

Нормализация отношений

Лекция 5

Доцент каф. ВТиКГ Данилова Е.В.

ПЛАН

1. Функциональная зависимость
2. Понятие нормализации
3. Нормальные формы

Функциональная зависимость

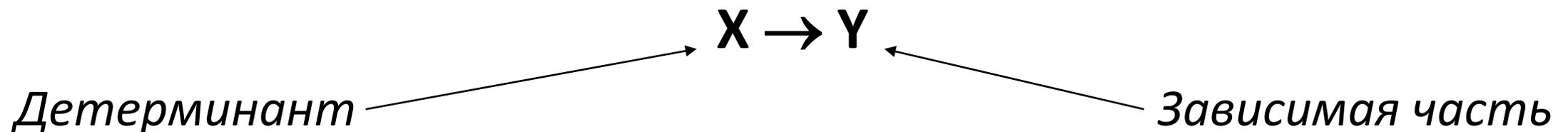
1

Определение функциональной зависимости

Пусть r – отношение, а X и Y – произвольные подмножества множества атрибутов отношения r .

Тогда Y функционально зависимо от X (X функционально определяет Y) тогда и только тогда, когда каждое значение множества X отношения r связано точно с одним значением множества Y отношения r .

Иначе говоря, если два кортежа отношения r совпадают по значению X , они совпадают и по значению Y .



Подмножество 1

множества
атрибутов r

X

Подмножество 2

множества
атрибутов r

Y

Код студента	№ ДОГОВОРА	ФИО	ОЦЕНКА	ГОД РОЖДЕНИЯ	СЕРИЯ ПАСПОРТА	НОМЕР ПАСПОРТА	НАЗВАНИЕ ОЦЕНКИ
1	АЛ-0034	Иванов И.И.	5	1998	0405	456756	отлично
2	АЛ-0035	Сидоров С.С.	4	1999	0533	859372	хорошо
3	АЛ-0036	Сидоров С.С.	4	1999	0533	473829	хорошо
4	АЛ-0037	Петров П.В.	3	1998	0405	473829	удовл.
5	АЛ-0038	Маркова А.И.	5	1997	0742	859372	отлично
6	АЛ-0039	Ржевская Ю.А.	4	1997	0742	946627	хорошо

X – ОЦЕНКА, Y – НАЗВАНИЕ ОЦЕНКИ

$X \rightarrow Y$

ОЦЕНКА \rightarrow НАЗВАНИЕ ОЦЕНКИ

Подмножество 1
множества
атрибутов r

X

Подмножество 2
множества
атрибутов r

Y

Код студента	НОМЕР ДОГОВОРА	ФИО	ОЦЕНКА	ГОД РОЖДЕНИЯ	СЕРИЯ ПАСПОРТА	НОМЕР ПАСПОРТА	НАЗВАНИЕ ОЦЕНКИ
1	АЛ-0034	Иванов И.И.	5	1998	0405	456756	отлично
2	АЛ-0035	Сидоров С.С.	4	1999	0533	859372	хорошо
3	АЛ-0036	Сидоров С.С.	4	1999	0533	473829	хорошо
4	АЛ-0037	Петров П.В.	3	1998	0405	473829	удовл.
5	АЛ-0038	Маркова А.И.	5	1997	0742	859372	отлично
6	АЛ-0039	Ржевская Ю.А.	4	1997	0742	946627	хорошо

X – {НОМЕР ДОГОВОРА, ФИО}, Y – {СЕРИЯ ПАСПОРТА, НОМЕР ПАСПОРТА}

$X \rightarrow Y$

{НОМЕР ДОГОВОРА, ФИО} \rightarrow {СЕРИЯ ПАСПОРТА, НОМЕР ПАСПОРТА}

Замечание 1

Детерминант X и зависимая часть Y являются множествами атрибутов (следует из определения функциональной зависимости).

Если множество содержит только один атрибут, оно называется **одноэлементным множеством**.

Поэтому скобки исключаются и символическая запись функциональной зависимости на слайде 5 имеет следующий вид:

ОЦЕНКА → НАЗВАНИЕ ОЦЕНКИ

Замечание 2

Если X является потенциальным ключом отношения r , то все атрибуты Y отношения r должны обязательно быть функционально зависимыми от X (следует из определения потенциального ключа):

Код студента	НОМЕР ДОГОВОРА	ФИО	ОЦЕНКА	ГОД РОЖДЕНИЯ	СЕРИЯ ПАСПОРТА	НОМЕР ПАСПОРТА	НАЗВАНИЕ ОЦЕНКИ
1	АЛ-0034	Иванов И.И.	5	1998	0405	456756	отлично
2	АЛ-0035	Сидоров С.С.	4	1999	0533	859372	хорошо
3	АЛ-0036	Сидоров С.С.	4	1999	0533	473829	хорошо
4	АЛ-0037	Петров П.В.	3	1998	0405	473829	удовл.
5	АЛ-0038	Маркова А.И.	5	1997	0742	859372	отлично
6	АЛ-0039	Ржевская Ю.А.	4	1997	0742	946627	хорошо

Код студента \rightarrow {Номер договора, ФИО, Оценка, Год рождения, Серия паспорта, Номер паспорта, Название оценки}

Замечание 3

Функциональные зависимости отражают взаимосвязи, обнаруженные между объектами предметной области и являются дополнительными ограничениями, определяемыми предметной областью. Функциональная зависимость – семантическое понятие. Она возникает, когда по значениям одних данных в предметной области можно определить значения других данных.

