

データベース学習のための関係代数アニメーション

佐々木 透^{1,a)} 寺田 実^{1,b)}

概要：現代のリレーショナルデータベースは問い合わせ言語として SQL を用いるソフトウェアの形式で存在しているが、情報工学においてこのデータベースを理論的に扱う関係代数という分野が存在する。関係代数が実用的に用いられることはないが、データの表現、操作、正規化理論などを離散数学的に扱うもので情報系の大学では講義のカリキュラムとして組込まれている場合が多い。本研究では、関係代数におけるデータ表現形式の可視性の高さに着目し、インタラクティブに関係代数のデータ操作、及び正規化の可視化をするソフトウェアを実装した。

1. 背景

意味のあるデータを表現する方式としては JSON や XML などの形式が存在するが、最も実用的に用いられている形式はリレーショナルデータベースである。リレーショナルデータベースでは図 1 のように、表現したいデータを有限の属性の組み合わせからなる項として 1 行に記述するテーブル形式という表現方法を導入しており、極めて可視性が高いものとなっている。情報工学においてリレーショナルデータベースを扱う理論は関係代数と呼ばれ、現代のデータベースで用いられるデータ操作やデータベースの正規化を数学的に厳密に議論するものであり、情報系の大学で講義のカリキュラムとして組み込まれていることが多い。

関係代数の学習過程においてテーブル同士の演算に関する項目が存在し、講義の形式については参考書やスライド、板書を用いるなど静的に視覚化したテーブルが用いられている。よって学習者は演算前のテーブルと演算後のテーブルを見て演算内容をイメージの中で補完するという学習方法を強いられる。この問題を解決するために本研究では関係代数の学習支援を目的とし、テーブルの演算内容を動的なアニメーションによって可視化するソフトウェアを実装した。

2. 関係代数

2.1 テーブル表現

関係代数とはエドガー・F・コッドによって確立され、集合論における直積を用いてデータを表現する理論体系である。例では COURSES というテーブルは ID と NAME と

いう属性を持ち、1 つの行はその 2 属性の直積の部分集合である関係によって数学的に表現される。

2.2 関係代数の演算

テーブルを操作する言語はいくつか種類が存在するが、コードは関係代数においてテーブルを操作するための 8 演算を [1] の中で定義し、特に次の 5 つの演算と同等の能力を持ち合わせている言語をリレーショナル完備であると定義している。

- 和演算: 2 つのテーブルのデータの和集合を返す
- 差演算: 2 つのテーブルのデータの差集合を返す
- 直積演算: 2 つのテーブルのデータの直積を返す
- 選択演算: 1 つのテーブルについて、指定した条件に合致するデータを取り出す
- 射影演算: 1 つのテーブルについて、指定した列のみを切り出したテーブルに分解する

上記以外に共通演算、結合演算、商演算の 3 つの演算も定義されているが、挙げた 5 つの演算を組み合わせることで実現が可能なので説明は省略する。また商演算のみ、演算としての重要度が低いため本研究では実装の対象外としている。関係代数演算は全てテーブルに対する演算であり、式の評価結果もテーブルであるというような再帰性がある。

COURSES		STUDENTS			
ID	NAME	ID	FIRST_NAME	LAST_NAME	AGE
1	Database	1	Maurizio	Cembalo	28
2	Programming	2	Mario	Rossi	32
3	Algebra	3	Michela	Esposito	23
4	Networks	4	Gianni	Verdi	19
5	Statistics				

図 1 関係代数におけるテーブル表現の例

¹ 電気通信大学

^{a)} s1731077@edu.cc.uec.ac.jp

^{b)} terada.minoru@uec.ac.jp

2.3 関数従属性

ある属性 A の持つ値が決定されると、別の属性 B の値も決定されるような従属性が存在する場合、B は A に関数従属すると表現する。従属される属性は単一でなくてもはならないが、従属させる属性は複数の属性の組み合わせでもない。関数従属性は、データベースの正規系を定義する際に用いられる概念である。

3. 関連研究

関係代数の視覚化に関する研究はないものの、SQL によるデータベースの動きを学習支援としてアニメーション化するという研究はいくつか存在する。Cembalo らは大学生を対象として SQL の構文の教育支援を目的とした SAVI[2] を実装した。SAVI はデータベース上のテーブルから SQL の SELECT 文によって取得したい結果のテーブルが取り出される様子をアニメーションによって可視化し、SELECT 文の WHERE 句、JOIN 句、GROUP BY 句をサポートしている。Kristin は SAVI 同様に SQL の理解支援を目的として SELECT 文のアニメーションを行う viSQLizer[3] を実装した。[3] ではデータベースの講義の担当講師へ事前アンケートを実施しており、その中でアニメーションによる可視化ツールを併せて学習に用いることは有効であるという回答を得ている。評価実験では講師が参考書を用いて講義をしたグループと、viSQLizer を用いて独学したグループに SELECT 文のみを対象とした SQL のテストを実施し、平均点で 1% 程度上回っている。

