Requirement

Specification

## Author: anp28

## Date: 15/02/2015

## Version: 0.1

## Status: Draft

# Overview of Proposed System

The aim of the project I have selected is to compare two DNA sequences and display the results. The results must be in a format that the general public will be able to understand and help them easily identify how different or similar one strand is compared to the other. The types of media I will be using to convey this information will be in the form of images, audio and animations. The target audience for this project is the general public in a learning environment such as a school, college or university as such the device for this program will be a portable school laptop. The files sizes of information that I will be working with can range from a few hundred characters to thousands and meaning that the file could be potentially megabytes in size. This means that the program will have to be robust enough that it can handle a lot of processing.

The program will have to read in two DNA sequences from text files. The DNA sequences in these text files will be in FastA format. The program will read the header of the file and extract a title that will help the users understand what the DNA sequence is. This will be displayed to the user above an image.

The program will then use an alignment algorithm to align the two sequences. This is so that both sequences are starting at the same point when being analysed later by the program. To align the sequences I have three methods to choose from. I have the code for the Needleman, Smith-Waterman and my own code based on a mix of Levenshtein distance algorithm and the Smith-Water algorithm. I shall do tests to see which is more efficient as the memory available on the laptops will be of lower quality than the one it is built on.

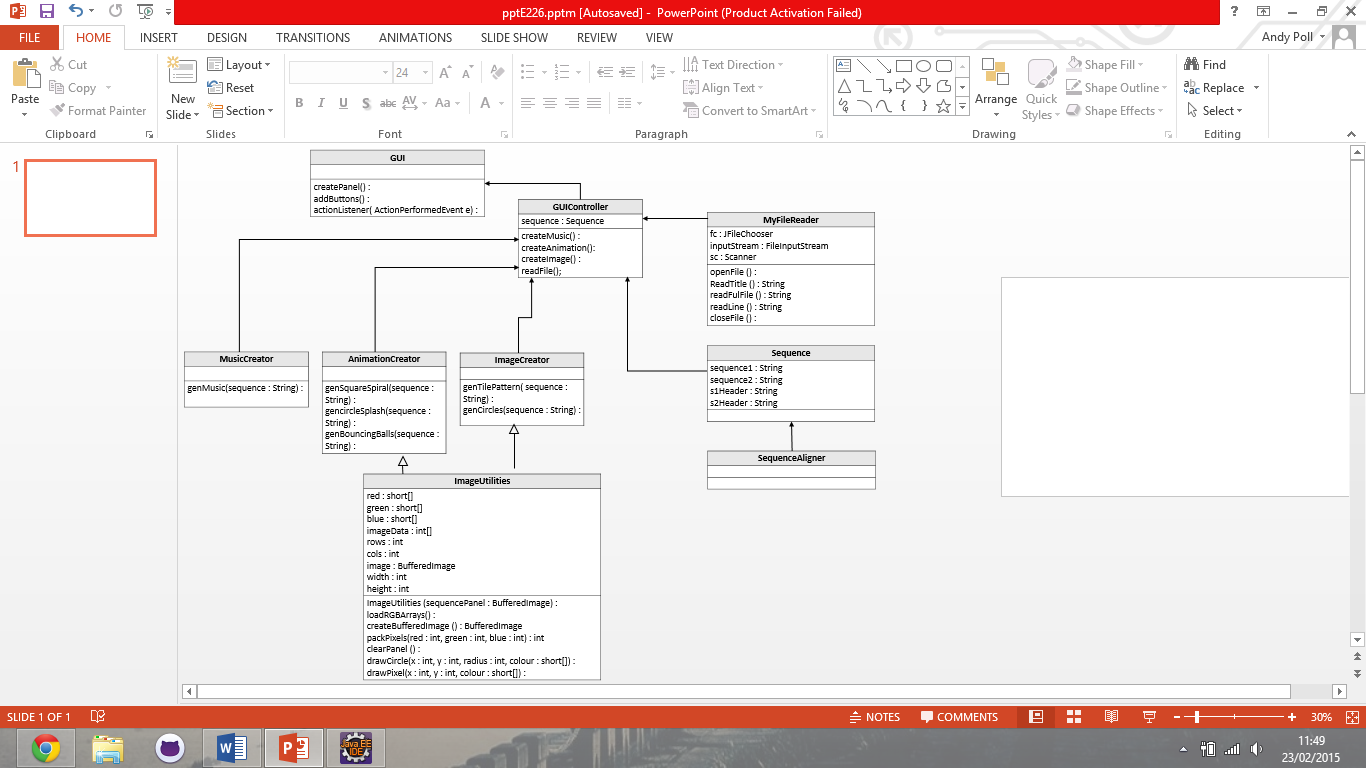
After the sequences are aligned the user will be able to select the method in which the information is displayed. The user can select one of three methods, display as an image, display animations or to play music. To do this there will always be buttons displayed for the user to select the sequence from a file and for the user to select which type of display they would like. To generate an image the program will draw circles and squares with colour, size and location dependent on the information received to graphs. The results will then be displayed in adjacent panels for the user to review. My design takes a similar approach to the animations by creating a splash effect and using the sequence to alter the effect, such as changing speed or colour. Another idea could be balls bouncing around the panel and changing the speed, colour, amount and direction. I can use this method and the splash effect to generate the music, using the colour and speed for the frequency and pitch.

To summarise the program must.

* Accept two DNA sequences from files in the FASTA format.
* Align the two sequences so that the starting point of each is the same
* Generate images, animations and music from the two sequences.
* Display or play the resulting output data from the generation process.
* Be able to run at an acceptable speed on the intended device.
* Able to read in different files.

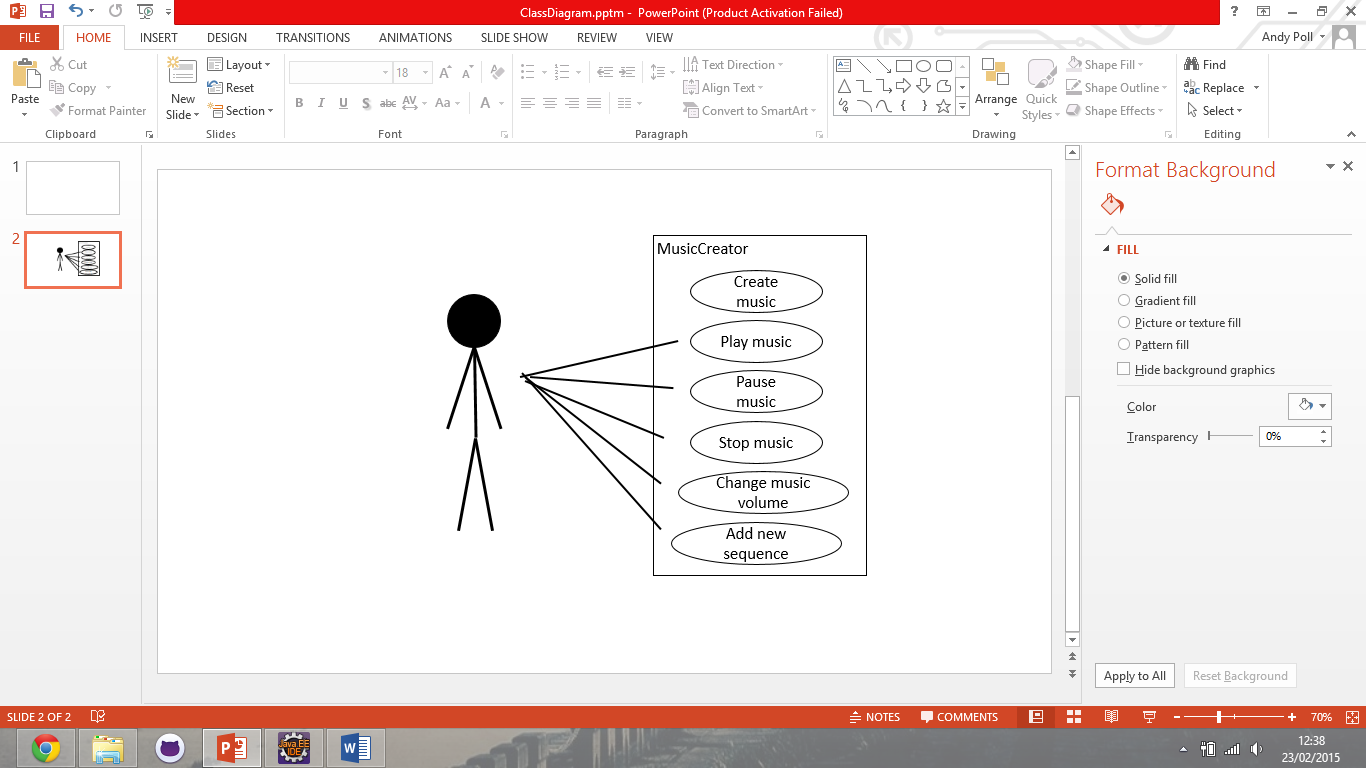
## Language

# Class Diagram

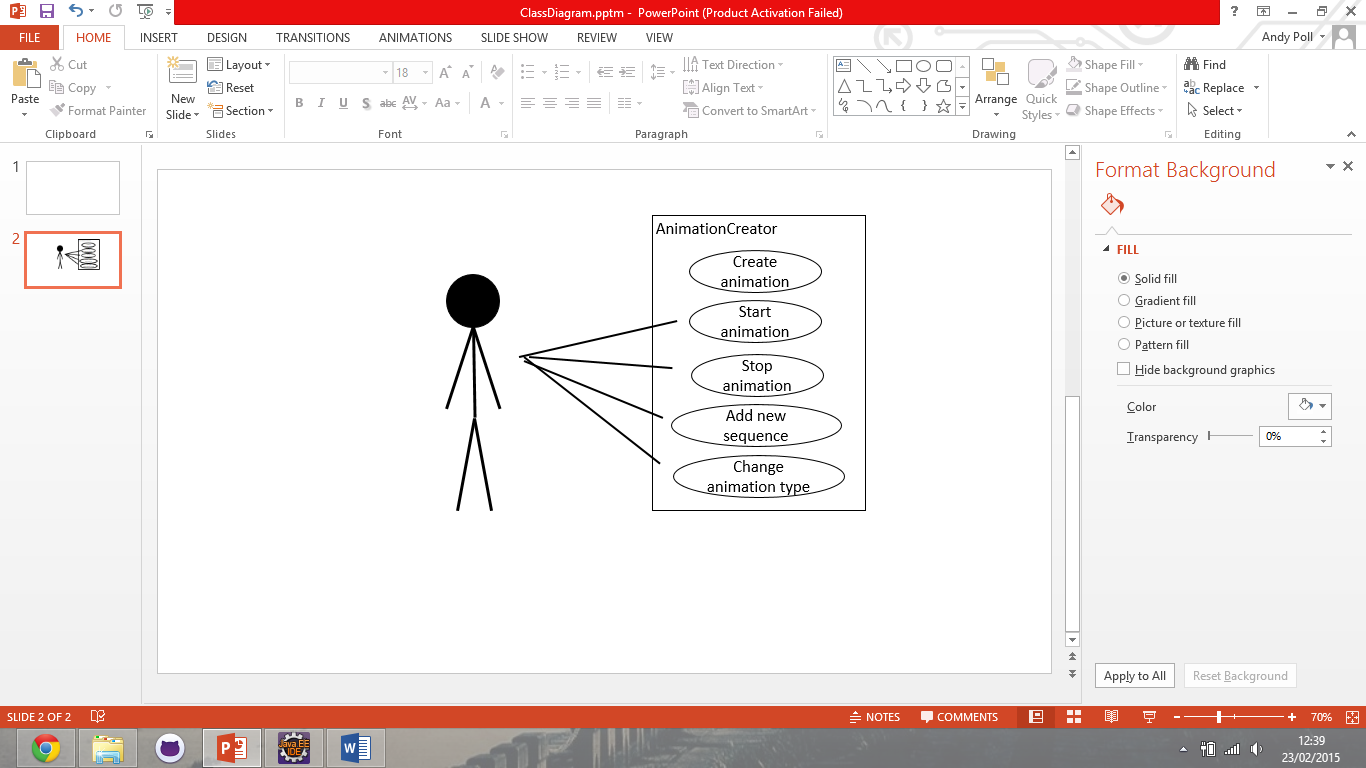


## Description

# Use Case Diagram



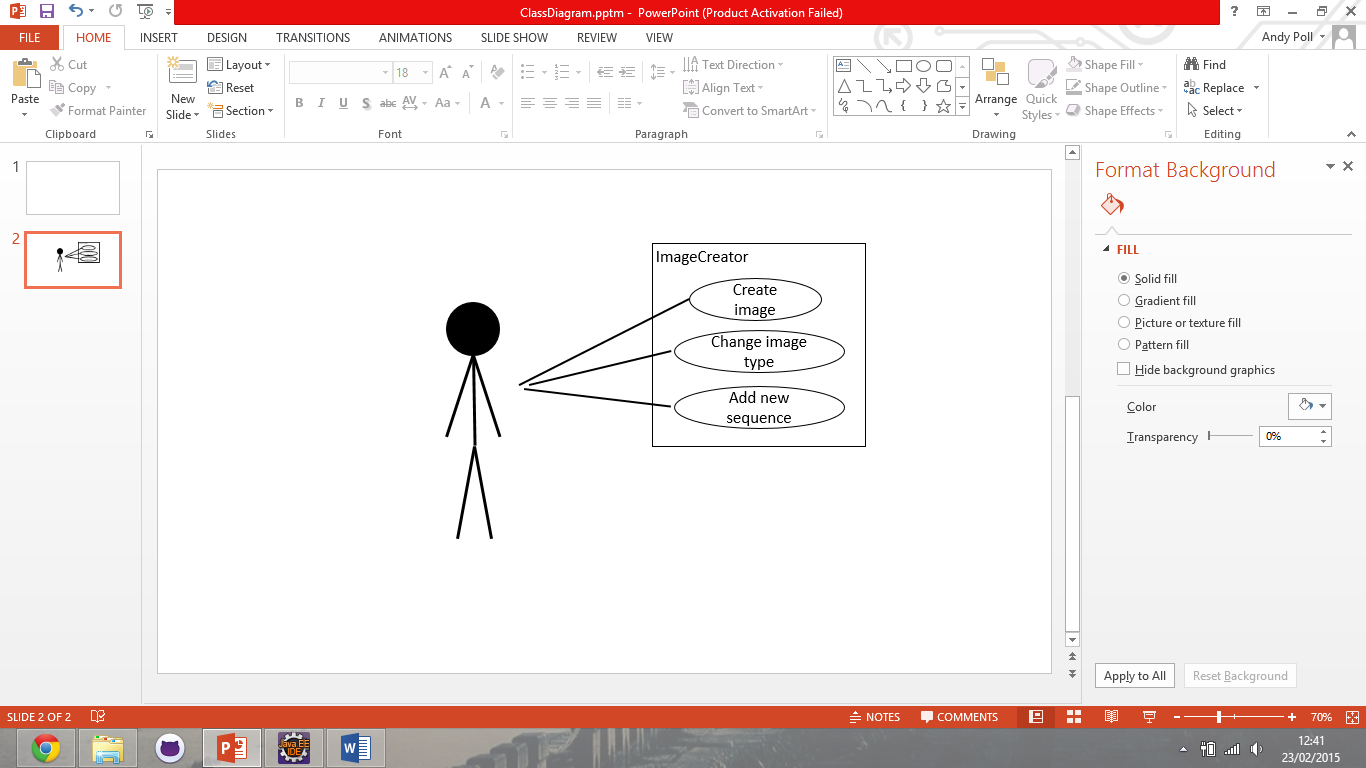
## MusicCreator description



## AnimationCreator description

## 

## Sequences description



## ImageCreator description

# User Interface Design

Sequence 1 title

Sequence 2 title

Open file

Open file

Display

Type

# Test Plan

# Sequence Diagram