Замечание 4

Полный набор функциональных зависимостей, выполняющихся для отношения, может быть очень большим. Необходимо использовать методы сокращения обширного множества функциональных зависимостей до некоторых допустимых размеров.

Виды функциональных зависимостей

Тривиальные

Нетривиальные

Тривиальная функциональная зависимость – функциональная зависимость, которая не может не выполняться.

Пример:

{Номер договора, ФИО} → Номер договора

Функциональная зависимость является тривиальной тогда и только тогда, когда правая часть ее символической записи является подмножеством (не обязательно строгим) левой части.

Все остальные зависимости относятся к **нетривиальным**.

Виды функциональных зависимостей

Полная

Частичная

Полная функциональная зависимость – функциональная зависимость неключевого атрибута (группы неключевых атрибутов) от составного ключа, при этом неключевой атрибут (группа неключевых атрибутов) не находится в функциональной зависимости ни от какой части составного ключа.

Частичная функциональная зависимость – зависимость неключевого атрибута (группы атрибутов) от части составного ключа.

Пример полной функциональной зависимости

Составной ключ

НОМЕР ДОГОВОРА	ФИО	Год рождения	ОЦЕНКА	СЕРИЯ ПАСПОРТА	НОМЕР ПАСПОРТА	НАЗВАНИЕ ОЦЕНКИ
АЛ-0034	Иванов И.И.	1998	5	0405	456756	отлично
АЛ-0035	Сидоров С.С.	1999	4	0533	859372	хорошо
АЛ-0036	Сидоров С.С.	1999	4	0533	473829	хорошо
АЛ-0037	Петров П.В.	1998	3	0405	473829	удовл.
АЛ-0038	Маркова А.И.	1997	5	0742	859372	отлично
АЛ-0039	Ржевская Ю.А	1997	4	0742	946627	хорошо

{Серия паспорта, Номер паспорта} → {ФИО, Год рождения}

Пример частичной функциональной зависимости

Составной ключ

Артикул препарата (PK)	Упаковка	Стоимость	Название препарата	Дата поставки	id_Изготовитель (PK)	Название изготовителя	Год основания
P48	блистер	200	парацетамол	21.02.2021	1	Биофарма	2000
P49	картон	150	парацетамол	17.03.2020	1	Биофарма	2000
P55	блистер	160	панадол	21.02.2021	1	Биофарма	2000
P56	стекло	300	панадол	25.02.2020	2	Гедеон	1901
P57	порошок	120	панадол	12.04.2018	2	Гедеон	1901
I12	стекло	20	йод	13.05.2019	3	Росбио	2005
I13	ампулы	200	йод	12.04.2018	3	Росбио	2005
P48	блистер	190	парацетамол	15.03.2021	2	Гедеон	1901
P56	стекло	290	панадол	21.02.2021	3	Росбио	2005
P56	стекло	280	панадол	16.09.2020	4	Дальфарма	2012
I13	ампулы	150	йод	30.05.2022	1	Биофарма	2000
P57	порошок	175	панадол	01.07.2020	1	Биофарма	2000
...

{Артикул препарата} → {Упаковка, Название препарата}

{id_Изготовитель} → {Название изготовителя, Год основания}

Виды функциональных зависимостей

Транзитивная

Транзитивная функциональная зависимость

- Пусть r является отношением, и $\{A, B, C\}$ – множество атрибутов этого отношения. Если имеются функциональные зависимости $A \rightarrow B$ и $B \rightarrow C$, то в этом случае выполняется функциональная зависимость $A \rightarrow C$ (при условии, что атрибут A функционально не зависит ни от атрибута B , ни от атрибута C), которая называется **транзитивной** функциональной зависимостью (говорят, C зависит от A транзитивно).

Отношение «ЗАДОЛЖЕННОСТИ»

ФИО СТУДЕНТА	ДИСЦИПЛИНА	ПРЕПОДАВАТЕЛЬ
Иванов И.И.	Алгебра	Лисицина Ю.М.
Петров П.П.	Алгебра	Лисицина Ю.М.
Петров П.П.	Информатика	Груздев А.П.
Сидоров С.Г.	Информатика	Груздев А.П.
Игнатова П.В.	Базы данных	Ивлеев Г.И.
Воронова А.М.	Алгебра	Лисицына Ю.М.

Транзитивная зависимость: ФИО СТУДЕНТА \rightarrow \rightarrow ПРЕПОДАВАТЕЛЬ
(ФИО СТУДЕНТА \rightarrow ДИСЦИПЛИНА,
ДИСЦИПЛИНА \rightarrow ПРЕПОДАВАТЕЛЬ)

- Замыкание множества зависимостей
- Транзитивная зависимость
- Правила вывода новых функциональных зависимостей на основе заданных (аксиомы Армстронга)
- Правила рефлексивности, дополнения, транзитивности
- Правила самоопределения, декомпозиции, объединения, композиции
- Общая теорема объединения (Дарвен)
- Замыкание множества атрибутов
- Подмножество замыкания
- Неприводимые множества зависимостей
- Покрытие множества
- Состав замыкания
- Неприводимый эквивалент множества зависимостей и др.

Нормализация

2

Понятие нормализации

Нормализация – процесс, позволяющий гарантировать эффективность структур базы данных [Ролланд].

Нормализация – процесс назначения атрибутов сущностям [Роб].

Нормализация – процесс замены отношения некоторым набором его проекций, составленным таким образом, чтобы обратное соединение этих проекций позволяло вновь получить исходное отношение [Дейт, стр. 489].

Нормализация – процесс последовательного приведения заданного набора отношений к некоторой все «более желательной» форме [Дейт, стр. 460].

