

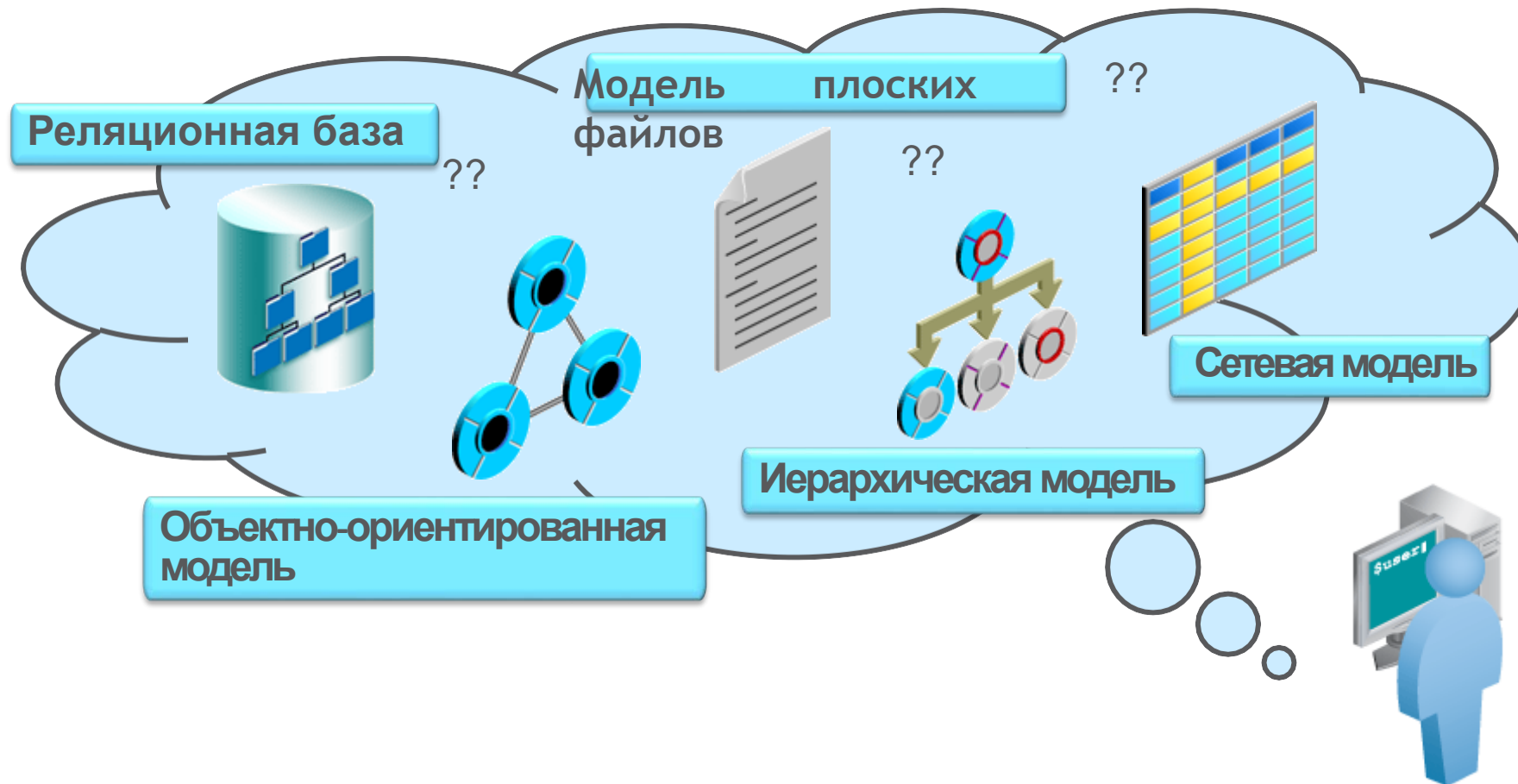
Практическое занятие №2

Типы моделей баз данных.

Нормализация отношений: приведение к первой нормальной форме, определение функциональных зависимостей.



Типы моделей баз данных



Пример модели плоского файла

- Сведения о книгах и авторах хранятся в одной таблице, это приводит к повторению значений данных.

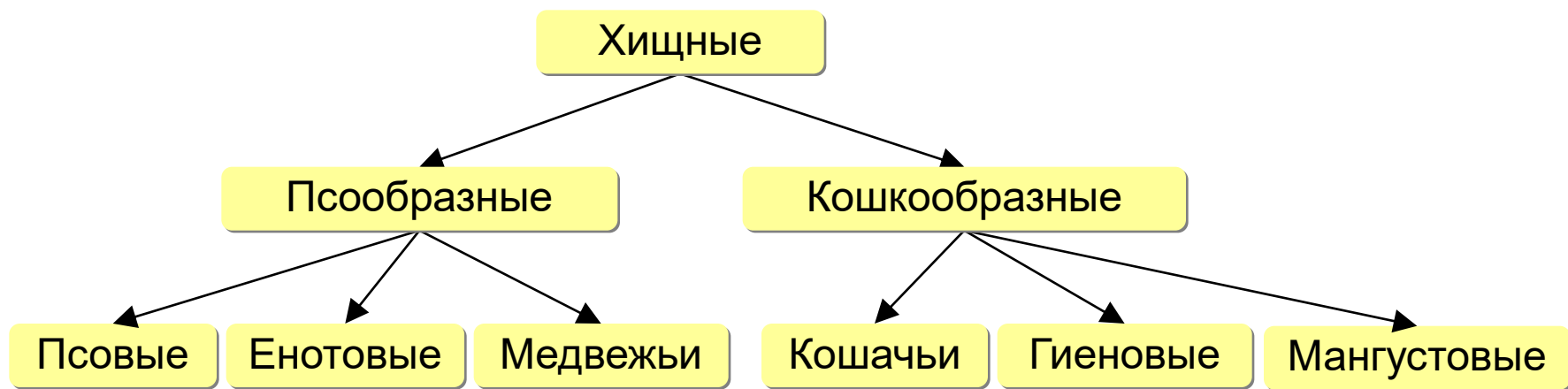
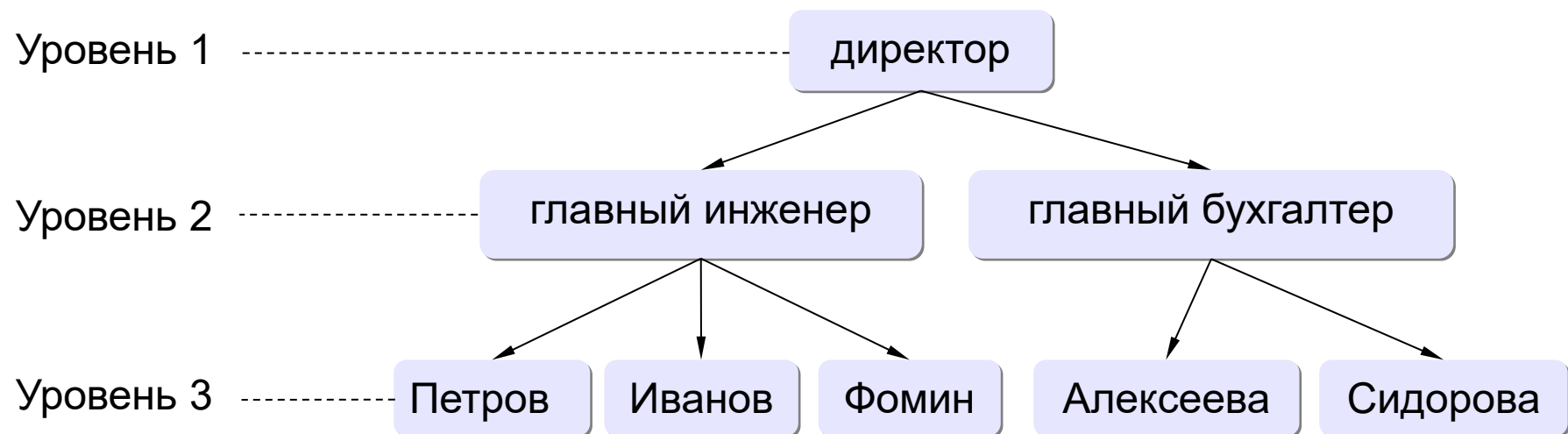
	AUTHOR_ID	AUTHOR_NAME	TITLE
Record 1	AD0001	Oscar Wilde	A Vision
Record 2	AD0002	Leo Tolstoy	War and Peace
Record 3	AD0003	Oliver Goldsmith	Citizen of the World
Record 4	AD0003	Oliver Goldsmith	The Deserted Village

