**Software Requirements Specification**

for

Real Neat Application

Version 0.3

Prepared by Anahit Hovhannisyan, Joshua Strandert, Nick Passantino

Instructor: Alfred Nehme

Course: Capstone

Date: 10/9/19

Revisions

|  |  |  |  |
| --- | --- | --- | --- |
| Version | Primary Author(s) | Updates Made | Date Completed |
| 0.1 | Anahit Hovhannisyan, Joshua Standerfer,  Nick Passantino | Created first version of document. Most sections for entire document completed | 10/22/19 |
| 0.2 | Anahit Hovhannisyan, Joshua Standerfer,  Nick Passantino | Small changes to requirements made |  |
| 0.3 | Nick Passantino | Several sections reworded for clarity. Some product functionalities removed (since it was decided those functionalities should be dropped). UI picture sample updated. | 4/28/20 |

**Table of Contents**

**Table of Contents ii**

**Revision History ii**

**1.** **Introduction 1**

1.1 Purpose 1

1.2 Product Scope 1

1.3 References 1

1.4 Definitions 1

**2.** **Overall Description 1**

2.1 Product Perspective 1

2.2 Product Functions 1

2.3 User Classes and Characteristics 2

2.4 Design and Implementation Constraints 2

2.5 Assumptions and Dependencies 2

**3.** **Specific Requirements 2**

3.1 User Interfaces 2

3.2 Functional Requirements 2

3.4 Performance Requirements 3

3.5 Software Quality Attributes 9

**Appendix B: Analysis Models 10**

**Appendix C: Requirement Traceability Matrix 10**

# 1. Introduction

## 1.1 Purpose

This document is intended to discuss and describe the functions of the [RNA] software. It will outline the goals the software will attempt to achieve, as well as the scenarios the software might be useful in.

## 1.2 Product Scope

This software attempts to provide visualization tools for the 2D conformation of a given RNA sequence. By providing a stable, desktop-executable program, this software aims to offer a quick, intuitive platform for visualization of RNA structures across operating systems.

**1.3 References**

*Sequence File Formats: FASTA and SEQ*. rna.urmc.rochester.edu/Text/File\_Formats.html.

## 1.4 Definitions

* Ribonucleic Acid: Abbreviated as RNA, is a molecule ubiquitous in biological functions. RNA is constructed from chains of nucleotides. An RNA sequence may be similar to the following: AUCGCGUAGCAUGC.
* Nucleotides: Often referred to as ‘bases’ in the context of RNA, nucleotides are specialized molecules which form the structure of RNA. This document primarily deals with 4 nucleotides: A, G, U, C. To understand how RNA forms structures, it is vital to know that nucleotides form pairs, also known as base pairs.
* Base Pairs: A set of two complementary nucleotides which form hydrogen bonds. A and U form a pair, C and G form another. The pairing AG, or UC cannot exist. This restriction is one of the driving forces of RNA structure formation, and is responsible for the complex forms that can be created. The bond of the CG is stronger than the bond between AU. Note: Bases adjacent to each other in the chain are not base pairs, any base can be adjacent to any other.
* Hydrogen Bonds: Strong electrostatic attraction between hydrogen, and some other species. The formation of a hydrogen bond is not considered a chemical reaction, there is no change in chemical formula. To give an example, hydrogen bonds are responsible for water tension, yet we would not consider the entire mass of a lake to be a single molecule.
* Free Energy (Gibbs Energy): A common method for predicting RNA shape includes calculating the minimum free energy present in the structure. It is sufficient for this document to liken free energy to gravitational potential energy; if a branch is heavy enough to fall off to break off of a tree, it will. In the same way, if the electrostatic forces between some set of unpaired bases is greater than the forces holding some other set of base pairs together, the molecule will rearrange itself until the energy of the unpaired bases is the lowest it can be.

# 2. Overall Description

## 2.1 Product Perspective

This product is a desktop application that will help researchers looking at unknown RNA sequences figure out what the sequences look like *in vitro.* This program will, instead of creating a new algorithm for modelling to add on to the existing slew of models, create an easy way to compare all the models side by side, giving the researcher an easy way to decide which model is likely to be the correct one. The algorithms the program uses to calculate the RNA structure will be supplied in the form of plug-ins. Several will be made by our team during development but any user will be able to create an algorithm to supply to the program.

## 2.2 Product Functions

* The user will be able to input a string to represent their RNA sequence, or upload a FASTA sequence.
* The user will be able to choose their preferred algorithm(s) for RNA modelling - a brief description of each algorithm will be provided.
* The program will render 2D models based off the feedback of the chosen algorithms, and display the different models in different window tabs.
* The user will be able to rerun their modified sequence using their chosen algorithms.
* The user will be able to craft their own custom plug-in from an RNA sequencing algorithm they have access to, by following the plug-in architecture specifications of the application.

## 2.3 User Classes and Characteristics

* RNA Biologists: These are people who are competent in the laboratory and are highly knowledgeable about RNA and molecular biology. They have used computers and various applications before, and have likely tried out free internet modelling tools. They are not computer programmers. These people want the product to work quickly and accurately, without any hassle of trying to figure out add-ons or plug-ins or extra features.
* Molecular Biology Students: These are people who have some molecular biology knowledge, and a basic understanding of what they expect the program to do. If they have downloaded this application instead of using a free internet source, they are likely doing some research of their own about RNA. These people are not programmers, and are not interested in any complications of the program. They want the program to work quickly, accurately and be easy to understand.

