

表作成によるスキーマの自動生成

5 K-7

加藤 隆康† 遠山 元道‡

†慶應義塾大学 理工学研究科 管理工学専攻

‡慶應義塾大学 理工学部 情報工学科

1 はじめに

関係データベースにおけるスキーマは、本来ならばER図の作成や正規化などの手順に基づき設計を行う。しかし、データベースの専門知識のないユーザーには、この作業は困難である。昨今ではパソコン上の表計算ソフトが普及し、一般ユーザは情報管理のツールとしてRDBよりもこちらを選択することが多い。その理由の一つとして、RDBで複数の表に分解しなければならない情報を一括して直観的に配置できることが挙げられる。しかしながら、これらは拡張性などにおいて著しい制約を受けてしまう。

そこで、ユーザーに表計算ソフトと同様に直観に基づくイメージ(表)を作成してもらい、その得られた表の入れ子構造を利用して、自動的にスキーマを生成する試作システムを作成した。本稿では、そのスキーマ生成の方法についての概要を報告する。

2 従来の論理設計

データベースシステムの目的は、実世界の一部をコンピュータ内に実現することである。実世界は無限の情報とこれに基づいた多種多様な活動から構成され、すべてコンピュータ内に実現することはできない。したがってデータベースを構築するとき、実世界のモデルが必要となる。

設計段階で、モデルの把握がきちんとしていないと有効なデータベースが得られないので、ER(Entity-Relationship)モデルなどを用いて慎重に行なうことが多い。また実世界の変化を正しくデータベースに反映させるためリレーションを正規化する必要がある。

る。正規化の利点としては以下の2つがあげられる。

1. 更新の異常を防ぐ。各リレーションを実世界の基本的な事実と対応づけることにより、事実単位に更新を行う。
2. 冗長性を減少し、記憶コストを低減する。

このようなモデル化や正規化には専門的知識が必要となるため、一般ユーザには敬遠されやすいのが現状である。

3 表の入れ子構造を利用した正規化の手法

実世界には様々な表が存在するが、本研究ではその入れ子構造をもとにリレーションを分解する事により、適切なスキーマ群を作成する。例えば次のような表の入れ子構造は、(店名(売場(従業員)(商品)))というリスト構造で表現できる。

店名	売場	従業員	商品
渋谷店	玩具	佐野 史郎	UNO
		内藤 和男	将棋
			人生ゲーム
	文具	鈴木 雅之	ボールペン
		和田 真弓	万年筆

得られたリストを入れ子により分割し、それぞれに主キー、外部キーを与えると、次のようなスキーマを得る。

スキーマ1 (主キー (1), 店名)
 スキーマ2 (主キー (2), 外来キー (1), 売場)
 スキーマ3 (主キー (3), 外来キー (2), 従業員)
 スキーマ4 (主キー (4), 外来キー (2), 商品)

この例で得られたスキーマは、第三正規形である。一般に、生成されたスキーマが正規形条件を満たすかどうかは、元となった表の構造に依存する。

Automatic Generation of Schema from a Hierarchical Table Structure.

KATO Takayasu, TOYAMA Motomichi
 Department of Administration Engineering, Faculty of Science and Technology, Keio University.

4 実装

本研究では表からのスキーマ生成支援システムを作成し、CASD(Computer Aided Schema Design)と名付けた。CASDは対話的かつグラフィカルなシステムであり、ユーザに表作成のための直観的な操作環境を提供する。図1は、CASDシステムの初期画面である。

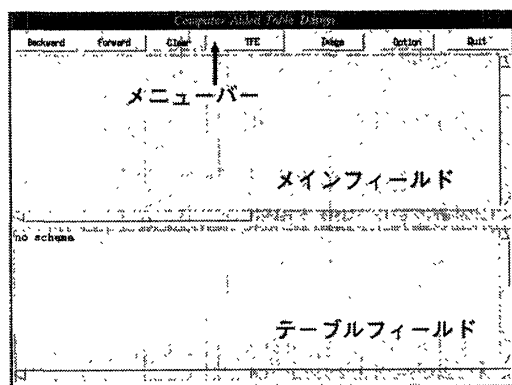


図1: CASDの初期画面

ユーザはメインフィールド上で表の概略イメージを作成する事により、テーブルフィールドに適切なスキーマを得る。表の項目をキーボードで入力し、データの反復の方向を指示する事で表の入れ子構造が決定される。表のイメージの作成は主にポインティングデバイスを用いるため、ユーザは簡単に現実の表のイメージを画面上に実現できる。概略イメージは、図2のように表示される。

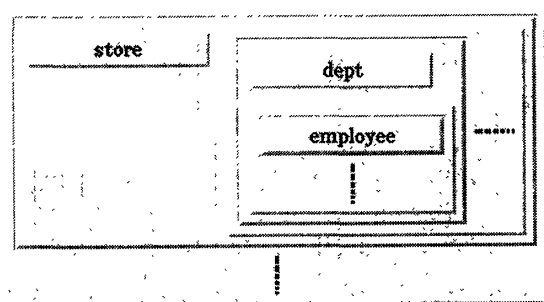


図2: 概略イメージ

5 評価・検討

本研究では、実世界に存在する様々な表から専門家が設計したスキーマと、本研究で試作したシステムにより得られたものとを比較した。その結果、専門知識のないユーザでもスキーマ設計を容易に行なう事ができた。ただし、名簿や履修表などの単純な表については、正しいスキーマを得られたが、参照関係が複雑な表については、少々の手直しが必要であった。(図3)

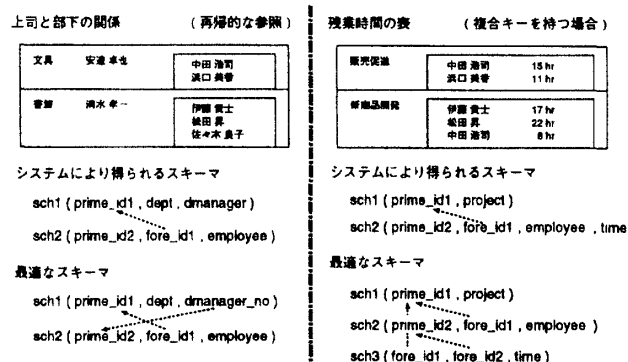


図3: 参照関係が複雑な表

6 おわりに

本研究では利用者が印刷物などで目にする表をそのままに近い形で入力し、これから入れ子構造を抽出してスキーマ設計を行う。過去にも入れ子構造に注目してスキーマ設計を行う研究[1]は行われており、これらと本質的に差があるとはいえない。しかし、一般利用者にとって、目前の表の階層構造を正しく認識することは想像以上に困難であり、本システムはこの点に着目して利用者の負担を軽減した。結果として、広く一般的なユーザにもデータベースが身近なものになることが期待される。

参考文献

- [1] Y. Kambayashi, T. Furukawa, H. Yamamoto: Realization of Nested Relation Interfaces for Relational and Network Databases, in Nested Relations and Complex Objects in Databases (Lecture Notes in Computer Science 361) 1987: 217-228