ICS 321 Data Storage & Retrieval — Mock Final Exam Name:_____

Answers to some of these questions are available at the following URL even though they may not be entirely correct.

`http://pages.cs.wisc.edu/~dbbook/openAccess/thirdEdition/solutions/ans3ed-oddonly.pdf`

Q1. **[Ex. 5.3]** The following relations keep track of airline flight information:

Flights(flno: integer, from: string, to: string, distance: integer, departs: time, arrives: time, price: real)
Aircraft(aid: integer, aname: string, cruisingrange: integer)
Certified(eid: integer, aid: integer)
Employees(eid: integer, ename: string, salary: integer)

Note that the Employees relation describes pilots and other kinds of employees as well; every pilot is certified for some aircraft, and only pilots are certified to fly. Write each of the following queries in SQL.

1. Find the names of aircraft such that all pilots certified to operate them have salaries more than $80,000.
2. For each pilot who is certified for more than three aircraft, find the eid and the maximum cruisingrange of the aircraft for which she or he is certified.
3. Find the names of pilots whose salary is less than the price of the cheapest route from Los Angeles to Honolulu.
4. For all aircraft with cruisingrange over 1000 miles, find the name of the aircraft and the average salary of all pilots certified for this aircraft.
5. Find the names of pilots certified for some Boeing aircraft.
6. Find the aids of all aircraft that can be used on routes from Los Angeles to Chicago.
7. Identify the routes that can be piloted by every pilot who makes more than $100,000.
8. Print the enames of pilots who can operate planes with cruisingrange greater than 3000 miles but are not certified on any Boeing aircraft.
9. A customer wants to travel from Madison to New York with no more than two changes of flight. List the choice of departure times from Madison if the customer wants to arrive in New York by 6 p.m.
10. Compute the difference between the average salary of a pilot and the average salary of all employees (including pilots).
11. Print the name and salary of every nonpilot whose salary is more than the average salary for pilots.
12. Print the names of employees who are certified only on aircrafts with cruising range longer than 1000 miles.
13. Print the names of employees who are certified only on aircrafts with cruising range longer than 1000 miles, but on at least two such aircrafts.
14. Print the names of employees who are certified only on aircrafts with cruising range longer than 1000 miles and who are certified on some Boeing aircraft.

Q2. **[Ex. 9.6]** Consider again the disk specifications from Ex. 9.5 (see Ex.17), and suppose that a block size of 1024 bytes is chosen. Suppose that a file containing 100,000 records of 100 bytes each is to be stored on such a disk and that no record is allowed to span two blocks.

1. How many records fit onto a block?
2. How many blocks are required to store the entire file? If the file is arranged sequentially on the disk, how many surfaces are needed?
3. How many records of 100 bytes each can be stored using this disk?
4. If pages are stored sequentially on disk, with page 1 on block 1 of track 1, what page is stored on block 1 of track 1 on the next disk surface? How would your answer change if the disk were capable of reading and writing from all heads in parallel?
5. What time is required to read a file containing 100,000 records of 100 bytes each sequentially? Again, how would your answer change if the disk were capable of reading/writing from all heads in parallel (and the data was arranged optimally)?
6. What is the time required to read a file containing 100,000 records of 100 bytes each in a random order? To read a record, the block containing the record has to be fetched from disk. Assume that each block request incurs the average seek time and rotational delay.

Q3. **[Ex. 12.5]** Consider again the schema with the Sailors relation:

Sailors(sid: integer, sname: string, rating: integer, age: real)

Assume that each tuple of Sailors is 50 bytes long, that a page can hold 80 Sailors tuples, and that we have 500 pages of such tuples. For each of the following selection conditions, estimate the number of pages retrieved, given the catalog information in the question.

1. Assume that we have a B+-tree index T on the search key 'Sailors.sid', and assume that IndexHeight(T) = 4, IndexLeafPages(T) = 50, LowestKey(T) = 1, and HighestKey(T) = 100,000.

   a) $\sigma_{Sailors.sid<50{,}000}(Sailors)$

b) $\sigma_{Sailors.sid=50{,}000}(Sailors)$

2. Assume that we have a hash index T on the search key 'Sailors.sid', and assume that IndexHeight(T) = 2, IndexLeafPages(T) = 50, LowestKey(T) = 1, and HighestKey(T) = 100,000.

   a) $\sigma_{Sailors.sid<50{,}000}(Sailors)$
   b) $\sigma_{Sailors.sid=50{,}000}(Sailors)$

Q4. **[Ex. 16.3.]** Consider a database with objects X and Y and assume that there are two transactions T1 and T2. Transaction T1 reads objects X and Y and then writes object X. Transaction T2 reads objects X and Y and then writes objects X and Y.

1. Give an example schedule with actions of transactions T1 and T2 on objects X and Y that results in a write-read conflict.
2. Give an example schedule with actions of transactions T1 and T2 on objects X and Y that results in a read-write conflict.
3. Give an example schedule with actions of transactions T1 and T2 on objects X and Y that results in a write-write conflict.
4. For each of the three schedules, show that Strict 2PL disallows the schedule.
