

ЛЕКЦИЯ 4. РАЗЛИЧНЫЕ ПРЕДСТАВЛЕНИЯ О ДАННЫХ В БАЗАХ ДАННЫХ. ОСНОВНЫЕ ЭТАПЫ ПРОЕКТИРОВАНИЯ БАЗ ДАННЫХ

В лекции рассматриваются различные представления о данных в базах данных. Описываются модели данных (внешнее представление, концептуальная модель, структура хранения) и основные этапы проектирования базы данных. Рассматривается жизненный цикл проектирования базы данных.

Ключевые термины: концептуальные требования пользователя, концептуальная модель, модель данных СУБД, логическая модель, обобщенное представление о данных, внешнее представление, трехуровневая архитектура базы данных, проектирование баз данных, основные этапы проектирования базы данных.

Цель лекции: Показать существование различных представлений о данных (различных моделей) у разных групп лиц, работающих с данными. Рассмотреть отражение этих представлений в трехуровневой архитектуре базы данных (внешний уровень, концептуальный уровень, внутренний уровень), сформулировать достоинство трехуровневой архитектуры. Выделить основные этапы проектирования базы данных как процесса построения вышеуказанных моделей.

4.1. Различные представления о данных в базах данных

Создание базы данных предполагает интеграцию данных, предназначенных для решения нескольких прикладных задач разных пользователей. Соответственно, при интеграции данных должны учитываться требования к данным каждого пользователя, основанные на его представлении о данных и связях между ними. Далее эти требования должны обобщаться в единое представление, которое и будет служить основой для построения единой базы данных (рис. 4.1).

Обобщение представлений всех пользователей о данных называется концептуальной моделью (схемой) БД. Концептуальная модель представляет информационное описание предметной области с учетом логических взаимосвязей, поэтому её еще называют инфологической (информационно-логической) моделью. В модели отсутствуют какие-либо понятия, связанные с ЭВМ, памятью ЭВМ, способами размещения данных в памяти ЭВМ, и, по сути, это модель только предметной области.

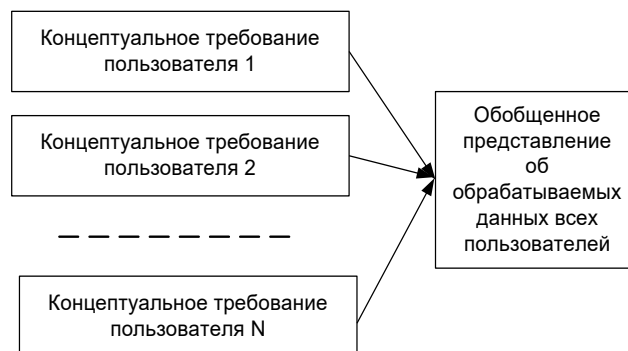


Рис. 4.1. Обобщение представления пользователей о данных

Как уже отмечалось, для создания базы данных и работы с ней используется система управления базами данных. *Каждая конкретная СУБД поддерживает определенный вид данных (форматов записей и отношений), называемый моделью данных СУБД.*

Следующий этап разработки базы данных предполагает выбор представления концептуальной модели с помощью модели данных конкретной СУБД. *Полученное таким образом представление концептуальной модели называется логической моделью БД. Или другими словами, логическая модель – это концептуальная схема, специфицированная в языке конкретной СУБД.* Логическая модель представляет данные и элементы данных вне зависимости от их содержания и среды хранения. Далее разработчик системы средствами СУБД отображает полученную логическую модель БД в память ЭВМ и определяет методы доступа. Полученное представление данных в памяти ЭВМ называется внутренним представлением или структурой хранения. Прикладные программы работают с логической моделью, причем каждому пользователю представляется подмножество этой логической модели (подсхема), отражающее его представление о предметной области. Каждая прикладная программа «видит» и обрабатывает только те данные, которые необходимы именно ей.

Соответствующее «видение» данных прикладными программами (пользователями) представляет собой внешние представления. Взаимосвязь вышеуказанных моделей изображена на рис.4.2.

