

CS465 Databases and Scripting Languages

Midterm

1. You may not use any electronic devices.
 2. Answer the questions on a separate sheet of paper. Put your name on **each** answer sheet. Put your section (CS465 or CS565) on your first answer sheet.
 3. You must answer all of the questions. There are four questions. The exam may be a long exam for some of you. Scan the questions and answer the easy ones first.
 4. Your answers should be as concise as possible.
 5. Graduate students only should attempt 1.e and 2.c. There is no extra credit for attempting these problems if you are an undergraduate
 6. Undergraduate students will be scored out of 100 points, graduate students out of 110 points (e.g., $105/110 = 95$)
 7. Good luck!
-

Questions 1-2 refer to the following description and relations for a medium-sized library that has about 20 branches:

- a. The library has several branches throughout the city. Each branch is given a branch number, which is unique.
- b. Each branch is allocated staff, which includes a manager. The manager is responsible for the day-to-day running of a given branch. Each member of staff is given a staff number, which is unique.
- c. The library system maintains a catalog of all unique titles that it has in its system. The data held on a title is the catalog number, title, publisher, author, and genre. The catalog number uniquely identifies each title. A title is given a genre such as Sports, Mystery, Romance, Fantasy, Biography, etc..
- d. Each branch has a collection of books. The data held on a book is the book number, catalog number, title, publisher, author, genre, and loan status. The book number uniquely identifies each book. In some cases, there are several copies of a book at a branch, and the individual copies are identified using the book number. The loan status indicates whether a specific copy of a book is in the library ('library'), at the bindery ('bindery'), or loaned to a member ('loaned').
- e. The library has members, who must first register as a member of a local branch. The data held on a member is the first and last name, address, and the date that the member registered at a branch. Each member is given a member number, which is unique.
- f. Members may borrow books. The data held on each book loan is the loan number, the first and last name and number of the member, the book number, title, and the dates the book is loaned out and is due. The loan number is unique.
- g. Members may place a request for a book that is currently out on loan (i.e., all of the branch's copies are currently on loan). The data held on a request is the member's first and last name, member's number, title, catalog number, and author of the book, date requested, status of the request, and the branch from which the book will be picked up. The status could be waiting or in-transit, meaning the book is in transit to the pick up branch. A member may not have two requests for the same book, and a request is deleted as soon as it is filled.

Hence the relation will never have two requests for the same book by the same member.

Here are the relations that the database designer has chosen for this database:

Branch(branchNo, address, phoneNo, manager): The value of the manager field is a staffNo
Staff(staffNo, branchNo, firstname, lastname, position, salary)
Catalog(catalogNo, publisher, title, author, genre)
Book(bookNo, catalogNo, loanStatus, branchNo)
Member(memberNo, branchNo, firstname, lastname, address, dateRegistered)
Loan(loanNo, memberNo, bookNo, dateLoaned, dateDue)
Request(memberNo, catalogNo, dateRequested, status, branchNo)

1. **SQL (24 points CS465, 30 points CS565):** Write the following SQL queries:
 - a. List the firstname, lastname, salary, and position of any staff who make more than 50,000 dollars per year and whose position is either 'checkout clerk' or 'reference librarian'.
 - b. List the firstname, lastname, salary, and branchNo of all managers.
 - c. List details of the books checked out by member 'Smiley Smith' (firstname 'Smiley', lastname 'Smith'). The information you should print for each book should be book title, genre, dateLoaned, and dateDue. Do not print out the member's number or name.
 - d. List the number of requests for each member in the Request relation. For each member, list the member's number and the number of requests as a column labeled numRequests (you can do this without using a view).
 - e. **(CS565 students only)** Write an assertion constraint that no member may have more than 20 books checked out (i.e., on loan). Assume the constraint is listed outside of any relation.
2. **Relational Algebra (20 points CS465, 24 points CS565)**
 - a. Give a three sentence or less English-language description of the following relational algebra query that a non-computer scientist could understand:

