部分グラフ同型判定のためのデータ依存回路

指導教官 市川周一　学籍番号 003317　山本 昌治

1　はじめに

部分グラフ同型判定は化学物質の構造活性推測など幅広い応用を持つが, 一般に NP 完全で計算は困難である．部分グラフ同型判定の専用回路には，Ullmann の提案 [1] と小西の提案 [2] があるが，どちらの提案も回路規模の制限により実装可能な頂点数には限界がある．先行研究 [3] では，Ullmann の提案回路をデータ依存回路で実現することにより回路規模が減少することを示したが，実装には至っていない．

本研究では, 部分グラフ同型判定のためのデータ依存回路を設計・実装し，ソフトウェア実装より高速な判定が可能であることを示す．評価には Xilinx XC2V3000 FPGA を搭載したボード YDK MIRE-MULTI3000 を用い，Ullmann の回路と小西の回路の 2 つについて，1. 回路規模，2. 実行時間，3. 回路生成時間 を評価した．

2　部分グラフ同型判定

グラフ G_α, G_β が与えられたとき，G_α が G_β の部分グラフと同型か判定する問題を部分グラフ同型判定問題という．V_α, V_β はそれぞれ G_α, G_β の頂点集合とする．以下，グラフの頂点数を $p_\alpha = |V_\alpha|$, $p_\beta = |V_\beta|$ とおく．部分グラフ同型判定問題は $v_i \in V_\alpha$ から $w_j \in V_\beta$ への部分グラフ同型になる写像を見つける一種の探索問題である．この問題の専用回路は，写像を探索木を用いて列挙する探索木巡回部と，探索木の各ノードで部分グラフ同型が成立するための必要条件を満たすか調べる枝刈り部で構成される．

3　データ依存回路

論理回路の入力が定まると，論理の簡単化によりゲートを削除できる (図 1)．あるゲートの出力が定数に決まれば，それを入力とするゲートも簡単化されるので，ゲートの削減は再帰的に適用される．論理段数が減少することにより動作周波数の向上も期待できる．記憶素子の入力が定数である場合，その記憶素子も削除できる．このようにして実現された回路は入力データに依存した実現になるため，以下では “データ依存回路” と呼ぶ．

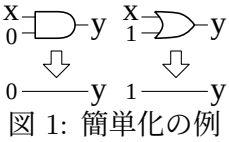

図 1: 簡単化の例

データ依存回路では入力グラフ毎に回路を生成するため，回路生成時間 T_{gen} が必要となり，総実行時間 T は実行時間 T_{exec} と T_{gen} の合計となる．T_{gen} には，入力グラフが与えられてからデータ依存回路が FPGA 上に構成されるまでの CAD ツールの処理時間が含まれる．

4　設計

Ullmann の提案回路 (Uo) は先行研究 [3] の *INDEP* を XC2V3000 に移植したものである．

Ullmann のデータ依存回路 (Ud) は先行研究 [3] の *BOTH2* を XC2V3000 に移植したものである．

小西の提案回路 (Ko) には，既存の実装 [2] が存在する．3 章で説明した論理の簡単化は，順序回路に含まれる内部状態 (記憶素子) を越えて適用できない．既存の実装 [2] では，枝刈り部が順序回路で実現されていたため，そのままではデータ依存回路に適していない．本研究では，枝刈り部を組合せ論理回路で新たに設計した．

小西のデータ依存回路 (Kd) は Ko をもとにしたデータ依存回路である．入力グラフ G_α, G_β の隣接行列 $A = [a_{ij}]$ $(1 \le i, j \le p_\alpha)$, $B = [b_{ij}]$ $(1 \le i, j \le p_\beta)$ は，Ko の枝刈り部を構成する組合せ論理の入力である．Kd は，a_{ij}, b_{ij} をグラフ G_α, G_β に応じた定数に置き換えて生成した回路で，a_{ij} と b_{ij} の記憶素子が不要になる．

