

配列アラインメント高性能化のための
ナノフォトリック・デバイスを用いたアーキテクチャ

浅井里奈

平成 28 年 12 月

情報知能工学科

配列アラインメント高性能化の為の ナノフォトリック・デバイスを用いたアーキテクチャ

浅井里奈

内容梗概

配列アラインメントナノフォトリックス高性能化

目次

| | | |
|-----|----------------------------|---|
| 1 | はじめに | 1 |
| 2 | 配列アラインメントアーキテクチャとその高性能化 | 3 |
| 2.1 | 配列アラインメントの基本原理 | 3 |
| 2.2 | CMOS によるレースロジック実装 | 3 |
| 2.3 | 従来技術の課題 | 3 |
| 3 | ナノフォトニック・デバイスを用いたレースロジック回路 | 4 |
| 3.1 | 光デバイスについて | 4 |
| 3.2 | 設計選択肢 | 4 |
| 3.3 | 提案回路 | 4 |
| 4 | 評価・考察 | 5 |
| 4.1 | 評価 | 5 |
| 4.2 | 考察 | 5 |
| 5 | おわりに | 6 |
| | 謝辞 | 7 |
| | 参考文献 | 8 |

第 1 章

はじめに

現在，生物学の分野において生物配列 (DNA の塩基配列とタンパク質アミノ酸配列) の文字列処理 (配列情報解析) が注目されている [1,2]．DNA やタンパク質はユニットと名付けられた単位の物質が一行に並んだ高分子である．ここでいうユニットとは，DNA においては 4 種の核酸，タンパク質においては 20 種類のアミノ酸である．それぞれのユニットを文字とし DNA やタンパク質の配列を単なる文字列だとみなして処理をしてもある種の本質は失われないという考えに基づき，文字列処理をすることで生物配列の解析を行なっている．DNA の塩基配列やタンパク質アミノ酸配列の研究は，バイオインフォマティクスの最重要課題の一つとして取り組まれてきた．配列情報解析の重要な対象であるゲノム塩基配列は，すでに 200 種類以上が決定されており，さらに多くの解析が進行中であるといわれている [3]．生物配列の文字列処理の中で，DNA 配列中に同じ順序で並んでいるユニットのパターンを見つける配列アラインメントがある [4]．アラインメントとは，複数の配列を入力として配列要素の間に最適な対応関係を求める処理であり，配列の類似性判定に応用できるものである．配列アラインメントは生物学において重要な手法であり，計算機を用いた処理の高速化は従来より多くの研究がなされてきた [5–9]．

一方，ナノフォトニクスと呼ばれる新しい光素子技術が注目を集めている．このナノフォトニクスを用いて機能を実現したデバイスをナノフォトニック・デバイスという．ナノフォトニック・デバイスは光速度で演算を実現できる素子として注目されており，パターン検出機構や加算器などのアーキテクチャの検討がされている []．

本研究では DNA 配列のグローバルアライメントを高速化する手段として，ナノフォトニクス・デバイスに着目した．本論文の構成は以下の通りである．第 2 章で配列アラインメントの基本原則，及び CMOS を用いたアラインメントアーキテクチャと問題点について説明する．第 3 章では，光デバイスの基礎事項と共にナノフォトニクスの基本事項を説明し，設計選択肢を整理した上で回路の提案を行う．第 4 章では提案した回路に対して評価と考察を行い，第 5 章でまとめを行う．

第 2 章

配列アラインメントアーキテクチャとその 高性能化

本章では、

2.1 配列アラインメントの基本原理

2.2 CMOS によるレースロジック実装

2.3 従来技術の課題

第 3 章

ナノフォトニック・デバイスを用いたレースロジック回路

3.1 光デバイスについて

3.2 設計選択肢

3.3 提案回路

第 4 章

評価・考察

4.1 評価

4.2 考察

第 5 章

おわりに

謝辞

本研究の進行および本論文執筆にあたりまして，懇切丁寧なご指導を頂きました井上弘士教授に心より感謝申し上げます．本論文執筆にあたり，多大なご指導を頂きました小野貴継助教に心より感謝申し上げます．本研究を行うにあたり，多大なご指導を賜りました日本電信電話株式会社 物性科学基礎研究所 ナノフォトンクスセンタ 主幹研究員 新家昭彦様に感謝の意を表するとともに厚く御礼申し上げます．本研究に際し，データの提供を頂きました京都大学大学院 情報学研究科 通信情報システム専攻 石原亨准教授に心より御礼申し上げます．本論文執筆にあたり多大な指導頂きました，井上研究室大学院修士1年磯部聖氏に深く感謝致します．

最後に，井上研究室の皆様の御意見，御厚意に感謝の意を表します．

参考文献

- [1] 浅井潔ほか. 配列情報と確立モデル (「分子生物情報学の新展開」). 人工知能学会誌, Vol. 15, No. 1, pp. 35–42, 2000.
- [2] 後藤修. マルチプルアラインメントは生体高分子情報の交差点. 生物物理, Vol. 38, No. 2, pp. 52–56, 1998.
- [3] 浅井潔ほか. バイオインフォマティクス: 7. ポストゲノム時代の配列情報解析. 情報処理, Vol. 46, No. 3, pp. 271–276, 2005.
- [4] 須戸里織, 吉見真聡, 三木光範. バイオインフォマティクスゲノム配列から機能解析へバイオインフォマティクスゲノム配列から機能解析へ, 2005. 電子情報通信学会技術研究報告. CPSY, コンピュータシステム, Vol. 111, No. 328, pp. 35–40, 2011.
- [5] 須戸里織, 吉見真聡, 三木光範. Gpu を用いた 3 次元 smith-waterman 法の高速化手法の提案. 電子情報通信学会技術研究報告. CPSY, コンピュータシステム, Vol. 111, No. 328, pp. 35–40, 2011.
- [6] 宗川裕馬, 伊野文彦, 萩原兼一ほか. 統合開発環境 cuda を用いた gpu での配列アライメントの高速化手法. 情報処理学会研究報告計算機アーキテクチャ (ARC), Vol. 2008, No. 19 (2008-ARC-177), pp. 13–18, 2008.
- [7] Edans Flavius de O Sandes and Alba Cristina MA de Melo. Smith-waterman alignment of huge sequences with gpu in linear space. In *Parallel & Distributed Processing Symposium (IPDPS), 2011 IEEE International*, pp. 1199–1211. IEEE, 2011.

-
- [8] Yu Liu, Yang Hong, Chun-Yuan Lin, and Che-Lun Hung. Accelerating smith-waterman alignment for protein database search using frequency distance filtration scheme based on cpu-gpu collaborative system. *International journal of genomics*, Vol. 2015, , 2015.
- [9] 伊野文彦, 小谷裕基, 萩原兼一ほか. Gpu グリッドによる高速な塩基配列アライメント. 情報処理学会研究報告ハイパフォーマンスコМПユーティング (HPC), Vol. 2007, No. 80 (2007-HPC-111), pp. 73–78, 2007.