

Algorytm Smitha-Watermana - poszukiwanie optymalnych lokalnych dopasowań sekwencji

Abstrakt

Algorytmy dopasowania sekwencji znajduje swoje zastosowanie m. in. w bioinformatyce do poszukiwań dopasowań sekwencji nukleotydów i aminokwasów. Algorytm Smitha-Watermana należy do podgrupy rozwiązań zajmujących się tzw. dopasowaniem lokalnym. W poniższym dokumencie przedstawiono opis prób zrównoleglenia tego algorytmu z wykorzystaniem technologii CUDA.

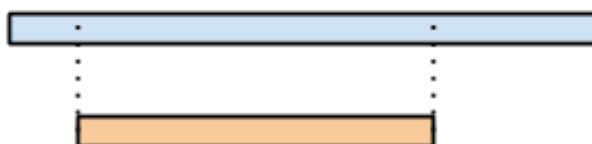
Przedstawienie problemu

Problem dopasowania sekwencji przyjmuje na wejściu dwa ciągi znaków. W ogólnym przypadku, ciągi te mogą składać się z liter dowolnego alfabetu. W przypadku zastosowań bioinformatycznych zazwyczaj ten alfabet jest relatywnie niewielki (np. czteroznakowy "TGAC").

Problem tej klasy można interpretować na dwa sposoby. Istnieją rozwiązania analizujące:

- dopasowanie globalne
- dopasowanie lokalne

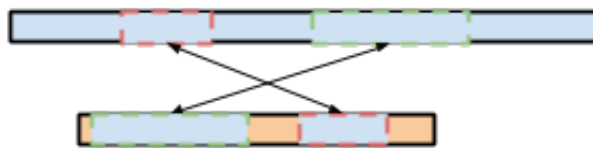
W przypadku dopasowania globalnego dwa ciągi porównywane są wzdłuż całej sekwencji. Takie rozwiązanie jest przydatne przy jednodomenowych białakach. Schemat takiego dopasowania przedstawiono poniżej:



Rysunek 1: Schemat dopasowania globalnego

Lokalny typ dopasowania polega na rozszerzeniu możliwości algorytmów pierwszego typu o zdolność do zauważenia podobieństw w małych obszarach. Dla przykładu pewne sekwencje mogą być zamienione kolejnością. Ten typ rozwiązania znajduje zastosowanie w analizowaniu białek wielodomenowych. Poniżej przedstawiono schemat dopasowania tego typu.

Vestibulum faucibus ex dolor, non suscipit enim sagittis eget. Cras et condimentum elit. Integer porta, quam ac posuere efficitur, magna mauris finibus nisi, at egestas augue diam sed risus. Ut nec nisi quis nunc sagittis volutpat at viverra eros. Donec ut porttitor orci. Mauris eget eleifend neque, id condimentum tortor. Nunc lobortis quam mi, aliquet varius dolor tempus non. Lorem ipsum dolor sit amet, consectetur adipiscing elit.



Rysunek 2: Dopasowanie lokalne a globalne

Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Nam sed auctor ex. Praesent tempus ipsum sit amet eros tempus, id ornare nulla vulputate. Nullam lobortis mi leo, ac tincidunt est volutpat dapibus. Pellentesque faucibus maximus suscipit. Duis vitae turpis in nisi condimentum finibus quis a arcu. Donec magna massa, elementum rutrum ligula at, dignissim blandit sem. Vivamus est nisl, aliquam ac tellus ac, egestas consectetur erat. Morbi commodo dui non ipsum vehicula dapibus.

Algorytm Smitha-Watermana w ujęciu sekwencyjnym

Algorytm bazujący na programowaniu dynamicznym umożliwiając poszukiwanie optymalnych lokalnych dopasowań sekwencji. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Cras sed velit nec nibh varius suscipit. Curabitur nisi purus, porttitor in urna ut, tincidunt aliquam leo. Integer sit amet nisi egestas, congue tellus nec, gravida tellus. Nulla facilisi. Quisque ultrices sem sed arcu mattis, in eleifend lectus imperdiet. Donec ullamcorper cursus tortor, in interdum sem imperdiet non. Aliquam sit amet viverra nisl, vitae ullamcorper dolor. Pellentesque varius ex a urna blandit, ac volutpat sapien feugiat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nulla posuere ligula ligula, ut ullamcorper nulla tempor vitae.

Vestibulum faucibus ex dolor, non suscipit enim sagittis eget. Cras et condimentum elit. Integer porta, quam ac posuere efficitur, magna mauris finibus nisi, at egestas augue diam sed risus. Ut nec nisi quis nunc sagittis volutpat at viverra eros. Donec ut porttitor orci. Mauris eget eleifend neque, id condimentum tortor. Nunc lobortis quam mi, aliquet varius dolor tempus non. Lorem ipsum dolor sit amet, consectetur adipiscing elit.

Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Nam sed auctor ex. Praesent tempus ipsum sit amet eros tempus, id ornare nulla vulputate. Nullam lobortis mi leo, ac tincidunt est volutpat dapibus. Pellentesque faucibus maximus suscipit. Duis vitae turpis in nisi condimentum finibus quis a arcu. Donec magna massa, elementum rutrum ligula at, dignissim blandit sem. Vivamus est nisl, aliquam ac tellus ac, egestas consectetur erat. Morbi commodo dui non ipsum vehicula dapibus.

