NAME Key EMAIL _____@kent.edu

Design In-Class Practicum (100 points total)

1. (30 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. Since we are not concerned with identifiers at this point, notation to designate the relationship as ID-Dependent (or not) is not necessary.*

Example: *A college is creating a database to keep track of the students enrolled in specific sections of a course. Students can take no sections or may take any number of sections. Each section has anywhere from zero to an unlimited number of students.*

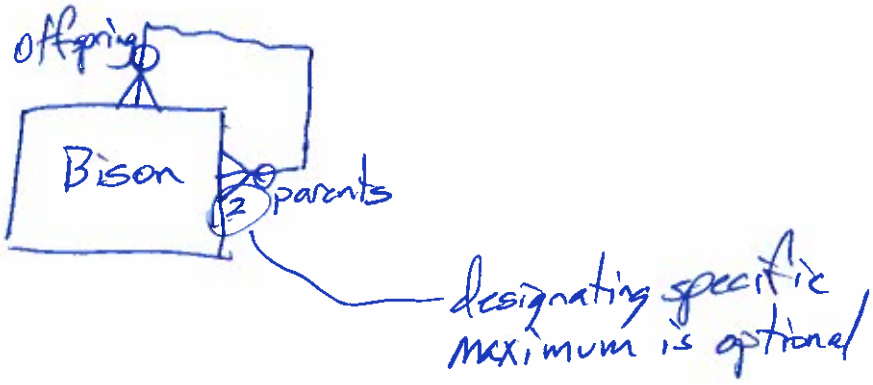
Example Answer:

a. In a university, faculty members are appointed to exactly one and only one university department. There are university departments that have no faculty assigned to it (e.g. administrative departments). Each department may have many appointed faculty members.

b. In a hospital database we are storing information about individual patients and the genetic diseases they have been diagnosed as having. Each disease may be related to many patients. They want the database to list all known genetic diseases whether or not they have identified a patient who has that disease. Each included patient may be diagnosed with zero genetic diseases or many genetic diseases.

c. A database is being created to store all US Supreme Court nominees and the results of their Senate votes[1]. They will store ALL Senate Vote results in the database (non-Supreme Court votes are being stored for future use in the database). Basic information about the voting session and about the nominee are also to be stored. Each Supreme Court Nominee has, at most, one Senate Vote result. Some nominees exit the process before a Senate Vote is conducted. Each Senate Vote instance will be related to, at most, one supreme court nominee.

d. Yellowstone National Park is creating a database to keep track of the genetics of a herd of endangered bison (also erroneously called American Buffalo). The purpose is to ensure a minimum level of genetic diversity exists in the herd. To accomplish this, genetic testing is conducted on each calf to identify the male parent. The database must store the known parental information of each animal. As you would expect, each animal is related to a maximum of two parents. Not all animals have been tested so some animals have no known parents and some only have identified the maternal parent. Each bison may be the (known) parent of no other bison or may be the parent of many bison.



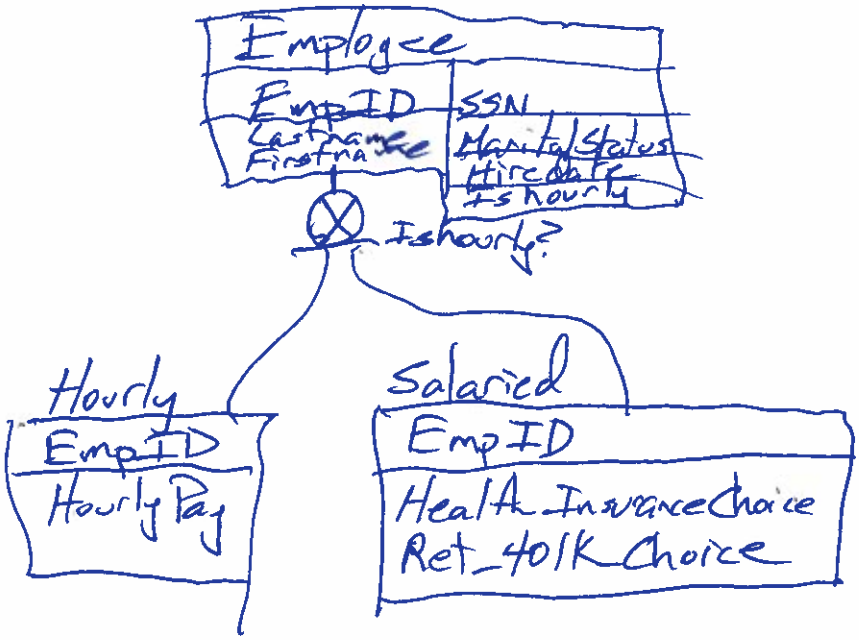
[1] Membership in the United States Supreme Court requires being nominated for the position and then being confirmed by a vote of the US Senate (congress).

