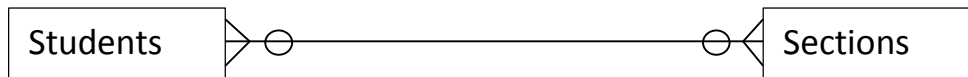


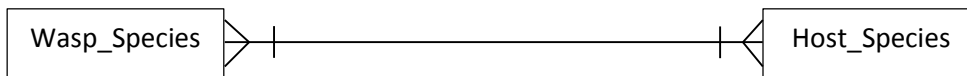
1. (32 points) The following set of brief scenarios describes the relationships between various entity classes. Draw the entity classes (only show entity name – no attributes or identifiers) and the relationship making sure to add notation (use crow's foot notation) that shows reasonable minimum and maximum cardinalities given the implied scenario. All scenarios will only have one relationship. The one relationship may be between a maximum of two entity classes (binary) or between an entity class and itself (recursive). *If you make any assumptions beyond what is stated be sure to note that in your answer.*

Example: *Students can take no sections or may take any number of sections. Each section has anywhere from zero to an unlimited number of students.*

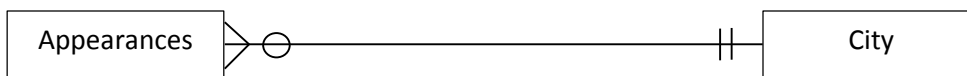
Example Answer:



a. A researcher is studying parasitical (parasitoid) wasp species and the species of animals (usually an insect or arachnid) that host the wasp's parasitical larva¹. Each species of wasp must be related to one or more host species, each host species must be the target of one or more wasp species.

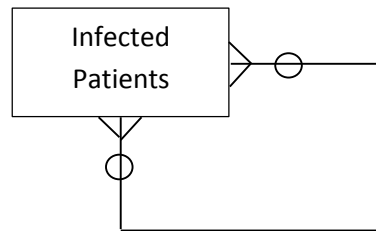


b. A musical group is creating a database of its appearances in various cities. Each city could have zero or any number of appearances. Each appearance would be in one and only one city.

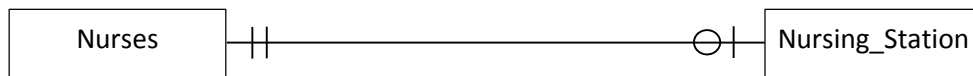


¹ More information not needed to answer the question: Inside the host the egg hatches into a larva or larvae. The larva feed inside the host until ready to pupate; by then the host is generally either dead or moribund. Depending on its species, the parasitoid then may eat its way out of the host or remain in the more or less empty skin. In either case it then generally spins a cocoon and pupates. As adults, parasitoid wasps feed primarily on nectar from flowers. https://en.wikipedia.org/wiki/Parasitoid_wasp

c. The Center for Disease Control (CDC) is investigating a new virus. They have created a database that includes data on infected patients. They want to determine whether there are any patterns of infection that can be attributed to the infected patients being employed in the same organization as another infected patient. They need to store which infected patients work with each other as a beginning to this investigation. An infected patient may list zero to many other infected patients as co-workers. An infected patient may be listed by zero to many other infected patients as one of their co-workers.

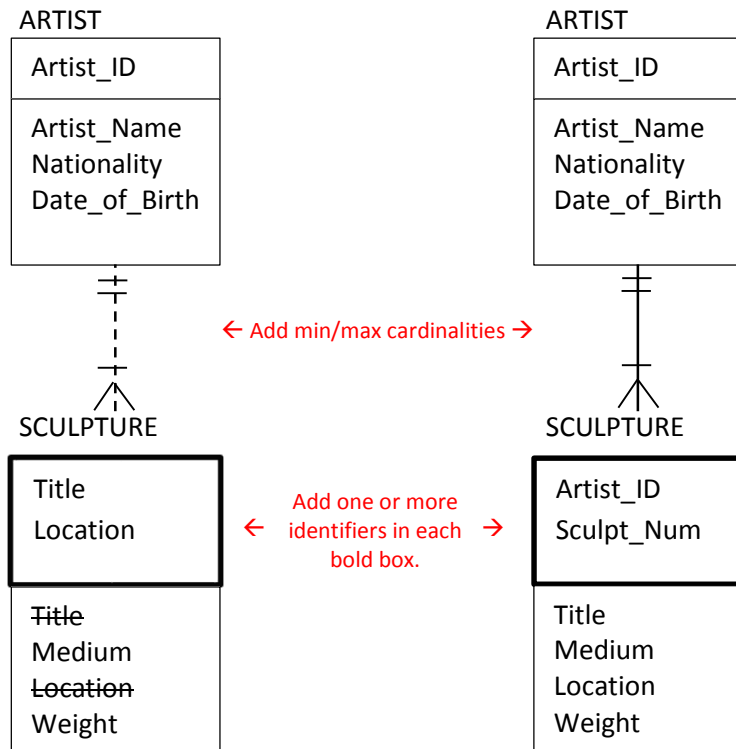


d. A hospital stores data about its nurses and their relationships to nursing stations in the hospital. Although nurses are related to nursing stations in several ways we are interested here in determining which nurse is the primary administrator of specific nursing stations. By policy, a nurse can only oversee a maximum of one nursing station. Each nursing station, likewise, can only have one primary administrator. All nursing stations must also have at least one primary administrator. Of course, many of the nurses in the database are not administrators at all.