本研究は SQL ではなく関係代数の理解を目的としているが、SQL と関係代数の演算の能力は関係代数の方が高い。SAVI が実装している演算能力は 5 つの関係代数演算の内直積演算、選択演算、射影演算、の 3 つ、viSQLizer は 2 つと、データベースに対する操作の能力はリレーショナル完備ではない。また、本研究ではテーブル内の属性同士の従属関係を設定することができ、単にデータベースの動きの学習ではなく正規化の学習支援に用いることが可能である。

4. 提案システム

本システムは起動時に以下の 2 つのウィンドウが立ち上がる。起動には Java 7.0 の実行環境が必要。

- アニメーションウィンドウ
- テーブル一覧ウィンドウ

また更にアニメーションウィンドウは 3 つのエリアに分割される

- 関係代数式入力エリア
- アニメーション描画エリア
- ログ出力コンソール

4.1 アニメーションウィンドウ

図 3 の (A) に配置している関係代数式入力エリアは数

SELECT * FROM exam returned the following table:

course_code	student_no	year	score
IMT4003	100902	2014	5
IMT4032	99123	2013	3
IMT4032	100902	2013	0

course_code	student_no	year	score
IMT4003	100902	2014	5
IMT4032	99123	2013	3

図 2 viSQLizer[3] による SELECT 文のアニメーション例

式を入力として受け付け、実行ボタンを押すと入力された代数式を構文解析し、(C) のアニメーション描画エリアにテーブル同士の演算を描画する。関係代数は演算結果自体がテーブルであるという再帰性があるので、演算結果のテーブルをさらに別の演算に用いることも可能である。

4.1.1 ログ出力コンソール

関係代数エリアの下部の (B) にログを出力するコンソールを設置してある。ここには主に次のようなログやエラーを出力する。

- 今までに実行した関係代数の演算式
- 関係代数式のパースエラー出力
- 演算の条件を満たしていないテーブル同士の演算を行った場合のエラー出力

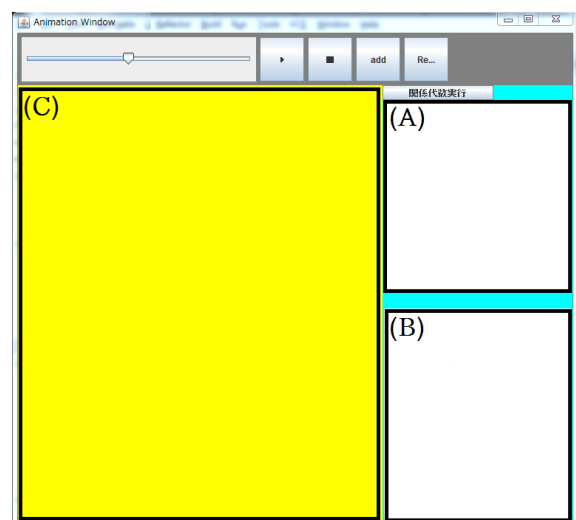


図 3 アニメーションウィンドウ

product	price
テレビ	198000
洗濯機	59800
餅つき機	29800

"PRICE"

customer	product	quantity	total
A商店	テレビ	3	594000
Bマート	テレビ	10	1980000
Bマート	洗濯機	5	299000
C社	餅つき機	1	29800

"ORDER"

図 4 テーブル一覧ウィンドウ

- テーブルを分解した時に分解後のテーブルが正規系になっているか出力

4.2 テーブル一覧ウィンドウ

図 4 に示すテーブル一覧ウィンドウでは、現在演算に使えるテーブルを一覧表示する。主キーはヘッダー色を濃く表示し、属性同士の従属関係は矢印で表す。初回起動時には開発者が用意したテーブルを一覧表示するが、アニメーションウィンドウにて新しい演算が行われた場合や、学習者が新たにテーブルを定義した場合は一覧ウィンドウを更新して新規追加されたテーブルを表示する。

5. 使い方

5.1 関係代数式の入力

ユーザはテーブル一覧画面に存在するテーブルから演算に使いたい任意のテーブルを選んで関係代数式を記述する。入力後に実行ボタンを押すとアニメーション描画エリアにテーブルの初期状態から演算結果の状態までを可視化するアニメーションが描画され、以降ユーザによる操作が発生するまで無限にループする。

5.2 アニメーションの実行制御

アニメーションの実行中、ユーザは停止ボタンを押下して任意の時点で停止させることが可能である。さらに停止状態でスクロールバーをドラッグ操作することによりアニメーションを制御し、再生や逆再生が出来る。

5.3 テーブルのユーザ定義

ユーザは開発者がデフォルトで用意したテーブルのほかに、自分でテーブルを定義することが可能である。add ボタンを押下することにより、図 5 に示すテーブルを追加する GUI ウィンドウが立ち上がり、ユーザはテーブル名、属性、従属性、テーブルの持つデータを記述して定義することが出来る。

