

CS465 Databases and Scripting Languages
Midterm
Fall 2018

1. This exam is closed-note, closed book.
2. You may not use any electronic devices.
3. Circle the appropriate answer(s) for the multiple choice/answer questions.
4. The last two pages of the exam gives the relations you should use for many of the exam questions. I recommend ripping these pages off the exam so that you can easily refer to them.
5. You must answer all of the questions.
6. Good luck!

Multiple Choice (42 points): Choose the best answer from the following choices. Circle only one answer! If two answers seem like they might work, choose the best of the two answers.

1. What is the name of the integrity constraint that specifies that a primary key must be non-null?
 - a. Entity integrity
 - b. Referential integrity
 - c. General integrity (i.e., business constraint)
 - d. Primary integrity
2. Suppose that we do not allow a library member to borrow more than 10 books at one time. What is the name of this type of constraint?
 - a. Entity integrity
 - b. Referential integrity
 - c. General integrity (i.e., business constraint)
 - d. Primary integrity
3. What is the formal significance of the relational calculus?
 - a. It is the formal basis for the DDL in relational databases
 - b. It is the formal basis used for eliminating update, insert, and delete anomalies by decomposing a universal relation into smaller relations
 - c. It is the formal basis for the DML in relational databases
 - d. It is the formal basis for mapping relations to files
4. What is the English-language meaning of the following relational algebra query:
$$\Pi_{\text{Distinct Member.firstname, Member.lastname}} (\text{Member} \bowtie_{\text{Member.memberNo=Loan.memberNo}} \text{Loan}) \bowtie_{\text{Loan.bookNo=Book.bookNo}} \text{Book} \bowtie_{\text{Book.catalogNo=Catalog.CatalogNo}} (\sigma_{\text{author='Stephen King'}} \text{Catalog})$$
 - a. Print members who are named 'Stephen King'
 - b. Print members who have at some point in time borrowed a book written by 'Stephen King'
 - c. Print members who have an overdue book written by 'Stephen King'
 - d. Print members who are currently borrowing a copy of a book written by 'Stephen King'

5. Why do we use B+ trees rather than balanced binary search trees to implement indices in database systems? In the following answers, assume that N is the number of records in the file and M is the branching factor of the B+ tree.
 - a. B+ trees require a maximum of 2 block accesses to retrieve any record from the database as opposed to $\log_2 N$ block accesses by a binary search tree
 - b. Even though B+ trees take more CPU time than a balanced binary search tree to locate a record, they require fewer disk accesses and hence take less wall clock time
 - c. Even though B+ trees require more disk accesses than a balanced binary search tree to locate a record, they require less CPU time and hence take less wall clock time
 - d. B+ trees are easier to implement than balanced binary search trees and their performance is comparable to balanced binary search trees so we prefer B+ trees for their simplicity.

6. Suppose users frequently write query the library's Loan relation to ask for all books loaned during a specified time interval (e.g., DateLoaned BETWEEN 2018-09-15 AND 2018-09-25). What type of index should be use for the attribute DateLoaned?
 - a. Join index
 - b. Range index
 - c. Secondary index
 - d. Cluster index

7. In the library's Loan relation, am I likely to want to place a secondary index on memberNo?
 - a. No, it is an artificial key without much meaning
 - b. No, the relation is too small for indexing
 - c. No, a query on this attribute will retrieve a significant fraction of the relation
 - d. Yes, it is the primary key of the loan relation
 - e. Yes, it will be frequently used in joins involving the loan relation
 - f. Yes, it will often be used to sort the results of a query

8. Am I likely to want to place an index on the branchNo attribute in the library's Branch relation?
 - a. No, it is an artificial key without much meaning
 - b. No, the relation is too small for indexing
 - c. No, it involves a long character string
 - d. Yes, it is the primary key of the branch relation and will be frequently used in joins
 - e. Yes, it is the foreign key of branch relation and will be frequently used in joins
 - f. Yes, it will often be used to sort the results of a query

