



УНИВЕРСИТЕТ
СИНЕРГИЯ

КАФЕДРА ИНФОРМАЦИОННЫХ СИСТЕМ

Архитектура информационных систем

Программа подготовки бакалавров по направлению
«Информационные системы и технологии»

Глушков Сергей Владимирович

Доцент, к.в.н., доцент

Тема 1. Введение в теорию баз данных

УЧЕБНЫЕ ВОПРОСЫ ТЕМЫ

1.1 История создания.

1.2 Основные понятия банков данных и знаний, информация и данные. Базы данных (БД). Принципы построения

1.3 Система управления базами данных

1.4 Централизованное и децентрализованное управление данными; достоинства и недостатки

1.5 Функции СУБД

1.1. История создания

Информационная система

представляет собой программно-аппаратный комплекс, обеспечивающий выполнение следующих функций:

1. надежное хранение информации в памяти компьютера;
2. выполнение специфических для данного приложения преобразований информации и вычислений;
3. предоставление пользователям удобного и легко осваиваемого интерфейса.

информационная система

**СОЗДАННАЯ С ЗАДАНЫМИ ЦЕЛЬЮ
И ФУНКЦИОНАЛЬНЫМИ ЗАДАЧАМИ**

ИМЕЯ В СВОЕМ СОСТАВЕ

ДОКУМЕНТЫ (МАССИВЫ ДОКУМЕНТОВ)

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Обработки данных и
документов

Управления

Автоматизации работы
должностных лиц

Поддержки принятия
решений

Экспертных систем

реализует

ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ,

каждый из которых может включать:

- Ввод информации из внешних и внутренних источников;
- Обработку входной информации и представление ее в удобном виде;
- Вывод информации для представления потребителям или передачи в другую систему;
- Обратную связь, т.е. коррекцию входной информации при использовании информации, переработанной должностными лицами (управленческим персоналом) или другой информационной системой.

Файлы и файловые системы

С точки зрения прикладной программы, файл — это именованная область внешней памяти, в которую можно записывать и из которой можно считывать данные

НОВЫЙ подход к управлению информацией реализован в рамках новых программных систем, названных впоследствии **Системами Управления Базами Данных (СУБД)**, а сами хранилища информации, которые работали под управлением данных систем, назывались базами или банками данных (БД и БнД).

Первый этап - базы данных на больших ЭВМ

В 1968 году была введена в эксплуатацию первая промышленная СУБД система IMS фирмы IBM.

В 1975 году появился первый стандарт ассоциации по языкам систем обработки данных — Conference of Data System Languages (CODASYL), который определил ряд фундаментальных понятий в теории систем баз данных, которые и до сих пор являются основополагающими для сетевой модели данных.

В дальнейшее развитие теории баз данных большой вклад был сделан американским математиком Э. Ф. Коддом, который является создателем реляционной модели данных. В 1981 году Э. Ф. Кодд получил за создание реляционной модели и реляционной алгебры престижную премию Тьюринга Американской ассоциации по вычислительной технике.

Первый этап развития СУБД связан с организацией баз данных на больших машинах типа IBM 360/370, ЕС-ЭВМ и мини-ЭВМ типа PDP11 (фирмы Digital Equipment Corporation — DEC), разных моделях HP (фирмы Hewlett Packard).

Особенности этого этапа развития выражаются в следующем:

1. Все СУБД базируются на мощных мультипрограммных операционных системах (MVS, SVM, RTE, OSRV, RSX, UNIX), поэтому в основном поддерживается работа с централизованной базой данных в режиме распределенного доступа.
2. Функции управления распределением ресурсов в основном осуществляются операционной системой (ОС).

Особенности этого этапа развития выражаются в следующем:

3. Поддерживаются языки низкого уровня манипулирования данными, ориентированные на навигационные методы доступа к данным.

4. Значительная роль отводится администрированию данных.

5. Проводятся серьезные работы по обоснованию и формализации реляционной модели данных, и была создана первая система (System R), реализующая идеологию реляционной модели данных.

Особенности этого этапа развития выражаются в следующем:

6. Проводятся теоретические работы по оптимизации запросов и управлению распределенным доступом к централизованной БД, было введено понятие транзакции.
7. Результаты научных исследований открыто обсуждаются в печати, идет мощный поток общедоступных публикаций, касающихся всех аспектов теории и практики баз данных, и результаты теоретических исследований активно внедряются в коммерческие СУБД.

