

2.2. Представление сущностей

Модель «сущность - связь». Сущность

Изучая модель «сущность-связь», нужно начать с главного фундаментального представления, которым является сущность. Классическое определение: как сказал Аристотель:

«Сущность – это единичное, обладающее самостоятельностью, в отличие от его состояний и отношений, которые являются изменчивыми и зависят от времени, места и связей с другими сущностями».

Аристотель.

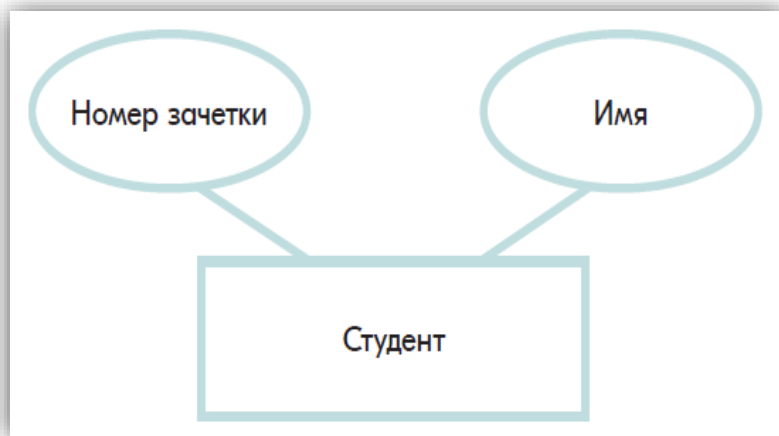
Но мы будем понимать под **сущностью** – **абстрактный или реальный объект определенного вида**, информацию о котором мы собираемся хранить (любой предмет или понятие: студенты, изучаемые предметы, преподаватели, автомобили, книги, банковские операции – всё, что угодно). У каждой сущности должно быть **уникальное имя**, к этому имени, в рамках базы данных, должна применяться одна и та же интерпретация. Набор одинаковых или схожих сущностей **образует множество сущностей**. Важное требование теории множеств – **множества не содержат дубликатов**. Экземпляры сущности должны быть отличимы друг от друга, от других экземпляров, то есть нам нужны какие-то идентификаторы. Идентификаторы являются свойствами сущностей, но кроме идентификаторов, у сущностей есть и другие свойства, которые мы называем атрибутами.

Атрибуты

Атрибут – это отдельная характеристика объекта. Каждый атрибут должен иметь **уникальное имя** (это имя должно быть уникально в рамках сущности). Каждый атрибут должен иметь **свой определенный тип данных**. Сущность может обладать **любым количеством атрибутов**, которое потребуется для ее описания. Значение любого атрибута должно быть **атомарным**.

Пример: представление сущности

Изображая на диаграмме, мы будем соединять сущность с ее атрибутами ненаправленными дугами (сущности мы обводим прямоугольниками, а атрибуты окружаем овалами).



Значение атрибутов берется из соответствующего множества значений, из области определения данного атрибута.

Для разных сущностей одним и тем же именем могут называться разные свойства (например, возьмем атрибут – цвет. Цвет бывает у автомобиля, у собаки, у книги и так далее, и для каждой сущности у атрибута цвет будет свое множество допустимых значений).

При моделировании может возникнуть вопрос: рассматривать ли какой-то элемент в качестве атрибута сущности, или в качестве отдельной сущности (например, адрес может быть атрибутом человека, если нас интересует только место прописки, но также адрес может быть самостоятельным объектом, если мы хотим хранить информацию о том, где человек живет, где он работает, где он бывает и так далее).

Идентификация сущностей

Идентификация – возможность отличимости.

Сущности идентифицируются специальным набором атрибутов, которые называются ключами.

Ключи

Ключ – это один или несколько атрибутов объекта, который позволяет однозначно отличить данный объект от других. Если мы храним информацию о студенте, то ключом может быть номер зачетки (мы знаем, что по номеру зачетки можно однозначно определить студента; может быть номер паспорта, а может быть какая-то сложная комбинация полей, например, фамилия, имя, отчество и дата рождения).

