CS 32 Project 4 Report

Siddarth Chalasani

# Known Bugs/Problems

As far as I am aware, my code does not have any bugs/problems.

# Big-O Requirements

As far as I am aware, my code satisfies all big-O requirements in the spec.

# Method Working

## Trie’s find() method

My trie’s find method checks if any of the root node’s children’s label matches the first character in the key. If it does, I call a helper function called findHelper on the applicable children that finds any matches recursively. Then, it returns a vector of matches.

## findHelper() pseudocode

My findHelper takes in the key, exactMatchOnly, a pointer to a Node and a reference to a vector of type ValueType as parameters. findHelper has two base cases:

1. The Node pointer is nullptr: it immediately returns
2. The key’s size is zero: this means that I have found the desired node, so it pushes back all the values in this node onto the vector of type ValueType, and then returns

If any of the node’s children’s label matches the first character in the key

Call findHelper on the key string starting at position 1, exactMatchOnly, a pointer to the child node, and the vector

return

Otherwise, if exactMatchOnly is false

Call findHelper on all of the children nodes with the key string starting at position 1, exactMatchOnly = true, and the vector

## GenomeMatcher’s findGenomeWithThisData() method psuedocode

Clears the vector of matches.

If any of the parameters are invalid

Return false

Create a vector of trieFragments (each of which contains the index of the genome in the vector and the position of the sequence in the genome) that match with the first minSearchLength characters of the fragment.

Sort this vector by the genome index, and by position in the case of a tie.

For all the trieFragments in the vector

For all the trieFragments from the same genome

Calculate the total match length

Find the trieFragment with the largest match length

If this match length is greater that minimum length

Push back a DNAMatch of the match length and this trieFragment’s genome name and position into the vector of matches

Return whether any matches were found