**IT 350**

**Midterm Exam**

**Kevin Tew**

**Due Monday Oct 17, 2016 11:59 pm**

*Please make your answers easy to read, especially be clear about where you got your information. Clearly you need to be very careful to provide references. Open Internet and Book. Closed neighbor.*

# NAME:\_\_Jacob Parry\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. (2) Define what a SQL aggregate function is and how it works?
   1. **A SQL aggregate function returns a single value, calculated from the values in a column. It groups the values of multiple rows on certain criteria to form a single value.**
   2. **http://www.w3schools.com/sql/sql\_functions.asp**
2. (5) Define Data and Information. What is their relationship? Why are these concepts important? Give an example that illustrates these concepts and their relationship.
   1. **Data: just bare facts. It could be a number, or word, or any kind of fact.**
   2. **Information: Data that has been processed into something meaningful.**
   3. **You can have a number 27, but that doesn’t mean much by itself. That would be data. But if you add a header to that number and say “Age”, then the 27 means something much more. It is now information.**
   4. **This was found from our text book.**
3. (10) Define the terms *conceptual model, logical model, physical model,* and *external model.*  Describe how they are related. Can the mappings between them be automated? Give an example from your experience in this class.
   1. **Conceptual Model: A model of a database that identifies all of the entities, attributes, and relationships that are to be represented in the database, and is an accurate reflection of the real world.**
   2. **Logical Model: The logical model is the model that is created when you take your conceptual model, and then tie it to a chosen DBMS so that that your conceptual model will work on the chosen system.**
   3. **Physical Model: The physical model is created by taking the logical model, and then adapting it the physical hardware and devices. There can be several, and the best mapping is then the physical model.**
   4. **External Model: The external model is then how the the users are going to use and think about the data that is going to be used.**
   5. **It is possible to have the mappings between them be automated, but it sometimes the auto-generated mappings need to be modified/edited/deleted because it is not always possible to understand what is the actual meaning is through the different models. The better the conceptual designs, the easier it is to map them.**
4. (5) Define and describe the differences between super keys, candidate keys, primary keys, foreign keys, and composite keys.
   1. **Super key: Any attribute or set of attributes that uniquely identify an entity.**
   2. **Candidate key: A super key that does not contain any subset of itself that is also a super key. It does not contain extra, not important attributes.**
   3. **Primary Key: The primary key is the chosen candidate key to identify an entity.**
   4. **Foreign key: An entity occasionally contains the primary key of a different entity. This different primary key is known as a foreign key.**
   5. **Composite key: A composite key is a primary/super key that is made up of more than one attribute.**
5. (5) Describe the function of the *natural join* relational operator. Give an example of a relational algebra expression containing a natural join. Show an example of what your expression does using example tables. What is this operator used for?
   1. **In a database, we often find ourselves wishing to find the rows of two tables that have the same values in the common column. By using a natural join, we can combine those two tables, using a common column as the thing that ties the two tables together.**
   2. **Student |x| Enroll**
   3. **We have two tables and a final table that houses the joined tables.**
      1. **Student (stuID, lastName, firstName)**
      2. **Enroll (stuID, classNo, grade)**
      3. **JoinedTable (StuID, lastName, firstName, classNo, grade)**
   4. **This operator is used, especially in relational databases, to minimize data duplication by allowing you to tie related tables together through a common denominator column.**
6. (5) Describe how the project (π) select (σ) are implemented in SQL. How could this cause confusion? These resulting SQL commands are a part of which SQL sublanguage?
   1. **Project operates on a single table, but it lets you select the columns that you want, not the rows.**
   2. **Select works on a single table, but it lets you select the rows that you want, not the columns.**
   3. **This can cause confusion because the two operators do very similar things, but in an opposite way. One selects the columns, and the other one selects the rows.**
   4. **DML**
7. (5) Describe how natural join is implemented in SQL syntax. What other join operations are available in SQL? Give examples of the syntax for each. This command is a part of which SQL sublanguage?
   1. **Natural Join**
      1. **SELECT \* FROM student JOIN enroll GIVING temp1;**
   2. **Inner Join**
      1. **SELECT \* FROM student INNER JOIN enroll ON student.stuID = enroll.stuID;**
   3. **Left Join**
      1. **SELECT \* FROM student LEFT OUTER JOIN enroll ON student.stuID = enroll.stuID;**
   4. **Right Join**
      1. **SELECT \* FROM student RIGHT OUTER JOIN enroll ON student.stuID = enroll.stuID;**
   5. **Full Join**
      1. **SELECT \* FROM student FULL OUTER JOIN enroll ON student.stuID = enroll.stuID;**
   6. **DML**
8. (5)
   1. (2)Give the SQL command to create the following schema:  
      Cust:{Firstname,Lastname,Street,City,State,Zip}.