Второй этап

Эпоха персональных
компьютеров

1. Все СУБД были рассчитаны на создание БД в основном с монопольным доступом
 2. Большинство СУБД имели развитый и удобный пользовательский интерфейс.
 3. В настольных СУБД отсутствовали средства поддержки ссылочной и структурной целостности базы данных.
- Наличие монопольного режима работы фактически привело к вырождению функций администрирования БД и в связи с этим — к отсутствию инструментальных средств администрирования БД.

4. И, наконец, последняя и в настоящий момент весьма положительная особенность — это сравнительно скромные требования к аппаратному обеспечению со стороны настольных СУБД. Вполне работоспособные приложения, разработанные, например, на Clipper, работали на PC 286.

Яркие представители этого семейства — очень широко использовавшиеся до недавнего времени СУБД Dbase (DbaseIII+, DbaseIV), FoxPro, Clipper, Paradox.

Третий этап

Распределенные базы данных

1. Практически все современные СУБД обеспечивают поддержку полной реляционной модели
2. Большинство современных СУБД рассчитаны на многоплатформенную архитектуру, развитие средств администрирования БД с реализацией общей концепции средств защиты данных.
3. все современные СУБД имеют средства подключения клиентских приложений, разработанных с использованием настольных СУБД, и средства экспорта данных из форматов настольных СУБД второго этапа развития.

Именно к этому этапу можно отнести разработку ряда стандартов в рамках языков описания и манипулирования данными начиная с SQL89, SQL92, SQL99 и технологий по обмену данными между различными СУБД, к которым можно отнести и протокол ODBC (Open DataBase Connectivity), предложенный фирмой Microsoft.

Именно к этому этапу можно отнести начало работ, связанных с концепцией объектно-ориентированных БД — ООБД. Представителями СУБД, относящимся к второму этапу, можно считать MS Access 97 и все современные серверы баз данных Oracle 7.3, Oracle 8.4 MS SQL 6.5, MS SQL 7.0, System 10, System 11, Informix, DB2, SQL Base и другие современные серверы баз данных, которых в настоящий момент насчитывается несколько десятков.

Перспективы развития систем управления базами данных

новая технология доступа к данным — *интранет*.
Основное отличие этого подхода от технологии клиент-сервер состоит в том, что отпадает необходимость использования специализированного клиентского программного обеспечения

1.2 Основные понятия банков данных и знаний, информация и данные. Базы данных (БД). Принципы построения

Банк данных (БнД) — это система
специальным образом
организованных данных — баз данных,
программных, технических, языковых,
организационно-методических
средств, предназначенных для
обеспечения централизованного
накопления и коллективного
многоцелевого использования данных.



Базы данных (БД). Принципы построения

Терминология, используемая в теории баз данных, на стадии проектирования и практической работы с базами данных.

Сведения о базах данных, как важнейшем компоненте информационных систем.

Общие принципы проектирования информационных систем.

Характеристика баз данных.

База данных (БД) — именованная совокупность данных, отражающая состояние объектов и их отношений в рассматриваемой предметной области.

Система управления базами данных
(СУБД) — совокупность языковых и
программных средств,
предназначенных для создания,
ведения и совместного использования
БД многими пользователями

Архитектура базы данных.