Естественный идентификатор, который свойственен хранимым атрибутам, удается найти не всегда. Если не удастся определить такой естественный ключ, то система может придумать искусственный ключ, который называется «суррогатным».

Три вида идентификации

- естественные ключи;
- ключи «по положению», например, географические или в порядке возникновения во времени;
- суррогатные, то есть искусственные ключи.

Суррогатные

Системе удобно, когда у каждого объекта есть свой уникальный идентификатор, который живет с объектом всю жизнь и не меняется в процессе работы системы. Обычно эти идентификаторы не связаны с реальным миром (например, номер зачетки, номер паспорта, ИНН, который дается человеку один раз) – это всё суррогатные идентификаторы, которые генерируются системой, и с точки зрения человека могут казаться ему естественными ключами, потому что мы можем помнить наизусть номер паспорта и пользоваться номером зачетки, и считать, что это естественные, свойственные нам атрибуты.

Естественные

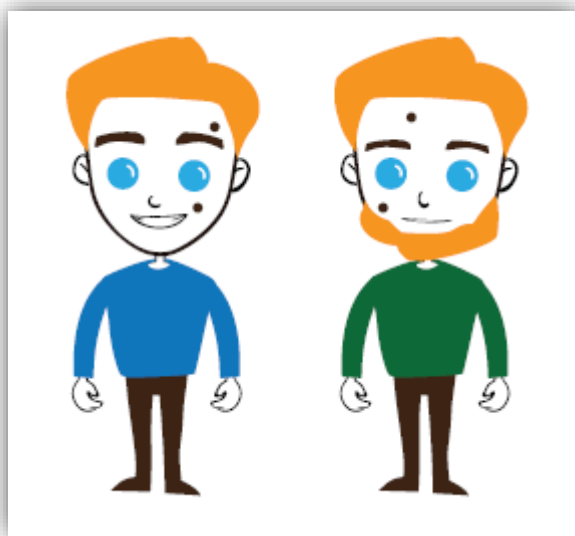
Но с естественными идентификаторами немного сложнее, потому что не всегда удастся найти такой атрибут или комбинацию атрибутов, которые позволяют однозначно определить объект.

Вспомним цитату из Бориса Годунова:

«А ростом он мал, грудь широкая, одна рука короче другой, глаза голубые, волосы рыжие, на щеке бородавка, на лбу другая».

(А. С. Пушкин, «Борис Годунов»)

Когда герой описывал таким образом Гришку Отрепьева, наверное, ему казалось, что таким образом можно однозначно определить человека по данным характеристикам.



Но если взглянуть на пример, представленный на рисунке, то мы, хоть и символично, но поймем, что по такому описанию найти человека в толпе

фактически невозможно. Этот пример говорит о том, что ни один естественный идентификатор не может быть абсолютно надежен (даже если мы рассматриваем номер паспорта, то номер паспорта может измениться у человека, потому что девушки выходят замуж, также в определенном возрасте паспорта меняют; зачетку тоже можно потерять), поэтому зачастую бывает очень трудно найти такой идентификатор, который проживет с объектом всю его жизнь ни разу не изменившись.

По положению

Очень часто в современных информационных системах мы храним данные, например, передаваемые множеством датчиков – это будут просто потоки цифр, которые могут идентифицироваться исключительно идентификатором датчика и каким-то моментом во времени, в который эти показания были переданы.

Пример: ключи в различных предметных областях

Студента можно однозначно идентифицировать по номеру зачетки.

Пассажира поезда – по номеру поезда, дате, номеру вагона и месту.

Гражданина России – по номеру российского паспорта.



Если среди атрибутов в сущности можно выделить несколько возможных ключей, то обычно выбирается наиболее важный, который чаще всего будет использоваться при поиске, этот ключ называется – *первичным*.