2. (24 points total)

a. (10 pts) Consider two entities: ARTIST and SCULPTURE. These entities are related to each other. ARTISTS can have one to many SCULPTURES in this database. Each SCULPTURE is always related to exactly one ARTIST. SCULPTURE is a weak entity in this relationship. Fill in the identifiers for SCULPTURE below (**bold boxes**) to show it as ID-Dependent or Non-ID-Dependent (use the relationship line format below to determine which to make ID-Dependent and Non-ID-Dependent and do not worry about the absence of 'rounded corners' on the weak entity box). You may move any of the non-identifier attributes listed for SCULPTURE to the identifier box if desired. Finally, insert the minimum and maximum cardinalities described on both of the relationships.



The identifier of SCULPTURE in the non-ID-Dependent diagram (left – dotted line) could be derived from the attributes of SCULPTURE as shown here or it could be a surrogate identifier that is created, such as Sculpture_ID.

The ID-Dependent diagram's SCULPTURE identifier must include Artist_ID (that is the definition of ID-Dependent) and also must include a created index (such as Sculpt_Num) or other attribute from SCULPTURE (Title would be a good choice).

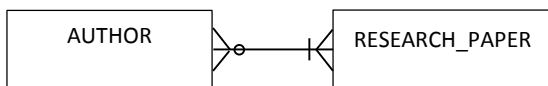
Grading:

Cardinalities: 2pts

Identifiers: 8pts (4 pts each)

b. (14 pts) For each of the two many-to-many relationships described below determine whether the relationship should be changed into an associative pattern (utilizing an associative entity). The many-to-many relationships are displayed for you and you should assume they are correct as shown (except in regards to the potential associative entity requirement). Each also lists a few attributes that would have to be stored somewhere in the data structure of the diagram. Some of the attributes may require it to be converted into an association pattern. First, determine if each relationship requires an association pattern or not. Next, below each diagram, **either** write “This requires no association pattern” **OR** redraw the relationship as the association pattern it requires. Note that you do not have to show the attributes on the diagram.

i. Authors have one to many research papers and each research paper can have zero (anonymous) to many authors. Attributes that need to be stored include: Author_Name, Paper_Publication_Date, Paper_Title.



“This requires no association pattern.”

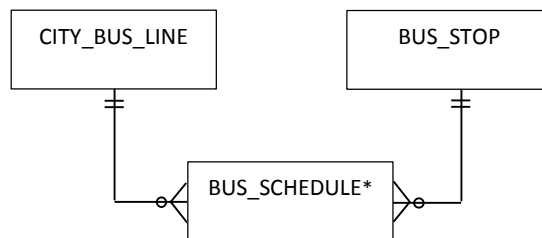
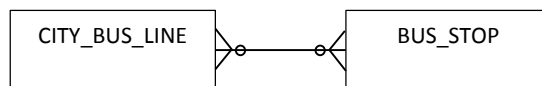
Author_Name can be stored in **AUTHOR**.

Paper_Publication_Date, and **Paper_Title** can be stored in **RESEARCH_PAPER** (the publication date would not vary by Author so it is not multivalued).

Grading:

- Identifying the correct association pattern (or lack of it) 3 pts.
- Drawing the diagram correctly 4 pts.

ii. A city has many bus lines. City bus lines (or routes) stop at zero to many bus stops. Each bus stop has zero to many bus lines stopping at it. Attributes that need to be stored include: Bus_Line_Name, Bus_Stop_Address, Scheduled_Arrival_Time.



* Name may vary.

Scheduled_Arrival_Time requires you to know both the bus line and the bus stop so this requires an association pattern. **Bus_Line_Name** can be stored on **BUS_LINE**, **Bus_Stop_Address** can be stored on **BUS_STOP**.

Grading:

- Identifying the correct association pattern (or lack of it) 3 pts.
- Drawing the diagram correctly 4 pts.

3. (46 points) draw an E-R diagram to represent this database.

- Where and if a relationship needs to be an association entity, show correctly on diagram.
- Show all minimum and maximum cardinalities.
- Do not show attributes for entities. Some attributes are given as examples in the problem so that you understand what is being stored and the structures they require. These should not be shown on diagram.
- Do not use any structures that imply multi-valued attributes would exist.
- List any non-obvious assumptions you made.
- *ID-Dependent vs. non-ID-Dependent relationships do not have to be distinguished for this diagram (will not be a grading point).*

Background (may help to understand the scenario but use the problem statement only in drawing your diagram):



This problem is inspired by a real organization (<http://transparentcalifornia.com/agencies/salaries/school-districts/>) and their real database found in the news.