Трехуровневая модель системы управления базой данных, предложенная ANSI

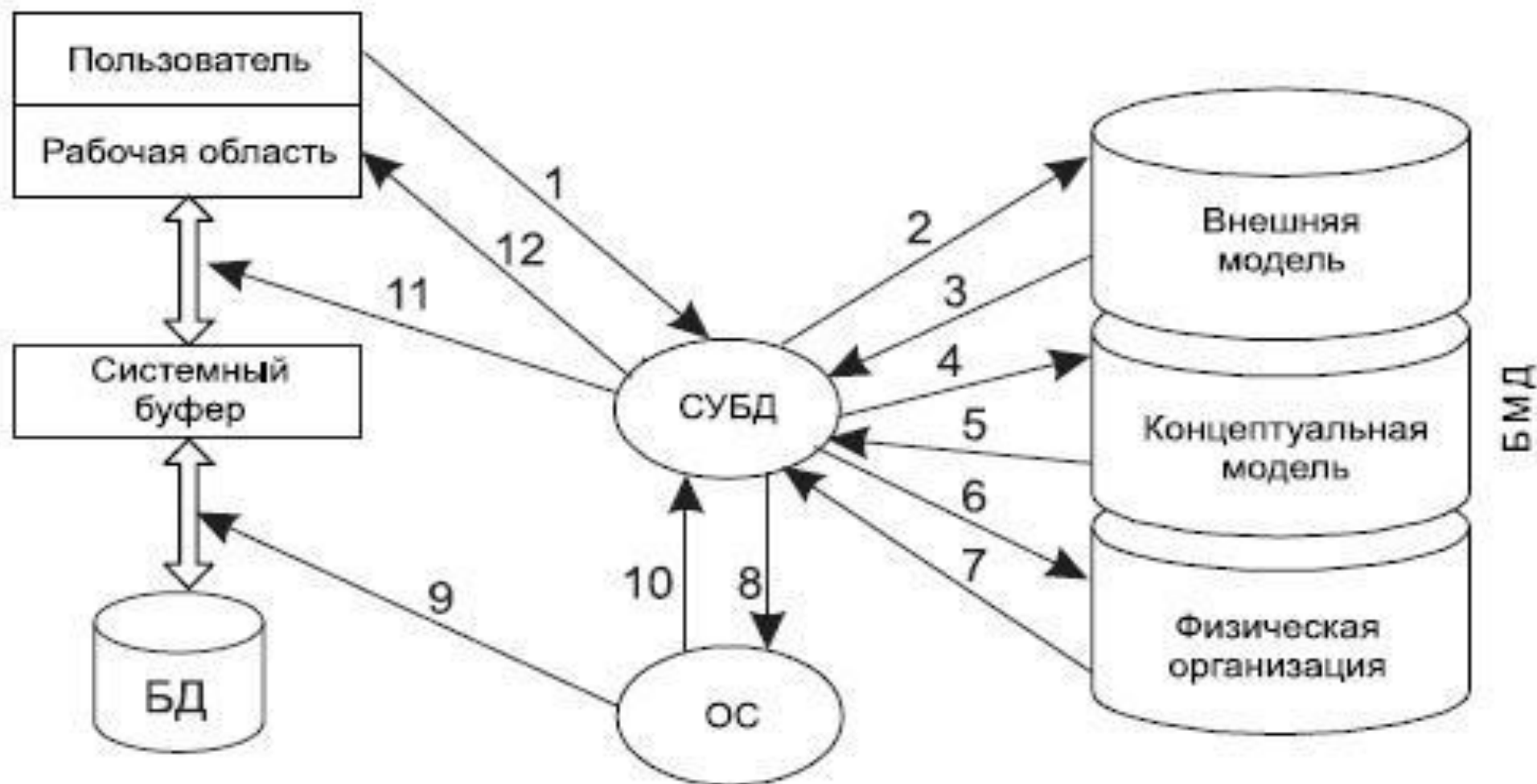


Схема прохождения запроса к БД



Как любой программно-
организационно-технический комплекс,
банк данных существует во времени и
в пространстве. Он имеет
определенные стадии своего развития:

1. Проектирование.
2. Реализация.
3. Эксплуатация.
4. Модернизация и развитие.
5. Полная реорганизация

основные категории пользователей

Конечные пользователи

**Администраторы банка
данных**

*Разработчики и
администраторы
приложений*



Классификация моделей данных

Понятие "**данные**" в концепции баз данных — это набор конкретных значений, параметров, характеризующих объект, условие, ситуацию или любые другие факторы. Данные не обладают определенной структурой, данные становятся информацией тогда, когда пользователь задает им определенную структуру, то есть осознает их смысловое содержание. Поэтому центральным понятием в области баз данных является понятие модели. Не существует однозначного определения этого термина, у разных авторов эта абстракция определяется с некоторыми различиями, но тем не менее можно выделить нечто общее в этих определениях.

Модель данных - это некоторая абстракция, которая, будучи приложенной к конкретным данным, позволяет пользователям и разработчикам трактовать их уже как информацию, то есть сведения, содержащие не только данные, но и взаимосвязь между ними.



Иерархическая модель данных

Иерархическая модель данных является наиболее простой среди всех даталогических моделей. Исторически она появилась первой среди всех даталогических моделей: именно эту модель поддерживает первая из зарегистрированных промышленных СУБД IMS фирмы IBM.

Основными информационными единицами в иерархической модели являются: *база данных (БД), сегмент и поле.*

Поле данных -минимальная, неделимая единица данных, доступная пользователю с помощью СУБД.