Zarys technologii CUDA

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Cras sed velit nec nibh varius suscipit. Curabitur nisi purus, porttitor in urna ut, tincidunt aliquam leo. Integer sit amet nisi egestas, congue tellus nec, gravida tellus. Nulla facilisi. Quisque ultrices sem sed arcu mattis, in eleifend lectus imperdiet. Donec ullamcorper cursus tortor, in interdum

sem imperdiet non. Aliquam sit amet viverra nisl, vitae ullamcorper dolor. Pellentesque varius ex a urna blandit, ac volutpat sapien feugiat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nulla posuere ligula ligula, ut ullamcorper nulla tempor vitae.

Vestibulum faucibus ex dolor, non suscipit enim sagittis eget. Cras et condimentum elit. Integer porta, quam ac posuere efficitur, magna mauris finibus nisi, at egestas augue diam sed risus. Ut nec nisi quis nunc sagittis volutpat at viverra eros. Donec ut porttitor orci. Mauris eget eleifend neque, id condimentum tortor. Nunc lobortis quam mi, aliquet varius dolor tempus non. Lorem ipsum dolor sit amet, consectetur adipiscing elit.

Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Nam sed auctor ex. Praesent tempus ipsum sit amet eros tempus, id ornare nulla vulputate. Nullam lobortis mi leo, ac tincidunt est volutpat dapibus. Pellentesque faucibus maximus suscipit. Duis vitae turpis in nisi condimentum finibus quis a arcu. Donec magna massa, elementum rutrum ligula at, dignissim blandit sem. Vivamus est nisl, aliquam ac tellus ac, egestas consectetur erat. Morbi commodo dui non ipsum vehicula dapibus.

Model zwróńoglenia Algorytmu Smitha-Watermana

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Cras sed velit nec nibh varius suscipit. Curabitur nisi purus, porttitor in urna ut, tincidunt aliquam leo. Integer sit amet nisi egestas, congue tellus nec, gravida tellus. Nulla facilisi. Quisque ultrices sem sed arcu mattis, in eleifend lectus imperdiet. Donec ullamcorper cursus tortor, in interdum sem imperdiet non. Aliquam sit amet viverra nisl, vitae ullamcorper dolor. Pellentesque varius ex a urna blandit, ac volutpat sapien feugiat. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Nulla posuere ligula ligula, ut ullamcorper nulla tempor vitae.

Vestibulum faucibus ex dolor, non suscipit enim sagittis eget. Cras et condimentum elit. Integer porta, quam ac posuere efficitur, magna mauris finibus nisi, at egestas augue diam sed risus. Ut nec nisi quis nunc sagittis volutpat at viverra eros. Donec ut porttitor orci. Mauris eget eleifend neque, id condimentum tortor. Nunc lobortis quam mi, aliquet varius dolor tempus non. Lorem ipsum dolor sit amet, consectetur adipiscing elit.

Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Nam sed auctor ex. Praesent tempus ipsum sit amet eros tempus, id ornare nulla vulputate. Nullam lobortis mi leo, ac tincidunt est volutpat dapibus. Pellentesque faucibus maximus suscipit. Duis vitae turpis in nisi condimentum finibus quis a arcu. Donec magna massa, elementum rutrum ligula at, dignissim blandit sem. Vivamus est nisl, aliquam ac tellus ac, egestas consectetur erat. Morbi commodo dui non ipsum vehicula dapibus.

Wnioski

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Cras sed velit nec nibh varius suscipit. Curabitur nisi purus, porttitor in urna ut, tincidunt aliquam leo. Integer sit amet nisi egestas, congue tellus nec, gravida tellus. Nulla facilisi. Quisque ultrices sem sed arcu mattis, in eleifend lectus imperdiet. Donec ullamcorper cursus tortor, in interdum sem imperdiet non. Aliquam sit amet viverra nisl, vitae ullamcorper dolor. Pellentesque varius ex a urna blandit, ac volutpat sapien feugiat. Pellentesque habitant morbi tristique

senectus et netus et malesuada fames ac turpis egestas. Nulla posuere ligula ligula, ut ullamcorper nulla tempor vitae.

Vestibulum faucibus ex dolor, non suscipit enim sagittis eget. Cras et condimentum elit. Integer porta, quam ac posuere efficitur, magna mauris finibus nisi, at egestas augue diam sed risus. Ut nec nisi quis nunc sagittis volutpat at viverra eros. Donec ut porttitor orci. Mauris eget eleifend neque, id condimentum tortor. Nunc lobortis quam mi, aliquet varius dolor tempus non. Lorem ipsum dolor sit amet, consectetur adipiscing elit.

Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. Nam sed auctor ex. Praesent tempus ipsum sit amet eros tempus, id ornare nulla vulputate. Nullam lobortis mi leo, ac tincidunt est volutpat dapibus. Pellentesque faucibus maximus suscipit. Duis vitae turpis in nisi condimentum finibus quis a arcu. Donec magna massa, elementum rutrum ligula at, dignissim blandit sem. Vivamus est nisl, aliquam ac tellus ac, egestas consectetur erat. Morbi commodo dui non ipsum vehicula dapibus.

Źródła

Łukasz Ligowski, Witold Rudnicki - AN EFFICIENT IMPLEMENTATION OF SMITH WATERMAN ALGORITHM ON GPU USING CUDA, FOR MASSIVELY PARALLEL SCANNING OF SEQUENCE DATABASES

E. Banachowicz Bioinformatyka - wykład monograficzny