

審査の結果の要旨

氏名 齊藤 太郎

XML(Extensible Markup Language)は記述の容易さから広く使われるデータ形式であるが、XMLでは木構造という汎用性の高いデータの表現方法を用いるため、データモデル（データの持つ意味）とは直接関連のないノードの親子関係までもを内包してしまうという問題があった。そのため、従来のXMLデータの問い合わせでは、データの意味とは関連のないノードの親子関係まで具体的に指示する必要があり、このことは、利用者にとって、XMLデータをデータベースシステム上で管理する上での障害となっていた。

本論文は、全7章から構成されており、データモデルとは直接関連のないXML中の木構造を、関数従属性(Functional Dependency)という概念を用いて抽象化（XML構造の純化）することで、XMLデータの管理、問い合わせを容易にする手法を提案している。その実現のために、問い合わせ代数を設計し、その代数の上での問い合わせアルゴリズム、効率の良い索引構造などを開発し、計算機実験を通じて性能を実証している。

第1章の導入部では、XMLデータを扱う上での現実的な問題点について、バイオインフォマティクスでの例などを用いながら言及し、パス構造を用いない問い合わせの必要性、関数従属性をXMLに導入する意義、またそれらを実現するためのデータベース設計について解説している。

第2章では、様々な形態を持った木構造をまとめて扱うためにamoebaという構造を導入し、それをもとにXMLの問い合わせにおける代数演算の入出力となるamoeba domainを定義している。過去の研究では、固定された木構造の入出力を用いて代数演算を設計していたが、本論文では扱える木構造に柔軟性を持たせたという点で新規性がある。

第3章では、XMLに関数従属性を導入し、与えられた関数従属性の定義に対応するXMLデータの取りうる木構造を定めている。また、様々な構造を取りうるXMLデータを、データベース中から取り出すために、ノードの結合順序を入れ替えることにより、結果が等価で、より効率の良い問い合わせスケジュールに置き換えられることを証明している。関数従属性を用いて、データベース中のノードを一意に特定するためのキーを定義することができ、これを応用し、XMLデータ中のノードの同値類を定めるための偏在キー(ubiquitous key)を考案しており、XML内で木構造がどのように構成されているかという知識を持たない状態でも、問い合わせ文を書くことを可能にしている。これらの貢献により、XMLデータの煩雑な木構造にとらわれず、関数従属性によって定義されるデータモデルにのみ着目して、より直感的な問い合わせが行えるようになっている。

第4章では、第3章までの内容を実装するために、amoeba構造をXMLデータベースから効率良く取り出すためのアルゴリズムを2種類紹介しており、第5章では、B+木を用いて、XMLデータの木構造、パス構造の組み合わせを用いた問い合わせが可能な多次元索引を提案している。

第6章では、提案手法について計算機を用いた実験評価を行っており、第7章で、関連研究と結論について述べている。また、XMLにおいて関数従属性を使うことにより、データの加工操作を必要とするバイオインフォマティクス分野などのデータ管理が容易になることを例示している。

なお、本論文の内容は、森下真一との共同研究によるものであるが、論文提出者が主体となって分析および検証をおこなっているもので、論文提出者の寄与が十分であると判断する。

よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。