PANTZERFAUST

PRODUCT PLANNING

29/04/2016

Introduction

Currently, there does not exist a good visualization tool to look at mutations of DNA strings. Our goal is to create such a visualization tool in order to assist the research to Tuberculosis. Last year an attempt was also made, but there were some flaws which kept the researchers from using those tools. We are doing a new attempt and hopefully will bring everything a step further. In order to keep delivering we have created a planning, which we will discuss in this document.

High-level product backlog

Needs allignment with the product vision.

Roadmap

Releases are scheduled for each week (see table 1), with major releases in week 4 (milestone 1, first scaling version), week 6 (milestone 5, first well-scaling version) and week 10 (milestone 9, final version).

Milestone 1: Simple visualising, linking of components

- Visuals: Display phylogenetic graph with some zooming on 10 genomes.
- Algorithms: Have implemented some basic agorithms.
- Data: Basic parsing of data into the storage, link is working.

Milestone 2: Have more than one view on the data
Milestone 3: Be able to load 340 genomes.
Milestone 4: Finalized the Data server/api, multiple algorithms on the data should be working.
Milestone 5: Be able to load 1500 genomes.
Milestone 6: Be able to load 3000 genomes.
Milestone 7: Be able to load 6000 genomes.
Milestone 8: Finalizing UI
Milestone 9: Final release

Functional requirements

Must haves:

- As a user, I must be able to load any graph in correct gfa format in order to display that graph.
- As a user, I must be able to load any phylogenetic tree in correct nwk format in order to display the tree.
- As a user, I want a graph with 10 genomes to be loaded and displayed in under 5 seconds.
- As a user, I must be able to use a phylogenetic tree as navigation.
- As a user, I must be able to see the whole graph in the upper-most view.
- As a user, I must be able to zoom-in on the graph, such that a single mutation can be distinguished.
- As a user, I want the preprocessing of a graph containing up to 6000 genomes to be done in under 150 minutes.

Should haves:

- As a user, I should be able to see relevant data about mutations.
- As a user, I should be able to see useful<to be defined> statistics about the data.
- As a user, I should be able to use different views on the data, which all contribute to understanding the data better.

Could haves:

- As a user, I could be able to extend the program with extensions, which talk with a program with an API.
- As a user, I could be able to swap elements such as the parser in the program with another parser that implements the interface given by the program.

Won’t haves:

- The user will not be able to run the program in a distributed fashion by using for example hadoop.

Product backlog

User stories of features

It is hard to define these user stories, as we do not yet exactly know what we want. The features we came up with can be found below, but will most certainly be updated throughout the process.

User stories of defects

Not yet applicable

User stories of technical improvements

Say something about the parser perhaps?

User stories of know-how acquisition

What does this even mean?

Initial release plan:

Milestone 1: Simple visualising, linking of components

- Visuals: Display phylogenetic graph with some zooming on 10 genomes.
- Algorithms: Have implemented some basic agorithms.
- Data: Basic parsing of data into the storage, link is working.

Milestone 2: Have more than one view on the data

Milestone 3: Be able to load 340 genomes.

Milestone 4: Finalized the Data server/api, multiple algorithms on the data should be working.

Milestone 5: Be able to load 1500 genomes.

Milestone 6: Be able to load 3000 genomes.

Milestone 7: Be able to load 6000 genomes.

Milestone 8: Finalizing UI

Milestone 9: Final release

Definition of done

First, we would like to say that the product we are creating will never be done. Dna sequences itself are just way to complex to have any program using them being done. For example, loading or displaying it can always be improved. Moreover, 10 weeks is really tight for creating such a program. The goal of the program we are creating is to explore the possibilities for visualizing such sequences. Thus, the program is done if we find a way to visualize DNA sequences in a way that our visualization will help researchers to do their research. This can be achieved either by providing a full working program but also a proof of concept which can later be extended to fully satisfy all needs might will be good.

Glossary

Not yet applicable.