# CSC 370

# Activity Worksheet:
# Entity-Relationship Diagram (ERD) Basics

Sean Chester

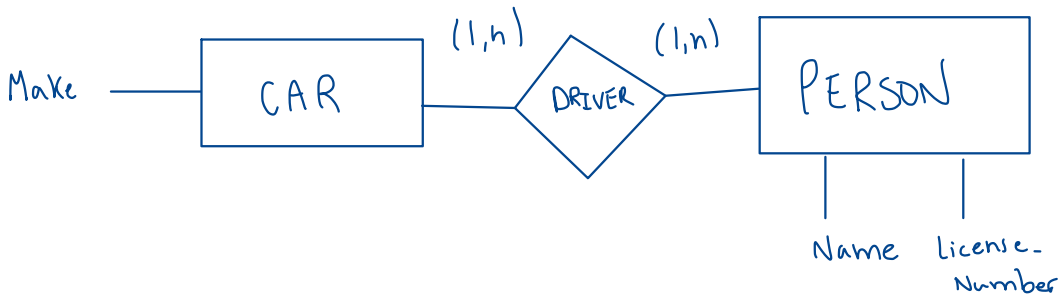
Fall 2022

## Notes

This worksheet provides a series of practice questions for constructing conceptual models and ERD's.
The style of the question varies and thus no model solution is provided for this worksheet.

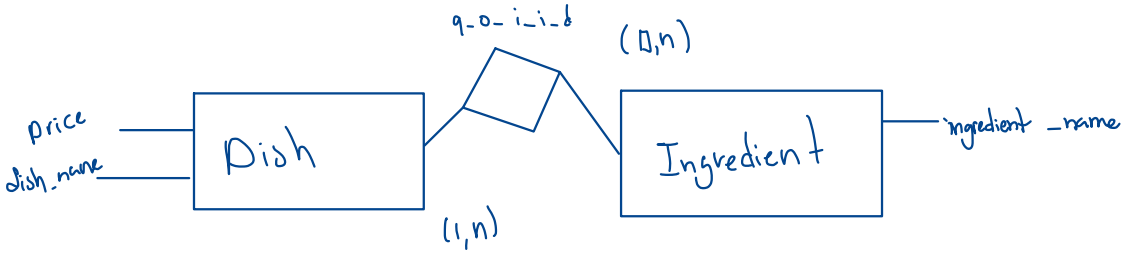
# Questions

1. Create an entity-relationship diagram that has two entity-sets, three attributes, and one (many-many) relationship.



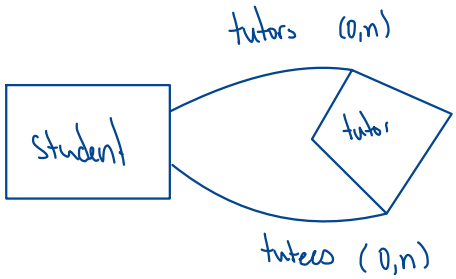

2. Create an entity-relationship diagram for a database of McDonald's delicacies that includes the following elements:

   Dish, Ingredient, *dish_name*, *ingredient_name*, *quantity_of_ingredient_in_dish*, *price*. Include and justify with one sentence multiplicities on any relationships.

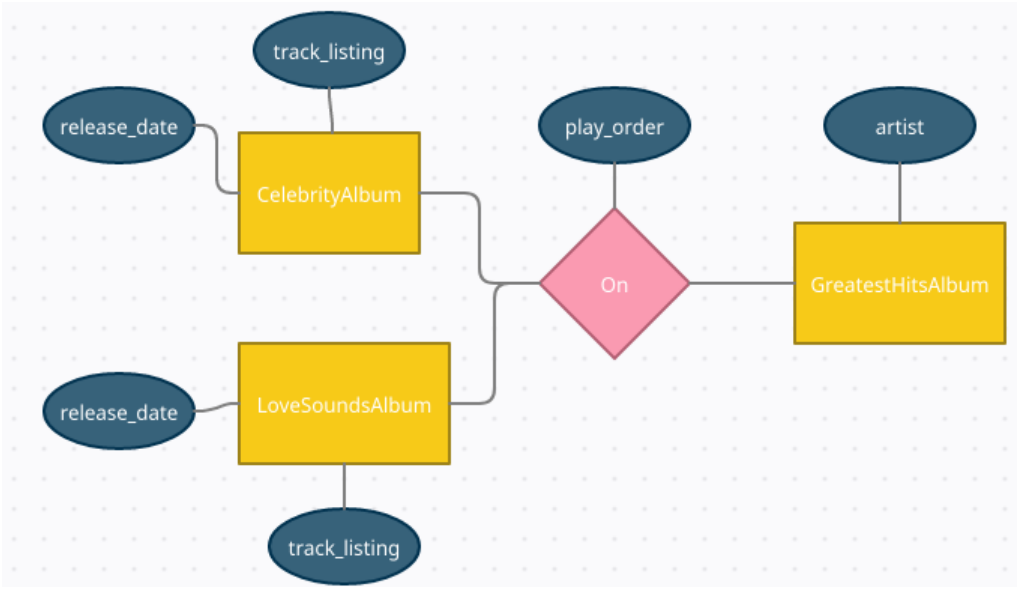



A dish must have at least 1 ingredient, but an unlimited amount of ingredients.