9. For the library's hold relation, does the following dependency represent a full functional dependency?

memberNo, catalogNo, dateRequested \rightarrow status, branchNo

- a. Yes, it uniquely determines the status of the hold and the branch from which the book should be picked up
 - b. No, it does not uniquely determine the status of the hold and the branch from which the book should be picked up
 - c. No, the memberNo and catalogNo by themselves uniquely determine the status of the hold and the branch from which the book should be picked up and hence dateRequested is unnecessary.
 - d. No, the memberNo and dateRequested by themselves uniquely determine the status of the hold and the branch from which the book should be picked up and hence catalogNo is unnecessary.
 - e. No, the dateRequested and catalogNo by themselves uniquely determine the status of the hold and the branch from which the book should be picked up and hence memberNo is unnecessary.
10. Which of the following SQL queries will return the list of members with overdue books (i.e., books that have not been returned and which are past the dateDue date)?

- a. `SELECT memberNo FROM Loan
WHERE dateDue < CURDATE() AND dateReturned = NULL`
- b. `SELECT memberNo FROM Loan
WHERE dateDue < CURDATE() AND dateReturned IS NULL`
- c. `SELECT memberNo FROM Loan
WHERE dateReturned > dateDue`
- d. `SELECT memberNo FROM Loan
WHERE CURDATE() > dateDue`

11. Which of the following queries would require set division?

- a. Find all people who are associated with the library as either staff or a member
- b. Find all members who are associated with the library who are both staff and a member
- c. Find all members who have visited all of the branches in the library system
- d. Find all members who have borrowed a book in the past 40 days

12. What is the English-language meaning of the following relational calculus query:
 $\{M.firstname, M.lastname \mid Member(M) \wedge (\exists H)(\exists C)(Hold(H) \wedge Catalog(C) \wedge M.memberNo = H.memberNo \wedge H.catalogNo = C.catalogNo \wedge C.title = 'Paradise Lost') \wedge ((CURRENT() - H.requestedDate) > 7)\}$

- List the names of all members who have borrowed the book titled “Paradise Lost” in the last 7 days
- List the names of all members who have placed a hold on the book titled “Paradise Lost” in the last 7 days
- List the names of all members who have a hold on the book titled “Paradise Lost” and have also borrowed a copy of the book for more than 7 days
- List the names of all members who have a hold on the book titled “Paradise Lost” and have waited more than 7 days for the book

13. In the relational algebra, how can I implement the set intersection operation, $R \cap S$, using only the fundamental relational algebra operations?

- $R - (R - S)$
- $R \cup (S - (R - S))$
- $S - (S \cup (R - S))$
- $(R \times S) - (R - S)$

14. When we use partial dependencies to convert a relation from 1st Normal Form to 2nd Normal Form and when we use transitive dependencies to convert a relation from 2nd Normal to 3rd Normal Form, we copy all of the attributes in the dependency to the new relation. Which of the following properties of the decomposition process are we observing when we perform this copying?

- Any constraint in the original relation is enforced in one of the smaller relations
- There must be no duplicated columns in any of the smaller relations (i.e., given any pair of relations from the set of smaller relations, there is no shared column)
- Any tuple in the original relation can be re-constructed from corresponding tuples in the smaller relations
- Any column that contains a multi-valued attribute in the original relation is included as a column containing a multi-valued attribute in one of the smaller relations.

Multiple Answer (11 points): Choose all answers that are appropriate from the following choices.