Иерархическая модель

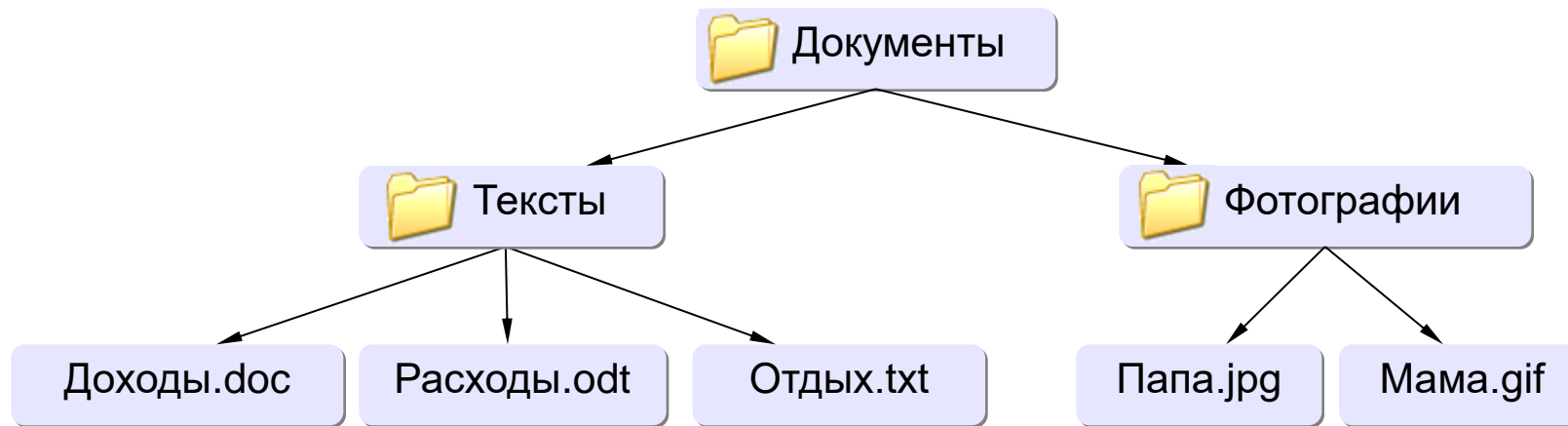
Иерархическая модель данных — представление базы данных в виде древовидной (иерархической) структуры, состоящей из объектов (данных) различных уровней.



