

Возможности использования функциональных методов и алгебраических структур при анализе данных в информационных системах

Суин Илья Алексеевич

*Смоленский государственный университет
магистрант*

Аннотация

В статье рассмотрены возможности использования функциональных методов и алгебраических структур при изучении информационных систем и баз данных в частности. Автором определены проблемы современных баз данных, а также приведены примеры использования методов функционального анализа и алгебраических систем при исследовании информационных систем. Актуальность статьи состоит в перспективности применения данных методов анализа структуры и семантики современных информационных систем.

Ключевые слова: базы данных, функциональный анализ, алгебраические системы, информационно-коммуникативные технологии, информационные системы

The possibility of using functional techniques and algebraic structures in analysis of data in information systems

Suin Ilya Alekseevich

*Smolensk state University
Undergraduate*

Abstract

The article considers possibilities of the use of functional methods and algebraic structures in the study of information systems and databases in particular. The author defines problems of modern databases and leads examples of using methods of functional analysis and algebraic systems at research of information systems. The relevance of the article is the prospect of applying these methods of analysis structure and semantics of modern information systems. The relevance of the article are the perspectives of the application of these methods of analyzing the structure and semantics of modern information systems.

Keywords: database, functional analysis, algebraic systems, information and communication technology, information systems

На данный момент изучению компьютерных баз данных, как примера информационных систем, уделяется особое внимание [1, 2, 3]. Количество информации, которой должен оперировать человек современного мира,

неуклонно растет. В связи с этим возрастает объем информации, хранимой в информационных системах различного вида. Так увеличивается объем данных в информационных объектах обучающего характера, таких как энциклопедии, словари, учебные пособия. Аналогичные процессы происходят с информацией, которая имеет отношение к деятельности конкретных фирм и предприятий сферы бизнеса. Постоянно расширяются списки поставщиков, клиентов, товаров, возрастает объем отчетов различного вида и так далее.

Возможности современных информационно-коммуникационных технологий на сегодняшний день таковы, что они позволяют пользователю информационных систем свободно оперировать такими большими объемами информации. Так, при необходимости во всех вышеперечисленных случаях работы с информационными объектами человек может усваивать, запоминать и в должной мере применять результаты обработки информации, например, с помощью последних достижений интернет-технологий. В такой ситуации встает вопрос о систематизации, распределении информации и приведении её к необходимому пользователю формату.

Одним из вариантов решения данного вопроса является представление структуры информационной системы в виде базы данных. Базы данных, как известно, представляют собой совокупность взаимосвязанных, хранящихся вместе данных, при условии наличия условий, необходимых для оптимального использования одной или несколькими программами [4]. На этапе проектирования базы данных возникают такие вопросы как, где хранить базу данных, или каким образом будет представлена находящаяся в ней информация?

Например, если информационная система имеет внушительный объем, разумнее будет не помещать её всю в одну базу данных, а условно разделить на несколько элементов, разместить их на нескольких вычислительных машинах и получить, таким образом, распределенную базу данных. В случае если данные удобно представить в виде каталога – есть корневой объект, у него есть связи с объектами более низкого уровня, а они в свою очередь представляют собой совокупность объектов еще более низкого уровня – такую информацию удобнее всего будет хранить в иерархической базе данных. Если же каждый элемент информации высшего уровня может иметь одновременную связь с любым числом элементов следующего уровня, то для такой информационной системы будет создана сетевая база данных. И наконец, если информационная система представляет собой набор объектов одного типа и набора, то она найдет свое место в табличной или реляционной базе данных.

Остановимся более подробно на отдельных аспектах проектирования и эксплуатации реляционных баз данных, как одним из перспективных направлений хранения и обработки информации с применением методов функционального анализа и алгебраических систем.