Сегмент в терминологии Американской Ассоциации по базам данных DBTG (Data Base Task Group) называется *записью*, при этом в рамках иерархической модели определяются два понятия:

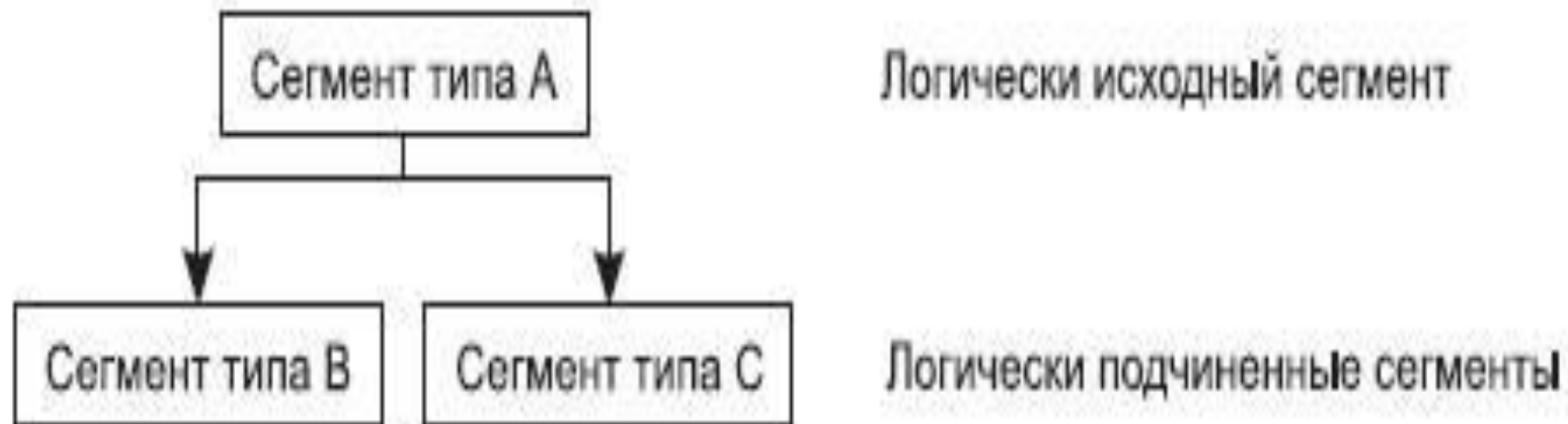
тип сегмента или *тип записи*

экземпляр сегмента или *экземпляр записи*.

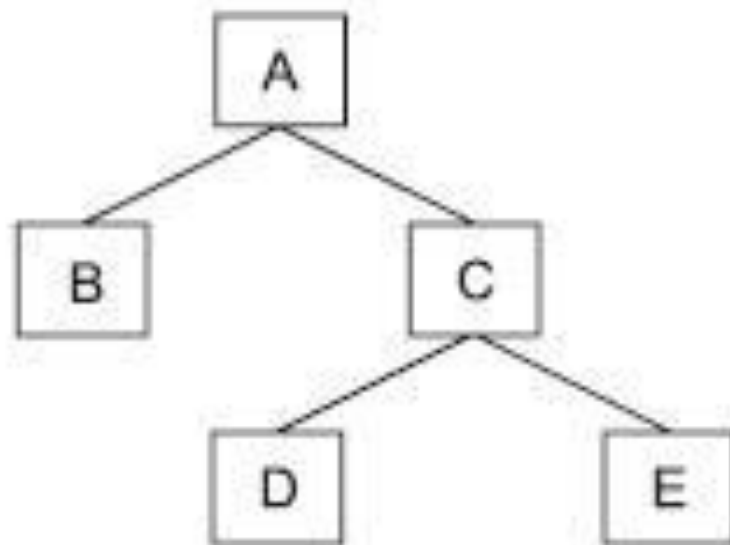
Тип сегмента — это поименованная
совокупность типов элементов данных,
в него входящих

Экземпляр сегмента образуется
из конкретных значений полей или
элементов данных, в него
входящих

