Abstract

Bla bla bla.

CHAPTER 1

Background

1.1 Introduction

1.1.1 Biological motivation

1.1.2 Problem definition

1.2 Overview of existing methods

1.2.1 Online methods

1.2.2 Indexed methods

1.2.3 Filtering methods

1.3 Related problems

1.3.1 Local similarity search

1.3.2 Dictionary search

1.3.3 Overlaps computation

CHAPTER 2

Online Methods

2.1 Myers' bit-vector algorithm

2.2 Banded Myers' bit-vector algorithm

2.3 Increased bit-parallelism using SIMD instructions

CHAPTER

3 Indexed Methods

3.1 Classic Full-Text Indices

3.1.1 Trie

3.1.2 Suffix trie and suffix tree

3.1.3 Suffix array

3.1.4 q-Gram index

3.2 Compressed Full-Text Indices

3.2.1 Burrows-Wheeler transform

3.2.2 FM-index

3.2.3 Rank dictionaries

3.3 Backtracking

3.3.1 Pruning methods

3.3.2 Multiple backtracking

CHAPTER

4 Filtering Methods

4.1 q-Gram filters

4.1.1 Exact seeds

4.1.2 Gapped seeds

4.2 Factor filters

4.2.1 Exact seeds

4.2.2 Approximate seeds

4.3 Suffix filters

CHAPTER 5

Read Mapping

5.1 Related work

5.1.1 Best mappers

5.1.2 All mappers

5.2 The Masai mapper

5.2.1 Single-end mapping

5.2.2 Paired-end mapping

5.3 The Masai 2 mapper

5.3.1 Single-end mapping

5.3.2 Paired-end mapping

5.3.3 Parallelization

5.3.4 Hardware acceleration

5.4 Experimental results

5.4.1 Comparison of filtration strategies

5.4.2 Rabema benchmark results

5.4.3 Variant detection results

5.4.4 Runtime results

5.5 Discussion

CHAPTER 6

String Similarity Search / Join

6.1 The competition

6.2 Our method

6.2.1 Implementation

6.2.2 Parallelization

6.3 Related methods

6.4 Experimental results

6.4.1 Search

6.4.2 Join

6.5 Discussion