15. What are the candidate keys for the Branch relation?

- branchNo
- branchName
- address
- phoneNo
- manager

16. What are the candidate keys for the Visit relation? **2 correct answers.**
- a. memberNo, branchNo
 - b. memberNo, branchNo, dateVisited
 - c. memberNo, branchNo, dateVisited, timeArrived
 - d. memberNo, branchNo, dateVisited, timeDeparted
 - e. memberNo, dateVisited, timeArrived
 - f. memberNo, dateVisited, timeDeparted
 - g. memberNo, dateVisited
 - h. memberNo, timeArrived
 - i. memberNo, timeDeparted
17. For the purposes of normalization, which full functional dependencies may be derived from the assumptions given for the library database. Note that the following set of answers does not include the complete set of full functional dependencies that could be derived from the assumptions. **5 correct answers**
- a. branchName -> branchAddress
 - b. manager -> branchName
 - c. dateRegistered -> memberNo
 - d. catalogNo -> publisher, title, author, genre
 - e. memberNo, bookNo -> dateLoaned, dateDue, dateReturned
 - f. dateLoaned, dateDue -> dateReturned
 - g. memberNo, bookNo, dateLoaned -> dateDue, dateReturned
 - h. memberNo, catalogNo -> dateRequested, status, branchNo
 - i. memberNo, dateVisited -> branchNo, timeArrived, timeDeparted
 - j. memberNo, branchNo, dateVisited -> timeArrived, timeDeparted
 - k. position -> salary
 - l. bookTitle -> branchNo
 - m. catalogNo -> bookNo, branchAddress
 - n. memberNo -> branchNo, dateLoaned, bookNo

Normalization (13 points): An agency called Instant Cover supplies part-time/temporary staff to hotels within the state of Tennessee. The timesheet table shown below lists the cumulative time billed by agency staff working at various hotels. The Employee Id is unique for every member of staff. The required fields are id, contractNo, hours, and empName.

id	contractNo	hours	empName	hotelNo	hotelCity
1135	C1024	16	Smith J	H25	Nashville
1057	C1024	24	Hocine D	H25	Nashville
1068	C1025	28	White T	H4	Knoxville
1135	C1025	15	Smith J	H4	Knoxville
1057	C1026	8	Hocine D	H25	Nashville

Assume that this relation has the following functional dependencies:

- id, contractNo \rightarrow hours
- id \rightarrow empName
- contractNo \rightarrow hotelNo
- hotelNo \rightarrow hotelCity

Answer the following questions.

18. What is the primary key for this relation?

- a. id
- b. id, contractNo
- c. contractNo
- d. id, contractNo, hours
- e. id, hours

19. This table is susceptible to various types of anomalies. Circle all of the following anomalies that apply (**4 correct answers**):

- a. You cannot insert a new contract unless it already has an employee with billable hours associated with it
- b. If you change the contractNo, you must change every instance of the contractNo or you will have an inconsistency
- c. If you update an employee's billable hours for a particular contract, you must change every instance of the employee's billable hours in the relation or you will have an inconsistency
- d. If you delete the last tuple containing an employee's timesheet, you lose information about the employee
- e. If you update an employee's name, you must change every instance of that employee's name in the relation or you will have an inconsistency
- f. You cannot insert a new employee unless the employee already has billable hours
- g. You cannot insert a new billing record unless it has a hotel city associated with it.

20. What functional dependency(s) would you use to convert this relation from 1st to 2nd normal form? Circle all functional dependencies that apply.
- a. id, contractNo → hours
 - b. id → empName
 - c. contractNo → hotelNo
 - d. hotelNo → hotelCity
21. What functional dependency(s) would you use to convert this relation from 2nd to 3rd normal form? Circle all functional dependencies that apply.
- a. id, contractNo → hours
 - b. id → empName
 - c. contractNo → hotelNo
 - d. hotelNo → hotelCity

ER Diagrams

22. **(4 points)** The Instant Cover database has three entities—contracts, employees, and hotels. Two of the relationships are 1) employees work on contracts, and 2) a hotel signs contracts to perform repairs. The ER diagram might be drawn as:



The numbers shown above represent missing multiplicities for the relationships. You may assume that:

- a. At least one employee works on each contract.
- b. Each hotel has at least one contract.
- c. A new employee may not work on any contracts.