Нормализация – процесс преобразования отношений базы данных к виду, соответствующему нормальным формам.

Нормальные формы

Процесс нормализации основывается на концепции «**нормальных форм**».

Отношение находится в определенной нормальной форме (НФ), если оно удовлетворяет заданному набору условий.

Виды нормальных форм:

- первая нормальная форма (1НФ)
- вторая нормальная форма (2НФ)
- третья нормальная форма (3НФ)
- нормальная форма Бойса-Кодда (НФБК)
- четвертая нормальная форма (4НФ)
- пятая нормальная форма (5НФ)
- доменно-ключевая нормальная форма (ДКНФ)
- шестая нормальная форма (6НФ)

Отношения В 1НФ (нормализованные)

Отношения В 2НФ

Отношения В 3НФ

Отношения В 4НФ

Отношения В 5НФ

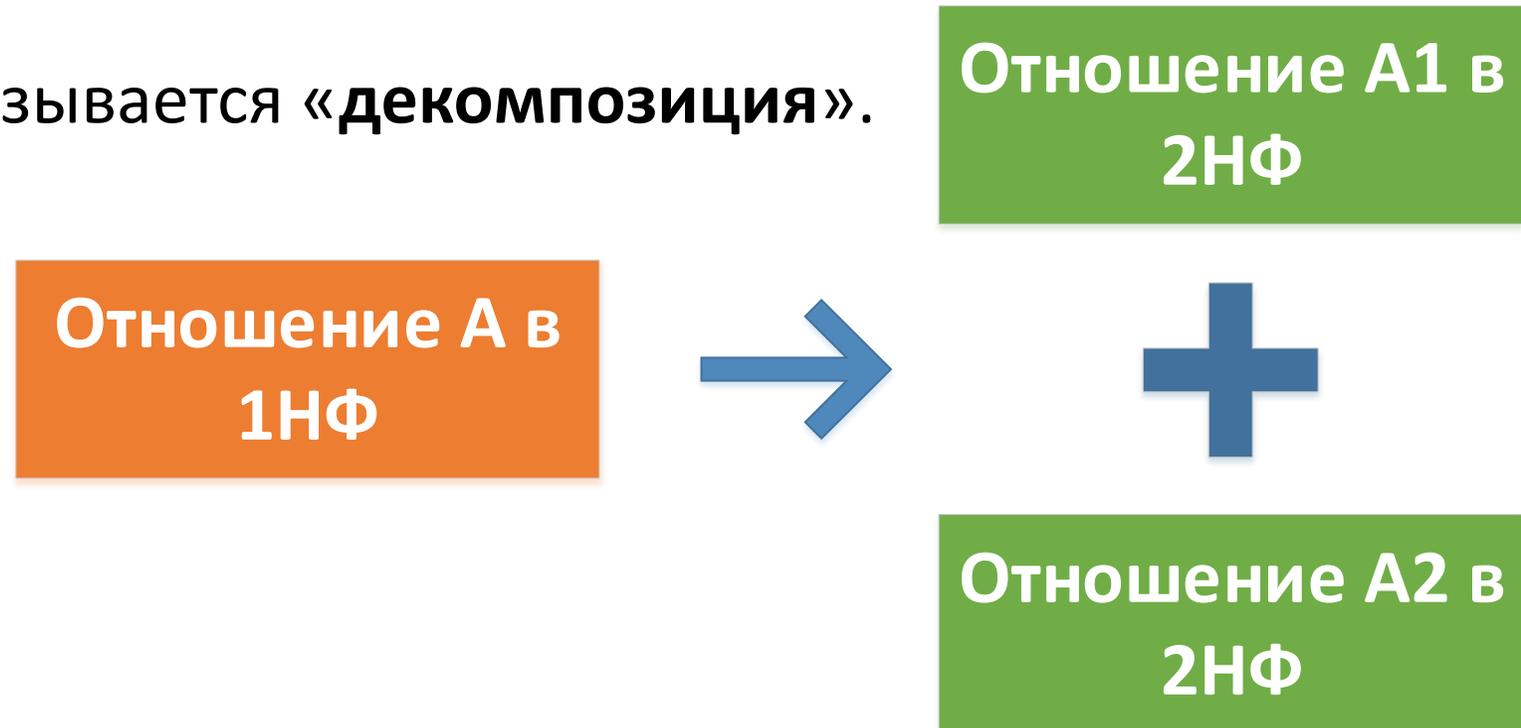
Отношения в НФ более
высокого порядка

Уровни нормализации

Декомпозиция

В результате нормализации отношение в некоторой нормальной форме (например, 1НФ) может быть преобразовано в несколько отношений в другой, более желательной нормальной форме (например, 2НФ).

