Processing palindromes with EERTREE data structure.

Timur Khazhiev
t.khazhiev@innopolis.ru
Innopolis University

Nikolai Kudasov as a Supervisor
n.kudasov@innopolis.ru
Innopolis University

ABSTRACT

This document represents a Data Science project proposal. It contains description of the project, the problem it required to solve, motivation and brief plan of execution of the project.

1 INTRODUCTION

For this course EERTREE was chosen. EERTREE or palindromic tree is an efficient data structure for processing palindromes in strings. During this course I will be working on analysis and research on this data structure especially from functional programming paradigm point of view.

2 WORK PLAN

Iteration I. 11.02 - 25.02
Analyze palindromic tree and similar solutions, algorithms and data structures. Find problems it solves. Define set of interesting operations over palindromic tree.

Iteration II. 25.02 - 11.03
Propose functional (persistent) version of palindromic tree. Analyze theoretical complexity (worst case and amortized) and compare it with original solution.

Iteration III. 11.03 - 25.03
Implement functional and original version. Compare them: show results for different use cases (benchmarks, real time/memory usage).

Iteration IV. 25.03 - 08.04
Working with one or more of the options:

- Stream fusion implementation
- Cache oblivious model optimization
- Linked and Array (vector) based implementation comparison

Iteration V. 08.04 - 22.04
Implement convenient interface, generalization.

Iteration VI. 22.04 - 06.05
Finalization of project: writing tests, documentation, formal proving, refactoring. Cabal contribution. Sum up results.