

氏 名	なか の けい すけ 中 野 圭 介
学位(専攻分野)	博 士 (理 学)
学 位 記 番 号	理 博 第 2964 号
学位授与の日付	平 成 18 年 3 月 23 日
学位授与の要件	学 位 規 則 第 4 条 第 1 項 該 当
研 究 科 ・ 専 攻	理 学 研 究 科 数 学 ・ 数 理 解 析 専 攻
学 位 論 文 題 目	An Implementation Scheme for XML Transformation Languages through Derivation of Stream Processors (ストリーム処理器の自動導出による XML 変換言語の実装法)
論文調査委員	(主 査) 助教授 長谷川 真人 教 授 齋 藤 恭 司 教 授 藤 重 悟

論 文 内 容 の 要 旨

XML (eXtensible Markup Language) は、タグとよばれる文字列を文章中に埋め込むことによって構造を与えられた構造化文書の汎用フォーマットである。XML 変換は、このような XML 形式のデータベースや文書を共有・交換するために重要な役割を果たしており、XML 変換言語とよばれる XML の木構造の変換に特化したプログラミング言語も数多く開発されている。しかし、これらの言語は、入力となる XML 全体の木構造をメモリに展開してその変換を行うため、メモリ空間や実行時間に多大な無駄が生じることが多い。

一方、XML ストリーム処理では、メモリに木構造を展開することなく、XML 文書を文字列として直接処理するため、メモリ空間の消費を大幅に節減することが可能である。また、入力全体を読み終える前に変換結果の出力が開始されるため、実行時間も短縮することが可能である。しかし、このようなストリーム処理は、複雑な状態遷移プログラムで記述する必要があり、プログラミングには大きな困難を伴う。

本論文では、木構造処理プログラムから、ストリーム処理器を自動的に導出する手法を提案する。これにより、既存の XML 変換言語と同様に、木構造処理として平易に書かれたプログラムから、メモリ空間や実行時間を浪費しない効率の良いストリーム処理器を、自動的に生成することが可能になる。

本論文で提案する手法は、属性文法の合成理論に基づいている。属性文法は、Knuth により提案された、文脈自由言語の意味の定義および言語変換を記述するための枠組みであるが、属性文法で与えられる言語変換は、その構成的な定義により、変換の合成が扱いやすいという著しい性質を持っている。ここでいう合成とは、言語 A から言語 B への言語変換と、言語 B から言語 C への言語変換を、それぞれ属性文法で与えたとき、それらを合成した言語 A から言語 C への変換を属性文法によって表現したものである。得られた合成結果は、言語 B を介さないため、もとのふたつの変換を適用するよりも効率的な変換となっている。本論文では、XML ストリームから木構造を導く変換と、プログラマが与える木構造から木構造への変換、および木構造から XML ストリームへの変換をそれぞれ属性文法で表現し、合成することで、木構造処理プログラムから、ストリーム処理器を自動的に導くことを可能にした。

ただし、既存の合成理論では、XML ストリームから木構造を導く変換に対応する属性文法を扱うことができない。本研究では、この困難を、著者の先行研究であるスタック付き属性文法のための合成理論を用いて解決することに成功した。

さらに、本論文では、今回提案する手法を実装し、既存の XML 変換言語との比較を行い、この研究に基づいた手法が、実際に効率的な処理を実現できることも示した。

論 文 審 査 の 結 果 の 要 旨

XML は現在最も普及している構造化文書形式であり、その効率的かつ明快な処理方法を与えることは、応用の豊富さを鑑みても、大変重要な課題である。

本研究は、この問題に対し、プログラム変換の理論および形式言語の理論のアイデアと枠組みを用いることで、エレガントかつ実用的な解決を提案した。その概要は、XML の木構造を処理するプログラムと、木構造とストリーム構造の間の変換を行うプログラムの合成結果から、効率的なストリーム処理を行う仕組みを抽出する、というものである。ここで鍵となる合成は、古典的な属性文法の合成の理論を、スタック付きの属性文法の合成の理論に拡張するという、著者自身の先行研究の結果を用いることによって可能となった。本論文では、さらに、本研究に基づく処理系が、既存のものと比較して効率が良いと考えられることを示す実験結果も報告している。

本論文の計算機科学における意義は、大きくわけて以下の二点にまとめられると考える。第一に、XML 処理という応用上重要な問題に対して、実用的な見地からみて顕著な貢献となっていること。本論文の成果は、今後、実用的な XML 処理系の開発に、多大な影響を与えることが期待される。第二に、プログラム変換論・形式言語論という、計算機科学の基礎的な理論を有効に駆使することが、実際的な問題に対しエレガントな解決を与えるという、非常に興味深い事例を示していることである。本論文の成果が刺激となって、これらの基礎研究のさらなる活性化につながるが大いに期待できる。

これらの貢献は、本研究が、計算機科学の応用と理論双方にまたがって仕事をしている、著者の広い見識と技量を示すものと考えられる。さらに付け加えると、本論文の先行研究における、スタック付き属性文法の合成に関する著者の基礎的な貢献も、高く評価できるものである。また、本論文に関連した著者の研究発表が日本ソフトウェア科学会高橋奨励賞を受賞するなど、著者の仕事の意義は、専門家の間でも、広く認められつつあるといえる。

よって、本論文は博士（理学）の学位論文として価値あるものと認める。また、論文内容とそれに関連した事項について試問を行った結果、合格と認めた。