

Purdue University
Computer Science Department
CS44800 Homework 3
Fall 2024

Reference: **Fundamentals of Database Systems by Elmasri & Navathe, 7th Edition**

Total Points: **3 points**

Submit via Gradescope

Provide enough detail to support your answers. If possible, use bullets to organize the ideas in your answer.

Question 1. (1.0 point)

Specify the following queries in **relational algebra (R.A.)** on the COMPANY database schema shown in Figure 5.5 in the Textbook. Do not use any other information from the database state on Figure 5.6 to build the queries.

- (a) List names of employees (Fname, Minit, Lname) who are directly supervised by a female supervisor.
- (b) List names of employees (Fname, Minit, Lname) who are NOT assigned to every project.
- (c) List names of employees (Fname, Minit, Lname) in Departments with locations in Houston or Sugarland.
- (d) List the names of department managers (Fname, Minit, Lname) who have no dependents.

Question 2. (1.0 point)

Consider SQL queries Q13 and Q4A in Chapter 6.

- (a) Draw the initial query tree for each of these queries, and then show how the query tree is optimized by the algorithm outlined in Section 19.1.
- (b) For each query, compare the initial query and the final query tree. Discuss the difference between each tree.

Question 3. (1.0 point)

Suppose that a disk unit has the following parameters: seek time $s = 20$ msec; rotational delay $rd = 10$ msec; block transfer time $btt = 1$ msec; block size $B = 2400$ bytes; interblock gap size $G = 600$ bytes. An EMPLOYEE file has the following fields: Ssn , 9 bytes; Last_name , 20 bytes; First_name , 20 bytes; Middle_init , 1 byte; Birth_date , 10 bytes; Address , 35 bytes; Phone , 12 bytes; Supervisor_ssn , 9 bytes; Department , 4 bytes; Job_code , 4 bytes; deletion marker, 1 byte. The EMPLOYEE file has $r = 30,000$ records, fixed-length format, and unspanned blocking. Write appropriate formulas and calculate the following values for the above EMPLOYEE file:

- (a) Calculate the record size R (including the deletion marker), the blocking factor bfr , and the number of disk blocks b .
- (b) Calculate the wasted space in each disk block because of the unspanned organization.
- (c) Calculate the transfer rate tr and the bulk transfer rate btr for this disk unit (see Appendix B in the textbook for definitions of tr and btr).
- (d) Calculate the average number of block accesses needed to search for an arbitrary record in the file, using linear search.
- (e) Calculate in msec the average time needed to search for an arbitrary record in the file, using linear search, if the file blocks are stored on consecutive disk blocks and double buffering is used.