

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 11 日現在

機関番号：62615

研究種目：基盤研究(C)

研究期間：2011～2014

課題番号：23500055

研究課題名(和文) グラフ構造を対象としたデータ統合問題解決のための文脈を保存する融合変換の開発

研究課題名(英文) A Context-Preserving Fusion Transformation for a Graph Query Language used in Data Integration

研究代表者

加藤 弘之 (KATO, Hiroyuki)

国立情報学研究所・コンテンツ科学研究系・助教

研究者番号：10321580

交付決定額(研究期間全体)：(直接経費) 2,400,000円

研究成果の概要(和文)：副作用のない純粋な関数型言語の分野で開発された、プログラム変換技術の一つである「融合変換」に、副作用を有する言語に適用可能なものに拡張するためのアルゴリズムを与えた。特にデータモデルとしてグラフ構造を対象とした言語を用いる。グラフ構造は、現実世界の実体を直接かつ自然に表現できるデータ構造であるからである。研究代表者が既に開発した木構造を対象とした副作用を有する言語であるXQueryはグラフ構造も扱うことができるため、この言語に対する健全な融合変換を与えた。また、完全性を示すためのスキーマのクラスと問合せ言語のクラスを示した。

研究成果の概要(英文)：In this research, context-preserving fusion has been developed. So far, fusion transformation works only for side-effect free languages. However, XQuery, which is a query language to XML data, has a side-effect when it is typically used in data integration settings. Most of real world entities can be represented in graph structures. XQuery can handle graph structures. In this work, a sound context-preserving XQuery fusion has been developed. Also, the class of the queries and the schema has been investigated to show the completeness of the fusion.

研究分野：データベースプログラミング言語

キーワード：関数型言語 融合変換 副作用 XQuery 最適化

1. 研究開始当初の背景

異なるスキーマのもとに存在している同じ意味を表すデータを統一的に扱う問題は、データ統合問題と呼ばれるデータベースの研究分野で長年取り組まれてきている問題である。データ統合問題において、対象とするデータモデルは関係モデルと Web 上での情報交換フォーマットとして開発されてきた XML データ (データモデルは木または森) であった。グラフ構造は、現実世界の実態をそのまま自然に表現できるデータ構造であるが、これまでグラフ構造を対象としたデータ統合問題に関する研究は存在しない。尚、データ統合問題においては、閉包性を達成するために、使われているデータモデルの問合せ言語を用いてスキーマの変換を記述する。

一般に、データ統合問題における問合せ処理では、スキーマ変更前のデータを検索することで冗長なスキーマ変更のオーバーヘッドの回避ができる。この手法は関数型言語の分野で開発された融合変換に基づく問合せの書換えで実現できる。ところが融合変換は、関係 DB の問合せ言語である SQL のように副作用のない言語上でしか成り立たない。スキーマ変更前のデータを検索した結果は、スキーマ変更後の構造を反映していないからである。このスキーマ変更後の構造に関する情報は文脈と呼び、文脈を保存する正しい融合変換の開発は、現実世界の問題であるデータ統合問題における問合せ最適化に役に立つ。

2. 研究の目的

本研究の目的はグラフデータモデルを対象とした副作用を有する問合せ言語に対する融合変換を開発することである。これにより、データ統合問題における多段のスキーマ変換が単純化することが可能となり、問合せの最適化につながる。

3. 研究の方法

(1). まず研究の対象とするグラフ構造を対象とする問合せ言語の調査を行った。具体的には、グラフ問合せ言語 UnQL が構造に関する副作用を有するように拡張可能かどうかの調査を行った。その結果、UnQL の大きな特徴である、双模倣不変を維持しながら副作用を追加することが難しいことがわかった。

(2). そこで、グラフ構造も扱うことができ、既存のデータを参照しながら新しい構造を作成することができかつ、実際に幅広く使われている言語として、XQuery を採用し、XQuery の融合変換および、その完全性について示すことにした。

