

## 論文審査の結果の要旨

氏 名 渋谷 哲朗

本論文は、タンパク質やDNAといった生物における塩基配列情報に対する比較及び検索のための高速なアルゴリズムの研究を行ない、さらに計算機実験を通じてそれらの性能を検証し、生物学上重要な実問題への応用も行なったものである。

本論文は七章からなり、第一章では、本論文で扱うアルゴリズム及び生物学上の諸問題の背景ならびに相互の関連性が概観されている。

第二章は、第三章以降においてよく触れられる問題とそれに対する従来のアルゴリズムについて解説を行なっている。

第三章は、複数の塩基配列を比較する重要な手法であるマルチプル・アラインメントを扱っている。この手法の計算上の最適解が生物学上最も最適な解であるとは限らないといわれている問題に対し、本章では最適解の代替解を提供する手法として、準最適解を代替の解として高速に列挙する手法と、使用するパラメータを変化させた場合の最適解の変化を効率的に列挙する手法を提案し、さらに、計算機実験を行なうことでこれらの手法の有効性も示した。

第四章は、cDNA 配列の集合を選択的スプライス集合と呼ばれる集合に分類する問題を扱っている。この分類はスプライスト・アラインメントという配列比較手法を用いることで正確に行なえるが、非常に計算量が大きい。これに対し、本章では、高速化のためのアルゴリズムを提案し、さらに実際のマウスの大規模cDNA ライブラリを用いた計算でも、実時間内に高精度な分類が可能であることを示した。

第五章は、大規模モチーフ検索に基づく遺伝子発見の全く新しい手法を提案している。この問題の従来の手法は統計に基づく手法と相同性検索による手法に分けられるが、本章の手法は、前者で困難な十分な学習セットのないような場合や、後者で困難な類似遺伝子が存在しない場合においても遺伝子を正確に推定できることを、計算機実験を通じて示した。さらに、モチーフ検索の高速アルゴリズムも提案することで、高速な遺伝子発見も実現した。

第六章では、RNA 構造の解析を行なう全く新しい手法を提案している。まず、RNA 配列中において同じ立体構造を持ち得るような部分文字列を捜し出すことができるように接尾辞木という配列比較のためのデータ構造を一般化し、それに対する高速構築アルゴリズムを提案した。また本章では、RNA 二次構造解析に応用できるものとして、木構造から様々なパターンを発見できるように接尾辞木を一般化したものに対しても、高速アルゴリズムを提案した。

第七章は、六章までに述べられた問題及び手法を総括し、さらに、今後取り組むべき課題についての展望を示している。

なお、本論文の第三章は今井浩氏と、第四章は Christian Schoenbach 氏と小長谷明彦氏と鹿島久嗣氏と、第五章は Isidore Rigoutsos 氏との共同研究であるが、論文提出者が主体となって分析及び検証を行なったもので、論文提出者の寄与が十分であると判断する。

したがって、博士（理学）の学位を授与できると認める。