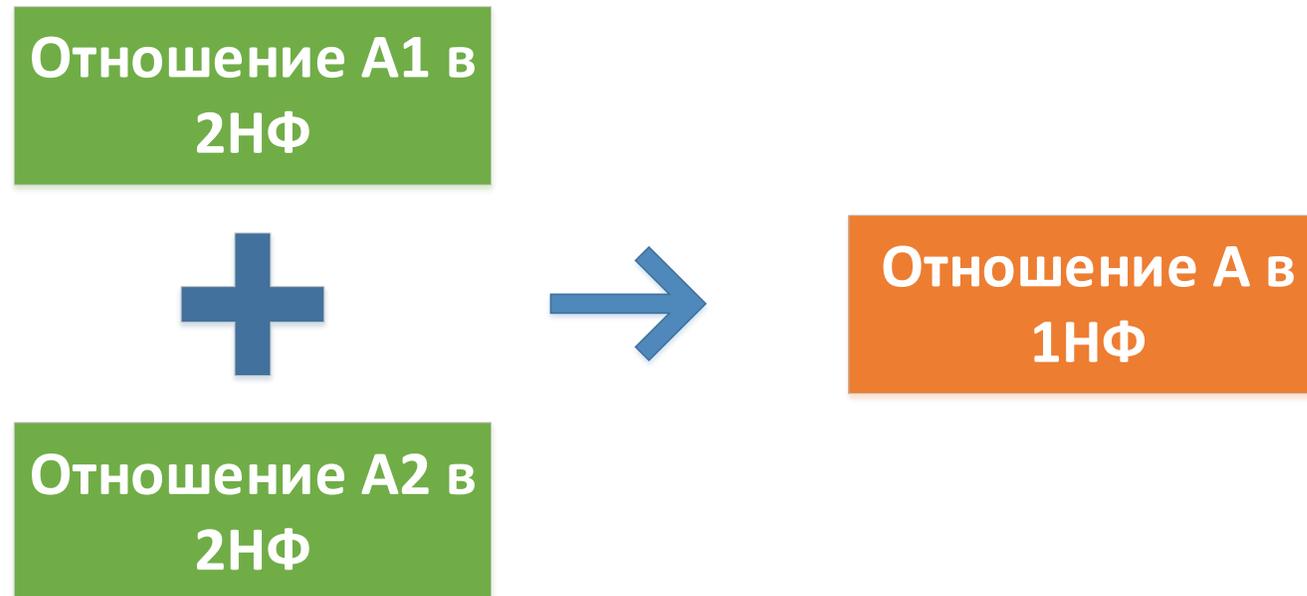
Данный процесс называется «**декомпозиция**».



Обратимость нормализации

Процедура нормализации отношений обратима \Rightarrow в процессе нормализации сохраняется информация (не происходит ее потеря).

Т. о. всегда можно использовать результат нормализации (например, множество отношений 2НФ) для обратного преобразования (в исходное отношение 1НФ).



Декомпозиция без потерь

Процедура нормализации предусматривает разбиение (декомпозицию) отношения на другие отношения, причем декомпозиция должна быть обратимой, т.е. выполняться без потерь информации.

Для получения эффективной структуры данных интерес представляют только те операции, которые выполняются без потерь информации.

Вопрос о том, происходит ли утрата информации при декомпозиции, связан с концепцией функциональной зависимости.

Отношение A

КОД СТУДЕНТА	ОЦЕНКА	ДИСЦИПЛИНА
1	5	Алгебра
2	4	Алгебра
3	4	Информатика
4	3	Информатика
5	5	Базы данных
6	4	Алгебра

Вариант 1 декомпозиции отношения A (декомпозиция без потерь)

Отношение A1

КОД СТУДЕНТА	ОЦЕНКА
1	5
2	4
3	4
4	3
5	5
6	4

Отношение A2

КОД СТУДЕНТА	ДИСЦИПЛИНА
1	Алгебра
2	Алгебра
3	Информатика
4	Информатика
5	Базы данных
6	Алгебра

Вариант 2 декомпозиции отношения A (декомпозиция с потерями)

Отношение A1

КОД СТУДЕНТА	ОЦЕНКА
1	5
2	4
3	4
4	3
5	5
6	4

Отношение A2

ОЦЕНКА	ДИСЦИПЛИНА
5	Алгебра
4	Алгебра
4	Информатика
3	Информатика
5	Базы данных
4	Алгебра

Необходимые условия декомпозиции без потерь

1. Декомпозиция при нормализации с точки зрения реляционной алгебры – это операция **проекции**.

Декомпозиция отношения без потерь – взятие одной или нескольких проекций исходного отношения так, чтобы эти проекции в совокупности содержали (возможно, с повторениями) все атрибуты исходного отношения.

Т. е. при декомпозиции не должны теряться атрибуты отношений.

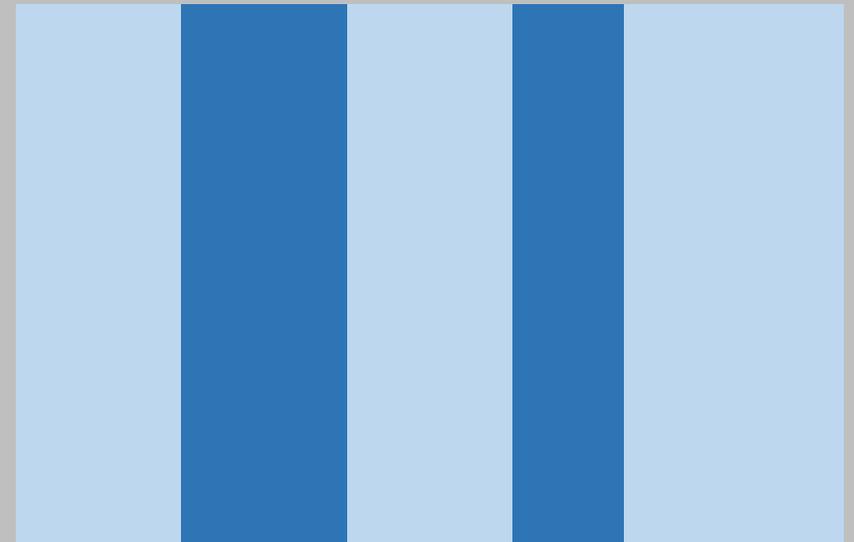
Проекция (PROJECT)

Проекцией отношения **A** с атрибутами **X, Y, ... , Z** называется отношение **A {X, Y, ... , Z}**, заголовок которого формируется из заголовка отношения **A** путем удаления всех атрибутов, не указанных в множестве **{X, Y, ... , Z}**, а тело состоит из всех кортежей **{X:x, Y:y, ... , Z:z}**, таких, что в отношении **A** присутствует кортеж со значением **x** атрибута **X**, **y** атрибута **Y** ... и **z** атрибута **Z**.