4. 研究成果

(1). グラフ問合せ UnQL を拡張し、構造に関する副作用を与えるようにするための、調査

を行った。具体的には、根付きのグラフ構造を対象とし、逆向きエッジやサイクルを表現するための特別なエッジの存在価値について、調査を行った。そのためにデータの静的構造とデータを操作するための代数構造が別のものとして設計することを目指した。これは、XPath の設計が、静的には木構造であるが操作する代数構造はグラフ構造を表現しているからである。その結果、入力順序を保持する変換は記述可能だが、入力順序を条件とした変換や、順序を変更するような順序に関する副作用は記述できないことが判明した。

(2). グラフ問合せ言語として、XQuery を対象とした場合、研究代表者によるさきの研究課題「基盤研究(C) 課題番号 20500043 XML データ統合問題解決のための XQuery の静的解析に基づく書換え手法の開発」において、問合せの中間結果に現れるエレメント構築子の削除による融合変換は健全ではあるが、完全ではなかった。そこで、多段の変換を単純化するためには、XQuery に対して、良く知られたモナド法則などを適用し、for-式の標準形を定義することで、in-式を経路式に限定することが可能となり、これまで削除できなかったエレメント構築子の削除に成功した。

(3). XQuery の書換えスキーマ情報を使うことの有効性とプロトタイプ実装に取り組んだ。まず、どのスキーマのクラスを用いるかの検討を行い、現実的に使われているスキーマのクラスとして、duplicate-free、covering、disjunction-freeなどを候補として、その記述能力、実際にどのくらい使われているかの調査を行った。その結果、これらのスキーマ間について記述能力に関しては包含関係が成り立たず incomparable であることがわかった。

現実的な使用状況については、duplicate-free かつ covering なものの使用がほとんどであるという報告があった[参考文献 1]。本研究では特に経路式に着目しその融合変換を用いた書換え手法に基づく最適化を提案していることから、本研究の成果が十分に適用可能でかつ現実的に幅広く使われているスキーマのクラスとして、duplicate-free DTD が適当であるとの結論を得た。プロトタイプ実装として、与えられた duplicate-free DTD とそれに妥当な XML document に対して、identity query を用意することで、本研究の成果にスキーマ情報を用いることが達成できた。

[参考文献 1]

Manizheh Montazerian, Peter T. Wood, and Seyed R. Mousavi, XPath Query Satisfiability is in PTIME for Real-World DTDs, 5th International XML Database Symposium, XSym 2007, LNCS4704,

Springer 2007, ISBN 978-3-540-75287-5.

(4). スキーマ情報を問合せ中に埋め込む手法として等価問合せを用いることを中心に研究をすすめた。具体的にはスキーマのクラスとして、データ交換などで広く使われている nested relational DTD を採用し、この DTD に妥当な文書全体について、等価問合せとの合成問合せを構築し、研究代表者が既に開発済みの副作用を考慮に入れた融合変換を適用することで、スキーマ情報を用いた従来よりも効率の良い問合せへと変換することが可能となった。本研究成果について興味を示している欧州の企業との情報交換なども行い、本研究成果の実用面での効果を確認できた。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 12 件)

Hiroyuki Kato, Soichiro Hidaka, Zhenjiang Hu, Keisuke Nakano and Yasunori Ishihara, Context-Preserving XQuery Fusion, Mathematical Structure in Computer Science (MSCS), 査読有, 25(4), 2015, pp.916-941, DOI:10.1017/S096012951300008X
加藤 弘之, 胡 振江, 日高宗一郎, 松田一孝, 高談闊論: 双方向変換の原理と実践 (解説論文), 日本ソフトウェア科学会論文誌コンピュータソフトウェア, 査読有, 31(2), 2014, pp.44-56, DOI:10.11309/jssst.31.2_44
Makoto Onizuka, Hiroyuki Kato, Soichiro Hidaka, Keisuke Nakano, and Zhenjiang Hu, Optimization for iterative queries on MapReduce, PVLDB, 査読有, 7(4), 2013, pp.241-252, <http://www.vldb.org/pvldb/vol7/p241-onizuka.pdf>
Soichiro Hidaka, Zhenjiang Hu, Kazuhiro Inaba, Hiroyuki Kato, Keisuke Nakano, GRoundTram: An integrated framework for developing well-behaved bidirectional model transformations, Progress in Informatics, Special Issue: Advanced Programming Techniques for Construction of Robust, General and Evolutionary Programs, 査読有, (10), 2013, pp.131-148, DOI:10.2201/NiiPi.2013.10.7
Kazuyuki Asada, Soichiro Hidaka, Hiroyuki Kato, Zhenjiang Hu and Keisuke Nakano, Parameterized Graph Transformation Languages with Monads, GRACE Technical Report