**CREATE TABLE Cust (**

**Firstname varchar(20),**

**Lastname varchar(20),**

**Street varchar (20),**

**City varchar (20),**

**State varchar (20),**

**Zip integer (10)**

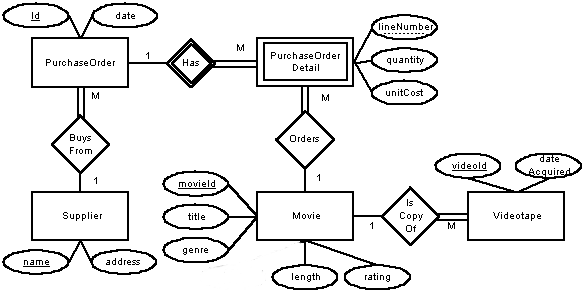
**);**

* 1. (2)Give the SQL command to add middleInitial to the above schema.

**ALTER TALBE Cust ADD middleInitial varchar(1);**

* 1. (1) These commands are part of which SQL sublanguage?

**DDL**



1. (15) Given the EER diagram above:
   1. (1) What are the attributes of *Videotape?*
      1. **videoID**
      2. **dateAcquired**
   2. (2) If a relational table were created to represent Videotape, what would be its key?
      1. **videoID**
   3. (2) What is/are the weak entity class(es)?
      1. **PurchaseOrderDetail**
      2. **Its relationship with PurchaseOrder is also weak.**
   4. (2) What are the constraints on the *Is Copy Of*  relationship?
      1. **videoID and movieID cannot be null**
   5. (3) Give a relational schema for *Movie, Is Copy Of,* and *Videotape.*
      1. ***Movie (movieID int primary key, title varchar, genre varchar, length int, rating int)***
      2. ***IsCopyOf (movieID foreign key, videoID foreign key)***
      3. ***VideoTape (videoID primary key, dateAquired dateTime)***

* 1. (3) Write an expression in the relational algebra that represents a table containing the title, genre, videoId and dateAccquired of all videos older than 1/2/2012.

π **title,genre,videoID,dateAcquired(σ dateAcquired>’1/2/2012(movie |x| videotape))**

* 1. (2) Write an equivalent SQL expression to f.

**SELECT a.title, a.genre, b.videoID, b.dateAcquired FROM movie AS a**

**JOIN videotape AS b**

**WHERE b.dateAcquired > ‘1/2/2012’;**

1. (20) For a-j assume R:{A,B,C,D,E,F,G,H,I,J} is the relation to which the FDs apply.

Note: Specify the violations completely by enumerating the FDs and how they violate the normal form rules. In the corrections specify the sequence of relations before and after each split starting from the universal relation R. If I can’t tell how you got a result, you didn’t get it, even if it is correct!

1. (2) With no other information given, what is the key of R?
   1. **A,B,C,D,E,F,G,H,I,J**
2. (2)DE→FGH, G→I is an example violation of which normal form?
   1. **1st nf**
3. (2)ABC→DEFG, A→HIJ is an example violation of which normal form?
   1. **2nd nf**
4. (2)With only the FDs in b, what is the key of R?
   1. **DE**
5. (2)With only the FDs in c, what is the key of R?
   1. **ABC**
6. How would you decompose R to remove the violation in b?
   1. **DEG 🡪 I**
   2. **DE 🡪 FH**
7. How would you decompose R to remove the violation in c?
   1. **A 🡪 HIJ**
   2. **BC 🡪 DEFG**

Given A→BEFGH, EF→GH, AC→IJ, J→D

1. What is the key of R
   1. **A**
2. What are the 2NF violations?
   1. **AC 🡪 IJ, J 🡪 D.**
3. What are the 3NF violations?
   1. **A 🡪 BEFGH, EF 🡪 GH**
4. Fix 2NF and 3NF violations.
   1. **AC 🡪 I**
   2. **ACJ 🡪 D**
   3. **A 🡪 B**
   4. **AEF 🡪 GH**

Bonus (0-5) In order to help me improve the course in the future, please answer the following questions relative to IT 350. You must really answer the questions to get the bonus points!

How is the material useful?

**The material is awesome and cool to learn. The SQL and learning how to normalize things is extremely useful in a real world setting as I have already had to do some of that for my job, but some of it (relational algebra) does not seem very applicable or important to know.**

What is your favorite aspect of the course?

**Learning best practices in how to create and set up databases, learning how to normalize, and learning about triggers. I feel that these are all very practical and useful things in the world, especially as I have had to do some of these before and now I feel like I can do them much better.**

What part of the course annoys you the most?

**The fact that the assignments sometimes don’t seem to correspond so much with the lecture. Sometimes we are supposed to do something in an assignment or in a lab before we have been taught how to do it, and I know that we can always “use google and figure it out”, but it is frustrating when I spend so much time on trying to figure something out for a deadline, and then feel like I still don’t understand what I was doing at all, only to have a lecture 3 weeks later clarify everything and I would have learned better from the assignment afterwards.**

What would you keep the same in the course?

**The fact that we are supposed to learn and do everything in multiple database environments. It is extremely useful and fun knowing how to set up and do things on multiple databases. Also, I would keep the project the same, or similar as I think it is a fun and productive way to use what we have learned.**

What would you change in the course?

**I would restructure the labs and the assignments so that they came after the lectures of the things that they were supposed to teach. It is annoying to me that a lecture only clarifies and teaches me something after I have bashed my head to try and understand something, and only have a partial understanding of it at that.**