

НАУЧНО-ИССЛЕДОВАТЕЛЬСКАЯ КОРПОРАЦИЯ ИТМО



Факультет программной инженерии и компьютерной техники

Информационные системы и базы данных.

Лабораторная работа №2

Вариант 312621

Выполнил:

студент группы Р33301

Нестеров Иван Алексеевич

Преподаватель:

Гаврилов Антон Валерьевич

г. Санкт-Петербург

2022 г.

Задание:

Привести описанное в лабораторной работе №1 отношение к 3NF, затем к BCNF, описать изменения функциональных зависимостей после нормализации, порассуждать о полезности денормализации.

1NF. Ни в одной таблице нет дублирующих строк, как минимум потому, что в каждом из них есть поле ID, которое является уникальным для каждой строки. В ячейках таблицы нет значений, которые, к примеру, можно разбить на несколько столбцов. Каждая из ячеек максимально "узко" описывает некоторый атрибут данных. А так как данные нельзя разбить на части, то они атомарны, что и необходимо для данной нормальной формы. Кроме того, как некоторое следствие из предыдущего пункта, можно указать, что ни в одной из ячеек таблиц, описывающих мою предметную область, нет составных значений вроде массива или списка. И разумеется, все данные в определенном столбце являются данными одного типа. Это следует из ограничений типов данных на столбец при создании таблицы и является справедливым утверждением по смыслу - в некоторый столбец одной из моих таблиц нельзя добавить данные, которые будут не схожи со своими "соседями по столбцу". Исходя из всего вышесказанного, можем сделать вывод что таблица нормализована к первой нормальной форме.

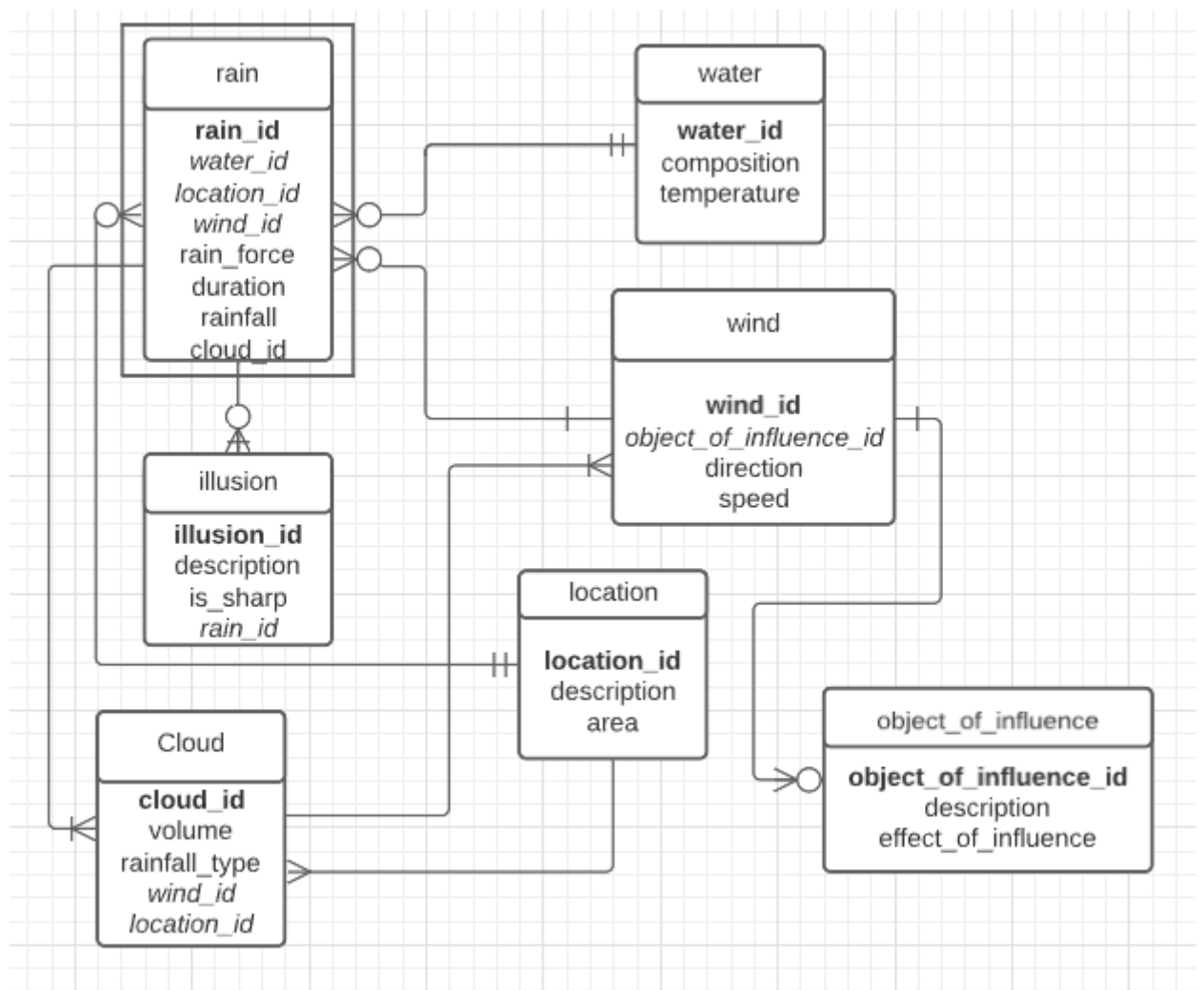
2NF. Таблицы действительно находятся в 1NF. Кроме того, как видно из диаграммы, каждая из сущностей имеет искусственный атрибут ID, который хранится в таблице, это позволяет однозначно определить все данные. В моей таблице нет составных ключей, поэтому нет необходимости проверять то, что каждая запись таблицы однозначно определяется лишь по полному ключу: нельзя идентифицировать какую-то запись лишь по части ключа. Но это требование я запомнил и в случае создания таблиц с составным ключом, буду использовать это знание.