Так как информация представлена в табличном виде, программист и пользователь могут столкнуться с такой проблемой как дублирование

информации. В частности это возникает в случае, если большое количество объектов обладает одинаковыми свойствами, которые будут продублированы в таблице, что увеличит объем информации исходной базы данных. Также может возникнуть проблема с обновлением данных в информационной системе в силу необходимости редактирования каждой записи [5].

Реляционная база данных лишена большей части этих недостатков, если она отвечает определенным требованиям. К ним относят требования нормализации, которые подразделяются на три формы, построенные по иерархической структуре. Форма низшего уровня достигается в случае, если ни одна из строк таблицы не содержит в любом своем поле более одного значения, а ни одно из полей однозначно идентифицирующую строку таблицы (первичные), не является пустым. Следующая форма включает в себя предыдущую и добавляет новый критерий. Все поля таблицы, не являющиеся первичными, с этими самыми первичными полями связаны полной функциональной зависимостью. В свою очередь, функциональная зависимость, определенная для таблицы, это зависимость при которой единственному первичному значению ставится в соответствие единственный набор описаний этого ключевого поля. Наконец последняя форма опять же включает в себя предыдущую со следующим дополнением – ни одно из полей, которое не является ключевым полем таблицы, не находится в функциональном отношении с другим не ключевым полем [6].

Если таблица удовлетворяет всем формам, то над ней можно проводить определенные операции [7]:

1. Пересечение таблиц с одинаковой структурой. Результат – выбираются те записи, которые находятся в обеих таблицах.

2. Объединение таблиц с одинаковой структурой. Результат – конкатенация (общая таблица: сначала первая, затем вторая).

3. Вычитание таблиц с одинаковой структурой. Результат – выбираются те записи, которых нет в вычитаемом.

4. Декартово произведение двух таблиц. Записи результирующей таблицы получаются путем объединения каждой записи первой таблицы с каждой записью второй.

5. Выборка (горизонтальное подмножество). Результат – отвечающие определенным условиям записи.

6. Проекция (вертикальное подмножество). Результат – отношение, содержащее часть полей из исходных таблиц.

Базы данных создаются не только для непосредственного хранения информации, они также открывают широкий спектр возможностей по анализу своего содержимого. Наиболее распространенным методом изучения содержимого баз данных является статистический анализ. В свою очередь реализация базы данных в программной среде не только значительно упрощает применение этого метода, но и открывает новые, более сложные, но и более эффективные способы взаимодействия с элементами баз данных. К таким способам относятся методы функционального анализа и алгебраические системы. Обе эти методологии целесообразно применять для

структурирования данных в информационных системах, в частности для оптимизации структур баз данных и выявления латентных, на первый взгляд, взаимосвязей между их различными компонентами.

Так, например, методы функционального анализа позволяют исследовать объекты многотабличных информационной системы не сами по себе, а изучать степень влияния элементов или групп элементов одной таблицы на элементы или группы элементов другой таблицы [8, 9]. Так же, относящиеся к функциональному анализу методология соответствия Галуа [10, 11] и математический аппарат имплицативных матриц [12, 13] могут быть применены в области образовательных технологий, в частности при формировании оптимальных стратегий обучения [14, 15, 16]. Применение данных методов обусловлено тем, что математические методологии остаются неизменными относительно рассматриваемой области знаний, что значительно расширяет область применения данных методов.

Так же, на базе информационной системы может быть введена алгебраическая система, как одноосновная, так и многоосновная. Введение алгебраической системы окажет существенную поддержку при разработке программного приложения, а именно инструментов изучения содержимого базы данных. Допустим, у нас есть база данных, элементами которой является условный текст, называемый в программировании строковой переменной. На этой базе данных можно задать двухосновную алгебраическую систему, имеющую в основе множество S (множества всех строковых переменных) и множество N_0 (множество натуральных чисел и 0). Также в этой алгебраической системе присутствует несколько операций:

1. «+»: $S \times S \rightarrow S$ – соединение двух строчных переменных в одну, путем непосредственного склеивания второй строковой переменной с первой.

