CSCE 2501 - Fundamentals of Database Systems
Assignment 4

This is an individual assignment. Rules governing academic integrity are strongly enforced without exception.

1. For each of the following relational schemas, Write the corresponding RA expression, RC expression and SQL query to retrieve the information required in the associated questions: **(24 points)**

   a. **Computer Store Database**

   Product(manufacturer, model)
   PC(model, speed, ram, harddrive, screen, price)
   Laptop(model, speed, ram, harddrive, screen, price)
   Printer(model, min_compatible_speed, is_colored, price)

   i. Retrieve the manufacturer and model of all the PCs that are priced between $500 and $800. **(3 points)**
   ii. Retrieve the models of PCs that are not manufactured by a company that also makes laptops **(3 points)**
   iii. Retrieve the model, speed and price of all the laptops that are compatible with a colored printers produced by the same manufacturer **(3 points)**
   iv. Using no aggregate function, retrieve the manufacturer that makes the PC with the fastest processor **(3 points)**

   b. Cooking Recipes Database

   recipe(name, kitchen)
   ingredients(recipe, food_item, grams)
   foodItem(food_item, type, calories_per_gram)
   stock(food_item, shop, price)

   i. Retrieve the food items, grams and total calories for preparing the recipe for “Mac & Cheese”. **(3 points)**
   ii. Retrieve all the recipes that contain both “Onions” and “Cheese”. **(3 points)**
   iii. Retrieve all the recipes that are gluten-free (contains no food items of type “Wheat”) **(3 points)**
   iv. Retrieve the food items and their price for all items of type “Dairy” sold at “Carrefour” **(3 points)**

2. Consider the following schema **(15 points)**

   Sailors(sid, sname, rating, age) *Reserves(sid, did, day)* *Boats(bid, bname, size)*

   Reserves.sid is a foreign key to Sailors and *Reserves.bid* is a foreign key to Boats.bid.
   Reserves contains 10,000 records with 40 records per page. *Sailors* contains 1000 records with 20 records per page. *Boats* contains 100 records with 10 records per page. There are 50 values for *Reserves.bid*, 10 values for *Sailors.rating* (1..10), 10 values for *Boat.size*, and 500 values for *Reserves.day*

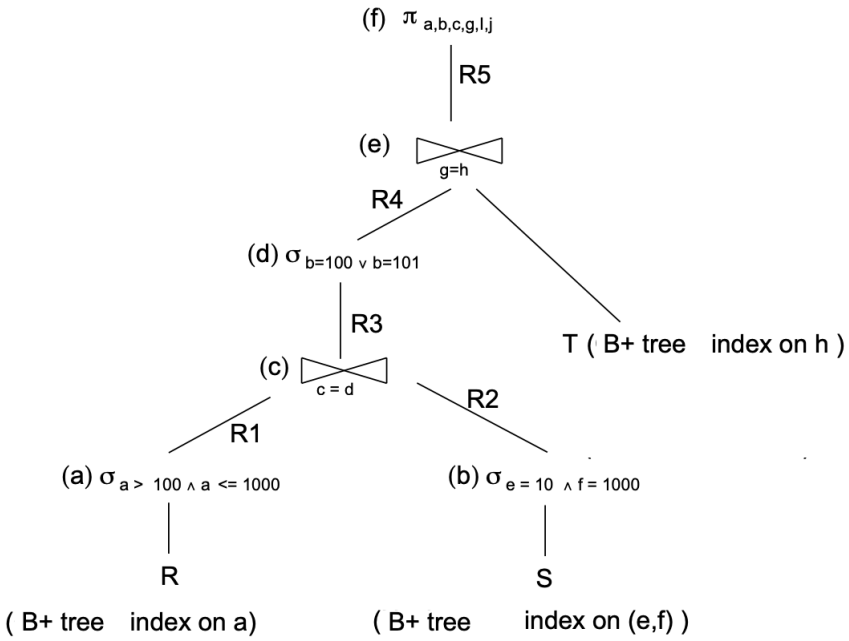
   a. Draw all possible left-deep join trees for the following query **(5 points)**

   ```
   SELECT S.sid, S.sname, B.bname
   FROM Sailors S INNER JOIN Reserves R ON S.sid = R.side
   INNER JOIN Boats B ON R.bid = B.bid
   ```

   b. Assuming a 50-frame buffer pool, evaluate the first join (bottom, left-most join) for each query plan you obtained in (a) and show which join algorithm works the best. Consider the Page-based nested loop Join, the Block-based nested loop Join, Sort-Merge Join, and Hash Join. **(10 points)**

2. Use the following schema and query plan to answer the corresponding questions **(11 points)**

R(a,b,c)	S(d,e,f,g)	T(h,i,j)
B(R) = 1000	B(S) = 1000	B(T) = 1000
T(R) = 10,000	T(S) = 10,000	T(T) = 10,000
Min(R,a) = 0	V(S,e) = 10	V(T,h) = 100
Max(R,a) = 9000	V(S,f) = 100	
V(R,b) = 100	V(S,d) = 50	
V(R,c) = 20	V(S,g) = 40	



a. Compute the cardinality of all intermediate relations labeled R1 through R5 and the final result **(6 points)**

b. Compute the cost of this query plan in terms of number of pages read from disk or written to disk. Assume that the B+ tree indexes on single columns cost 2.4 reads / lookup, the indexes on multiple columns cost 4 reads / lookup, and the buffer pool of the system can hold up to 252 frames **(5 points)**