## 2.4 Design and Implementation Constraints

* This program will be developed using the Qt framework, version 5.14.1, all modifications to the application will require using this version of the framework or a later version.
* The options available for RNA modelling are dependent on already created open source modeling programs, and our ability to interface with them.
* RNA sequences can be very, very long. A single file may be multiple gigabytes in size. This program may restrict the size of a file in order to limit memory overuse.
* This program will be available only in English
* This program does not need a high degree of security as it will hold no sensitive or personal information.

## 2.5 Assumptions and Dependencies

It is assumed that the user will be inserting real RNA sequences into the program. If random sequences are input, the structures will be equally non-sensible and may not produce any shape at all.

This product is dependent on the source code from other RNA modelling programs, and accuracy will improve with the updating of the underlying algorithms or the addition of new programs.

# 

# 

# 

# 

# 

# 

# 

# 

# 

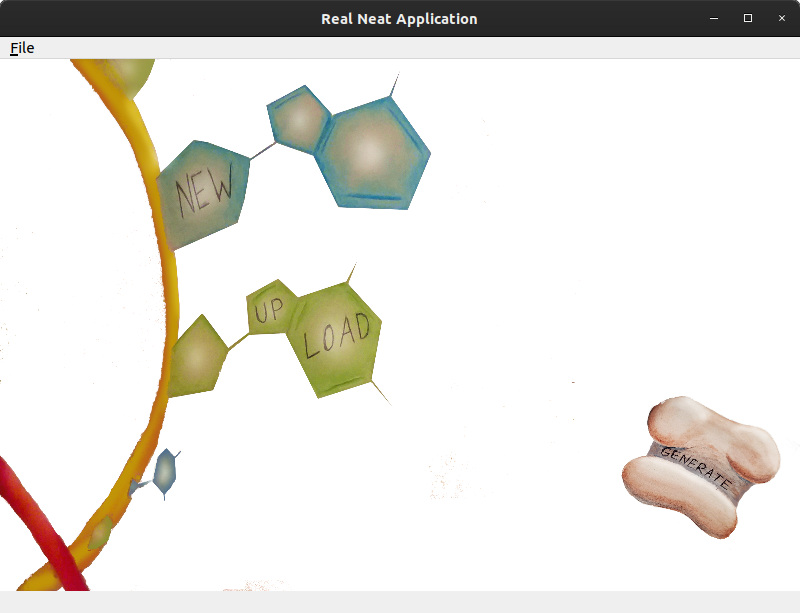
# 

# 

# 3. Specific Requirements

**3.1 User Interfaces**

The application will be a desktop application. The greeting screen will have a button for manually entering new sequences, and a file menu for uploading sequences from a file. In the file menu, there will be additional options, such as an option to access the algorithm menu and to specify the location to output any generated FASTA files to. In addition, there will be a button to generate the sequence once all the necessary information has been specified. Generated structures will appear in a separate tabbed window, with each tab holding one of the generated structures.



**3.2 Functional Requirements**

**3.2.1 Generating RNA Model**

Priority ratings specified out of 10.

1. Description

The program will be able to form a two-dimensional RNA model from a provided RNA sequence and have the model display on a viewing screen within the application.

**Priority Rating:** 10

1. Requirements

* The application will require a file which specifies an RNA strand (and is in a supported file type such as FASTA) or a sequence specified otherwise by the user (via the drag-and-drop function)

**3.2.2 Upload RNA File to Generate Model**

1. Description

The application will allow the user to upload an RNA sequence file (such as a .fa file) which it will use to generate the model.

**Priority Rating:** 9

1. Requirements

* If the file is corrupted, incomplete, or of the wrong type the program will fail to produce a model and will generate an error. No model will be displayed

**2.3.4 2D Model Algorithm Choice**

1. Description

There are several ways of generating probable 2D models of what an RNA strand would form into given a specific sequence. The program will provide the user with several choices of algorithms to choose from.

**Priority Rating:** 8

1. Requirements

* The user will have a choice of these models via a prompt which the user will be confronted with each time they try to generate a model.

**3.2.5 Estimating Model Generation Time**

1. Description

Given a supported RNA sequence file, the program will estimate the amount of time that the specified algorithm will take in order to complete the model.

**Priority Rating:** 2

1. Functional Requirements

* The application requires a selected algorithm and RNA sequence file in order to calculate the estimated time to completion
* If the time to complete the model generation is over 10 minutes, the user will be asked to confirm the operation and will be warned that it will take an extended amount of time

**3.2.6 Adding and Removing Plug Ins**

1. Description

Allow the user to select which plug ins to include or remove from their list.

**Priority Rating:** 2

1. Functional Requirements

* The application allows the user to add a plug in to their folder
* The application allows the user to remove a plug in from their format

**3.5 Software Quality Attributes**

Adaptability:

* Software should be able to accept FASTA files without any modifications to the file, and should be able to determine whether or not the file is suitable for RNA structural predictions.
* Software should be able to receive Connectivity Table files from packages.
* The application should be able to use any RNA sequencing algorithm plug-in that uses the specified plug-in interface

Maintainability:

* The algorithmic portion of the software should be independent of the UI/Desktop executable framework to prevent system updates from interfering with the operations.

Usability:

* The layout of the main structure display should be designed based off popular UI conventions to encourage intuitive usage.

**Appendix B: Analysis Models**

**Appendix C: Requirement Traceability Matrix**

Requirements Traceability Matrix (RTM)

|  |  |  |  |
| --- | --- | --- | --- |
| Req. # | Requirement | Remarks | Test Case Number |
| 3.2.1 | Generating RNA Model |  |  |
| 3.2.2 | Upload RNA File to Generate Model |  |  |
| 3.2.3 | 2D Model Algorithm Choice | Allow user to choose algorithm from supplied algorithm plug-in |  |
| 3.2.4 | Estimating Model job Time |  |  |