



UNIVERSITÀ DEGLI STUDI DI MILANO - BICOCCA

Scuola di Scienze

Dipartimento di Informatica, Sistemistica e Comunicazione

Corso di Laurea Magistrale in Informatica

# Algoritmi per la trasformata di Burrows-Wheeler Posizionale con compressione run-length, RLPBWT

**Relatore:** *Prof.ssa Raffaella Rizzi*

**Correlatore:**

**Tesi di Laurea Magistrale di:**

*Davide Cozzi*

*Matricola 829827*

**Anno Accademico 2021-2022**

# Abstract

# Indice

<b>1</b>	<b>Introduzione</b>	<b>3</b>
<b>2</b>	<b>Preliminari</b>	<b>4</b>
2.1	Motivazioni Biologiche . . . . .	4
2.2	Trasformata di Burrows-Wheeler . . . . .	4
2.2.1	Trasformata di Burrows-Wheeler run-length . . . . .	4
2.3	Trasformata di Burrows-Wheeler posizionale . . . . .	4
2.3.1	Implementazione originale . . . . .	4
2.3.2	Varianti . . . . .	4
<b>3</b>	<b>Metodo</b>	<b>5</b>
3.1	Introduzione agli strumenti usati . . . . .	5
3.2	Costruzione della RLPBWT . . . . .	5
3.2.1	Implementazione con divergence array . . . . .	5
3.2.2	Implementazione senza divergence array . . . . .	5
<b>4</b>	<b>Risultati</b>	<b>6</b>
4.1	Ambiente di benchmark . . . . .	6
4.2	Analisi temporale . . . . .	6
4.2.1	Confronto con implementazione originale e varianti . . . . .	6
4.3	Analisi spaziale . . . . .	6
4.3.1	Confronto con implementazione originale e varianti . . . . .	6
<b>5</b>	<b>Conclusioni</b>	<b>7</b>
5.1	Sviluppi futuri . . . . .	7
	<b>Bibliografia e sitografia</b>	<b>7</b>

# Capitolo 1

## Introduzione

# Capitolo 2

## Preliminari

### 2.1 Motivazioni Biologiche

### 2.2 Trasformata di Burrows-Wheeler

#### 2.2.1 Trasformata di Burrows-Wheeler run-length

MONI

### 2.3 Trasformata di Burrows-Wheeler posizionale

#### 2.3.1 Implementazione originale

Gli algoritmi di Durbin

#### 2.3.2 Varianti

PBWT multiallelica

PBWT con struttura LEAP

PBWT dinamica

PBWT bidirezionale

# Capitolo 3

## Metodo

### 3.1 Introduzione agli strumenti usati

### 3.2 Costruzione della RLPBWT

#### 3.2.1 Implementazione con divergence array

Algoritmo per match massimali

#### 3.2.2 Implementazione senza divergence array

Algoritmo per match massimali

# Capitolo 4

## Risultati

### 4.1 Ambiente di benchmark

### 4.2 Analisi temporale

#### 4.2.1 Confronto con implementazione originale e varianti

### 4.3 Analisi spaziale

#### 4.3.1 Confronto con implementazione originale e varianti

# Capitolo 5

## Conclusioni

### 5.1 Sviluppi futuri