Возвращает кортежи с сокращенным набором атрибутов.

A {X, Y, ... , Z}

Применение операции проекции приводит к получению «вертикального» подмножества, т.е. подмножества, полученного путем удаления всех атрибутов, не указанных в разделенном запятыми списке имен атрибутов, и последующего устранения дубликатов кортежей из множества оставшихся кортежей.



Отношение **A** – студенты, обучающиеся в ДВГУПС

ФИО	ГОД РОЖДЕНИЯ	КУРС	СРЕДНИЙ БАЛЛ
Иванов И.И.	1998	2	4
Петров П.П.	1997	3	3
Сидоров С.С.	1999	1	5
Никаноров В.Е.	1998	2	3,7
Завгородний А.А.	1998	2	4,5
Светличная С.В.	1998	2	5
Липатова У.В.	1998	3	4,9

Отношение **A** {КУРС, ГОД РОЖДЕНИЯ}

КУРС	ГОД РОЖДЕНИЯ
2	1998
3	1997
1	1999
3	1998

Необходимые условия декомпозиции без потерь

2. Обратная декомпозиции операция с точки зрения реляционной алгебры – операция **соединения**.

Декомпозиция отношения произведена без потерь, если отношение точно восстанавливается из полученных в процессе нормализации проекций при помощи операции **естественного соединения**.

Т. е. при декомпозиции не должны теряться и сами данные.

Соединение (естественное, natural JOIN)

Если отношения **A** и **B** имеют общий атрибут, то результат естественного соединения содержит только строки, в которых значения общего атрибута совпадают.

Связывает таблицы, выбирая только строки с общими значениями их общих атрибутов [Роб, Ролланд].

Теорема Хита (Теорема Хеза, Heath)

Пусть R является отношением, и $\{A, B, C\}$ – множество атрибутов этого отношения. Если имеется функциональная зависимость $A \rightarrow B$, то проекции $R_1 = R [A, B]$ и $R_2 = R [A, C]$ образуют декомпозицию без потерь.

Теорема Хита (Хеза) доказывает корректность процедуры нормализации (декомпозиция без потери информации).

Нормальные формы

3

Первая нормальная форма

3.1

Первая нормальная форма, 1НФ

Первая нормальная форма (1НФ) – это обычное отношение. Согласно определению, любое отношение автоматически уже находится в 1НФ.

Отношение находится в **1НФ** тогда и только тогда, когда каждый его кортеж содержит только одно значение для каждого из атрибутов.

Свойства отношений (свойства 1НФ):

- В отношении нет одинаковых кортежей.
- Кортежи не упорядочены.
- Атрибуты не упорядочены и различаются по наименованию.
- Все значения атрибутов атомарны (каждый кортеж содержит точно одно значение (соответствующего типа) для каждого атрибута).

Пример «отношения» не в 1НФ

Код студента	НОМЕР ДОГОВОРА	ФИО	ОЦЕНКА	СЕРИЯ ПАСПОРТА	НОМЕР ПАСПОРТА	НАЗВАНИЕ ОЦЕНКИ
1	АЛ-0034	Иванов И.И., 1998	5	0405	456756	отлично
2	АЛ-0035	Сидоров С.С., 1999	4	0533	859372	хорошо
3	АЛ-0036	Сидоров С.С., 1999	4	0533	473829	хорошо
4	АЛ-0037	Петров П.В., 1998	3	0405	473829	удовл.
5	АЛ-0038	Маркова А.И., 1997	5	0742	859372	отлично
6	АЛ-0039	Ржевская Ю.А., 1997	4	0742	946627	хорошо

Пример отношения в 1НФ

Код студента	НОМЕР ДОГОВОРА	ФИО	Год рождения	ОЦЕНКА	СЕРИЯ ПАСПОРТА	НОМЕР ПАСПОРТА	НАЗВАНИЕ ОЦЕНКИ
1	АЛ-0034	Иванов И.И.	1998	5	0405	456756	отлично
2	АЛ-0035	Сидоров С.С.	1999	4	0533	859372	хорошо
3	АЛ-0036	Сидоров С.С.	1999	4	0533	473829	хорошо
4	АЛ-0037	Петров П.В.	1998	3	0405	473829	удовл.
5	АЛ-0038	Маркова А.И.	1997	5	0742	859372	отлично
6	АЛ-0039	Ржевская Ю.А	1997	4	0742	946627	хорошо

Вторая нормальная форма

3.2

Вторая нормальная форма (2НФ)

- **Отношение находится во второй нормальной форме (2НФ) тогда и только тогда, когда отношение находится в 1НФ и нет неключевых атрибутов, зависящих от части сложного ключа (т.е. отсутствуют частичные зависимости).**

Если в отношении первичный ключ содержит всего лишь один атрибут, то оно автоматически будет находиться в 2НФ, если до этого было приведено к 1НФ.