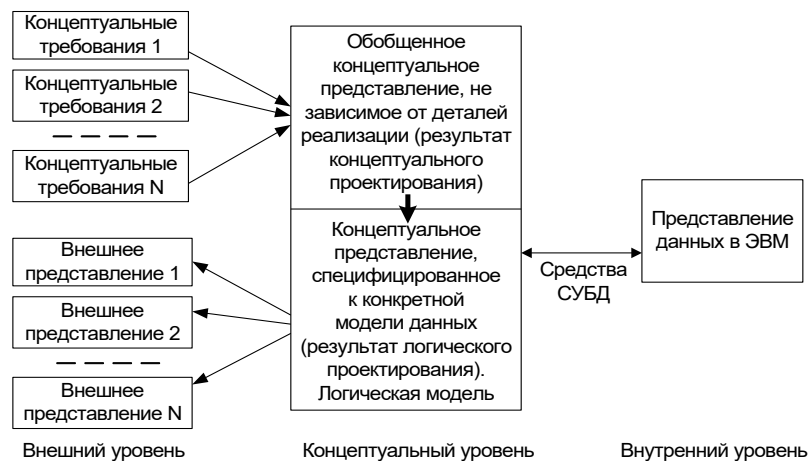


Рис. 4.2. Различные представления о данных в БД

На данной схеме выделены три различных уровня описания данных (внешний, концептуальный, внутренний). *Эти уровни формируют так называемую трехуровневую архитектуру ANSI/SPARC*, предложенную в 1975 г. Комитетом планирования стандартов и норм SPARC (Standards Planning and Requirements Committee) Национального института стандартизации США (American National Standards Institute – ANSI). Основная цель этой архитектуры состоит в отделении пользовательского представления о данных в базе данных от их физического представления. Использование таких представлений о данных позволяет обеспечить выполнение основного требования к БД – независимости программ и данных. При изменении прикладных программ может измениться соответствующее внешнее представление, но логическая модель данных не изменяется и, соответственно, не будут изменяться другие прикладные программы. При изменении внутреннего представления (структур хранения) логическая модель не изменяется, соответственно, не изменяются прикладные программы.

Использование соответствующих представлений также позволяет четко разграничить полномочия различных лиц, работающих с базой данных.

Соответствующие представления позволяют описать «видение» базы данных разными лицами, работающими с ней:

- внешнее представление – представление специалиста предметной области (пользователя);
- внешнее представление и логическая модель – представление прикладного программиста, разрабатывающего конкретное приложение для пользователя;
- логическая модель и внутреннее представление – представление системного программиста, администрирующего базу данных.

4.2. Основные этапы проектирования базы данных

Проектирование данных (базы данных) представляет собой процесс последовательного отображения исследуемых явлений реального мира в виде данных в памяти ЭВМ (рис. 4.3.).



Рис. 4.3. Общая схема проектирования

Конкретные явления реального мира, представляющие интерес для проводимого исследования, будем называть предметной областью.

Проектирование (моделирование) базы данных представляет собой многоэтапный процесс. Основные этапы этого процесса приведены на рис. 4.4.).

