

Информационная избыточность – термин из теории информации, означающий превышение количества информации, используемой для передачи или хранения сообщения, над его информационной энтропией.

Информационная энтропия – это мера неопределенности информации, неопределенность появления какого-либо символа.

Избыточность ведет к аномалиям.

Аномалии в базе данных – это проблемы связанные с обработкой информации, а точнее с удалением данных из базы данных, с модификацией данных в таблице базы данных и аномалия добавления данных в базу данных.

Три аномалии:

- Аномалия включения – это проблема, связанная с добавлением данных в базу данных;
- Аномалия модификации – это проблема, связанная с изменением данных в базе данных;
- Аномалия удаления – это проблема, связанная с удалением данных в базе данных.

Пример таблицы:

Код преподавателя	Преподаватель	Предметы	Код предмета
1	Иванов	Мат. ан.	1
2	Петров	Химия	2
3	Сидоров	Физика	3
4	Иванов	Дискретная математика	4

Аномалия включения:

Преподаватель «Иванов» дублируется это является аномалией включения. Чтобы избавиться от аномалии включения нужно разбить таблицу на две: таблица преподавателей и таблица предметов. Примерно это будет выглядеть так:



Здесь мы разделили общую таблицу, тем самым избавились от аномалии включения и от возникшей информационной избыточности, то есть от дублирования в базе данных. В принципе то, что мы сделали в данный момент – привели базу данных ко второй нормальной форме.

Аномалия модификации:

Следующая проблема, которая может возникнуть из-за избыточности базы данных – это проблема внесения изменений в таблицы базы данных или как ее еще называют – аномалия модификации.

В нашем примере проблема модификации могла бы возникнуть при попытке изменения фамилий преподавателей, например, если бы в этом списке была незамужняя женщина с фамилией Сидорова, то возможно, когда-нибудь она вышла бы замуж и поменяла фамилию, а оператору пришлось бы для каждой записи, в которой имелась фамилия Сидорова заменить на новую фамилию. Это довольно нудная работа. Каждая такая запись или строка таблицы базы данных называется кортежем.

Чтобы избавиться от аномалии модификаций и все связанные с ней проблемы мы можем прибегнуть к предыдущему способу, просто разбиваем одну большую таблицу на две маленьких. То есть, приводим базу данных ко второй нормальной форме или просто нормализуем.

И опять же, таким образом мы избавляемся от дублирования данных в базе данных.

Аномалия удаления:

Проблема удаления данных из базы данных – это еще одна проблема, которая появляется, если данные в базе избыточны ее еще называют аномалия удаления. Проблема удаления данных из базы данных заключается в том, что при удаление одной записи или кортежа из таблицы, относящейся к какому-либо из преподавателей, вместе с записью о преподавателе, из базы данных удалится вся информация о предмете, который вел этот преподаватель.

Решается проблема удаления данных из базы данных очень просто, нормализуем базу данных до второй нормальной формы, то есть разделяем таблицу на две.