

## 論文審査の結果の要旨

氏名 倉光 君郎

データ設計やデータベースの設計は、従来は関係データベースを基本的に用い、特定のプラットフォーム、組織の特殊事情、従来方の応用を強く意識して行われてきた。しかし、インターネットやユビキタスコンピュータのようなオープンな計算・データ処理環境が登場した現在では、個別な条件や事情をもとに最適化されたデータ設計は必ずしも有効であるとはいえない。本論文は、電子カタログ交換システムやスマートカード・チケットシステムの開発の実践から、どこでも誰でも使える新しいデータ表現としての **Nomadic Data** を提案している。

問題設定として、**Nomadic Data** に求められる要求は、二つに大別することができる。

1. (意味論的な相互運用性) データ作成者とデータ利用者は、組織に分離しており、相互運用に関する合意が必要であるが、それはアドホックであり進化し続ける。
2. (構造・様式的な相互運用性) データは、RDBMS や XML, スマートカードなどの様々なプラットフォームの上で利用される。

ここでは、意味論的な相互運用性に対して現実的な解を提供することを目指し、グローバルスキーマ方式を採用する。**Nomadic** データモデルは、柔軟で再利用性の高いグローバルスキーマを構築することを目指している。更に、そのデータモデルを、現在実用されている XML, RDBMS, スマートカード上に適用した場合の有効性や効率上のメリットを測定・議論し、**Nomadic Data** の有用性を示している。各章のアウトラインと貢献は以下のとおりである。

第1章は、本論文の中心的内容である **Nomadic Data** の考えに立ち至った動機付けとなる問題を、データベース研究の発展のなかで位置づけつつ、研究成果の概要を述べるとともに、成果の新規性を要約している。

第2章は、**Nomadic Data** の意味論(解釈)の基礎となる **T-Mapping** の定義をしている。**T-Mapping** の持つ翻訳的な性質を形式化するため、FOL ステートメントを用いて、翻訳の操作を公理化している。重要な主張は、(ステートメントの範囲から) **context-free-mapping** のクラスを導入した点である。また、**mapping** を **context-free** にするため、ドメイン分割(とドメインの正規形)の定義が与えられている。

第3章では、**Nomadic Data** モデルを形式的に定義している。新しいデータモデルを支える中核的なアイデアは、**mapping-friendly** である。スキーマは、従来の汎化・特化ではなく、**mapping** の再利用性(**context-free** の度合い)から部品化する。また、**Mapping** の複雑さから、データ表現の正規化する手法を提案し、例えば、この正規化手法を用いて (**check-in-date, check-out-date**) と (**check-in-date, night**) の表現上の優劣が判断できる。また、本モデルで作成されたグローバルスキーマは、自律分散的なスキーマ進化、値の柔

軟性、異なる名前スキームの運用などの要件を満たすことが示されている。

第4章では、Web Service におけるアドホックなデータ交換への Nomadic Data の適用が議論されている。この環境では、完全なグローバルスキーマの想定が成り立たず、一般に論理言語を用いたデータ統合の作業が必要となる。Nomadic Data は、データ表現の正規化のおかげで、簡単なグラフベースのマッチングだけで必要なデータ統合を導出することが可能となる。これは、RDF と XML Datatypes を用いた簡単な記述言語作成し、示されている。

第5章では、Nomadic Data が関係データベースをベースとして大容量の格納・検索（レポジトリ）を構築する上でも適している点を論じている。実際、Nomadic Data モデルのドメインの意味論に着目して、ストレージ分割する方法を開発している。これによって、明らかに意味的に関係ない部分を検索する無駄が省ける。DBLP のデータコレクション用いた実験結果により、従来の EDGE 方式に比べ、10 倍以上の効率化を達成できた。また、元々にリレーショナルストレージ自身に対しても、目立った性能劣化はなく、Null 値の多い属性では検索処理がより高速になることが確かめられている。

第6章では、スマートカード上の Nomadic Data を議論されている。明らかに、スマートカードでは、限られたリソースのもとでの最適化が主目的となる。関係データモデルをスマートカード上で用いると、Null 値の発生が無駄となる。実際、プロトタイプ実装したスマートカード DB 上で、Null 値と Nomadic data のストレージコストの対応を定量的に分析している。この結果、多品種少量データを蓄積する場合は、(スマートカード上では一般にこのケースが当てはまるが)、Nomadic Data は蓄積効率に優れていることが示された。

第7章では、本文の結論として研究成果の概要を再び列挙するとともに Nomadic Data が電子商取引、電子チケットやデジタル博物館等において実用に供される可能性を示唆している。

以上より要するに、本論文では、柔軟なグローバルスキーマの構築に優れ、RDBMS やスマートカードなど様々なプラットフォームに利用できるデータ表現である Nomadic Data を提案し、その意味論や構造・様式を厳密に定義するとともに、プロトタイプシステムを実装し、その有効・有用性を示した。よって、理学上の貢献が高く、博士（理学）の学位を授与するに十分であると、審査委員全員が一致して判断した。

なお、全ての章および付録の内容は論文提出者の指導教官である坂村健氏他との多くの共著論文をベースにしたのであるが、研究上の視点および技術的内容の大部分は、論文提出者自身のものであると判断されるので、論文提出者に学位を授与することに何ら問題はない。