2. (24 points)

a. (12 pts) A retail store chain stores information on all of its employees in a common database. Currently, the store has both hourly and salaried employees. An employee cannot be both hourly and salaried. Along with the information stored about every employee, they are also storing data on the hourly employees' hourly pay and the salaried employees' choices on health insurance and retirement. Although their current system of storing this in a single table (shown below) seems to work ok, they have noticed that there are many null values being created due to this structure and are appropriately concerned. They have hired you to redesign this structure so that the number of nulls created would be dramatically reduced. Draw your revised data structure below. Designate it as exclusive or inclusive and add an appropriate subtype discriminator. Note any assumptions that you make.

EMPLOYEE

EmpID
Lastname Firstname SSN MaritalStatus Hiredate HealthInsuranceChoice Ret_401K_Choice HourlyPay



b. (12 pts) We spent some time in class discussing the importance of 'line-item patterns' or 'transaction-processing patterns' (these are synonyms). A library is designing a line-item structure to capture the library's patrons' checkout transactions as part of its "library management and reporting system" database. The line item pattern must contain data relating to **Library Materials**, the **Checkout Transaction**, the **Checkout Line Items**, and the **Patron** borrowing the materials. On the diagram below, fill in the entity names and the most likely minimum and maximum cardinalities for the described scenario.

3. (46 points)

This problem was inspired by a real-world database problem (treezilla.org) found in the news:
http://www.geek.com/science/crowd-sourced-database-seeks-to-catalog-every-tree-in-britain-1559359/

What is Treezilla?

Treezilla is an exciting new platform for citizen science that everyone from school children to university students and the general public can get involved with. The idea is to map every tree in Britain.

An organization (Treezilla) is embarking on a project to catalog every tree in Britain and needs to design a database to hold the data. They will need to store basic data (species code, trunk diameter, GPS coordinates, etc.) about each tree in the database. In addition to the basic data stored about every tree, some trees require additional data specific to their situation. One of these is trees that have been forcibly modified by a municipality for power lines and such (we'll categorize these simply as "Modified"). They need to store the type of modification, the reason modified, and the date modified. A second category of trees requiring additional data are removed trees (for any reason). Trees that have been removed (categorized as "Removed") need to store their date removed, reason for removal, who removed (arborist, owner, municipality, natural process, etc.), and a yes/no attribute on whether the tree was replaced with a new tree. A tree can be both modified and removed. Modification records for removed trees are retained since the modifications may have contributed to the eventual removal.

The data for each tree is related to a list of all tree species. Each tree species can be related to zero to many cataloged trees and each tree is related to exactly one tree species. Additionally, each tree has status records associated with it. Each time a status is recorded it is added to the database rather than overwriting the existing record (so trends in status may be analyzed). Therefore, each tree may have zero (if its status has never been recorded) to many status records. Each status record belongs to one and only one tree. Finally, statuses need a relationship to a list of "Pests and Diseases". Each status record may be related to zero to many pests and diseases records. Each pest and disease record is related to zero to many statuses. When a status record is connected to a pest and disease record the database must also store the date diagnosed, last date examined, and whether the condition is still active or resolved.

Additionally, individuals or organizations may sponsor specific trees in the database. Each sponsor would sponsor one to many trees. Each tree may have zero to many sponsors. Each tree that is sponsored must also store the dates the sponsorship is effective along with any monetary maximums the sponsor is willing to supply – so an individual can limit their responsibility to a maximum of $500, $1000, or whatever amount they choose over the minimum sponsorship of $50.

Draw an E-R diagram to represent this database.

- If necessary, create an association pattern only if a relationship needs additional attributes other than foreign keys.
- 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. These do not need to 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).*

Name_____

Draw your diagram on this page. You may remove this page from the exam to make drawing easier. If separated make sure you write your name.