Удаление частичной функциональной зависимости

- Если в отношении между атрибутами существует частичная функциональная зависимость, то такие функционально зависимые атрибуты удаляются из него и помещаются в новое отношение вместе с копией их детерминанта

Пример частичной функциональной зависимости

Отношение «Поставка препаратов»

Артикул препарата (PK)	Упаковка	Стоимость	Название препарата	Дата поставки	id_Изготовитель (PK)	Название изготовителя	Год основания
P48	блистер	200	парацетамол	21.02.2021	1	Биофарма	2000
P49	картон	150	парацетамол	17.03.2020	1	Биофарма	2000
P55	блистер	160	<u>панадол</u>	21.02.2021	1	Биофарма	2000
P56	стекло	300	<u>панадол</u>	25.02.2020	2	Гедеон	1901
P57	порошок	120	<u>панадол</u>	12.04.2018	2	Гедеон	1901
I12	стекло	20	йод	13.05.2019	3	Росбио	2005
I13	ампулы	200	йод	12.04.2018	3	Росбио	2005
P48	блистер	190	парацетамол	15.03.2021	2	Гедеон	1901
P56	стекло	290	<u>панадол</u>	21.02.2021	3	Росбио	2005
P56	стекло	280	<u>панадол</u>	16.09.2020	4	Дальфарма	2012
I13	ампулы	150	йод	30.05.2022	1	Биофарма	2000
P57	порошок	175	<u>панадол</u>	01.07.2020	1	Биофарма	2000
...

{Артикул препарата} → {Упаковка, Название препарата}

{ id_Изготовитель } → {Название изготовителя, Год основания}

Отношение «Изготовители»

id_Изготовитель (PK)	Название изготовителя	Год основания
1	<u>Биофарма</u>	2000
2	<u>Гедеон</u>	1901
3	<u>Росбио</u>	2005
4	<u>Дальфарма</u>	2012
...

Отношение «Поставка препаратов»

Артикул препарата (PK)	id_Изготовитель (PK)	Стоимость	Дата поставки
P48	1	200	21.02.2021
P49	1	150	17.03.2020
P55	1	160	21.02.2021
P56	2	300	25.02.2020
P57	2	120	12.04.2018
I12	3	20	13.05.2019
I13	3	200	12.04.2018
P48	2	190	15.03.2021
P56	3	290	21.02.2021
P56	4	280	16.09.2020
I13	1	150	30.05.2022
P57	1	175	01.07.2020
...

Отношение «Препараты»

Артикул препарата (PK)	Упаковка	Название препарата
P48	блистер	парацетамол
P49	картон	парацетамол
P55	блистер	<u>панадол</u>
P56	стекло	<u>панадол</u>
P57	порошок	<u>панадол</u>
I12	стекло	йод
I13	ампулы	йод
...

Третья нормальная форма

3.3

Третья нормальная форма (3НФ)

- Отношение, которое находится в первой и во второй нормальных формах и не имеет атрибутов, не входящих в первичный ключ, которые находились бы в транзитивной функциональной зависимости от этого первичного ключа [Конноли, Бегг]

Отношение «СВЕДЕНИЯ ОБ ИЗУЧАЕМЫХ ДИСЦИПЛИНАХ»

КОД СВЕДЕНИЙ (РК)	ФИО СТУДЕНТА	КОД ДИСЦИПЛИНЫ	ДИСЦИПЛИНА
1	Иванов И.И.	17	Алгебра
2	Петров П.П.	17	Алгебра
3	Петров П.П.	54	Информатика
4	Сидоров С.Г.	54	Информатика
5	Игнатова П.В.	25	Базы данных
6	Воронова А.М.	17	Алгебра

Транзитивная зависимость: Код сведений → → Дисциплина

Код сведений → Код дисциплины

Код дисциплины → Дисциплина

Третья нормальная форма (3НФ)

Примечание: если в отношении имеется функциональная зависимость одного или нескольких неключевых атрибутов от другого неключевого атрибута, то в данном отношении присутствует транзитивная функциональная зависимость.

Отношение «СВЕДЕНИЯ ОБ ИЗУЧАЕМЫХ ДИСЦИПЛИНАХ»

КОД СВЕДЕНИЙ (РК)	ФИО СТУДЕНТА	КОД ДИСЦИПЛИНЫ	ДИСЦИПЛИНА
1	Иванов И.И.	17	Алгебра
2	Петров П.П.	17	Алгебра
3	Петров П.П.	54	Информатика
4	Сидоров С.Г.	54	Информатика
5	Игнатова П.В.	25	Базы данных
6	Воронова А.М.	17	Алгебра

Неключевой атрибут 1 → Код дисциплины → Дисциплина ← Неключевой атрибут 2

Транзитивная зависимость присутствует

Третья нормальная форма (3НФ)

- Отношение находится в третьей нормальной форме (3НФ), если оно, во-первых, находится в 2НФ, и, во-вторых, не содержит транзитивных зависимостей [Ролланд, стр. 99].
- Отношение находится в третьей нормальной форме (3НФ) тогда и только тогда, когда отношение находится во 2НФ и все неключевые атрибуты взаимно независимы [citforum.ru].

Удаление транзитивной зависимости

- Если в отношении обнаружена транзитивная зависимость, необходимо удалить участвующие в ней зависимые атрибуты. Затем следует создать новое отношение, в которое войдут удаленные атрибуты, а также детерминант – атрибут, от которого они зависят. Этот атрибут становится первичным ключом нового отношения.