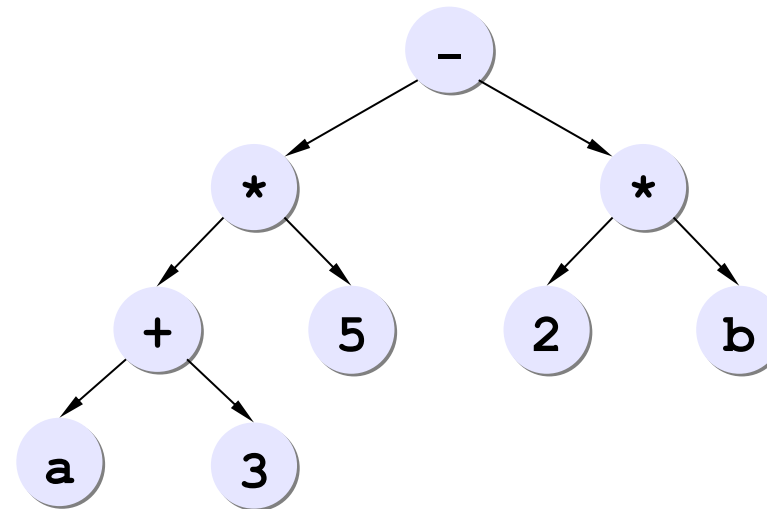
Иерархические модели



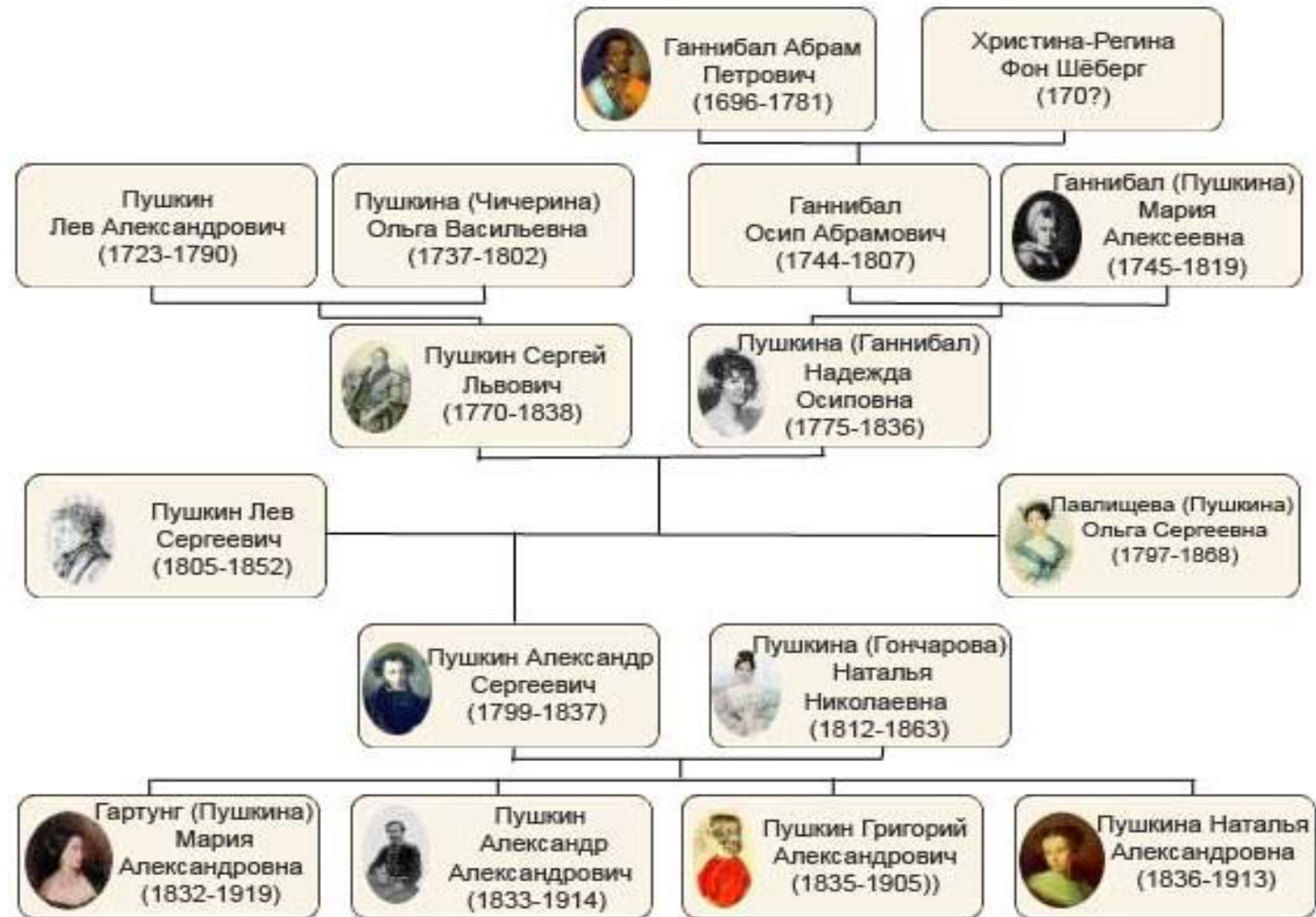
Иерархические модели

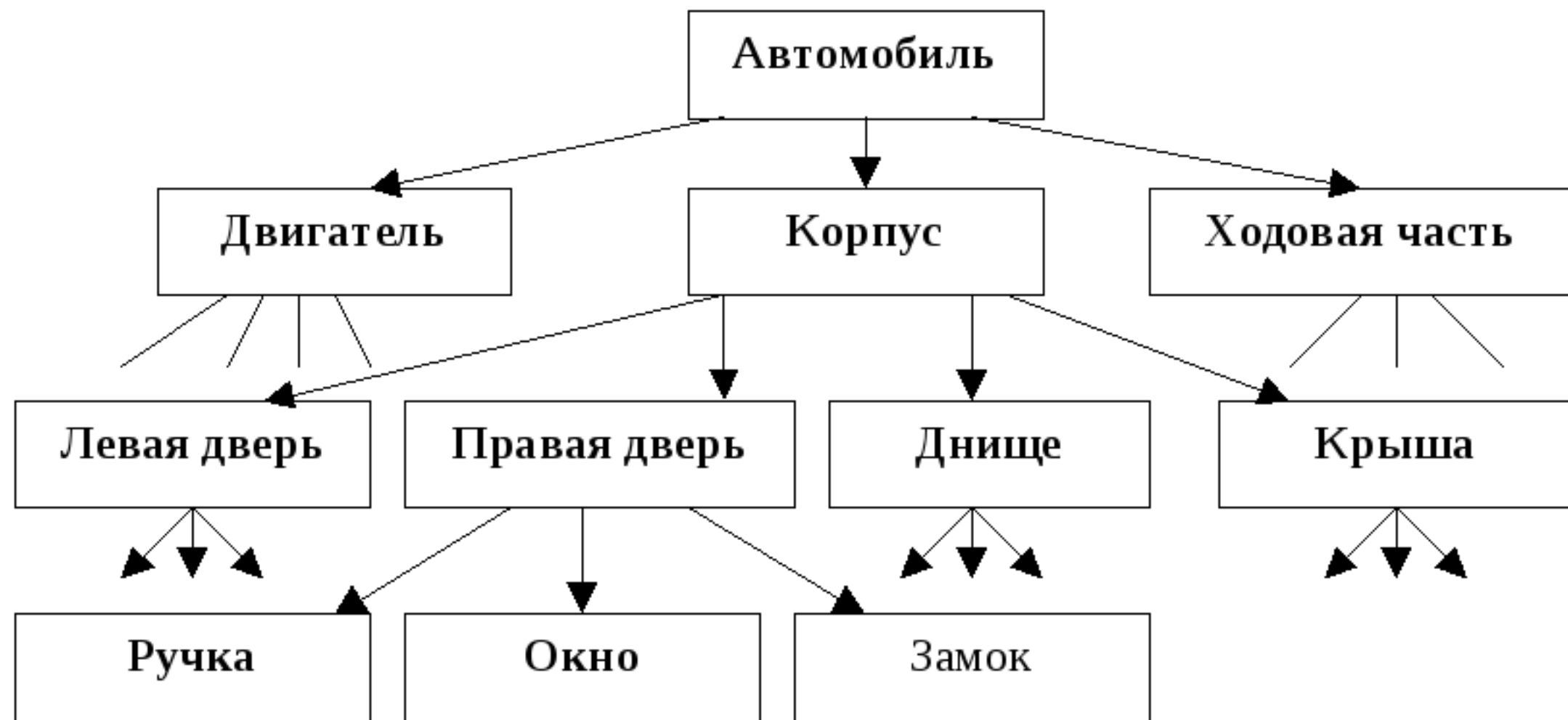


$(a+3) * 5 - 2 * b$



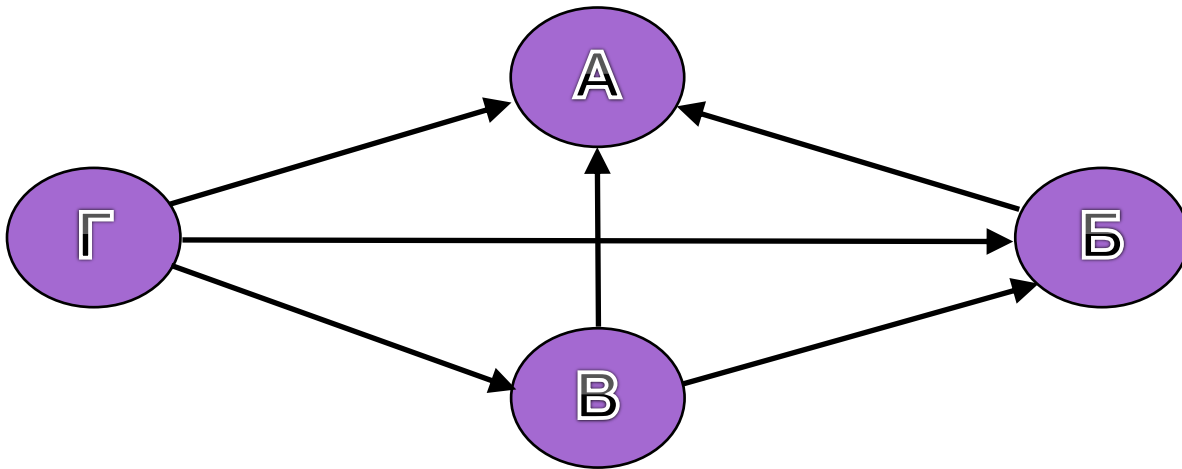
Генеалогическое дерево





Сетевая модель

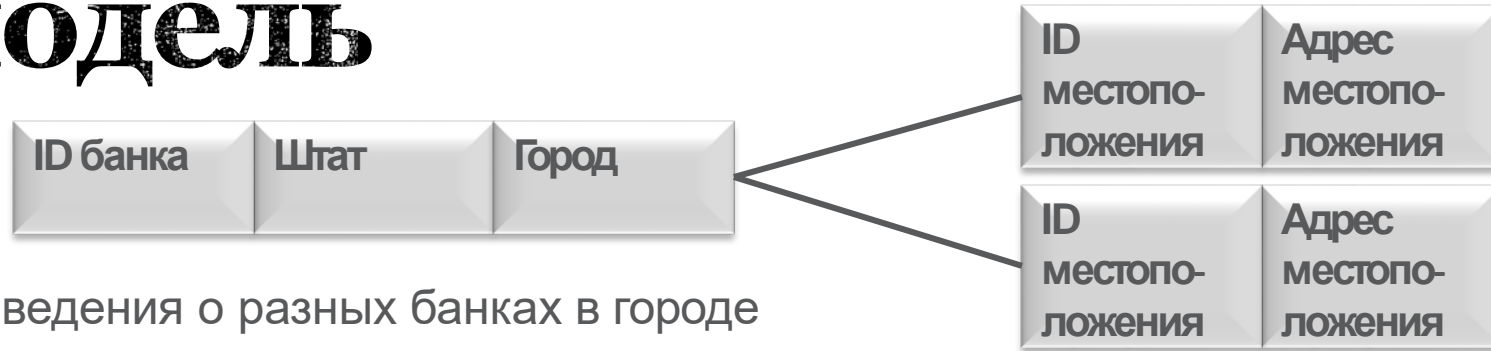
Сетевая БД – это набор узлов, в которых каждый может быть связан с каждым (схема дорог).



- Сетевая база данных является обобщением иерархической за счет допущения объектов, имеющих более одного предка. На связи между объектами в сетевой модели не накладывается никаких ограничений.