From the news article (<http://www.latimes.com/local/lanow/la-me-ln-database-public-school-20140723-story.html>):

“Last year, James Hammond, the superintendent of the Montclair-Ontario Unified School District in the Inland Empire, was paid \$492,077. Jonathan Eagan, the principal of a junior high school in the Bay Area city of Martinez made \$279,669. And 31 custodians at California public schools were paid more than \$100,000 in 2013. That is a sample of statistics found in a newly released online database that allows users to search and download detailed employee compensation figures for superintendents, teachers, principals and other staff members at school districts across the state.

The figures were added to Transparent California, which compiles compensation data for a variety of public sector employees. The education section of the website is composed of more than 581,000 individual compensation records from last year for about two-thirds of districts statewide.”

THE PROBLEM STATEMENT IS ON NEXT PAGE

Nothing written on this page will be graded.

3. Problem Statement:

The database generally focuses on storing data about different types of employees and their total annual pay and benefits. The system is designed to maintain multiple years of data.

This database stores data about administrators, teachers, and 'other' (custodial services, councilors, etc.) employees throughout all of California's school districts. Most of the data about each of these categories (administrators, teachers, others) is the same such as names, address, phone, etc. However, each category does have some unique data being stored about it. It is also noteworthy that an employee can serve as more than one type of employee. For instance, a teacher on an administrative track may serve as both a teacher and an administrative intern during the same year.

The system needs to keep track of the types and descriptions of pay and benefit items. Pay and benefit items can be categorized as regular pay, overtime pay, and benefits. Each of these have some common attributes with each other but each needs to also store some data that are unique to that type. A particular pay and benefit item can only be in one of the categories. For example, a pay item cannot be both regular and overtime pay simultaneously.

Annual pay is stored in a line item model. Every employee can be linked to zero (for a new employee) to many total pay and benefit records (one per year) and each of these records are related to one and only one employee. Each total pay and benefit record must be related to at least one pay and benefit line item and usually will be related to many. Each pay and benefit line item is related to exactly one total pay and benefit record. Pay and benefit items are related to zero to many pay and benefit line items and each of these line items are related to one and only one pay and benefit items.

Other requirements:

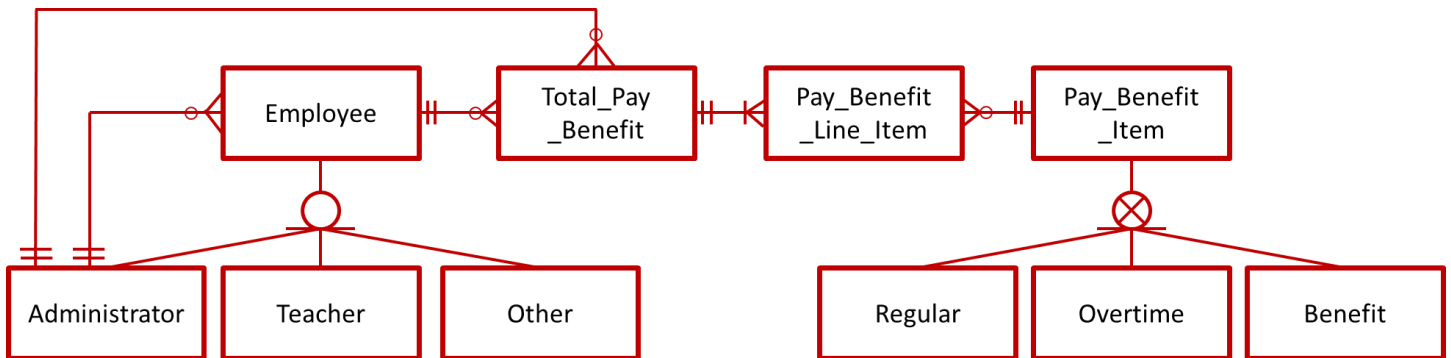
- Since these records are to be public, an administrator must review and signoff on the calculations for each total pay and benefit record. The administrator reviewing each record is to be stored in the database. An administrator can review zero to many total pay and benefit records and each record must be reviewed by one and only one administrator.
- The direct administrator of each employee (including most but not all of the administrators) must be stored. Only administrative categorized employees can serve as administrators. Each administrator can be the direct administrator² of zero to many employees and each employee can have one and only one direct administrator.

Nothing written on this page will be graded.

² A direct administrator would be directly above the employee in the organizational chart.

NAME _____

3. Draw your diagram for question three here. If you remove this sheet write your name where indicated.



Grading:

- **Entities (10):** 15 pts.
- **Has-a Relationships (5):** 10 pts.
- **Is-a Relationships (2):** 10 pts. (8 pts structure, 2pts exclusive/inclusive)
- **Cardinalities (20):** 10 pts. (0.5 each)