Отношение «СВЕДЕНИЯ ОБ ИЗУЧАЕМЫХ

Код сведений (PK)	ФИО студента	Код дисциплины (FK)
1	Иванов И.И.	17
2	Петров П.П.	17
3	Петров П.П.	54
4	Сидоров С.Г.	54
5	Игнатова П.В.	25
6	Воронова А.М.	17

Отношение «ДИСЦИПЛИНЫ»

Код дисциплины (PK)	Дисциплина
17	Алгебра
54	Информатика
25	Базы данных

Отношение «СВЕДЕНИЯ ОБ ИЗУЧАЕМЫХ ДИСЦИПЛИНАХ» в 2НФ → два отношения в 3НФ: «СВЕДЕНИЯ ОБ ИЗУЧАЕМЫХ ДИСЦИПЛИНАХ» и «ДИСЦИПЛИНЫ»

Нормальная форма Бойса- Кодда

3.4

Отношение «ДОКЛАДЧИКИ»

(Студенты уникально идентифицируются по «№ ЗАЧ. КН.» и по «№ СТУД. БИЛЕТА»)

№ ЗАЧ. КН.	№ ДОКЛАДА	№ СТУД. БИЛЕТА	ГРУППА
1254	3	А-45ПВД	921
2311	3	В-35ПМ	922
1402	5	А-54ИВТ	923
1254	4	А-45ПВД	921
1402	2	А-54ИВТ	923
2311	5	В-35ПМ	922

Функциональные
зависимости:

№ ЗАЧ. КН. → № СТУД. БИЛЕТА
№ СТУД. БИЛЕТА → № ЗАЧ. КН.

Потенциальные
ключи:

{№ ЗАЧ. КН.; № ДОКЛАДА}
{№ СТУД. БИЛЕТА; № ДОКЛАДА}

Составной РК

Отношение «ДОКЛАДЧИКИ»

№ ЗАЧ. КН.	№ ДОКЛАДА	№ СТУД. БИЛЕТА	ГРУППА
1254	3	А-45ПВД	921
2311	3	В-35ПМ	922
1402	5	А-54ИВТ	923
1254	4	А-45ПВД	921
1402	2	А-54ИВТ	923
2311	5	В-35ПМ	922

- **1НФ:** все значения атрибутов атомарны
- **2НФ:** отсутствуют неключевые атрибуты, зависящие от части сложного (составного) ключа
- **3НФ:** не содержит транзитивных зависимостей (все неключевые атрибуты взаимно независимы)

Отношение «ДОКЛАДЧИКИ»

№ ЗАЧ. КН.	№ ДОКЛАДА	№ СТУД. БИЛЕТА	ГРУППА
1254	3	А-45ПВД	921
2311	3	В-35ПМ	922
1402	5	А-54ИВТ	923
1254	4	А-45ПВД	921
1402	2	А-54ИВТ	923
2311	5	В-35ПМ	922

- **Аномалия обновления (UPDATE):** для замены значения атрибута «№ СТУД. БИЛЕТА» (например, при утере студенческого билета) необходимо отыскать все кортежи, соответствующие данному студенту. Иначе будет нарушена FD № ЗАЧ. КН. → № СТУД. БИЛЕТА, и база данных окажется в несогласованном состоянии.

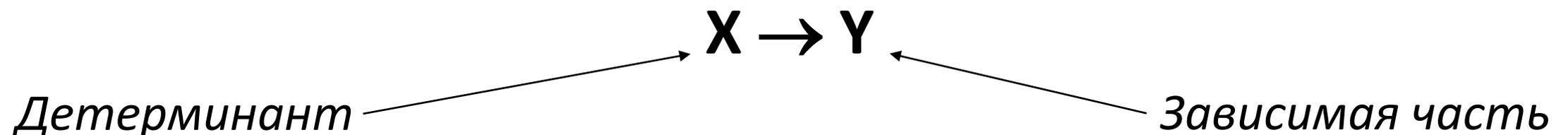
Причиной отмеченных аномалий является то, что в требованиях 2НФ и 3НФ не требовалась минимальная функциональная зависимость от первичного ключа атрибутов, являющихся компонентами других возможных ключей.

Проблему решает нормальная форма, которую исторически принято называть нормальной **формой Бойса-Кодда**, которая является уточнением 3НФ в случае наличия нескольких перекрывающихся возможных ключей.

Нормальная форма Бойса-Кодда

Отношение находится в **нормальной форме Бойса-Кодда**, если все его детерминанты являются потенциальными ключами [Ролланд, стр. 103].

Отношение находится в **нормальной форме Бойса-Кодда** тогда и только тогда, когда детерминанты всех его функциональных зависимостей являются потенциальными ключами [Дейт, стр. 480].



Возможный вариант декомпозиции отношения «ДОКЛАДЧИКИ» для приведения к НФБК

СТУДЕНТЫ

№ ЗАЧ. КН.	№ СТУД. БИЛЕТА
1254	А-45ПВД
2311	В-35ПМ
1402	А-54ИВТ
1254	А-45ПВД
1402	А-54ИВТ
2311	В-35ПМ

ДОКЛАДЫ

№ ЗАЧ. КН.	№ ДОКЛАДА	ГРУППА
1254	3	921
2311	3	922
1402	5	923
1254	4	921
1402	2	923
2311	5	922

Устранены аномалии (например, обновления)

Замечания

- Все отношения, находящиеся в НФБК, находятся в ЗНФ.
- НФБК – более сильная форма ЗНФ.
- Если в отношении имеется только один потенциальный ключ, то формы ЗНФ и НФБК эквивалентны.
- Отношение может находиться в ЗНФ и не находиться в НФБК только в том случае, когда в нем существует несколько потенциальных ключей и хотя бы один из них – составной.

