

## 木分割アルゴリズムを用いた XSLT 実行手法

# An XSLT Transformation Method Using a Tree Partitioning Algorithm

学籍番号：201521655

氏名：朱 亜婷

Yating ZHU

XML は Web のデータ形式における事実上の標準として広く用いられている。ここで、ある程度まとまった量の XML データを管理・蓄積する場合を考える。このような場合、DTD 等のスキーマ言語を用いて格納すべきデータの構造を前もって定義しておき、それに対して妥当なデータを作成・利用するのが一般的である。データを蓄積・保存する場合、XML は有用な言語であるが、ユーザに提示する際には HTML など他の形式に変換した方が都合がよいことが多い。このような変換を行うための変換用言語として、XSLT がよく使われている。

ここで、XML データに対して XSLT を用いて変換処理を行うことを考える。近年、サイズの大きな XML データが急速に増加しており、そのようなデータに対する変換処理の効率が悪化するという問題が生じている。これまで、サイズの大きい XML データの処理に関しては、XPath の実行手法については研究されているが、XSLT の実行手法に関する研究は行われていない。近年、XML データのサイズが増加する一方で、計算機プロセッサの性能は大幅に向上している。このため、この状況に応じた効率の良い XSLT 変換手法の重要性が増大している。Java などのプログラミング言語自体は複数スレッドの生成・実行に対応しているが、XSLT 処理系で複数スレッドの実行を考慮して処理効率の向上を図ったものは著者の知る限り存在しない。

そこで本論文では、XSLT のサブセットである木変換機を対象に、効率の良い変換手法を提案する。本手法では、DTD を用いて XML データを分割すべき箇所を求め、そこで変換処理を切り分けることで、XML データの変換処理を並列化し、処理効率の向上を図る。XML データと比較すると、DTD は十分に小さく、短時間で分割位置を決定できるという利点がある。提案手法の有効性を評価するため、変換規則数が異なる 4 つの木変換機を用いて評価実験を行った。その結果、提案手法を用いた場合、変換処理を並列化しない場合と比べて処理時間を短縮できることを示した。

研究指導教員：鈴木 伸崇

副研究指導教員：阪口 哲男