(GRACE-TR-2012-07), 査読無
2012-07, 2012, pp.3-23
中野圭介, 日高宗一郎, 胡振江, 稲葉一浩, 加藤弘之, 模倣に基づくグラフスキーマを利用したビュー更新可能性判定, コンピュータソフトウェア, 査読有
29(2), 2012, pp.174-192, DOI:10.11309/jssst.29.2_174
Soichiro Hidaka, Zhenjiang Hu, Kazuhiro Inaba, Hiroyuki Kato, Kazutaka Matsuda, Keisuke Nakano and Isao Sasano, Marker-directed Optimization of UnCAL Graph Transformations, 21st International Symposium on Logic-Based Program Synthesis and Transformation (LOPSTR 2011), Revised Selected Papers, Lecture Notes in Computer Science, 査読有, VOL.7225, 2012, pp.123-138, DOI:10.1007/978-3-642-32211-2_9
Soichiro Hidaka, Kazuyuki Asada, Hiroyuki Kato, Keisuke Nakano, Zhenjiang Hu, Towards Bidirectional Transformations on Ordered Graphs, Grace Technical Report (GRACE-TR-2011-07), 査読無
2011-07, 2011, pp.1-24
Yijun Yu, Yu Lin, Zhenjiang Hu, Soichiro Hidaka, Hiroyuki Kato, Lionel Montrieux, blinkit: Maintaining Invariant Traceability through Bidirectional Transformations, Technical Reports at the Center for Research in Computing and Technology, The Open University, 査読無, 2011(9), 2011, pp.1-14, DOI:10.1109/ICSE.2012.6227162
Soichiro Hidaka, Zhenjiang Hu, Kazuhiro Inaba, Hiroyuki Kato, and Keisuke Nakano, GRoundTram: An Integrated Framework for Developing Well-Behaved Bidirectional Model Transformations, Grace Technical Report (GRACE-TR-2011-05), 査読無
2011-05, 2011, pp.1-10
Soichiro Hidaka, Zhenjiang Hu, Kazuhiro Inaba, Hiroyuki Kato, Kazutaka Matsuda, Keisuke Nakano and Isao Sasano, Marker-directed Optimization of UnCAL Graph Transformations, GRACE Technical Report (GRACE-TR-2011-02), 査読無
2011-02, 2011, pp.1-20
Keisuke Nakano, Soichiro Hidaka, Zhenjiang Hu, Kazuhiro Inaba, Hiroyuki Kato, Simulation-based Graph Schema for View Updatability Checking of Graph Queries, GRACE Technical Report (GRACE-TR-2011-01), 査読無
2011-01, 2011, pp.1-20

〔学会発表〕(計 14 件)

