

ОБ ОДНОМ ПОДХОДЕ К КЛАССИФИКАЦИИ МОДЕЛЕЙ ДАННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

У статті досліджено загальні принципи формування формально-логічних моделей даних. Запропоновано розширити класифікацію моделей даних. Проведено дослідження щодо відповідності об'єктно-реляційного підходу до формування самостійної логічної моделі даних.

Ключові слова: інформаційна система, бази даних, реляційна модель даних.

Рассмотрены общие принципы формирования моделей данных в информационных системах. Предложено расширить существующую классификацию моделей. Проведено исследование соответствия объектно-реляционного подхода формированию на его основе логической модели данных.

Ключевые слова: информационная система, базы данных, реляционная модель данных.

The paper researched the general principles of formal logical data models. Proposed to expand the classification models. A research on correspondence with the object-relational approach to the formation of independent logical data model.

Keywords: information system, database, relational data model.

Введение.

Основные идеи информационных технологий XXI века базируются на концепции, согласно которой данные должны быть организованы в базы данных с целью адекватного отображения изменяющегося реального мира и удовлетворения информационных потребностей пользователей. Базы данных создаются и функционируют под управлением специальных программных комплексов, называемых системами управления базами данных (СУБД). Главная роль СУБД заключается в обеспечении пользователя инструментарием, позволяющим оперировать данными в абстрактных терминах, не связанных со способами их хранения в памяти.

Поддерживая целостную, централизованную структуру данных, информационные системы, основанные на СУБД, позволили избавиться от проблем избыточности и слабого контроля данных. Новые методы обращения к данным значительно упростили процесс

связывания элементов данных, что, в свою очередь, привело к расширению возможностей работы с данными. Основные характеристики систем управления базами данных упрощают процесс программирования и уменьшают необходимость дополнительной программной поддержки.

1. Основные проблемы выбора модели данных.

Современные универсальные системы управления базами данных, как правило, ориентируются на определенную модель данных – сетевую, иерархическую или реляционную. Это заставляет разработчиков либо искусственно подгонять свои потребности под возможности освоенной СУБД, либо осваивать новую с более подходящей моделью данных, что в общем случае, является довольно трудоемкой задачей [1].

Модели данных, как базовой составляющей любой информационной системы, уделяется особое внимание. Согласно Л. А. Калиниченко, в настоящее время модель данных трактуется «как совокупность методов и средств определения логической структуры базы данных и динамического состояния предметной области в базе данных» [2]. К. Дейт выделяет в модели данных три наиболее существенных компонента:

- совокупность средств определения допустимых структур данных;
- множество операций, применимых к допустимому состоянию базы данных для поиска или модификации данных;
- множества ограничений целостности, явно или неявно определяющих множество допустимых состояний базы данных [3].

Благодаря сочетанию указанных свойств, модель данных предоставляет средства описания данных, манипулирования и контроля целостности, выраженные в одном, или нескольких языках программирования баз данных. Задача определения функционально полного набора операций над типами данных требует тщательных исследований и не может быть признана решенной в настоящее время.

2. Цель исследований.

Статья посвящена исследованию общих принципов построения моделей данных, классов моделей данных, их применению и развитию в информационных системах. Предлагается расширить классификацию формально-логических моделей данных, определив новые подклассы для постреляционных моделей данных.

3. Классификация моделей данных.

Следуя предложениям Американского национального института стандартов [3], в настоящее время различают три уровня описания данных – концептуальный, внутренний и внешний.

Концептуальный уровень задает концептуальную модель, на основе которой для каждой базы данных создается конкретная концептуальная схема, содержащая общее описание базы данных, как ее представляет разработчик, учитывающий интересы всех пользователей. Внутренний уровень задается внутренней моделью и определяет взаимосвязь с операционной системой, физический уровень хранения и доступа к данным на внешних носителях. И, наконец, внешний уровень задается внешней моделью, на основе которой для каждой базы данных создается одна или несколько внешних схем (подсхем), описывающих определенные части базы данных, интересующие соответствующие группы пользователей. В [4] была предложена классификация моделей данных (рис.1).

Существует достаточно много моделей, которые нашли свое применение в теории и практике моделирования данных. Их разнообразие соответствует разнообразию областей применения и контингента пользователей.

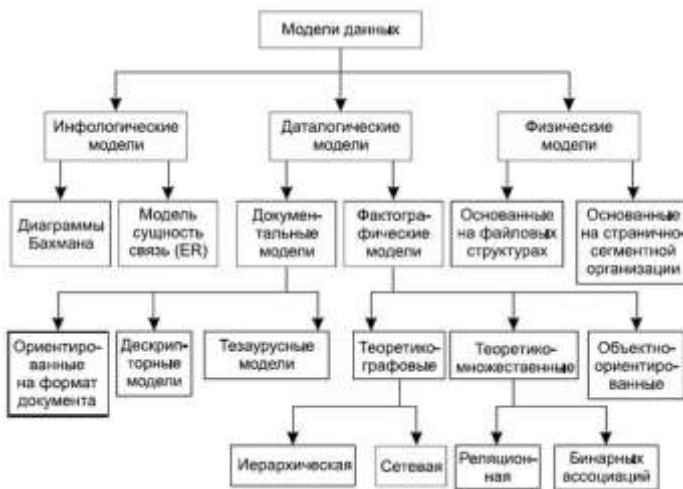


Рисунок 1. Общая классификация моделей

Различают теоретико-познавательный и прикладной характер моделирования данных. В первом случае, как правило, используется

традиционный аппарат построения логико-математических выражений. Здесь для спецификации данных, структур и ограничений целостности широко используются таблицы, граф-схемы, диаграммы. Во втором случае модель отображается на прикладном языке обработки данных, который является языком программирования БД. При этом структурная спецификация и операции по модификации структуры будут заданы языком определения данных (*Data Definition Language – DDL*), а операции над данными – языком манипулирования данными (*Data Manipulation Language – DML*). Что касается операций над данными, то они могут быть заданы как с помощью языка определения данных, так и с помощью языка манипулирования данными.

В связи с этим различают даталогические модели и инфологические. Первый класс моделей ориентирован на внутреннее представление, то есть, предназначены для непосредственной интерпретации концептуального представления в структуры хранения. Что касается второго класса – инфологических моделей, то это абстракции, дополненные аппаратом определения семантики, которые предназначены для использования на внешнем уровне представления данных [3].

Одним из классификационных признаков для моделей данных является мера типизации. В сильно типизированных моделях каждый концепт и, соответственно, каждое значение относится к определенной категории (домен, объект-сущность, объект-отношение и т.п.). В слабо типизированных моделях таких категорий может не существовать, или данные могут одновременно принадлежать нескольким категориям. К числу наиболее известных сильно типизированных моделей относятся реляционная модель данных и модель данных «сущность-связь».

Модели данных, поддерживаемые СУБД, могут быть разделены на несколько подклассов, различаемых по структурным характеристикам модели данных – способу представления объектов и их взаимосвязей в базе данных.

Определяющую роль в классификации играет понятийный базис модели данных, принятый в ней подход к представлению реальных сущностей, к категории объектов или связей между объектами.

Существенные различия между моделями данных проявляются в принятых методах определения семантики операторов манипулирования данными, в средствах обеспечивающих инвариантность заданных свойств типов объектов и типов связей между объектами в течение всего времени «жизни» базы данных. Рассмотрим общую классификацию моделей данных по их

структурным характеристикам, понимая, что модели данных сами по себе являются весьма сложными объектами и любая их классификация условна.

4. Исследование методов представления моделей данных.

Под моделью данных понимается способ абстракций, предназначенный для отображения структуры и семантики данных, который дает возможность реализовать интерпретацию данных в соответствии с заданными свойствами предметной области. В контексте систем баз данных модель данных – это способ формального представления данных на внешнем или концептуальном уровне. Необходимо отметить, что термин «модель данных» употребляется относительно логического представления данных. Что касается физического представления – используется термин «структура данных» [5].

Модели данных обычно представлены тремя основными функциональными компонентами: структурной компонентой, ограничениями целостности и операционной спецификацией.

Структурная часть определяет правила, в соответствии с которыми структурируются данные. Взаимосвязь элементов структуры может определяться их упорядоченностью или задаваться явно.

Формально структурную компоненту можно рассматривать как алгебраическую систему над множеством объектов разных типов с отношениями различной арности (степени), понимая под n -арным отношением между множествами $A_1 \dots A_n$ подмножество декартового произведения $A_1 \times \dots \times A_n$.

Взаимосвязь элементов структуры может быть определена следующим образом:

если $S = S_1 \dots S_k$ – это множество имен типов объектов (концептов), A_{S_i} – множество объектов типа S_i , тогда взаимосвязь, в предметной области между объектами типов S_1, \dots, S_n (не обязательно различных) может быть задана отношением $R(A_1, \dots, A_n)$. Последовательность S_1, \dots, S_n может представлять тип объекта-отношения.

Кроме того, прерогативой структурной части модели является установка принадлежности значений элементов данных доменам.

Компонента ограничений целостности позволяет специфицировать логические ограничения, накладываемые на данные. Формально ограничение целостности – это условие, которое для

данного множества объектов либо истинно, либо ложно. Заданные ограничения отвечают установленным свойствами предметной области. В качестве способов задания ограничений могут использоваться декларативные спецификации (например, объявление идентификатора объекта), формулы исчисления высказываний или исчисления предикатов первого порядка. Объекты данных, которые удовлетворяют заданным ограничениям, составляют допустимое состояние БД.