Четвертая нормальная форма

3.5

Четвертая нормальная форма (4НФ)

- Отношение находится в **четвертой нормальной форме (4НФ)** в том и только в том случае, когда оно находится в НФБК, и все многозначные зависимости являются функционально зависимыми с детерминантами – возможными ключами данного отношения [<http://citforum.ru>].
- Отношение приведено к 4НФ, если оно приведено к 3НФ и в нем отсутствуют многозначные зависимости [Роб, стр. 252].

Терминология для 4НФ

- Многозначная зависимость (multi-valued dependency, MVD) – Рон Фейджин, 1971г.
- Многозначная зависимость – обобщение понятия «функциональная зависимость»
- **Лемма Фейджина** о выполнении в отношении многозначной зависимости
- **Теорема Фейджина** о возможности декомпозиции отношения, содержащего многозначную зависимость, на два отношения без потерь (обеспечивает основу для декомпозиции отношений, удаляющей «аномальные» многозначные зависимости, с приведением отношений в четвертую нормальную форму).

Доменно-ключевая нормальная форма

3.6

Доменно-ключевая нормальная форма (ДКНФ)

Предложена Фейджином (Фейгином)

Не определяется в терминах функциональных зависимостей, многозначных зависимостей или зависимостей соединения

Отношение находится в **ДКНФ** тогда и только тогда, когда каждое наложенное на него ограничение является логическим следствием ограничений доменов и ограничений ключей, наложенных на данное отношение

Доменно-ключевая нормальная форма (ДКНФ)

Ограничение домена – это ограничение, предписывающее использование для определенного атрибута значений только из некоторого заданного домена.

Ограничение ключа – это ограничение, утверждающее, что некоторый атрибут или комбинация атрибутов представляет собой потенциальный ключ.

Пятая нормальная форма

3.7

Отношение находится в **пятой нормальной форме** (иначе — в проекционно-соединительной нормальной форме) тогда и только тогда, когда каждая *нетривиальная* зависимость соединения в нём *определяется* **потенциальным ключом** (ключами) этого отношения

Зависимость соединения $*\{A, B, \dots, Z\}$ *определяется* **потенциальным ключом** (ключами) тогда и только тогда, когда каждое из подмножеств **A, B, ..., Z** множества атрибутов является **суперключом** отношения

Условие «каждое из подмножеств **A, B, ..., Z** множества атрибутов является **суперключом** отношения» можно эквивалентно сформулировать так: «каждое из подмножеств **A, B, ..., Z** множества атрибутов включает некоторый **потенциальный ключ** отношения»

Свойства 5НФ

- Любое отношение в 5НФ автоматически находится также в 4НФ и, следовательно, во всех других нормальных формах
- 5НФ является окончательной нормальной формой (в контексте операций проекции и соединения)
- Рональд Фейгин в 1979 г. показал, что любая переменная отношения может быть подвергнута декомпозиции без потерь на эквивалентный набор переменных отношения в 5НФ, т.е. 5НФ всегда достижима. Однако К. Дейт отмечает, что задача определения всех зависимостей соединения для отношения может оказаться очень сложной, а по поводу отношения можно утверждать, что оно находится в 5НФ, только при условии известности всех его потенциальных ключей и всех его зависимостей соединения.

Свойства 5НФ

- Очень редко отношение, находящееся в 4НФ, не соответствует 5НФ. Это те ситуации, в которых реальные правила, ограничивающие допустимые комбинации атрибутов, никак не выражены в структуре отношения (см. пример ниже). В таком случае, если отношение не приведено к 5НФ, бремя обеспечения логической целостности данных отчасти перекладывается на приложение, отвечающее за добавление, удаление и изменения данных. В этом случае существует риск возникновения ошибок. Пятая нормальная форма исключает возникновение таких аномалий.

Шестая нормальная форма

3.8

Отношение находится в шестой нормальной форме тогда и только тогда, когда она удовлетворяет всем нетривиальным зависимостям соединения.

Отношение находится в 6НФ тогда и только тогда, когда она неприводима, то есть не может быть подвергнута дальнейшей декомпозиции без потерь. Каждая переменная отношения, которая находится в 6НФ, также находится и в 5НФ.

ОБЩАЯ СХЕМА ПРОЦЕДУРЫ НОРМАЛИЗАЦИИ

3.9

1. Отношение в 1НФ следует разбить на такие проекции, которые позволят исключить все функциональные зависимости, не являющиеся неприводимыми. В результате будет получен набор отношений в 2НФ
2. Полученные отношения в 2НФ следует разбить на такие проекции, которые позволят исключить все существующие транзитивные функциональные зависимости. В результате будет получен набор отношений в 3НФ
3. Полученные отношения в 3НФ следует разбить на проекции, позволяющие исключить все оставшиеся функциональные зависимости, в которых детерминанты не являются потенциальными ключами. В результате такого приведения будет получен набор отношений в НФБК
4. Полученные отношения в НФБК следует разбить на проекции, позволяющие исключить все многозначные зависимости, которые не являются также функциональными. В результате будет получен набор отношений в 4НФ
5. Полученные отношения в 4НФ следует разбить на проекции, позволяющие исключить все зависимости соединения, которые не определяются потенциальными ключами. В результате будет получен набор отношений в 5НФ

Список литературы

1. Коннолли, Томас; Бегг, Каролин. Базы данных. Проектирование, реализация и сопровождение. Теория и практика. 3-е издание. : Пер. с англ. — М. : Издательский дом "Вильямс", 2017. — 1440 с. : ил.
2. Дейт, К. Д. Введение в системы баз данных : Пер. с англ. / К. Д. Дейт. - 8-е изд. - Москва : Диалектика, 2019. - 1328 с. : ил.