

論文の内容の要旨

論文題目 Purifying XML Structures (XML構造の純化)

氏名 齊藤 太郎

木構造を持つXMLでは様々なデータ表現が可能である。本論文では、データベース理論でよく知られている「関数従属性 (Functional Dependency)」という概念がXMLのデータモデリングにおいて重要な役割を担うことを示す。これまでXML向けに提案されてきた関数従属性は固定されたパスに対して定義されていたため、XMLの構造に必要以上に強い制約を課すものであった。この問題に対して、我々は、パスではなくノードを基準に関数従属性を導入し、XMLデータの表現に自由度を与えることを試みた。この方式の問題点は、様々な構造で表現されたXMLデータをすべて問い合わせするコストが高くつかかもしれない、ということである。そのため、我々はamoeba join (アメーバ結合) という新しい問い合わせ手法を考案し、関数従属性を満たすXMLの構造に狙いを定め、それらを効率よく取り出す技術を開発した。これらの手法は、XMLの問い合わせにパスのようなXMLの構造を指定しなくても、目的となる木構造に関数従属性から自動的に定めることができるという点で、XMLの問い合わせに新しい観点をもたらすものである。複雑なXMLの構造をユーザーの目から隠し、関数従属性から構造を決めてしまうという、この発想の転換を「XML構造の純化」と呼ぶ。さらに、関数従属性から派生して定義されるキーを用いることで、異なった構造を持ったXMLデータを容易に集約することができるようになる。これらの技術は、複数の階層構造を持ったXMLデータを管理するための障害を取り除き、XMLデータベースの表現力を高めるのにも役立つ。また、問い合わせのために必要な、B+木上で実装可能なXMLの索引構造についても紹介する。実験では、提案された手法の有用性と、それが実現可能であることを示す。そして、本研究の成果が科学データの表現にどのように活用でき、今後、XMLデータの管理にどのようなインパクトをもたらすのかについて議論する。