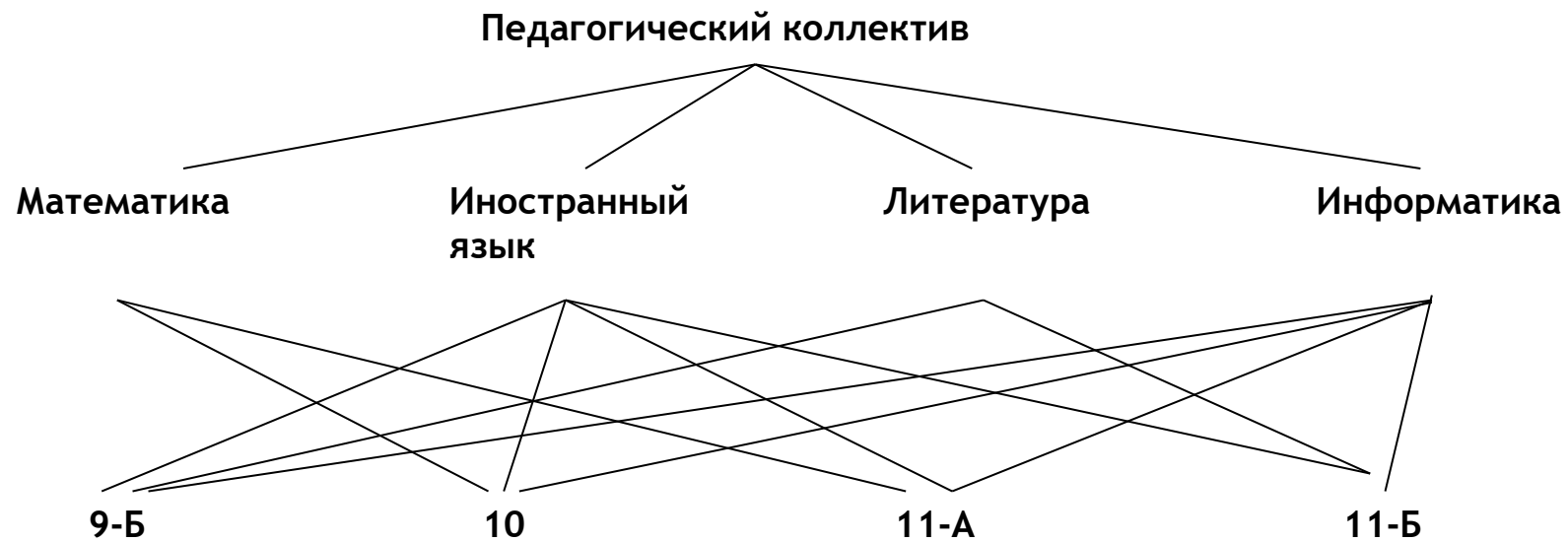
Сетевая модель



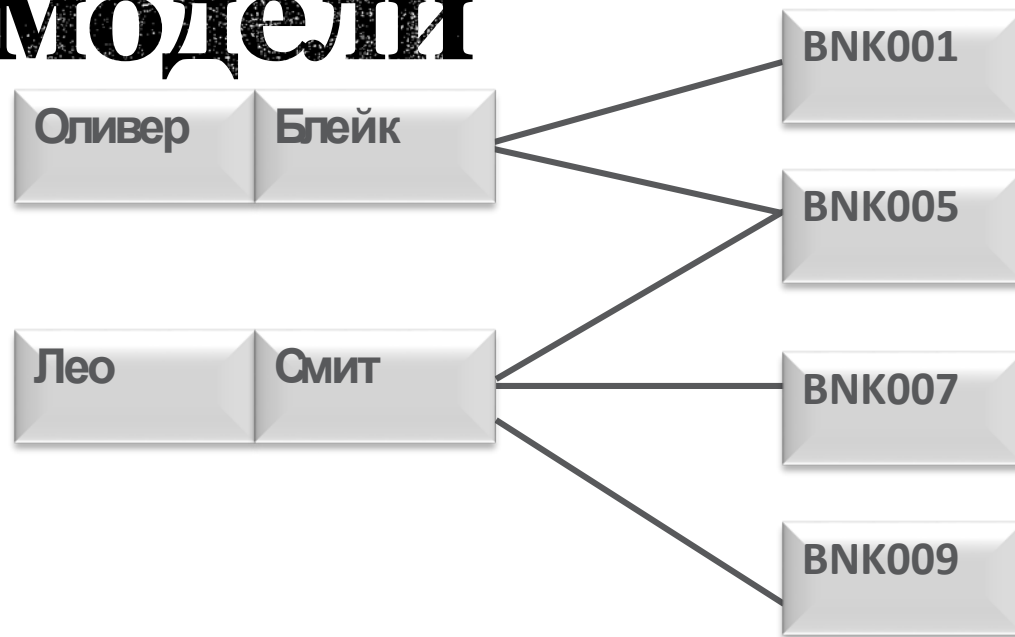
- Сетевая модель — это модель БД, которая обеспечивает гибкий способ представления объектов и их связей.
- Сетевая база данных состоит из набора записей, между которыми существуют связи (прямоугольники = поля, линии = связи).
- Каждая запись представляет собой набор полей, каждое из полей содержит только одно значение данных.
- Связь указывает на ассоциацию двух записей.

Примеры сетевой модели

1. Всемирная паутина глобальной компьютерной сети Интернет.
2. (см. ниже)

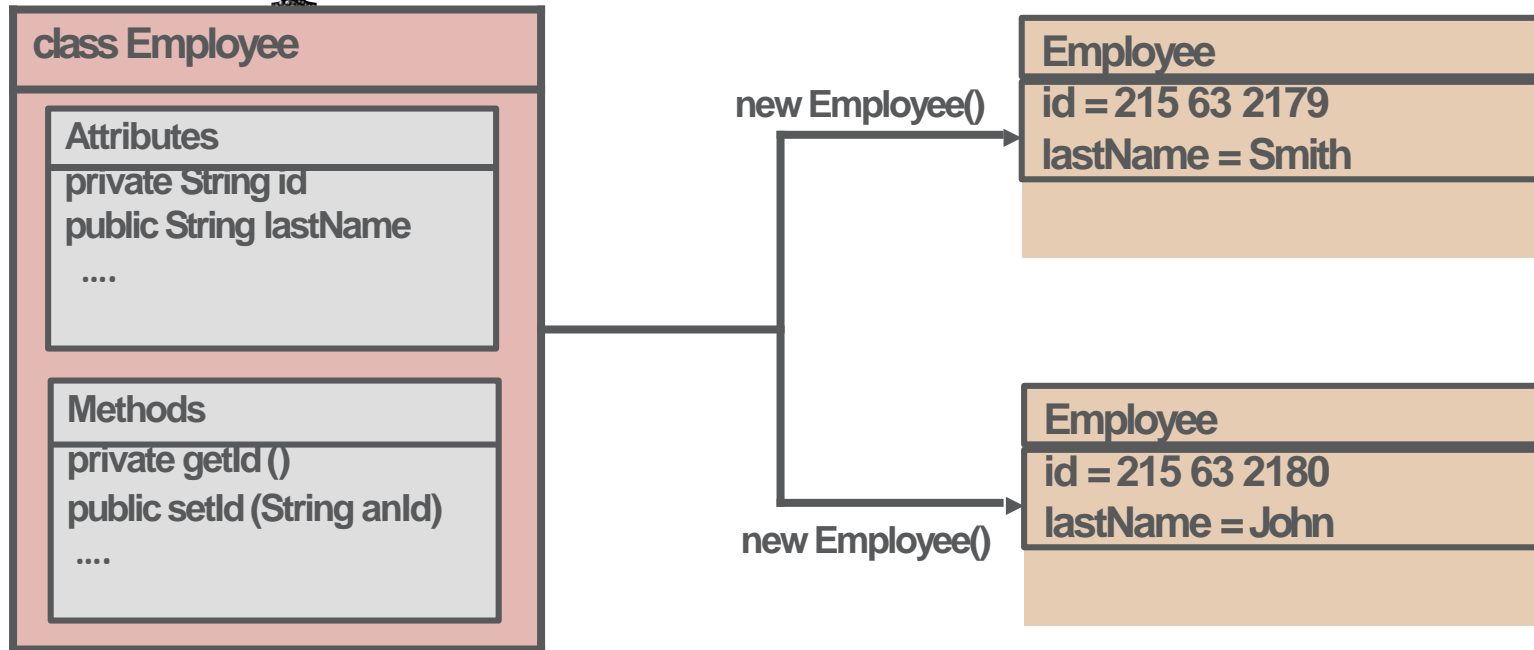


Пример сетевой модели



- Оливер Блейк имеет счета в двух банках: BNK001 и BNK005.
- Лео Смит имеет счета в трех банках: BNK005, BNK007 и BNK009.

Пример объектно-ориентированной модели



- Показаны два объекта Employee (Сотрудник), созданных из класса Employee
- Каждый из них имеет разные значения атрибутов id и lastName

Табличная (реляционная) модель

- Табличная база данных содержит перечень объектов одного типа, то есть объектов, имеющих одинаковый набор свойств. Такую базу удобно представлять в виде двумерной таблицы: в каждой ее строке последовательно размещаются значения свойств одного из объектов; каждое значение свойства - в своем столбце, озаглавленном именем свойства.
- **ПОЛЕ БАЗЫ ДАННЫХ** - это столбец таблицы, содержащий значения определенного свойства.
- **ЗАПИСЬ БАЗЫ ДАННЫХ** - это строка таблицы, содержащая набор значений свойств, размещенный в полях базы данных.
- Каждая таблица должна содержать одно ключевое поле, содержимое которого уникально для каждой записи в таблице.
- **КЛЮЧЕВОЕ ПОЛЕ** - это поле, значение которого однозначно определяет запись в таблице.
- Табличная модель данных может состоять из нескольких таблиц, которые связываются между собой ключом.



Пример табличной (реляционной) модели

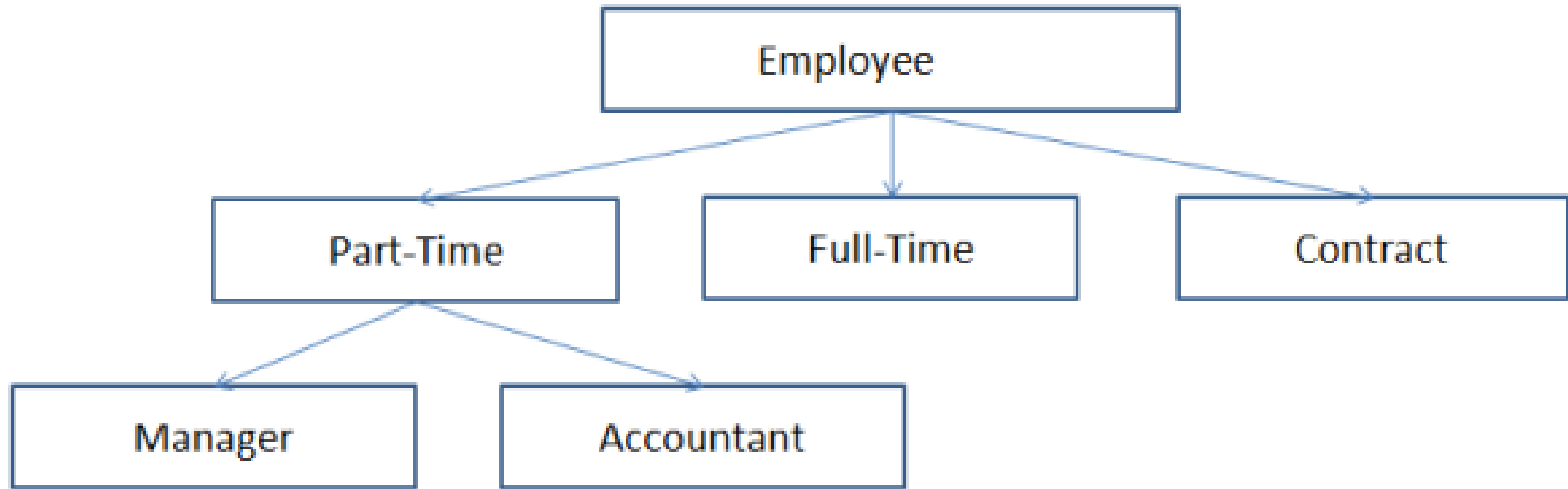
Тип данных Счетчик		Тип данных Текстовый	Тип данных Текстовый	Тип данных Числовой
↓		↓	↓	↓
Поля →	№ п/п	Название	Тип процессора	Оперативная память
Записи →	1	Compag	Celeron	64
	2	Dell	Pentium III	128
	3	IBM	Pentium 4	256

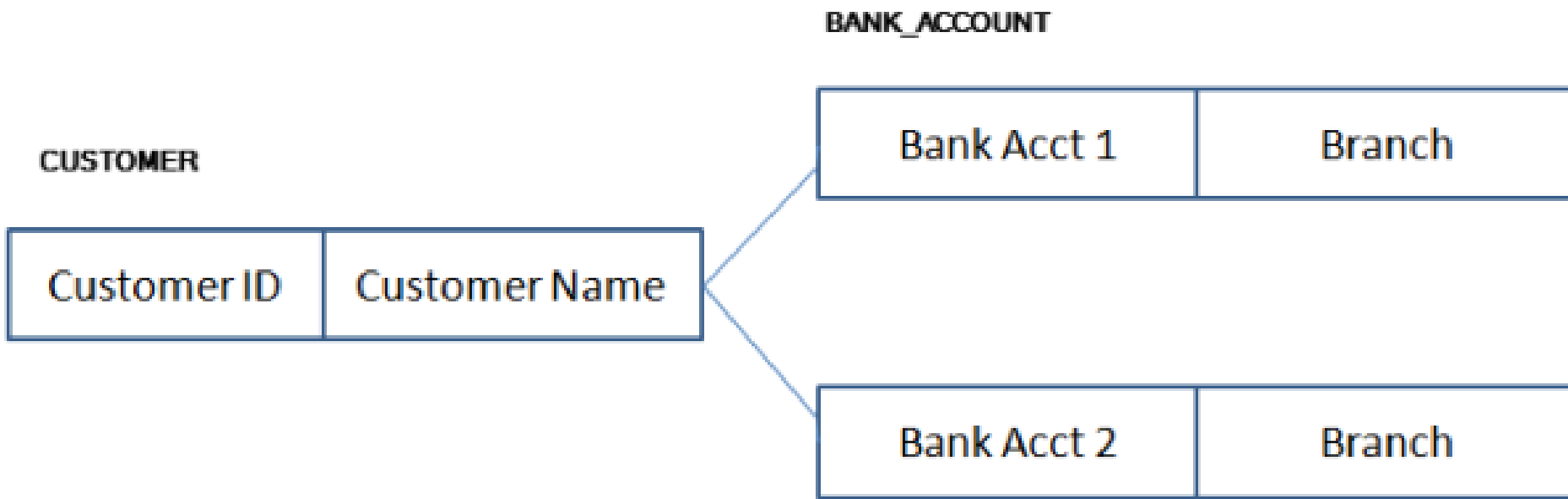


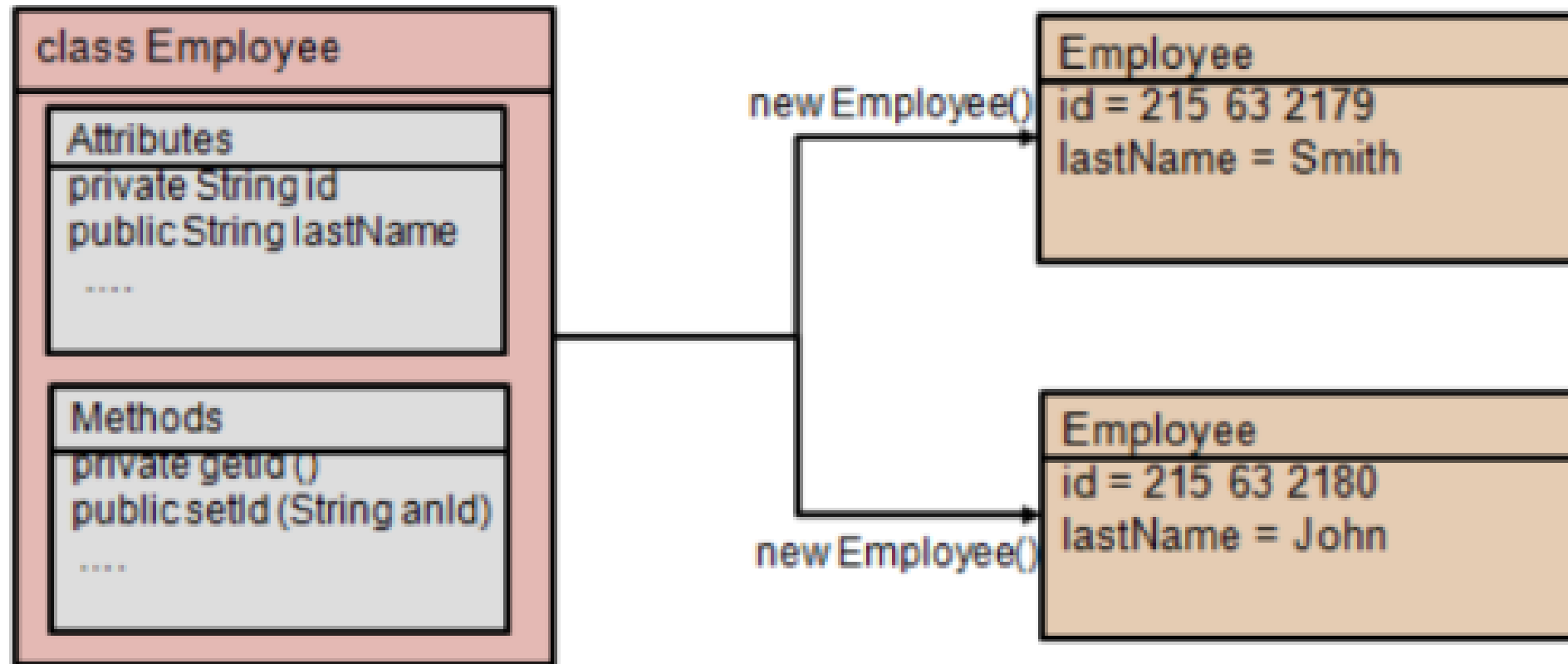
Упражнение 1

- Определить тип модели, представленный на слайдах.







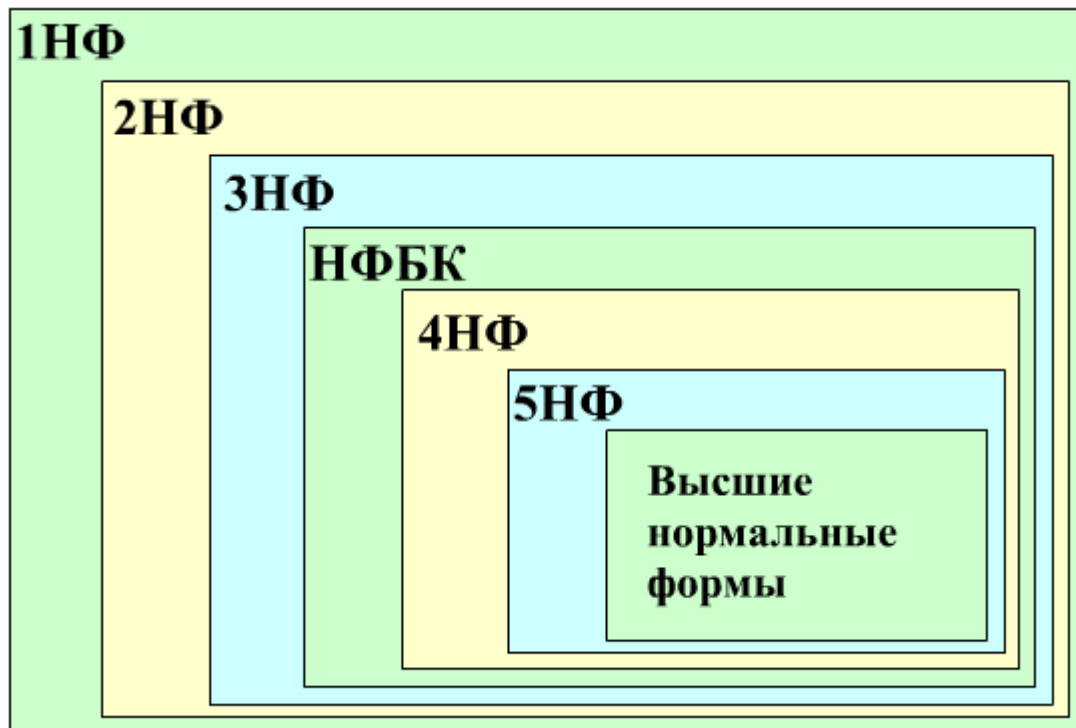


CUSTOMER

Account ID	Customer Name	Branch
A0001	Jeff Covey	Burlington Blvd
A0002	William Jake	Sheldon Park
A0003	Mary Schmidt	Notre Dam Street



Нормализация отношений. Приведение к 1НФ



Определение

Нормализация отношений - процесс построения оптимальной структуры таблиц и связей в реляционной БД (процесс уменьшения избыточности информации).



Цели нормализации

1. Обеспечить быстрый доступ к данным.
2. Исключить ненужное повторение данных.
3. Обеспечить целостность данных, т.е. чтобы при изменении одних объектов автоматически происходило соответствующее изменение связанных с ними объектов.



Понятие функциональной зависимости

Это базовое понятие для нормализации.

Пусть X и Y – произвольные наборы атрибутов отношения.

Y **функционально зависит от** X ($X \rightarrow Y$), если в каждый момент времени каждой совокупности значений набора атрибутов X соответствует не более чем одна совокупность значений набора атрибутов Y .

Иначе, функциональная зависимость Y от X означает, что если в любой момент времени известно значение X , то можно однозначно получить и значение Y .

Если при этом Y не зависит функционально от любого подмножества X (не совпадающего с множеством X), функциональная зависимость называется **полной**.

Если Y является подмножеством атрибутов X , функциональная зависимость называется **тривиальной**. Вообще, тривиальная зависимость – такая зависимость, которая справедлива при любых условиях (обусловлена структурой зависимости).



Понятие функциональной зависимости

Х называется **детерминантом** (левая часть функциональной зависимости) функциональной зависимости, Y называется **зависимой частью**. Функциональная зависимость определяется исходя из смысла хранимых данных.

Если атрибуты X составляют потенциальный ключ некоторого отношения R, то *любой* атрибут отношения R функционально зависит от X.

Примеры функциональной зависимости:

Отношение СТУДЕНТ (ФИО, Номер_зачетной_книжки, Номер_группы)

Функциональные зависимости (при условии, что не совпадений по ФИО, каждый студент имеет единственный номер зачетной книжки и числится в одной группе):

Номер_зачетной_книжки -> ФИО

Номер_зачетной_книжки -> Номер_группы



Аномалии в таблицах БД

Аномалией называется такая ситуация в таблице БД, которая приводит к противоречию в БД, либо существенно усложняет обработку БД. Причина - излишнее дублирование данных в таблице, которое вызывается *наличием функциональных зависимостей от неключевых атрибутов*.

Аномалии-модификации проявляются в том, что изменение одних данных может повлечь просмотр всей таблицы и соответствующее изменение некоторых записей таблицы.

Аномалии-удаления — при удалении какого либо кортежа из таблицы может пропасть информация, которая не связана напрямую с удаляемой записью.

Аномалии-добавления возникают, когда информацию в таблицу нельзя поместить, пока она неполная, либо вставка записи требует дополнительного просмотра таблицы.



Первая нормальная форма

Отношение, находящееся в первой нормальной форме (1НФ) должно отвечать следующим требованиям:

- **Все атрибуты атомарны** (все атрибуты являются простыми, т.е. не имеют компонентов). Иными словами, домен атрибута должен состоять из неделимых значений и не может включать в себя множество значений из более элементарных доменов.
- **Отношение представлено множеством неповторяющихся кортежей** (у отношения есть первичный ключ).
- В отношении нет повторяющихся групп (атрибутов с одинаковым смыслом).

В реляционной модели отношение всегда находится в 1НФ по определению понятия *отношение*.

Что касается различных *таблиц*, то они могут не быть правильными представлениями отношений и, соответственно, могут не находиться в 1НФ.



Примеры ненормализованных таблиц

Курс	Лектор	Дисциплина1	Дисциплина2
1	Яшин Г. А.		С++
2	Медведев А. Г.	СУБД	Web
3	Кудряшова В. Ф.	АСОИУ	ИС
4	Коротков Л. С.	Практика	

Повторяющиеся группы: атрибуты Дисциплина1, Дисциплина2 с одинаковым смыслом

Курс	Лектор	Дисциплины
1	Яшин Г.А.	С++
2	Медведев А. Г.	СУБД Web
3	Кудряшова В. Ф.	АСОИУ ИС
4	Коротков Л. С.	Практика

Атрибут
«Дисциплина»
состоит из
нескольких
значений



Порядок приведения к 1НФ

Для приведения к 1НФ следует:

1. Заполнить таблицы базы данных, используя большее количество повторяющихся данных.
2. Избавиться от сложных атрибутов.
3. Определить первичный ключ в каждой таблице.

При выполнении п.2 вначале необходимо избавиться от атрибутов с одинаковым смыслом, объединив каждую такую группу в один составной атрибут (например, объединить атрибуты Студент1, Студент2 в Студенты, Телефон1, Телефон2, Телефон3 - в Телефоны - но так получится сложный атрибут).