日高宗一郎, 浅田和之, 胡振江, 加藤弘之, 中野圭介, Structural Recursion for Querying Ordered Graphs, 18th ACM SIGPLAN International Conference on Functional Programming(ICFP 2013), 2013年9月27日, Hilton Boston Longan Airport Hotel, Boston, USA
浅田和之, 日高宗一郎, 加藤弘之, 胡振江, 中野圭介, A Parameterized Graph Transformation Calculus for Finite Graphs with Monadic Branches, 15th International Symposium on Principles and Practice of Declarative Programming(PPDP 2013), 2013年9月16日, Computer Science Faculty of the Complutense University of Madrid, Madrid, Spain
浅田和之, 日高宗一郎, 加藤弘之, 胡振江, 中野圭介, A Parameterized Graph Transformation Calculus for Finite Graphs with Monadic Branches, 日本ソフトウェア科学会第30回大会, 2013年9月13日, 東京大学本郷キャンパス, 東京都
Vo Huu-Phuc, 加藤弘之, 日高宗一郎, 胡振江, An Experimental Implementation of Self-adjusting Bidirectional Transformations, 日本ソフトウェア科学会第30回大会, 2013年9月12日, 東京大学本郷キャンパス, 東京都
加藤弘之, Query Reformulation in XML Data Integration Systems, 3rd Asian Workshop on Advanced Software Engineering, 2013年7月21日, Temasek Labs in National University of Singapore, Singapore
Soichiro Hidaka, Kazuyuki Asada, Zhenjiang Hu, Hiroyuki Kato, Keisuke Nakano, Structural Recursion for Querying Ordered Graphs, 第15回プログラミングおよびプログラミング言語ワークショップ(PPL 2013), カテゴリ1, 2013年3月6日, 御宿東鳳, 福島県
Kazuyuki Asada, Soichiro Hidaka, Hiroyuki Kato, Zhenjiang Hu, Keisuke Nakano, Structural Recursion on Ordered Graphs, 日本ソフトウェア科学会第29回大会, 2012年8月24日, 法政大学小金井キャンパス, 東京都
Yijun Yu, Yu Lin, Zhenjiang Hu, Soichiro Hidaka, Hiroyuki Kato, Lionel Montrieux, Maintaining Invariant Traceability through Bidirectional Transformations, 34th International Conference on Software Engineering(ICSE 2012), 2012年6月7日, The Kongresshaus, Zurich, Switzerland
Soichiro Hidaka, Zhenjiang Hu,

Kazuhiro Inaba, Hiroyuki Kato, and Keisuke Nakano, GRoundTram: An Integrated Framework for Developing Well-Behaved Bidirectional Model Transformations, 26th IEEE/ACM International Conference on Automated Software Engineering(ASE 2011), 2011年11月9日, The Oread, Lawrence, USA
Hiroyuki Kato, Incremental Graph View Maintenance, 2nd NII-PKU International Joint Workshop on Advanced Software Engineering, 2011年10月23日, ラフォーレ強羅, 神奈川県
篠埜 功, 胡 振江, 日高 宗一郎, 稲葉 一浩, 加藤 弘之, 中野 圭介, GRoundTramによるATLの双方向化の実現, 日本ソフトウェア科学会第28回大会, 2011年9月29日, 沖縄産業支援センター, 沖縄県
加藤弘之, 鬼塚真, 日高宗一郎, 中野圭介, 胡振江, 漸進的グラフビュー維持の枠組みの提案, 日本ソフトウェア科学会第28回大会, 2011年9月29日, 沖縄産業支援センター, 沖縄県
Kazuhiro Inaba, Soichiro Hidaka, Zhenjiang Hu, Hiroyuki Kato, Keisuke Nakano, Graph-transformation verification using monadic second-order logic, 13th International ACM SIGPLAN Symposium on Principles and Practice of Declarative Programming(PPDP2011) 2011年7月20日, University of Southern Denmark, Odense, Denmark
Soichiro Hidaka, Zhenjiang Hu, Kazuhiro Inaba, Hiroyuki Kato, Kazutaka Matsuda, Keisuke Nakano and Isao Sasano, Marker-directed Optimization of UnCAL Graph Transformations, 21st International Symposium on Logic-Based Program Synthesis and Transformation (LOPSTR 2011), 2011年7月19日, University of southern Denmark Odense, Denmark

〔図書〕(計 0 件)

〔産業財産権〕
出願状況(計 0 件)

取得状況(計 0 件)

〔その他〕
ホームページ等
<https://researchmap.jp/katohiroyuki/>

6. 研究組織

(1) 研究代表者

加藤 弘之 (KATO, Hiroyuki)

国立情報学研究所・コンテンツ科学研究

系・助教
研究者番号：10321580

(2)連携研究者
該当なし