5　評価

4 章の 4 つの設計を，回路規模，実行時間，回路生成時間について評価する．以下，頂点数を $p = p_\alpha = p_\beta$ とする．入力グラフには，各 p に対してランダムグラフを 100 組生成したものを用いた．ランダムグラフとは辺存在確率が与えられたグラフである．本評価では，G_α と G_β の辺存在確率はそれぞれ 0.3 と 0.6 とした．

入力グラフに依存した回路記述を生成する自作プログラム，論理合成 CAD Synopsys FPGA Compiler II 3.7.0, FPGA 実装 CAD Xilinx ISE 5.2i, を Athlon XP 2600+ PC (メモリ 1GB，Windows 2000 Pro. SP4) 上で実行し，処理時間の合計を回路生成時間 T_{gen} とした．実行時間 T_{exec} は，パラレル I/O ボード ADTEK aPCI-P31A を用いて判定終了信号を検出して測定した．性能の比較対象として Ullmann のアルゴリズムを C 言語にて実装し, Athlon XP 2600+ PC (メモリ 512MB，Red Hat Linux 9) 上で実行時間を測定した (Soft)．

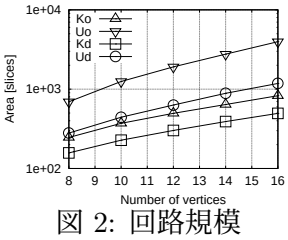

図 2: 回路規模

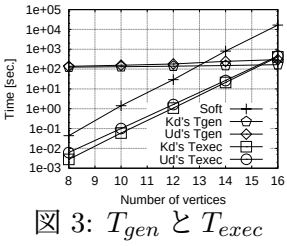

図 3: T_{gen} と T_{exec}

$p = 8 \sim 16$ のとき, Ud の回路規模は Uo の 40.3〜29.7%, Kd は Ko の 63.8〜59.9%であった (図 2)．Ud の T_{gen} は，Kd の 1.1〜1.7 倍と大きい (図 3)．一般に，T_{gen} は回路規模に依存する．Kd の回路規模が Ud より小さいので，Kd の T_{gen} は Ud より小さくなったと考える．Ud の性能は Soft の 7.0〜36.5 倍，Kd の性能は Soft の 15.7〜40.1 倍であった (図 3)．アルゴリズムの効率は入力に依存する [2]．$p \le 16$ の場合，Kd の方が効率的にアルゴリズムが働いたと考える．

6　おわりに

データ依存回路は，論理の簡単化により回路規模が減少し，同じ回路規模なら，より大きな p の回路が実装できる．しかし，データ依存回路では回路生成時間 T_{gen} が発生する．本研究では，入力にランダムグラフを用いて，部分グラフ同型判定のためのデータ依存回路を定量的に評価し，回路生成時間を含めてもソフトウェア実装の 28.6 倍高速に判定できることを確認した (図 4 の $p = 16$)．このときソフトウェア実装は 2.1GHz，データ依存回路は 25Hz で動作していた。

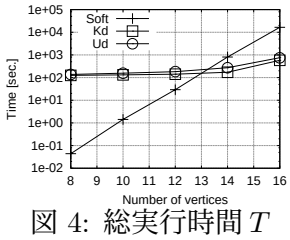

図 4: 総実行時間 T

参考文献

[1] J. R. Ullmann. An Algorithm for Subgraph Isomorphism. *J. ACM*, 23(1):31-42, 1976.

[2] 市川周一, L. ウドーン, 小西幸治. 部分グラフ同型判定アルゴリズムの FPGA による実装と評価. 情報処理学会論文誌ハイパフォーマンスコンピューティングシステム, 41(SIG5(HPS1)):39–49, 2000.

[3] S. Ichikawa and S. Yamamoto. Data dependent circuit for subgraph isomorphism problem. *IEICE Transactions on Information and Systems*, E86-D(5):796–802, 2003.