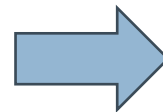
Далее возможны 2 способа к приведению к 1НФ:

- Первый способ предполагает декомпозицию строки со сложным атрибутом: для каждого атомарного значения сложного атрибута создается своя строка. Полученная в результате таблица будет содержать атомарные значения для каждого из атрибутов. Все строки полученной таблицы будут различны, т.е. можно будет выделить первичный ключ (см. следующий слайд).



1 способ: пример приведения к 1НФ

Курс	Лектор	Дисциплины
1	Яшин Г.А.	С++
2	Медведев А. Г.	СУБД
		Практика
3	Кудряшова В. Ф.	АСОИУ
		ИС
4	Коротков Л. С.	Практика



<u>Курс</u> (ПК)	<u>Лектор (ПК)</u>	<u>Дисциплина</u> (ПК)
1	Яшин Г.А.	С++
2	Медведев А. Г.	СУБД
2	Медведев А. Г.	Практика
3	Кудряшова В. Ф.	АСОИУ
3	Кудряшова В. Ф.	ИС
4	Коротков Л. С.	Практика

Отношение в 1НФ



Порядок приведения к 1НФ: 2 способ

- Второй способ предполагает, что один атрибут или группа атрибутов назначаются ключом ненормализованной таблицы, а затем сложные атрибуты (или их группы) удаляются из таблицы и помещаются в отдельную (подчиненную) таблицу вместе с копиями ключа из исходной таблицы.

<u>Курс (ПК)</u>	Лектор
1	Яшин Г.А.
2	Медведев А. Г.
3	Кудряшова В. Ф.
4	Коротков Л. С.

<u>Курс (ПК, ВК)</u>	<u>Дисциплина (ПК)</u>
1	С++
2	СУБД
2	Практика
3	АСОИУ
3	ИС
4	Практика

Отношение в 1НФ: 2 таблицы



2 способ: другой пример

Shipment
<u>Ship Registration Number</u>
<u>Departure Data</u>
Ship Name
Origin
Destination
1st Consignment
2st Consignment
3st Consignment
Customs Value
Ship Capacity
Customs Declaration

Рассмотрим ненормализованное отношение SHIPMENT (ОТГРУЗКА). Оно содержит повторяющиеся группы, представляющие массив значений, 1st Consignments, 2st Consignments, 3st Consignments (партии грузов).

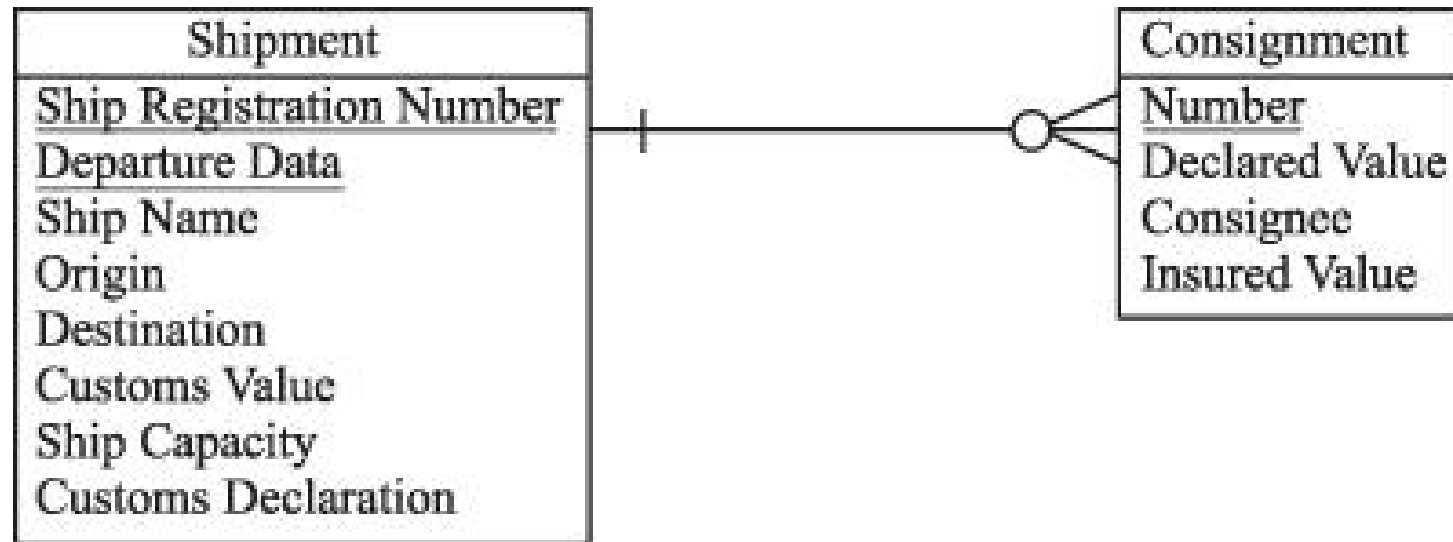
Атрибуты, характеризующие партию грузов (показаны на следующем слайде), - Consignee (грузополучатель), Insured Value (застрахованная стоимость) и Declared Value (объявленная стоимость), - повторяются для каждой такой партии.

Для такого отношения следует ввести бизнес-правило, требующее, чтобы груз состоял не более чем из трех партий, так как четвертую партию вставить в этом отношении некуда.



2 способ: другой пример

Приведение отношения SHIPMENT к 1НФ заключается в изъятии данных о партиях груза из отношения SHIPMENT и создании для них связанного подчиненного отношения CONSIGNMENT (ПАРТИЯ_ГРУЗА). Для такого представления сущности SHIPMENT не требуется вводить упомянутое ограничительное бизнес-правило.



Упражнение 2.

1. Создайте отношение (таблицу), добавьте в него гипотетические данные (в том числе, повторяющиеся).
2. Выполните нормализацию до первой нормальной формы (нужно выделить составной первичный ключ, запрещается вводить суррогатный ключ!).
3. Найдите и обоснуйте функциональные зависимости между атрибутами.
Номер варианта вычислите по формуле: $\text{Номер_по_журналу} \% 20 + 1$, % - операция взятия остатка от деления.

Варианты заданий

1. Предметная область «Птицеферма». Требуется хранить информацию о продаже скупщикам птицы и о содержании птиц. Птицеферма (Наименование_вида_птицы, Цена_мяса_данного_вида_птицы_за_кг, Номер_птичника, Наименование_птичника, Адрес_птичника, Вид_птичника, Вид_птиц_в_птичнике, Вместимость_птичника, ФИО_директора_птичника, Телефон_директора1, Телефон_директора2, Телефон_директора3, Дата_продажи, Наименование_скупщика, Статус_скупщика, Количество_проданных_птиц_данного_вида, Общая_стоимость). Вид птичника выбрать из списка: брудергауз, батарейный цех, акклиматизатор, репродукторный, взрослый. Статус скупщика выбрать из списка: физическое лицо, юридическое лицо.
2. Предметная область «Организация выступлений». Требуется хранить информацию о выступлении коллективов с различными номерами на площадках. Выступления (Название_номера, Тип_номера, Название_коллектива, Дата_основания_коллектива, ФИО_руководителя_коллектива, Контактный_номер_руководителя1, Контактный_номер_руководителя2, Количество_участников, Телефон_ответственного_за_площадку, Дата_время_выступления, Стоимость_входного_билета_и_тип_билета). Название типа номера выбрать из списка: танец, песня, юмор, оригинальный. Тип билета: обычный, льготный.
3. Предметная область «Прокат спортивного инвентаря». Требуется хранить информацию о прокате клиентам спортивного инвентаря. Прокат (ФИО_клиента, Телефон_клиента, Паспорт_клиента, Инвентарный_номер, Вид_инвентаря, Цена_аренды_за_сутки, Цена_аренды_за_трое_суток, Цена_аренды_за_неделю, Остаточная_стоимость_инвентаря, Штраф_за_сутки_просрочки, Штраф_за_неделю_просрочки, Дата_Время_выдачи_инвентаря, Длительность_проката_в_днях, Фактическая_дата_возврата_инвентаря). Название вида выбрать из списка: лыжи, коньки, велосипед и т.д.



