

# 論文審査の結果の要旨

氏名 細谷 晴夫

インターネットの普及により、WWW は非常に多くのユーザに利用され、情報社会のインフラストラクチャーとして生活にとってさえ不可欠な存在になっている。WWW のデータ記述言語として最も利用されているのが HTML (Hyper Text Markup Language) である。HTML は言語仕様が簡潔であったため、WWW 黎明期において非常に勢いで広がり、また初期の段階における需要をカバーするには十分であった。しかし、おびただしい数のデータが HTML により記述されるに至り、HTML の言語仕様の素朴さは、拡張性に乏しい物足りなさとして映るようになってしまった。

HTML では `<title> Thesis </title>` のような `<>` で囲まれたタグを用いて、データを入れ子型に構造化しながら記述してゆく。HTML の不自由な点の一つは、タグ名を自由に定義できないことである。また、住所録のようなデータを記述する場合には、構造化の順序を規定する文法を記述してから、データの中身を書けるようにできれば、データが文法に従って書かれているか自動的にチェックできる。しかし HTML には文法を記述する機能が備わっていない。これらの要望を実現する形で最近 WWW コンソーシアムの標準化委員会が提案したのが XML (eXtensible Markup Language) であり、技術的には XML 処理言語の設計が研究の焦点となっている。

本論文の第1章では、XML 言語に関する簡潔な解説をした後、例題を用いながら非専門家にも研究成果がわかるように腐心している。

第2章では、XML 処理言語の核となる機能を2つ、すなわち「正規表現型」と「正規表現パターンマッチ」を提案している。正規表現型は、スキーマ言語によく見られる「繰り返し (\*)」や「選択 (|)」などの正規表現の記法を型システムに採り入れる。さらに、「意味論」的に定義された部分型をサポートすることを試みている。そして、この部分型がもつ柔軟性が、XML を用いたソフトウェアをスムーズに進化させるのに必要であることを論じている。正規表現パターンマッチは、従来の（関数型言語にあるような）パターンマッチ機構に、上記の繰り返しや選択の演算子を追加したものである。これら演算子を使うことによって、任意の長さを持つ部分木「列」をマッチすることができるようになるので、列中の任意の位置に一拳に飛んで、そこから欲しいデータを抽出するようなパターンが簡潔に記述できるという利点がある。本論文では、これら2つの機能を使った言語の例として XDuce 言語を設計し、その核言語を定式化し、型健全性を証明することで、言語の理論的基盤を与えている。本章の XDuce は巧妙に設計されており、計算不能な性質を生み出しかねない言語機能は注意深く避けた組み立てをしながらも、実用性を損なわない配慮が秀逸である。

第3章では、部分型の判定問題を計算量理論的に解析している。この判定問題は、正規ツリーオートマトンの包含判定問題に簡単に帰着できるが、この問題自体が指数時間完全という高い計算量をもつことが知られている。したがって理論的には、

すべての場合に効率的なアルゴリズムを設計することは困難であることが示唆される。しかし、多くの現実例で効率的に動作するソフトウェアが切望されており、次章でこの問題の解決に取り組んでいる。

第4章では、実際の入力に対しては効率的に動作するアルゴリズムを開発し、意表を突く良い結果を出している。Aiken と Murphy の集合制約の解消アルゴリズム考察し、それを拡張するアプローチを取っているものの、十分独創的であり、(1) 完全性の証明、(2) XML処理で発生する典型例に対応した実装技法の提案、(3) 予備実験がなされている。予備的実験では、現実例をいくつか試した結果、実用上申し分ない効率で動作することを確認している。XML が広がるにつれて、ここで提案されたアルゴリズムが広く世の中で利用されてことが期待できる。

第5章では木構造を操作するための便利な言語機能を提示している。正規表現パターンマッチを型付きプログラミング言語に用いる際、パターンマッチに出現する変数の型を自動的に計算することにより、過度な型注釈をユーザに要求することを回避できる。そのために、パターンマッチを取り巻く文脈からパターンの変数へ、型情報を伝播するような型推論アルゴリズムを開発している。このアルゴリズムは、まず型とパターンをツリーオートマトンへ変換し、ツリーオートマトンの閉包演算(和・積・差)を使って型情報の伝播を行なっている。技術的に困難な点は、繰り返し演算子と、選択演算子の「第一マッチ則(先に出現する選択肢が優先的にマッチする性質)」が相互作用し、これによって解析の停止性と精度を同時に保証することが難しくなるということである。本論文では、閉包演算を明示的に含んだような型の表現を導入することにより、この問題点を解決している。技術的に困難な問題に解答を与えた点で評価できるだけでなく、このアルゴリズムも広く普及することが期待できる。

なお本論文は Benjamin Pierce, Jerome Vouillon 氏との共同研究であるが、論文提出者が主体となって提案・分析・評価を行ったもので、論文提出者の寄与が十分であると判断する。

したがって、博士(理学)を授与できると認める。