



UNIVERSITÀ DEGLI STUDI DI SALERNO

Dipartimento di Informatica
Corso di Laurea Triennale in Informatica

TESI DI LAUREA

**Il ruolo dell'informazione negativa nei
confronti alignment-free fra stringhe:
integrazione di un algoritmo di Minimal
Absent Words nel tool SMART**

RELATRICE

Prof.ssa **Rosalba Zizza**

Università degli Studi di Salerno

CANDIDATO

Antonio Gravino

Matricola: 05121 07161

Anno Accademico 2021-2022

Abstract

Le stringhe sono un potente strumento matematico che si confà ad un spettro ampio e versatile di problematiche, fra cui la bioinformatica. I confronti fra stringhe, fra le altre cose, sono attori centrali nell'evidenziare similitudini delle sequenze genomiche in termini dei loro pattern costituenti.

In letteratura, la maggior parte delle soluzioni algoritmiche al problema dei confronti fra stringhe risultano computazionalmente onerose e prevedono tecniche di confronto allineate; d'altro canto, recenti considerazioni accademiche hanno evidenziato le potenzialità dell'informazione negativa, intesa in questa tesi come *absent words*, nel contesto di confronti alignment-free.

La tesi ha l'obiettivo di integrare un algoritmo alignment-free di confronti fra sequenze che fa uso di *absent words* nel tool String Matching Algorithms Research Tool (SMART), al fine di dimostrare la qualità del risultato.

Forniamo dunque gli elementi matematici fondamentali per la comprensione dell'algoritmo in esame; né elaboriamo, quindi, una codifica prima in pseudocodice e poi in linguaggio di programmazione C; e, infine, procediamo ad integrare il suddetto algoritmo nel tool SMART per ottenere una visione completa della sua resa qualitativa rispetto ad altre soluzioni ben note.

Indice

1	Introduzione	3
---	--------------	---

1 Introduzione

[..]

Questo primo capitolo consiste esclusivamente in una sintetica introduzione del lavoro svolto. Nel Capitolo Due, vengono presentate le componentistiche matematiche, con opportuni riferimenti accademici, utilizzate dall'algoritmo di confronto fra stringhe in esame, o dal tool SMART. Nel Capitolo Tre, si fornisce - a linee generali - una descrizione sommaria del problema del pattern matching e del ruolo dell'informazione negativa in tale ambito di studi. Nel Capitolo Quattro, si procede ad un'approfondita presentazione dell'algoritmo in esame, tralasciando i dettagli implementativi che verranno affrontati con cura nel Capitolo Cinque. Nel Capitolo Sei, viene presentato a livello generale il tool SMART, definendone gli obiettivi e descrivendone le caratteristiche. Il Capitolo Sette rappresenta il risultato originale di questo lavoro di ricerca, fornendo dettagli circa l'integrazione dell'algoritmo in esame nel tool SMART, delinando gli aspetti implementativi e teorici della soluzione ottenuta. L'ultimo Capitolo consiste in una serie di considerazioni conclusive su lavoro svolto. Infine, viene presentata la bibliografia impiegata per la stesura di questa tesi di laurea.