5.3.1 テーブル名

テーブル名を入力する。構文解析の都合上、テーブル名

図 5 テーブル追加ウィンドウ

は半角アルファベットと数字の組み合わせとする。例ではテーブル名として"R1"という名前を付けている。

5.3.2 属性 (テーブルヘッダー)

テーブルヘッダーは属性名をカンマ区切りで記述する。命名規則はテーブル名と同様。例では 5 つの属性を定義している。

5.3.3 従属関係の記述

テーブルヘッダーに記述した属性の内、従属させるカラムと従属するカラムをそれぞれ記述する。例では、"species" が "ID" に関数従属するという従属性を定義している。

5.3.4 データ

テーブルの保持するデータをカンマ区切りで入力する。複数のデータを入力したい場合は改行して同じように行を入力する。例では、3 つのデータを入力している。

5.4 アニメーションのリセット

演算やテーブルの追加を行って学習を進めていると不要な結果テーブルやログがコンソールエリアに蓄積されてゆく。その場合はリセットボタンを押すことで開発者が用意したデフォルトテーブル以外のテーブルと、ログを削除できる。

6. 評価実験

学生を被験者とする評価実験を予定している。被験者は本システムを用いて関係代数と正規化について学習したグループと、本学のデータベース論の講義で用いられている参考書 [4] を用いて学習したグループに分け、こちらが用意する問題を解いて正答率を比較する。

問題は関係代数、正規化をどちらもデータベーススペシャリストの午前 II 問題から出題する予定である。

7. まとめ

リレーショナルデータベースの学習支援は SQL に関わる研究が大半で関係代数に焦点があてられたものはほぼ存

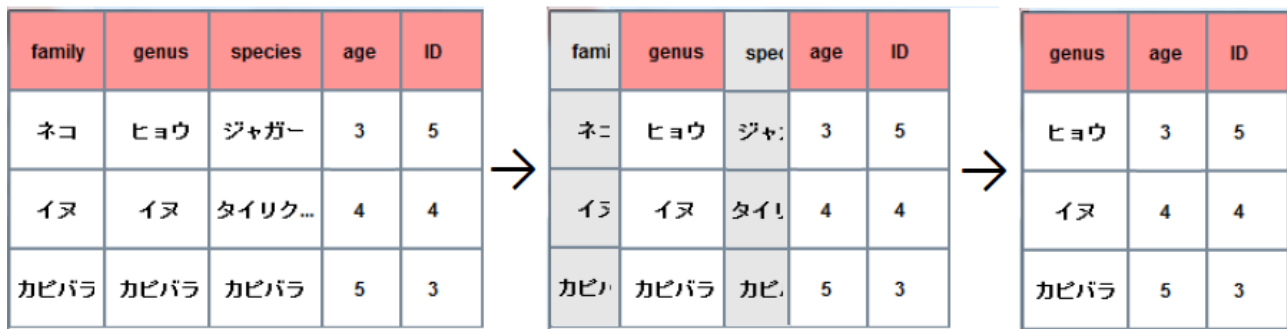


図 6 射影演算 $R1[\text{genus}, \text{age}, \text{ID}]$ のアニメーション例

在しない。関係代数演算は実用的に用いられることはないが、情報工学系の学部で講義として扱う大学は多く、特に正規化に関する厳密な理論の習得は現代のデータベースの設計に大いに役に立つ。

8. 今後の課題

使いやすさと学習範囲の2つの観点から、今後に追加が望ましい機能として以下のものが考えられる。まず、使いやすさに関して、学習者が自分でテーブルを定義するための機能を提供しているが、学習の都度入力させるのは不要な時間を費やしてしまう。これを解決するために、予め外部ファイルに任意のテーブルを保存して読み込ませる等の機能が挙げられる。学習に関して、本システムは関数従属性の実装までしているが関係代数には多値従属性という概念が存在し、これを実装すればさらに高次の正規系を学習範囲とすることができる。

参考文献

- [1] E. F. Codd(1972), “Relational Completeness of Data Base Sublanguages” *Database Systems*:65-98, Prentice Hall and IBM Research Report RJ 987, San Jose, California.
- [2] Maurizio Cembalo, Alfredo De Santis, Umberto Ferraro Petrillo, “SAVI: a new system for advanced SQL visualization” *SIGITE '11 Proceedings of the 2011 conference on Information technology education*, Pages 165-170, West Point, New York, USA — October 20 - 22, 2011.
- [3] Folland, Kristin Annabel Torjussen(2016) “viSQLizer: An interactive visualizer for learning SQL” (Master’s thesis). Norwegian University of Science and Technology (NTNU), Norway.
- [4] 増永良文 (2015). リレーショナルデータベース入門 [新訂版]. サイエンス社.