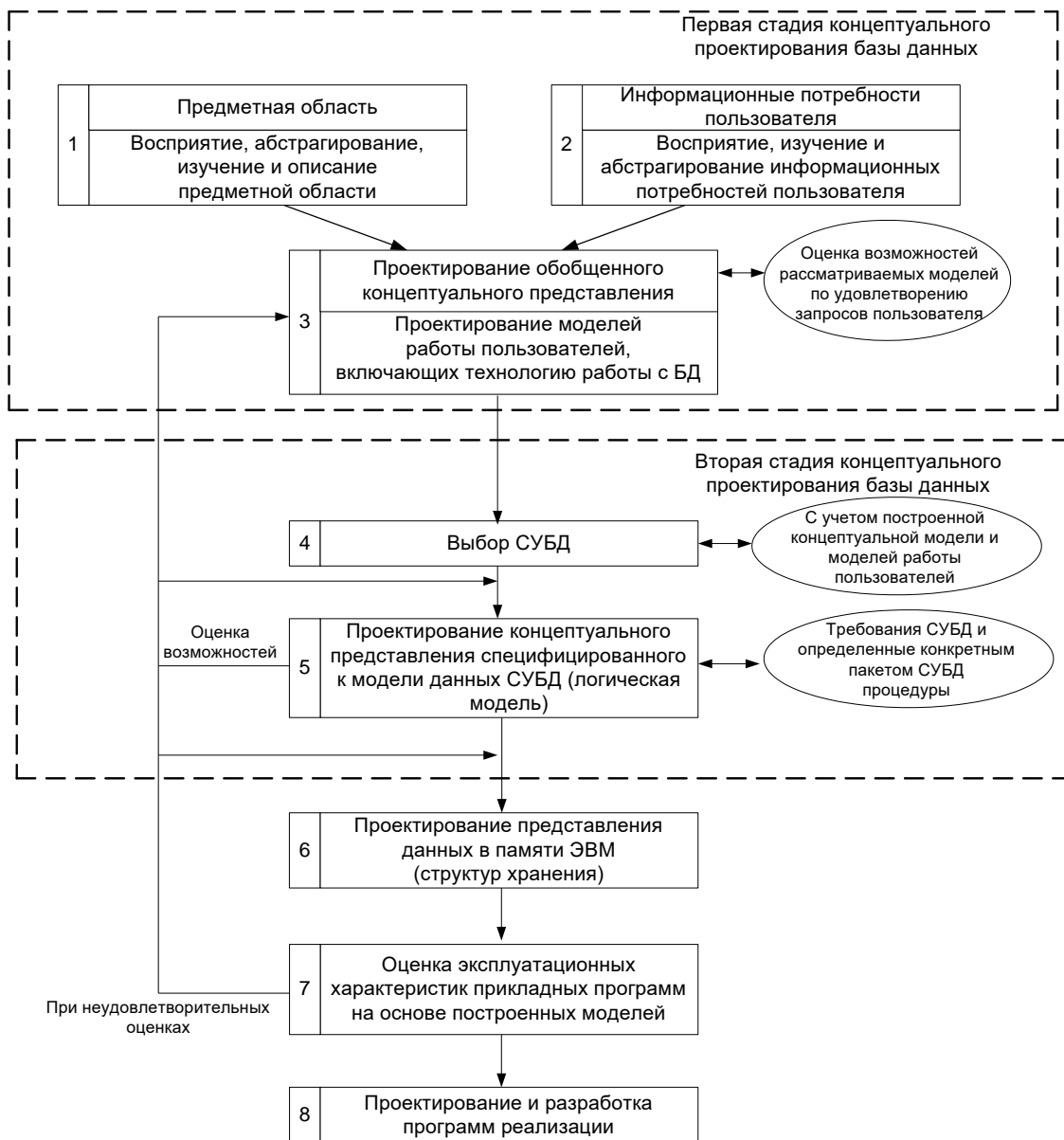


Рис. 4.4. Этапы проектирования базы данных

Подробно действия, отраженные на приведенном рисунке, будут рассмотрены в следующих лекциях. Здесь дадим лишь краткие комментарии к соответствующим блокам.

В блоках 1,2 необходимо особое внимание обратить на слово «абстрагирование». Имеется ввиду, что проектирование базы данных нужно вести не под конкретный документ, обрабатываемый пользователем, и не под конкретные действия пользователя с этим документом, а под обобщенный (абстрактный) образ документов и обобщенные (абстрактные) действия пользователей. Например, рассматривать документ не с конкретными числами строк и столбцов, а с абстрактными числами n и m ; вместо требуемого пользователем поиска по конкретному полю (например, фамилии) рассматривать поиск по любому полю и т. д. Это очень важно, так как конкретные формы документов и действия пользователей при работе с ними достаточно часто изменяются. В этом случае при проектировании базы данных под конкретные формы документов и конкретные

действия придется перепроектировать базу данных. что связано с существенными временными и стоимостными затратами.

Очень важным является выбор СУБД (блок 4), от которого в значительной степени зависит работоспособность построенной базы данных. Проблема выбора СУБД уже обсуждалась в лекции 3. Заметим здесь, что выбор СУБД зависит от количества форм документов, от сложности связей между данными, от объема обрабатываемых данных, от количества пользователей, работающих с БД и т. д.

Ранее отмечалось, что отображение логической модели базы данных в структуру хранения (представление данных в памяти компьютера) осуществляется системой управления базой данных. Тем не менее, во многих СУБД для повышения эффективности функционирования базы данных представляется возможность выбора ряда параметров, управляющих представлением данных в памяти компьютера. Выбор таких параметров и подразумевается в блоке 6.

Заметим, что очень важно при проектировании базы данных делать оценки ее возможной работоспособности. Так, по завершении проектирования обобщенного концептуального представления нужно попытаться оценить необходимое число производимых операций с элементами моделей при реализации возможных запросов пользователей. При невозможности в рамках построенной модели ответить на какой-то запрос пользователя или при значительном числе производимых при этом операций (что приведет к невозможности реализации соответствующего запроса в реальном масштабе времени) необходим возврат по схеме рис. 4.4. на шаг назад (построение более эффективного обобщенного концептуального представления). Аналогичные оценки необходимо делать и при завершении других этапов проектирования (блоки 5, 7). При этом возможен возврат назад на один или несколько шагов. Так, например, при проектировании логической модели (блок 5) не удастся достичь адекватного представления концептуальной модели средствами модели данных СУБД. В этом случае необходимо либо вернуться на шаг назад и выбрать другую СУБД, либо вернуться к блоку 3 и изменить вид концептуальной модели. Так же, если полученные при реализации блока 7 оценки эксплуатационных характеристик не отвечают требованиям пользователя, возможны пересмотры всех ранее полученных решений (блоки 7, 6, 5, 4, 3). Кроме этого, необходим возврат на проектирование обобщенного концептуального представления при изменении внешних требований пользователей, а также при выявленных ошибках проектирования.

Краткие итоги: Рассмотрены различные представления о данных в базах данных - модели обрабатываемых данных (внешнее представление, концептуальная модель, структура хранения). Представлено отражение этих представлений в трехуровневой архитектуре базы данных (внешний уровень, концептуальный уровень, внутренний уровень), сформулировано достоинство трехуровневой архитектуры. Описаны основные этапы проектирования базы данных как процесса построения вышеуказанных моделей и жизненный цикл проектирования базы данных (создание, апробация, исправление ошибок и улучшение характеристик, опытная эксплуатация).