

部分グラフ同型問題を解くための外部記憶アルゴリズム

An External Memory Algorithm for Subgraph Isomorphism

学籍番号：201521649

氏名：王 志華

Zhихua WANG

部分グラフ同型問題は、グラフに関する問題の中で最も基本的かつ重要な問題の一つであり、画像における特徴量抽出、化学式における類似構造の検出など様々な分野に応用できる。これまで、部分グラフ同型問題に対して、多くのアルゴリズムが提案されている。一方、近年では大規模なグラフデータが大幅に増加しているため、主記憶のサイズを超えるデータを処理する必要性が生じている。これまで提案されてきた部分グラフ同型問題を解くためのアルゴリズムは、データ全体を主記憶に格納して処理することを前提としているため、データが主記憶に収まらない場合には適用困難である。

そこで、本研究では、サイズが大きく全体を主記憶上で処理できないグラフデータも処理可能な、部分グラフ同型問題を解くための外部記憶アルゴリズムを提案する。ここで、外部記憶アルゴリズムとは、外部記憶アルゴリズムとは、データにアクセスするタイミングをアルゴリズム側で制御し、大規模データを効率よく処理するためのアルゴリズムの総称である。提案アルゴリズムは、次のような二つのアルゴリズム（アルゴリズム1、アルゴリズム2）として構成する。アルゴリズム1では、データグラフを先頭から一部分ずつ主記憶に読み込み、パターングラフとマッチングを行い、完全解と部分解を求める。ここで、パターングラフと同じ形になるグラフは完全解と呼び、パターングラフと部分的に一致し、現時点で完全解になるか否か不明のグラフを部分解と呼ぶ。アルゴリズム2では、アルゴリズム1で得られた部分解に対して、部分解がなくなるまで「拡大」操作を行う。ここで、パターングラフとデータグラフのノードとエッジを参照し、部分解において不明だったノードとエッジの情報を追加するという処理を「拡大」操作と呼ぶ。「拡大」操作の後、完全解になったものは解として出力し、完全解にならなかったものは削除する。

評価実験を行い、データサイズを変化させたとき得られた完全解数と部分解数、実行時間及び主記憶使用量を評価した。その結果、概ね所望の効率でアルゴリズムが動作することが確認された。

研究指導教員：鈴木 伸崇

副研究指導教員：若林 啓