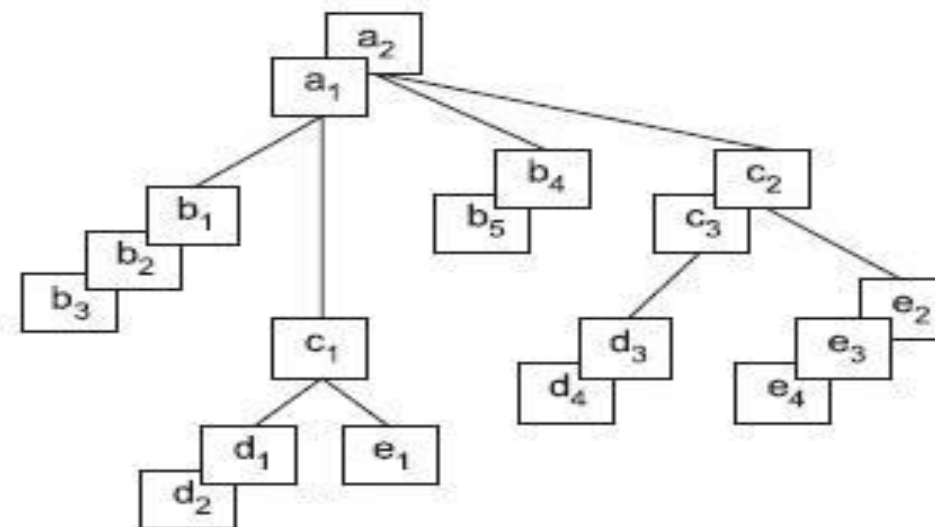
Ключом называется набор элементов данных,
однозначно идентифицирующих экземпляр сегмента



Пример иерархических связей между
сегментами



Пример структуры иерархического дерева



Пример двух экземпляров данного

Язык описания данных иерархической модели

Выделяют:

- языковые средства описания данных
(DDL, Data Definition Language)
- средства манипулирования данными
(DML, Data Manipulation Language).

Каждая физическая база описывается набором операторов, определяющих как ее логическую структуру, так и структуру хранения БД. Описание начинается с оператора DBD (Data Base Definition):

DBD Name = < имя БД>, ACCESS = < способ доступа>

Определено 5 способов доступа:

HSAM — *hierarchical sequential access method* (иерархически последовательный метод),

HISAM — *hierarchical index sequential access method* (иерархически индексно-последовательный метод),

HDAM — *hierarchical direct access method* (иерархически прямой метод),

HIDAM — *hierarchical index direct access method* (иерархически индексно-прямой метод),

INDEX — индексный метод.

Далее идет описание наборов данных, предназначенных для хранения
БД

В системе может быть несколько физических БД (ФБД), но каждая из них описывается отдельно своим DDD и ей присваивается уникальное имя. Каждая ФБД содержит только один корневой сегмент. Совокупность ФБД образует концептуальную модель данных.

Представление внешней модели называется логической базой данных и определяется совокупностью блоков связи данного приложения с физическими БД, входящими в концептуальную схему БД. Блок связи — *PCB, program communication block* — описывает связь с одной физической БД

Совокупность блоков РСВ образует
полное внешнее представление
данного приложения, называемое
"блоком спецификации программ" (
PSB, program specification block).

Сетевая модель данных

Стандарт сетевой модели впервые был определен в 1975 году организацией CODASYL (Conference of Data System Languages), которая определила базовые понятия модели и формальный язык описания.

Базовые объекты модели :

- элемент данных;
- агрегат данных
(типа *вектор* и агрегат типа *повторяющаяся группа*);
- запись;
- набор данных.

Агрегат типа вектор соответствует линейному набору элементов данных.

*Например, агрегат **Адрес** может быть представлен следующим образом:*

Адрес			
Город	Улица	дом	квартира

Агрегат типа повторяющаяся группа соответствует совокупности векторов данных.

Например, агрегат **Зарплата** соответствует типу повторяющаяся группа с числом повторений 12.

Зарплата	
Месяц	Сумма

Записью называется совокупность агрегатов или элементов данных, моделирующая некоторый класс объектов реального мира.

Понятие записи соответствует понятию "сегмент" в иерархической модели. Для записи, так же как и для сегмента, вводятся понятия *типа записи* и *экземпляра записи*.

Набор

Набором называется двухуровневый граф, связывающий отношением "один-ко-многим" два типа записи.



Язык описания данных в сетевой модели

описание базы данных — области
размещения;

описания записей — элементов и агрегатов
(каждого в отдельности);

описания наборов (каждого в отдельности).

Внешняя модель при сетевой организации данных поддерживается путем описания части общего связного графа.

