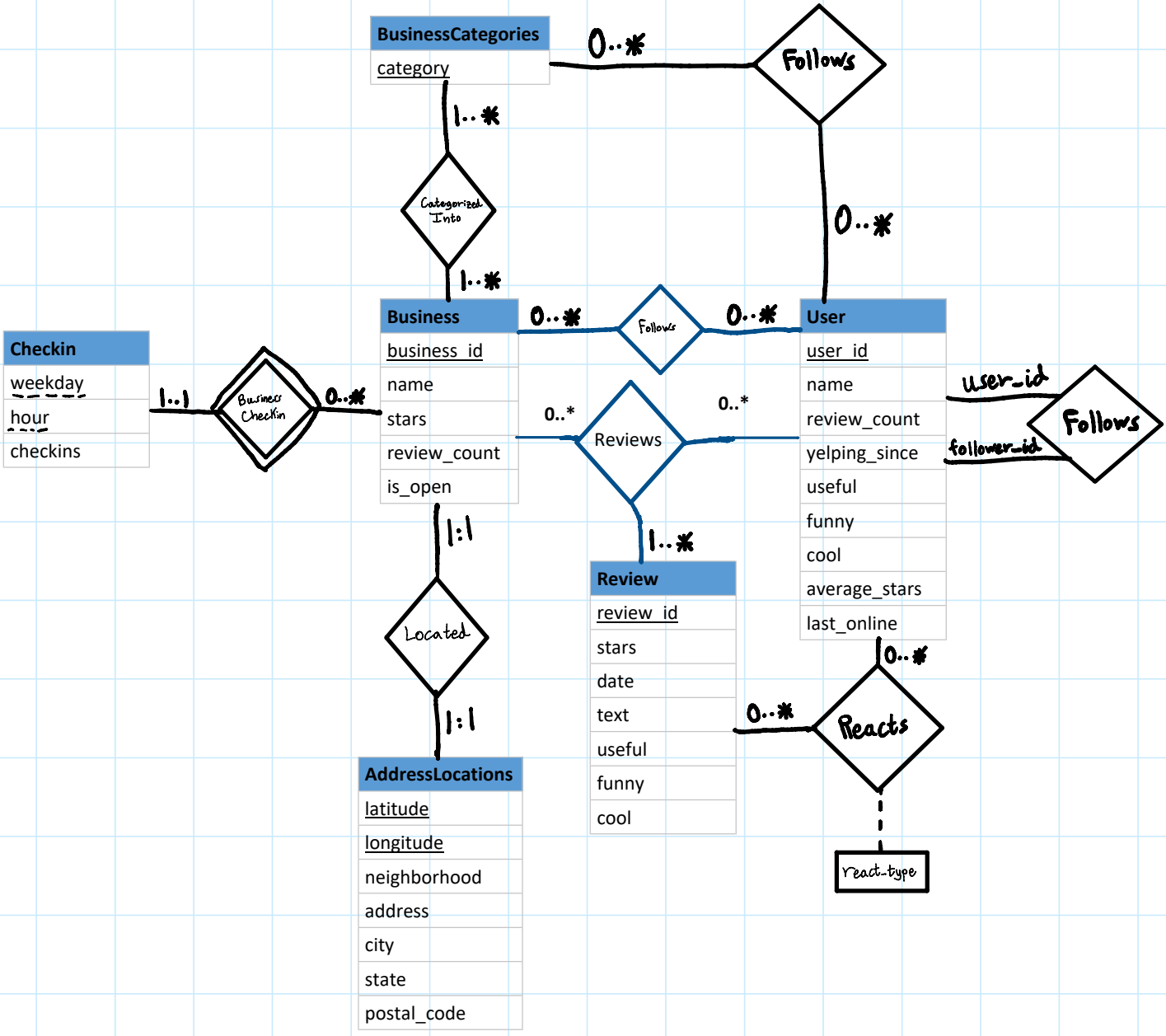
Yelp Database Project

Sunday, March 29, 2020 3:39 PM

o Entity-Relationship Model:



For our project, we designed a database for the Yelp social network based on the Kaggle dataset found in our README. Yelp is a social network where users can post reviews about businesses, follow other topics or users, view a feed of reviews and business details, and react to other users' reviews.

The ER Model above shows the structure of the Yelp social network database in a real world problem space. Each entity contains basic information about the item it is describing. For example, Business contains the name attribute and the number of reviews it has received (review_count). Another important detail is that the useful/funny/cool attributes in the Review and User entities are slightly different. In User, these attributes describe the number of reactions the user has received on all of their reviews. However, in Review, these attributes describe the number of reactions a review has received. Another notable entity is Checkin because it is a weak entity defined by a business, with the discriminator attributes weekday and hour; A Checkin defines how many people entered the Business at a certain time of day.

As for the relations in the ER model, the Review ternary relation shows that a User can review a Business; The Review has a large number of attributes and it can be reacted to as well (by a User), which is why it is a separate entity. Moreover, a user can follow other users, multiple businesses, and multiple categories of businesses (e.g. food or automotive). Furthermore, a Business identified by a unique business_id must have one and only one location, identified by the latitude and longitude in AddressLocations. The CategorizedInto relation shows that a Business can be a part of multiple categories and a BusinessCategory can include multiple businesses.

When mapping the ER model to the relational schema, the "Follows" relations are converted into tables in the schema, containing the primary keys of the joining entities as the composite primary key and the only attributes. These relations become the UserFollowers, BusinessFollowers, CategoryFollowers tables in our SQL schema. Similarly, the "Reacts" relation becomes the ReviewReacts table.