学位論文審査の要旨

博士の専攻分野の名称　博士 (情報科学)　　氏名　栗田 和宏

審査担当者　主　査　教　授　有村 博紀
副　査　教　授　吉岡 真治
副　査　教　授　Zeugmann Thomas

学位論文題名

Efficient Enumeration of Substructures in Sparse Graphs
（疎なグラフにおける部分構造の効率良い列挙）

インターネットやクラウドコンピューティングに代表される情報通信技術の急速な発達にともなって, 社会と人間の活動に関する大規模な離散構造データが蓄積されつつある. 本論文では, 複数の対象をリンクでつないだグラフとして表現された大規模離散構造データから, その中に含まれる興味深い部分構造を網羅的に見つけ出すための列挙アルゴリズムについて研究する. このような部分構造の列挙アルゴリズムは, 情報科学において古くから研究されており, これまでに連結部分グラフや, マッチング, 部分木, パス, 閉路といった部分グラフの族に対して, 効率良いアルゴリズムが開発されている.

データマイニング・機械学習・情報検索といった列挙アルゴリズムの実世界応用においては, サイズの大きな入力グラフに対して短い時間で全ての部分構造を列挙できる計算性能の良い列挙アルゴリズムが望ましい. 一般に, アルゴリズムの計算性能評価には, 入力サイズに対する総計算時間の増大を見積もる入力依存時間計算量を用いる. しかし, 列挙では, 最悪時には入力グラフのサイズ n の指数的な数の解があるので, どのような列挙アルゴリズムでもすべての解を出力するには, 入力サイズの指数時間を要してしまう. アルゴリズム同士の計算効率の優劣を比較するのは難しい. しかし, 実際の応用では, 最悪時の見積もりよりも著しく小さい場合も多く, そのような解の数が小さい場合には短い時間で終了するようなアルゴリズムが望ましい. そのため, 本論文では, 全ての解を出力する際に, 解一つあたりの計算時間を比較する出力依存計算量を用いて列挙アルゴリズムの効率を評価する.
数十頂点程度の分子構造データから, 数百万頂点以上の SNS のリンクデータにおよぶ実世界のグラフデータの多くは, 可能な頂点の組み合わせよりも, それが含むリンクの数が著しく小さいという特性, すなわち, いわゆるグラフの疎性 (sparsity) をもつことが知られている. グラフアルゴリズムの分野では, 入力グラフがある性質をもつか判定したり, ある性質をもつ部分グラフを一つ見つけるといった伝統的な計算タスクにおいて, 入力グラフの疎性を利用した高速なアルゴリズムの設計が盛んに研究されている. このようなグラフ疎性にはさまざまな尺度があるが, 本論文では, 次の尺度に着目する.

(a) グラフの内周 g: グラフがもつ最短の閉路の長さ.
(b) グラフの次数 d: グラフがもつ頂点の最大次数.
(c) グラフの縮退数 k: グラフの任意の部分グラフがもつ最小次数の最大値.

これらの尺度は, それぞれ, グラフの異なるタイプの疎性をとらえている. 例えば,(a) の内周 g が定

数のとき, 任意の頂点から g の半分程度の距離の部分グラフでは, 枝数が頂点数の線形になる. 一方,(b) の次数 d や (c) の縮退数 k が定数ならば, そのグラフの枝数は頂点数の線形で上から抑えられる.

本論文において, 著者は, このようなグラフの疎性が, どのように列挙アルゴリズムの性能に影響するかを考察している. 具体的には, 次の 4 つの問題について考察し, 入力グラフまたは列挙される部分グラフが疎性を持つ場合に, 計算効率の良い列挙アルゴリズムを与えている. 以下では, 入力グラフの頂点数を n とし, 辺数を m で表す.

第 3 章では, その中のどの二つの辺も互いに別の辺で連結されていないような辺集合である誘導マッチング列挙問題を考察する. 著者は, 入力グラフの内周が小さな場合, 具体的には, 長さ 4 の閉路を含まない場合に, 解一つあたりならし $O(1)$ 時間で解を列挙できるアルゴリズムを与えている.

第 4 章では, 著者は, 内周制約部分グラフの列挙問題を考察し, 出力される部分グラフが小さな閉路を持たないときに, 一般の場合より効率良く, 解一つあたりならし $O(n)$ 時間で解を列挙できるアルゴリズムを与えている.

第 5 章では, 入力グラフの全ての頂点にある辺で接しているような頂点集合である支配集合の列挙問題に取り組む. 著者は, 縮退数が k のグラフに対して, ならし $O(k)$ 時間の列挙アルゴリズムを与えた. さらに, 内周が 9 以上のとき, この問題がならし $O(1)$ 時間で解けることを示している.

第 6 章では, 二部グラフ, かつ, 長さ 6 以上の誘導サイクル, すなわち大きさが 6 以上の「穴」を持たないような部分グラフである弦二部誘導部分グラフの列挙問題を考察する. 著者は, 入力グラフのパラメータである縮退数 k, グラフの次数 Δ, グラフ中の最大二部クリークサイズ t がすべて定数の場合に, 解一つあたりならし $O(1)$ 時間で解を列挙できるアルゴリズムを与えている.

これを要するに、著者は、離散アルゴリズムにおける種々の部分構造列挙問題を考察し, 種々の疎性パラメータに着目し, これらのパラメータが定数で限定された疎なグラフにおいて, 部分構造の効率良い列挙アルゴリズムの設計と実装が可能であるという新知見を得たものであり, 情報科学における離散アルゴリズムと大規模データ処理分野において貢献するところ大なるものがある。よって著者は北海道大学博士 (情報科学) の学位を授与される資格あるものと認める.