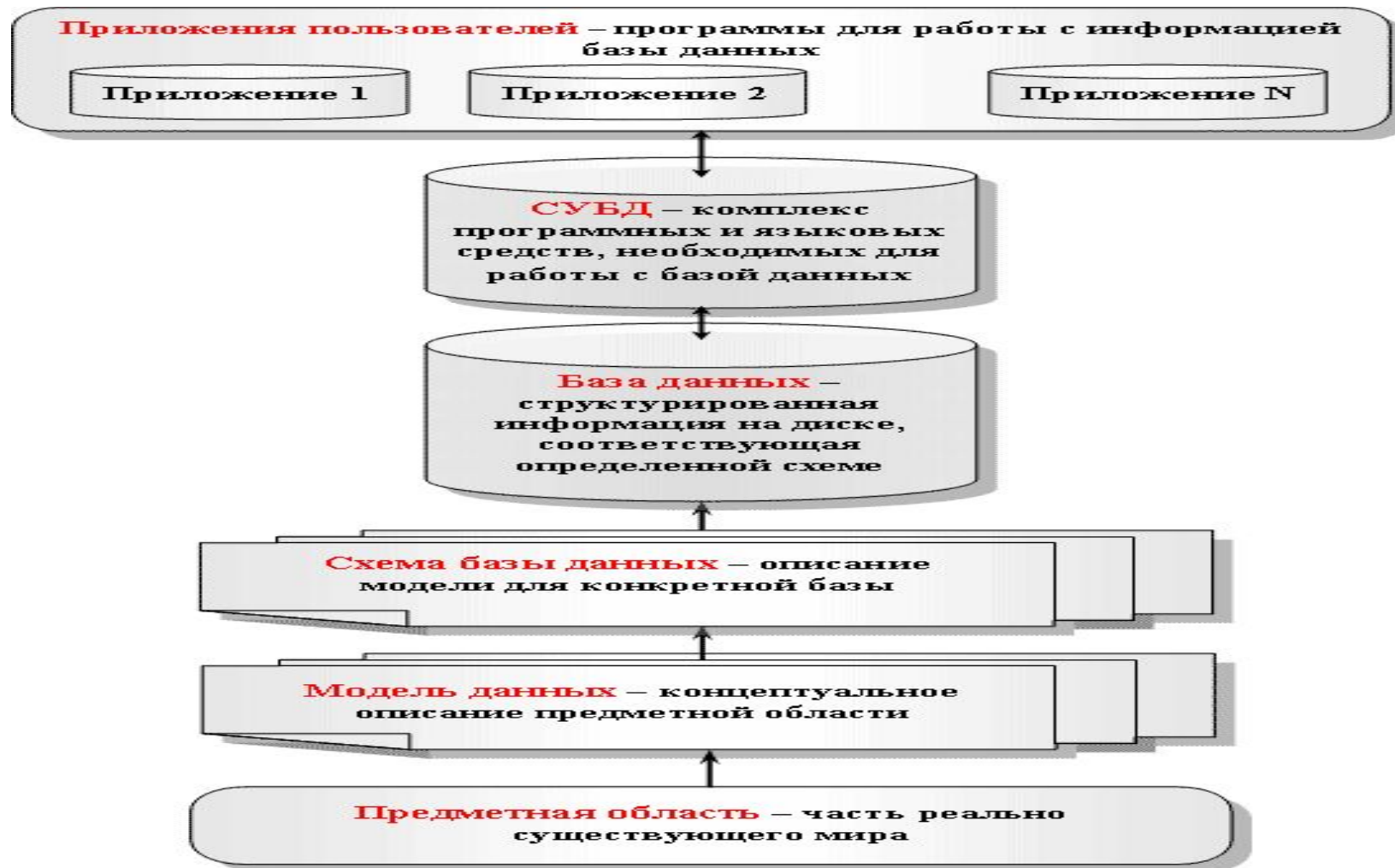
Язык манипулирования данными в сетевой модели

Все операции манипулирования данными в сетевой модели делятся на *навигационные операции* и *операции модификации*.

Навигационные операции осуществляют перемещение по БД путем прохождения по связям, которые поддерживаются в схеме БД. В этом случае результатом является новый единичный объект, который получает статус *текущего объекта*.

Операции модификации осуществляют как добавление новых экземпляров отдельных типов записей, так и экземпляров новых наборов, удаление экземпляров записей и наборов, модификацию отдельных составляющих внутри конкретных экземпляров записей

Терминология



Реляционные БД