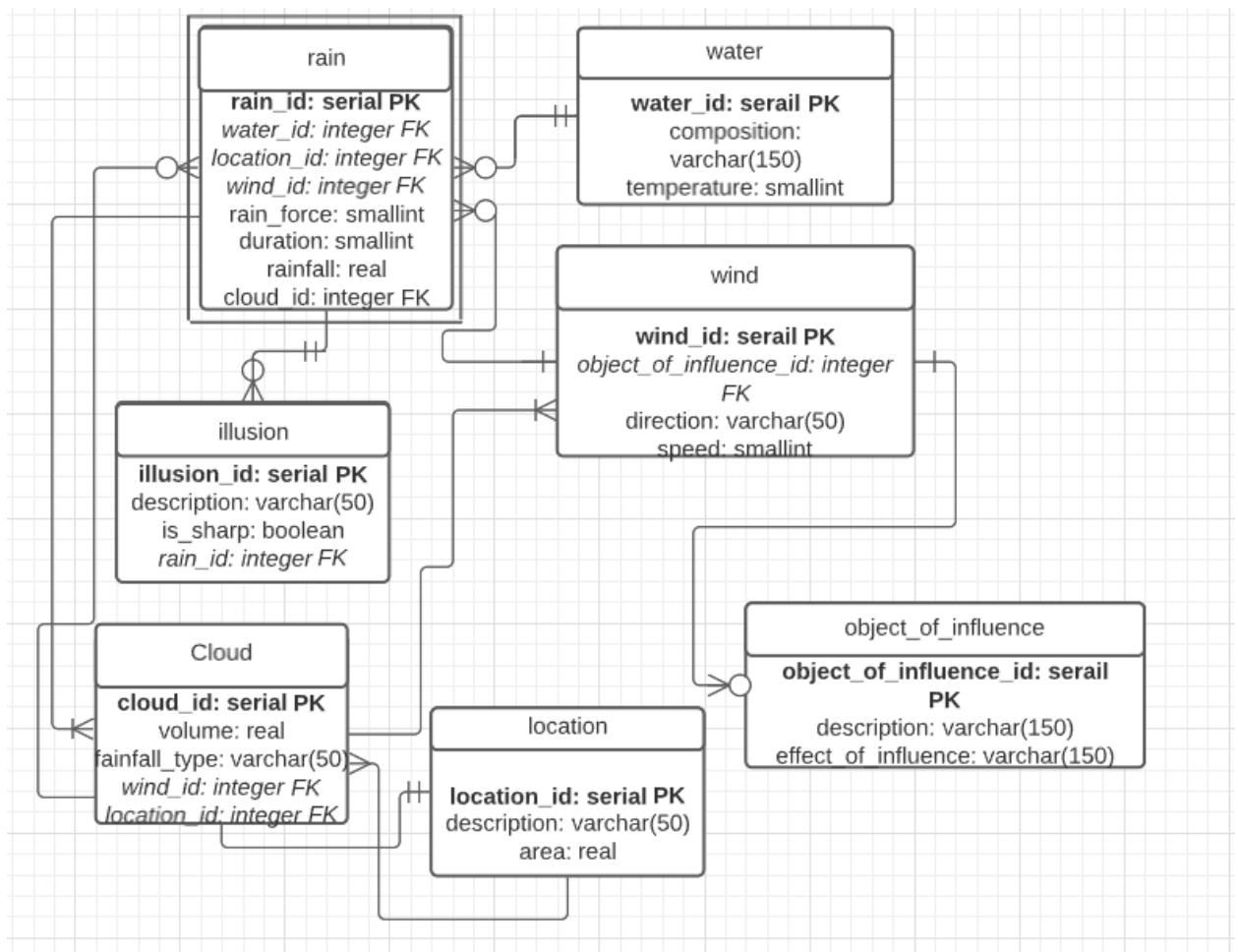
3NF. Подвергнув таблицу дальнейшему анализу, заметим также, что все неключевые столбцы таблицы лишь хранят данные, для которых они предназначены. Мы не можем по значению какого-то столбца однозначно определить, о какой записи шла речь. К примеру, значения зачастую могут дублироваться в различных записях таблицы (одна и та же скорость у нескольких ветров, несколько дождевых облаков, температуры схожей температуры). Для цели однозначного определения подойдет только ее ключ, корректность которого доказана, так как таблицы уже находятся в 2NF. Если

BCNF. Приятная особенность спроектированной мной таблицы – это то, что в ней нет составных ключей. Конечно, в определенных ситуациях они полезны,

но для описания моей предметной области я посчитал их избыточными. А значит, мои таблицы, находясь в 3NF, «автоматически» находятся и в BCNF.

Отразим полученную 3NF (и BCNF, следовательно, тоже):





Порассуждаем теперь о денормализации. Избыточными можно сделать данные об описании облаков. Сейчас данные приведены к BCNF, однако можно, в теории допустить, ситуацию, когда под физически одно существующее облако будет создано несколько записей в соответствующей таблице. Предметная область, как мне кажется, может допускать такую опцию, так как несколько людей, увидев одно и то же облако, могут по-разному его оценить. Над одним человеком это облако проливается дождем и кажется большим, а другой человек, смотря на него издалека, может посчитать, что оно меньше по объему, а не чувствуя дождь, решить, что облако не дождевое. Локация тоже может быть определена не совсем корректно, ведь человек не спутник. Поэтому в данной предметной области такая денормализация допустима, однако если реализовать предметную область для некоторой синоптической системы, то, конечно, такое «разночтение облаков» станет недопустимым.