кодът е коментиран с JavaDocs и има допълнителни коментари във методите, което може да послужи за генерирането на пълна документация.

Приложението е тествано с различни примерни текстови файлове, както и с unit тестове с JUnit4 на по-ранни етапи.

3. Инструкции за компилация и изпълнение.

Във файлът на уеб услугата се задава път до директорията която съдържа текстови файлове с генетични секвенции. Можете да въведете път до папката TestSamples, която идва с проекта. Разширението е .txt, при други разширения е възможно да се получи грешка в четенето на файла. Стартира се уеб услугата от FileReaderWS, след което тя вече е достъпна за клиентското приложение. Стартира се клиентското приложение от Main.

4. Използвани технологии

Java 8 – SDK version 1.8(144)

GlassFish 5.0.0

JavaFX 2.0

IDE – IntelliJ Idea Ultimate 2020.1.4

OS – Windows 10