

遺伝子配列の解析における相同性検索の FPGA 実装

藤井 惇司[†] 渡邊 誠也[†] 名古屋 彰[†]

[†] 岡山大学大学院自然科学研究科

1. はじめに

近年、ソフトウェアによる遺伝子配列の解析に要する処理時間の短縮が求められており、ハードウェアを用いた高速化の研究が試みられている。遺伝子配列の相同性検索では、動的計画法を用いたアルゴリズムが知られており、特にローカルアラインメントを求めるための Smith-Waterman アルゴリズムに関しては、FPGA (Field Programmable Gate Array)化の研究例[1]がある。一方で、グローバルアラインメントを求める Needleman-Wunsch (NW) アルゴリズムに関しては、扱う配列の範囲が広がることから、ハードウェア化の研究はあまりなされていない。そこで本研究では、NW アルゴリズムの FPGA への実装を試みることにした。

2. ハードウェアでの NW アルゴリズム実現手法

提案するハードウェアの構成を図1に示す。

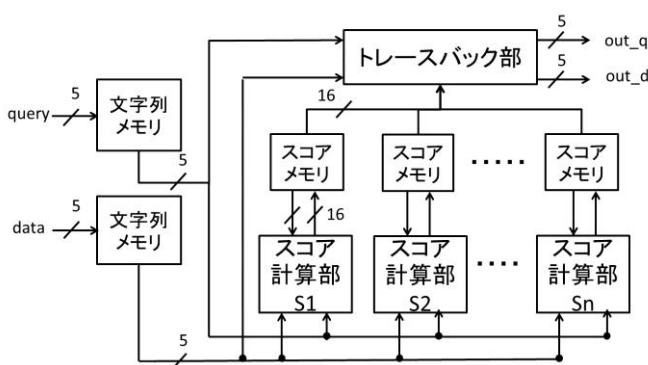


図 1. 提案ハードウェアの構成

未知の配列(query)と既知の配列(data)の遺伝子配列を文字列メモリに保存し、data の文字数分用意したスコア計算部を並列動作させることによって処理時間を削減する。スコアメモリに保存したスコアからトレースバックを行い、アラインメントを求める。メモリには FPGA に搭載されているブロック RAM を用いることでアクセス時間を短縮する。

3. 提案ハードウェアの実装方法

ハードウェア記述言語である SFL[2]を用いて提案回路を実装した。この際、対象とする query と data の文字数に応じた回路の SFL 記述を自動的に生成するツールを作成し、生成した SFL 記述を Verilog HDL に変換することで、FPGA への実装を実現した。図 2 に、実装までの流れを示す。

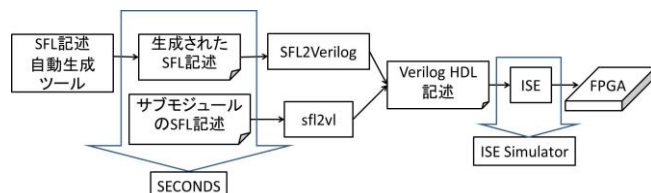


図 2. 実装までの流れ

4. 評価

FPGA 上の回路規模はスコア計算部の数にほぼ比例する。ターゲットとしたデバイス (XC7VX980T-2FFG1926) では、query と data が 1500 文字ずつ程度までならば実装可能であると見積りを得ることができた。

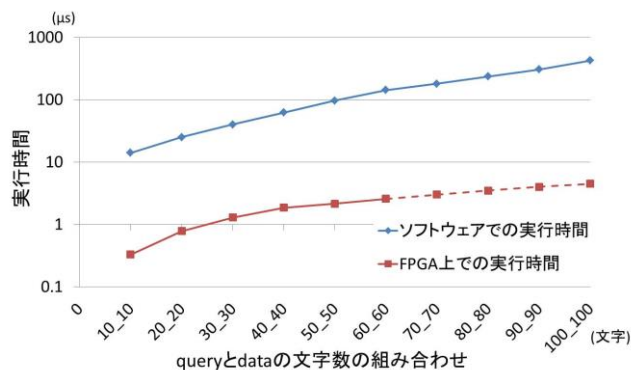


図 3. NW アルゴリズムの実行時間の比較

(破線部は予測値)

図 3 に実行時間を示す。60 文字ずつに対応した処理で、FPGA 上に実装した回路の性能は、ソフトウェア (Core i3-2100 上で実行) の約 50 倍となった。対象文字数が増加するほど、性能的に有利となると予測できる。

5. 今後の課題

より長大な遺伝子配列に対応した回路の実装には、さらに工夫が必要である。

参考文献

[1] O. Straasli, W. Yu, D. Strenskei, and J. Maltby, "Performance Evaluation of FPGA-based Biological Application," Cray Users Group, 2007.
 [2] 小栗 清, 中村 行宏, 野村 亮, 名古屋 彰, "主要なハードウェア記述言語の特徴と標準化状況 SFL," 情報処理, vol. 33, no. 11, pp. 1256-1262, 1992.