3. Provide an entity-relationship diagram with one entity set and two relationships.




4. A colleague provides you with the following conceptual design for a Justin Timberlake discography database. While remaining sensitive, explain what might be some weaknesses in the design. Moreover, provide an improved conceptual design in the form of an ERD.




Duplication of release_date and track_listing

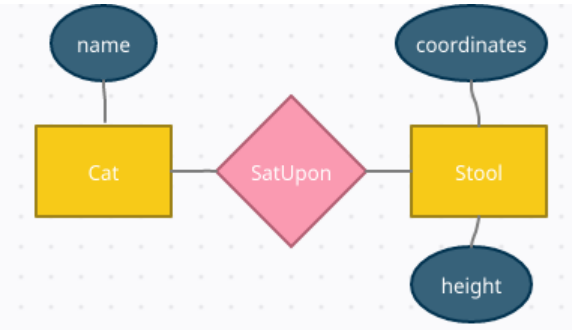
Artist only on Greatest Hits Album

On is vague + Attribute play_order

# Solutions

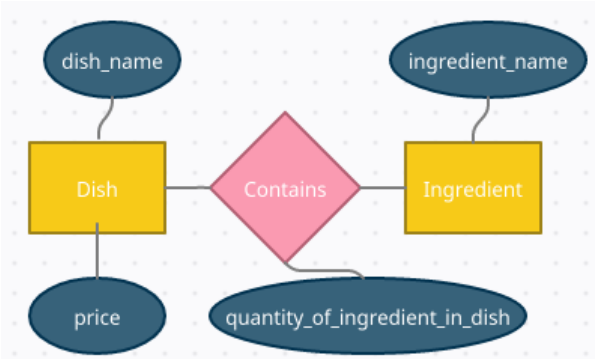
## Question 1

There are infinitely many solutions to this question. It should contain two rectangles, one diamond, and three ellipses. The ellipses can be connected to anything (except each other), the diamond should (probably, except for **very** creative solutions) connect the the rectangles, and hopefully the overall design makes sense (i.e., is "faithful").




## Question 2

There are fewer degrees of freedom here, as the placement of attributes (lower case with underscores) should make sense with respect to the entity sets and relationships.

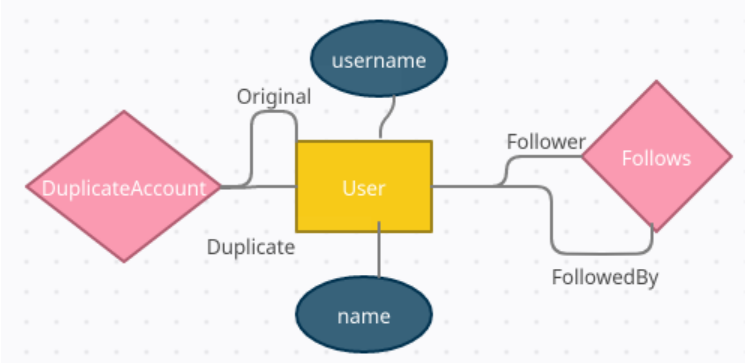



The multiplicity here is many-many because there exists at least one dish (e.g., a Big Mac) that has more than one ingredient (burger and bun) and at least one ingredient that appears in more than one dish (e.g., copious salt appears in everything).

Observe that *quantity_of_ingredient_in_dish* is an attribute of the relationship because it could vary depending on to which dish the ingredient is added and which ingredient is included in the dish.

## Question 3

Here, again, the solution space is infinite, though the "shape" of the ERD is fixed. Moreover, the question is a bit artificial. Nonetheless, an example would be:




Observe that DuplicateAccount might not be many-many depending on assumptions of whether duplicate accounts are all merged into the same original account.

## Question 4

It's always good to point out the aspects that are good:

- *play_order* is a property of the relationship, not the albums, and this is one of the trickier aspects to get correct. So that's well done.

However, there are several things worth reconsidering:

- Is the "artist" attribute necessary? The database only records Justin Timberlake songs. Maybe this represents artists featured on the track, in which case it's not really a property of the Greatest Hits album, but of each song on it?
- The CelebrityAlbum and LoveSoundsAlbum entity sets have the exact same set of attributes. This is a good indicator that either they are subclasses of a broader entity set or, as is more likely in this case, they are actually entities from the same entity set. It might be better to have an Album entity set and make "Celebrity" and "LoveSounds" instances thereof.
- The *track_listing* attribute might be difficult to use effectively. It might be better to use a separate Track entity set and a relationship between Album and Track to indicate the track listing. In fact, we could even use this for the Greatest Hits album. Generally, we don't want attributes that are vectors instead of scalars.
- It isn't common to consider an album as being "on" a greatest hits album, but rather just individual tracks. There could be some tracks from each album that are not on the Greatest Hits album; so, it would be advantageous to show this finer granularity.

A proposed alternative is: