

分散 XML に対する XSLT 実行手法

水本 弘貴

XML(Extensible Markup Language) は Web 上の標準的なデータ形式として幅広く用いられている。近年、一つの組織が扱うデータ量が大幅に増加していること、更にまた地理的または管理上の要因から XML データを複数の断片に分割して管理した方が都合が良いことなどから、複数サーバに分散させて管理することの必要性が増加している。分散 XML に関しては、XPath(XML Path Language)の実行手法やスキーマ設計については研究されているが、XSLT(XML Stylesheet Language Transformations)の実行手法に関する研究は行われていない。

分散 XML に対する現行の XSLT 実行手法は、複数に分割された XML データを一つのサーバで統合し XSLT を実行するというものである。XSLT はサイズの大きい XML データに対して実行効率が著しく悪化することが知られている。そのため上記のような実行手法では、分割配置された個々の XML データのサイズが小さいものであっても統合したデータのサイズは大きいので、XSLT の実行効率が著しく悪化してしまう。そこで本論文では、分散 XML に対して効率よく XSLT を実行する実行手法を考案する。

考案手法では分散 XML の配置状況に応じて各サーバ上で並行して XSLT を実行することで、特定のサーバへの処理の集中を回避し、処理の効率化を図っている。

XSLT はチューリング完全であり、XSLT の動作を完全に解析した上で効率の良い分散実行手法を立てることは困難である。そこで本論文では、XSLT の機能を XSLT の基盤となる下降型木変換機(top-down tree transducer)の機能に制限している。また、分散 XML は DTD の外部実体参照宣言を利用し実現している。

本考案手法を Ruby で実装し評価実験を行った。その結果、考案手法は既存手法と比較して、十分な実行時間の短縮が認められる結果が得られた。

(指導教員 鈴木伸崇)