4. Предметная область «Результаты соревнований по спортивному ориентированию». Требуется хранить информацию о результатах участия спортсменов в соревнованиях по спортивному ориентированию. Соревнования (ФИО_участника, Год_рождения, Номер_телефона_участника, Возрастная_группа_участника, Пол_участников_группы, Минимальный_возраст, Максимальный_возраст, Название_коллектива_и_год_создания, ФИО_тренера, Телефон_тренера1, Телефон_тренера2, Дата_время_старта, Дата_время_финиша, Результат_прохождения_трассы). Результат прохождения трассы выбрать из списка: финишировал, сошел с дистанции, дисквалифицирован.
5. Требуется хранить информацию о выступлении спикеров на семинарах. Семинар (ФИО_спикера, Год_рождения_спикера_и_город_рождения_и_страна_рождения, Судейская_категория_спикера, Оплата_за_час_семинара, Тема_семинара, Дата_время_начала_семинара, Максимальное_количество_участников, Адрес_площадки, Вместимость_площадки, Телефон_ответственного_за_площадку1, Телефон_ответственного_за_площадку2). Судейскую категорию выбрать из списка спортивный судья 3-ей категории (ССЗК), СС2К, СС1К, ССВК.
6. Предметная область «Приют для животных». Требуется хранить информацию о животных, сделанных им прививках и усыновлении животных. Приют (Кличка_животного, Вид_животного, Примерный_возраст_и_цвет_животного, Пол, Дата_поступления_в_приют, Наименование_вида_прививки, Препарат, Регулярность_прививания, Дата_прививки, ФИО_усыновителя_и_Телефон1_усыновителя, Телефон2_усыновителя, Адрес_усыновителя, Дата_усыновления).
7. Предметная область «Сотовая компания». Требуется хранить информацию о детализации звонков абонентов. Сотовая компания (ФИО_абонента, Паспорт_абонента, Адрес_абонента, Телефон1_абонента, Телефон2_абонента, Название_тарифа_и_описание_тарифа, Стоимость_минуты_разговора, Количество_ГБ, Дата_время_заключения_договора, Дата_время_начала_разговора, Продолжительность_разговора_в_минутах).



8. Предметная область «Животноводческая ферма». Требуется хранить информацию о размещении скота на ферме и продаже скота покупателям. Ферма (Наименование_вида_скота, Средний_вес_скота_данного_вида, Цена_мяса_данного_вида_скота_за_кг, Номер_и_вид_помещения, Вместимость_помещения, Вид_скота_в_помещении, Ответственный_за_помещение, Наименование_покупателя, Телефон1_покупателя, Телефон2_покупателя, Адрес_покупателя, Статус_покупателя, Дата_продажи, Количество_проданных_голов_скота_данного_вида, Общая_стоимость). Вид помещения выбрать из списка: стойло, хлев, загон, конюшня. Статус выбрать из списка: физическое лицо, юридическое лицо.
9. Предметная область «Швейная мастерская». Требуется хранить информацию о продаже выполненных швеями изделий. Мастерская (ФИО_швеи_и_дата_рождения_швеи, Разряд, Адрес, Стаж, Название_изделия, Тип_изделия, Размер, Цвет, Трудочасы, Номинальная_цена, Название_материала, Стоимость_за_1_м2, Максим_допустимое_количество_оборотов_во_время_стирки, Максим_допустимая_температура_стирки, Процент_усадки, Дата_изготовления_изделия, Розничная_цена) Тип изделия выбрать из списка: плечевые изделия (платье, блузка, пиджак, плащ, куртка) и поясные изделия (юбка, брюки, шорты).
10. Предметная область «Туристическая фирма». Требуется хранить информацию о продажах менеджерами туров клиентам. Фирма (ФИО__и_адрес_клиента, Паспорт, Загранпаспорт, Телефон1, Телефон2, Телефон3, ФИО_менеджера, Паспорт_менеджера, Телефон_менеджера, Дата_приема_на_работу, Страна_назначения, Город_назначения, Отель, Тип_номера, Стоимость_тура_для_одного_человека, Дата_продажи, Кол-во_человек, Дата_отъезда, Дата_приезда)
11. Предметная область «Доставка пиццы». Требуется хранить информацию о курьерской доставке пиццы клиентам. Доставка (ФИО_и_телефон_клиента, Почта1, Почта2, ФИО_курьера, Паспорт, Адрес_курьера, Телефон1_курьера, Телефон2_курьера, Название_пиццы, Описание, Среднее_время_приготовления, Калорийность, Стоимость, Дата_время_заказа, Адрес_доставки, Кол-во_пицц)
12. Предметная область «Парикмахерская». Требуется хранить информацию об услугах, предоставляемых клиентам. Парикмахерская (Дисконтная_карта1, Дисконтная_карта2, ФИО_и_дата_рождения_клиента, Пол, Телефон, ФИО_мастера, Дата_рождения_мастера, Стаж, Специализация, Наименование_услуги, Цена, Продолжительность, Описание, Дата_время_записи)



13. Предметная область «Каршеринг». Требуется хранить информацию о поездке клиентов на автомобиле. Каршеринг (Гос._номер, Марка_автомобиля, Модель_автомобиля, Год_выпуска, ФИО_и_номер_водительского_удостоверения_клиента, Адрес, Телефон1, Телефон2, Номер_банковской_карты, Название_тарифа, Тип_страховки, Цена_за_час, Дата_время_начала_поездки, Дата_время_окончания_поездки) Тип страховки выбрать из списка: ОСАГО, КАСКО, без страховки.
14. Предметная область «Театр». Требуется хранить информацию о расписании спектаклей и участвующих в них актерах. Театр (Название_спектакля, Жанр1, Жанр2, Возрастная_категория, Автор_пьесы, Режиссер, Дата_премьеры, ФИО_и_стаж_актера, Дата_рождения, Адрес, Звание, Оклад, Роль, Дата_время_представления, Стоимость_билета)
15. Предметная область «Медицинский центр». Требуется хранить информацию о посещениях врачей пациентами. Медцентр (ФИО_и_категория_врача, Специализация, Стаж, Дата_последнего_повышения_квалификации, ФИО_пациента, Дата_рождения_пациента, Адрес, Телефон1, Телефон2, Номер_страхового_полиса, Название_услуги, Описание, Предварительная_подготовка, Стоимость, Дата_время_посещения, Кабинет1, Кабинет2, Кабинет3)
16. Предметная область «Выставка кошек». Требуется хранить информацию о процессе проведения выставки кошек. Выставка (Кличка, Возраст, Порода, ФИО_и_телефон_хозяина, Адрес_площадки, ФИО_ответственного, Цена_входного_билета, Вместимость, Максимальное_количество_животных, Дата_начала_выставки, Дата_окончания_выставки, ФИО_куратора, Стоимость_участия, Номер_места_расположения_кошки)
17. Предметная область «Рок-фестиваль». Требуется хранить информацию о расписании выступлений музыкальных групп на фестивале. Фестиваль (Название, Жанр, Количество_участников, ФИО_и_номер_менеджера, Название_сцены, Площадь_сцены, ФИО_ответственного_за_сцену, ФИО_работника_сцены, Телефон1, Телефон2, Должность, Оклад, Дата_время_начала_выступления, Длительность, Гонорар)
18. Предметная область «Фитнес клуб». Требуется хранить информацию о продаже абонементов на занятия фитнесом. Фитнес клуб (ФИО_и_дата_рождения_клиента, Адрес, Телефон1, Телефон2, Тип_абонемента, Описание, Стоимость, Срок_действия, Дата_время_покупки_абонемента, Дата_время_персональной_тренировки, Стоимость_тренировки, ФИО_тренера)
19. Предметная область «Клуб ролевиков». Требуется хранить информацию об участии ролевиков в играх клуба. Клуб ролевиков (ФИО_и_Никнейм, Дата_рождения, Номер_телефона1, Номер_телефона2, Номер_телефона3, Номер_телефона4, Дата_время_начала_игры, Длительность, Куратор, Стоимость_участия, Имя_героя_ролевика_в_игре, Раса, Описание_расы)
20. Предметная область «Сбор грибов». Требуется хранить информацию о продажах грибов грибниками в пункты приема. Сбор грибов (Адрес_пункта_приема1, Адрес_пункта_приема2, Адрес_пункта_приема3, Контактный_телефон1, Контактные_телефон2, Контактный_телефон3, ФИО_ответственного1, ФИО_ответственного2, ФИО_ответственного3, ФИО_и_телефон_грибника, Название_гриба1, Род_гриба1, Название_гриба2, Род_гриба2, Название_гриба3, Род_гриба3, Стоимость_за_кг, Дата_время_продажи, Масса_грибов).