2. «L»: $S \rightarrow N_0$ – перевод символьной переменной в числовую – вычисление количества символов в строковой переменной.

3. «P»: $S \times S \rightarrow N_0$ – определение позиции строки второго S в первом S .

4. «C»: $S \times N_0 \times N_0 \rightarrow S$ – копирование части строки начиная с первого N_0 до второго N_0 .

5. «I»: $S \times S \times N_0 \rightarrow S$ – вставка второго S в первый S , начиная с позиции N_0 .

Помимо операций в данную систему мы можем ввести отношения или предикаты. Например, отношение эквивалентности, как отношение посимвольного совпадения строковых переменных. Наличие данной системы позволяет программе использовать для анализа данных таблиц привычные операции, что значительно повышает эффективность работы. Очевидно, что данная система не является единственной. То, как будет выглядеть алгебраическая система, зависит, прежде всего, от задач, которые ставит перед собой программист.

Итак, на данный момент, в силу постоянного увеличения объема информации, большое количество пользователей нуждается в качественном

анализе информации. Чтобы у пользователя была такая возможность, разрабатываются новые методы работы с содержанием баз данных. Таким образом, можно утверждать, что использование методов функционального анализа и алгебраических систем является перспективным направлением.

Библиографический список

1. Баженов Р. И., Лопатин Д. К. О применении современных технологий в разработке интеллектуальных систем // Журнал научных публикаций аспирантов и докторантов. 2014. № 3 (93). С. 263-264.
2. Козлов С. В. Использование соответствия Галуа как инварианта отбора контента при проектировании информационных систем // Современные информационные технологии и ИТ-образование. 2015. Т. 2. № 11. С. 220-225.
3. Максимова Н. А. Моделирование информационно-образовательной среды учебного заведения // Концепт. 2016. № 5. С. 195-200.
4. Гринченко Н. Н., Гусев Е. В., Макаров Н. П. Проектирование баз данных. СУБД Microsoft Access. Учебное пособие. М., 2004. 240 с.
5. Ульман Д. Д., Уидон Д. Реляционные базы данных. М., 2014. 384 с.
6. Харрингтон Д. Проектирование объектно-ориентированных баз данных. М.: ДМК-Пресс, 2012. 272 с.
7. Ребекка М. Р. Основы реляционных баз данных. М., 2010. 390 с.
8. Козлов С.В. Особенности организации и анализа данных в информационных системах // NovaInfo.Ru. 2017. Т. 1. № 62. С. 9-14.
9. Козлов С.В. Интерпретация инвариантов теории графов в контексте применения соответствия Галуа при создании и сопровождении информационных систем // International Journal of Open Information Technologies. 2016. Т. 4. № 7. С. 38-44.
10. Козлов С.В. Применение соответствия Галуа для анализа данных в информационных системах // Траектория науки. 2016. Т. 2. № 3 (8). С. 18.
11. Парватов Н. Г. Соответствие Галуа для замкнутых классов дискретных функций // Прикладная дискретная математика. 2010. №2(8). С. 10-15.
12. Козлов С. В. Использование математического аппарата импликативных матриц при создании и сопровождении информационных систем // International Journal of Open Information Technologies. 2017. Т. 5. № 12. С. 16-23.
13. Козлов С. В. Информационные системы как инструмент функционального анализа потоков данных // NovaInfo.Ru. 2017. Т. 1. № 61. С. 25-30.
14. Козлов С. В. Применение методов функционального анализа при формировании оптимальных стратегий обучения школьников // Международный журнал экспериментального образования. 2016. № 3-2. С. 182-185.
15. Козлов С.В. Особенности обучения школьников информатике в профильной школе // Научно-методический электронный журнал Концепт. 2014. № 1. С. 31-35.

16. Киселева О.М. Реализация принципа индивидуализации образовательного процесса с использованием программы «Траектория обучения» // Современные научные исследования и инновации. 2014. № 5-2 (37). С. 41.