The example data in the table gives you additional information you can use to figure out the multiplicity of the relationships. For each number, enter the letter associated with its multiplicity:

- | | |
|---------|----------|
| a. 1..1 | 1) _____ |
| b. 0..1 | |
| c. 1..0 | 2) _____ |
| d. 1..* | |
| e. *.* | 3) _____ |
| f. 0..* | |
| g. *..0 | 4) _____ |
| h. *..1 | |

23. **(5 points)** If you were creating an ER diagram for the library database, what entities would you use? Circle all entities that apply. **(5 correct answers)**

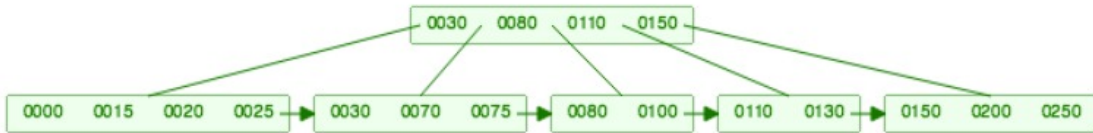
- a. Branch
- b. Manager
- c. Catalog
- d. Book
- e. Visit
- f. Loan
- g. Hold
- h. Genre
- i. Member
- j. Position
- k. Publisher
- l. Staff
- m. Author

24. **(7 points)** If you were creating an ER diagram for the library database, what relationships would you draw? **7 correct answers**

- a. Staff manages branch
- b. Branch employs staff
- c. Staff has position
- d. Library has branches
- e. Position defines salary
- f. Branch provides books
- g. Members place a hold on a book in the library's catalog and specify the branch from which the book will be picked up.
- h. Member visits a branch
- i. Catalog describes a book
- j. Member borrows a book
- k. Catalog specifies a title and a genre
- l. Book title determines branch
- m. Member registers at branch

26. (6 points) Show what the following B+ tree looks like after 10 is inserted into it.

- Assume that $M = 5$ and $L = 4$.
- When you split a node and one of the two nodes must contain an odd number of keys, put the odd number of keys in the leftmost node.



You are given the following information about a medium-sized library that has approximately 20 branches:

- a. The library has several branches throughout the city. Each branch is given a branch number, which is unique. Each branch also has a branchName, address, phone number, and a manager. The branchName, address, and phone number also uniquely identify a branch. A manager is allowed to manage multiple branches.
- 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. Staff members have positions, such as 'reference librarian' or 'checkout clerk'. A position is not the same as being a manager. For example, a member may be listed as 'reference librarian' and may also be listed as being the branch's manager in the branch relation.
- c. A staff position uniquely determines a salary (e.g., all checkout clerks make the same salary).
- d. 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. There may be multiple copies of each book described in the catalog at one or more branches.
- e. Each branch has a collection of books. The data held on a book is the book number, catalog number, title, publisher, author, genre, loan status, and the branch where the book is located. 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'). More than one branch may have copies of a book as indicated in bullet d (i.e., the BookNos would be different but the catalogNos would be the same).
- f. The library has members. The data held on a member is the first and last name, address, and the date that the member joined the library. Each member is given a member number, which is unique.
- g. The library keeps track of visits by members. The data held on each visit includes the member's number, the branch number visited, the date the branch was visited, and the time at which the member entered and departed the branch. A member may visit multiple branches on the same day, or the same branch two or more times on the same day. The times of an individual member's visits may not overlap.
- h. Members may borrow books and they may borrow multiple books on the same day. A member may not have two copies of a book checked out at the same time, but they may have checked out the same book for two or more non-overlapping periods of time. 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, the dates the book is loaned out and is due, and the date the book is returned. The loan number is unique. If the member has not yet returned the book, then the date returned field is NULL. The library does not delete loan tuples from the database, so that it has a historical record of who has borrowed which books from the library.
- i. Members may place a hold 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 holds for the same book, and a hold is deleted as soon as it is filled. Hence the relation will never have two holds for the same book by the same member.

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

Branch(branchNo, branchName, 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, firstname, lastname, address, dateRegistered)
Visit(memberNo, branchNo, dateVisited, timeArrived, timeDeparted)
Loan(loanNo, memberNo, bookNo, dateLoaned, dateDue, dateReturned)
Hold(memberNo, catalogNo, dateRequested, status, branchNo)