$$\Pi_{\text{Distinct Member.firstname, Member.lastname}} \left(\text{Member} \bowtie_{\text{Member.memberNo}=\text{Loan.memberNo}} \left(\sigma_{\text{CURDATE between dateLoaned and dateDue}} (\text{Loan}) \right) \right)$$
 - b. Using relational algebra notation, write a query that prints the first and last names of all members who belong to the branch at '1678 Cardiff Rd.'
 - c. **(CS565 students only)** Give a three sentence or less English-language description of the following relational algebra query that a non-computer scientist could understand:

$$\Pi_{\text{Distinct Catalog.title, Catalog.author}} \left(\text{Catalog} \bowtie_{\text{Catalog.catalogNo}=\text{Book.catalogNo}} \text{Book} \bowtie_{\text{Book.bookNo} = \text{Loan.bookNo}} \left(\sigma_{\text{CURDATE between dateLoaned and dateDue}} (\text{Loan}) \right) \bowtie_{\text{Book.branchNo} = \text{Branch.branchNo}} \left(\sigma_{\text{branchNo}='B003'} (\text{Branch}) \right) \right)$$

3. **Normalization (28 points):** The table below shows a sample booking for a guest at the XYZ chain of hotels:

HotelNo	GuestNo	Hotel Name	Hotel City	Hotel Zip	Guest Name	Guest City	Guest Zip	Date From	Date To	RoomNo	Room Type	Room Price
3	232	Hilton	San Diego	83835	Brad VZ	Knoxville	37996	2012-11-28	2012-12-02	635	King	89.99

You may make the following assumptions about the data:

- A hotel number uniquely identifies a hotel's name and zip code
- A zip code uniquely identifies a city for both hotels and guests (not true in the real world, but true in our fantasy world)
- A guest number uniquely identifies a guest's name and zip code
- A room number and a hotel number uniquely determine a room type and price
- A guest number and date from uniquely determine a hotel number, date to, and room number.

Answer the following questions:

- Give an example of the following types of anomalies:
 - insert
 - update
 - delete
- What are the functional dependencies for this relation?
- What is the primary key for this relation?
- Show how you would convert this relation to 2nd normal form, and show which functional dependencies you would use to create each new relation.
- Show how you would convert the relations from 2nd to 3rd normal form, and show which functional dependencies you would use to create each new relation.
- What is the name for the type of functional dependency used to convert a relation to 2nd normal form?
- What is the name for the type of functional dependency used to convert a relation to 3rd normal form?

- 4. Database Design (28 points):** Consider the following description of an investment firm:
- The firm has a number of offices. Each office has a number of investment advisors who service clients. Clients own stocks that pay dividends.
 - An investor has a name, a unique investor number, a single investment advisor and owns zero or more stocks.
 - Each investor's holding of a stock contains information about the stock's ticker symbol (a 3 or 4 character code that uniquely identifies the stock, such as IBM or APPL), the stock's name (e.g., International Business Machines), the number of shares of stock owned by the investor, and the dividend per share paid by the stock. We only keep track of an investor's cumulative holding—we do not care about the individual transactions that led to this holding.
 - an investment advisor has a name, a unique advisor number, one or more clients that the advisor services, and an office to which the advisor belongs.
 - an office has an office number and an address.
 - an investment advisor is associated with only one office
 - an office will have 1 to 20 advisors,
 - the dividend per share is determined by the stock (e.g., all holdings of IBM pay the same dividend, such as \$.05 dividend per share). Stocks are owned by 1 or more clients.

Answer the following questions:

- Define the entities and attributes for each entity. An answer like the following is fine:
 Hotel(hotelNo, hotelName, city)
 Guest(guestNo, guestName, guestAddress)
 ...
 In other words, don't bother justifying why you chose the entities, just list them and their attributes.
- Identify the primary and foreign keys for each entity. Please don't underline the attributes from the previous answer. Instead follow the format of the following example answer:

Entity	Primary Key	Foreign Key(s)
Hotel	hotelNo	----
Room	hotelNo, roomNo	hotelNo
Booking	hotelNo, roomNo, dateFrom	hotelNo, roomNo, guestNo
- Draw an ER diagram for this investment firm that shows entities, relationships, and multiplicity of relationships.