

## 審査の結果の要旨

氏 名 オット ザーシャ

ヒトゲノム計画の完了（レファレンスとなる全ヒトゲノム塩基配列の決定）を契機に浮かび上がってきた生命科学の次の焦点の一つは、ゲノムにコードされた遺伝子産物間の相互作用ネットワークの解明である。特に遺伝子発現レベルの相互依存関係ネットワーク（遺伝子ネットワーク）を、DNAマイクロアレイ実験などに代表される網羅的な実験によって得られるデータから推定する研究が、現在生命情報科学（バイオインフォマティクス）のホットな課題になっている。この種の研究では、実験データから導かれる評価関数に対して、膨大な組み合わせの依存関係の候補を探索する必要があるが、この探索問題はNP困難である。そこで本論文では、指数関数的な時間で超指数関数的な探索空間中の最適ネットワークを発見することのできるアルゴリズムを、いわゆる動的計画法に基づき開発したことを報告している。これによって、遺伝子数が20個程度の場合なら厳密解を得ることが可能になった。次に、本論文の著者は生物学的に妥当な制約条件（各遺伝子を制御する親遺伝子の数に上限をもたせる）の下では遺伝子数が35個程度でも最適解を探索可能にする方法も提唱した。さらに、発現パターンがほとんど同じである遺伝子をグループにまとめることで、遺伝子が100個程度の場合でもある程度現実的な探索を行う方法を示している。これらの結果は遺伝子ネットワークのモデル化の方法によらないので、モデル化方法の評価にも応用可能である。

一方、現実の解析では単に理論上の最適解を求めるだけでは十分ではない。なぜなら、同程度のスコアをもつ解が多数存在し、スコア付けの方法や実験データが完全でないからである。そこで、著者は最適解に近いネットワークをスコアの順に数え上げる方法を提案し、著者がネットワークモチーフと呼んでいるいくつかのネットワークの共通構造を抽出するアルゴリズムと組み合わせることで、推定された遺伝子ネットワークが、枯草菌や酵母などの現実の生物の既知ネットワークと有意に一致することを明らかにしている。この方法を巨大な遺伝子群からなるネットワークに適用する場合には、遺伝子ネットワークの密度が高い部分を、生物学の知識やクラスタリングなどによって抽出することにより可能であることも示されている。これらの研究の結果、世の中に存在する遺伝子ネットワーク推定手法の有効性とその限界が明らかにされつつあり、これは今後のこの分野の一層の発展を促すものと思われる。

以上、見てきたとおり、本論文の著者は、生命情報科学における非常に重要かつホットな課題に取り組み、顕著な成果を上げているといえる。実際、この論文に報告されている成果の多くはすでにいくつかの査読付き国際学会に採択されているか、現在投稿中である。また、本論文の内容とは直接の関係はないが、著者はRNAスプライシングにおいて、いかに巨大なイントロンがうまく切り出されるかという問題に対して、intra-splicingというユニークな仮説を提唱しており、生命情報科学において別の側面からも貢献をしている（Ott et al., Pacific Symposium on Biocomputing 2003, 339-350）。審査会議上では、ここに述べてきたような、生物学の実際上の問題をきっちり解くという著者の姿勢を積極的に評価しようという声が多かった。反面、これらの仕事は、純粋に情報科学のアルゴリズム研究という観点からみた場合、抽象化の面で少し弱いと言わざるを得ないが、それは著者の基本的な研究へのスタンスの問題であり、そのことをもってして学位の要件を満たさないと考えるのは不当であるという結論に至った。

よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。