В операционной компоненте выделяются две составляющие: операции над данными и отображение моделей (в связи с этим иногда выделяют четыре компоненты моделей данных). Множество операций над данными определяет допустимые действия над данными, реализуя возможность получения новых множеств и отношений на основе существующих в базе данных.

Следует отметить, что операции над данными часто приводят к результату с типом, отличным от типов операндов. Таким образом, вся операция над данными сопровождается отображением моделей [5].

Как показано в [6, 7], объектно-реляционная модель данных (или обобщенная реляционная модель) отвечает всем требованиям к построению формальной модели данных. В связи с этим, предлагается расширить подкласс теоретико-множественных моделей путем объектно-реляционной модели данных. Так как термин «модель данных» употребляется относительно логического представления данных, то класс «физические модели» предлагается не включать в данную классификацию.

Таким образом, классификация моделей данных может быть представлена следующим образом:

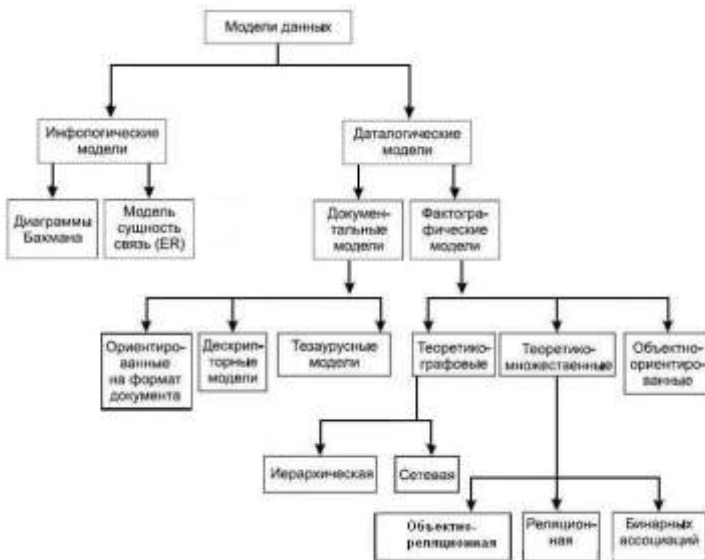


Рисунок 2. Классы моделей данных информационных систем

5. Выводы.

Современные СУБД представляют ограниченный набор моделей данных. Чаще всего используются наиболее универсальные, известные, получившие широкое применение модели данных. К таким можно отнести реляционную модель данных как базис различных СУБД.

Появление новых моделей или расширение существующих обусловлено тем, что правила предметной области требуют расширить функциональные или операционные возможности предоставляемой модели.

Путем синтеза различных моделей определялся целый ряд пост-реляционных моделей, которые являются предметом обсуждения многих исследователей. К таким моделям относятся объектно-ориентированные и объектно-реляционные модели. Согласно приведенным выше общим принципам построения моделей, объектно-реляционная модель является самостоятельной, независимой моделью данных, отвечает всем требованиям, предъявляемым к модели, и может быть выделена в отдельный подкласс, дополняя теоретико-множественные модели данных.

ЛИТЕРАТУРА:

1. Цикритзис, Д., Лоховски, Ф. Модели данных [Текст] : пер. с англ. М. : Финансы и статистика, 1985. – 344 с.
2. Калиниченко, Л. А. Методы и средства интеграции неоднородных баз данных [Текст]: монография / Л. А.Калиниченко. – М. : Наука, 1983. – 424 с.
3. Дейт, К. Введение в системы баз данных [Текст] : пер. с англ. – М. : Издательский дом «Вильямс», 2001. – 1072 с.
4. Карпова, Т.С. Базы данных: модели, разработка, реализация [Текст] : учебное пособие / Карпова Т.С. – СПб.: Питер, 2002. – 304 с.
5. Буслік М.М Моделі і структури даних [Текст] : Навч.посібник / М.М. Буслік. – Харків: ХТУРЕ, 2000. – 136 с.
6. Чапланова, Е.Б. Об одном подходе к построению объектно-реляционной модели данных [Текст] / Н. В. Касаткина, С. С. Танянский, Е. Б. Чапланова // Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. – К. : ВІКНУ, 2009. – випуск 20. – С. 141–146.
7. Касаткина, Н. В. Методы хранения и обработки нечетких данных в среде реляционных систем [Текст] / Н. В. Касаткина, С. С. Танянский, В. А. Филатов // Автоматика. Автоматизація. Електротехнічні комплекси та системи. – Херсон : ХНТУ, 2009. – вип. 2 (24) . – С. 84–90.