UNIVERSITY OF MANNHEIM
School of Business Informatics and Mathematics

Data and Web Science Group
Prof. Dr. Heiko Paulheim

Database Technology

Exercise 9: Query Processing

8.1. Sort-merge

Apply the sort-merge algorithm to the following relation (sort it based on the first attribute). Show the runs created on each pass assuming that only one tuple fits in a block and memory holds at most 3 blocks.

kangaroo	17
wallaby	21
emu	1
wombat	13
platypus	3
lion	8
warthog	4
zebra	11
meerkat	6
hyena	9
hornbill	2
baboon	12

8.2 Optimized relational-algebra

Consider the following bank database:

branch(branch_name, branch_city, assets)
customer(customer_name, customer_stree, customer_city)
loan(loan_number, branch_name, amount)
borrower(customer_name, loan_number)
account(account_number, branch_name, balance)
depositor(customer_name, account_number)

Write an efficient relational-algebra expression that is equivalent to the following SQL query:

```
SELECT T.branch_name
FROM branch T, branch S
WHERE T.assets > S.assets AND S.branch_city=”Brooklyn”
```

Justify your choise.

8.3. Query costs

Let relations r1(A, B,C) and r2(C, D, E) have the following properties: r1 has 20,000 tuples, r2 has 45,000 tuples, 25 tuples of r1 fit on one block, and 30 tuples of r2 fit on one block. Estimate the number of block transfers and seeks required, using each of the following join strategies for r1 ⋈ r2:

a. Nested-loop join.
b. Block nested-loop join.
c. Merge join.
d. Hash join.

8.4. Query costs 2

Let r and s be relations with no indices, and assume that the relations are not sorted. Assuming infinite memory, what is the lowest-cost way (in terms of I/O operations) to compute r ⋈ s? What is the amount of memory required for this algorithm?

8.5. Index usage

Consider again the bank database of from exercise 9.2. Suppose that a B+-tree index on branch city is available on relation branch, and that no other index is available. List different ways to handle the following selections that involve negation:

a. $\sigma_{\neg(branch\_city<\text{"Brooklyn"})}(branch)$
b. $\sigma_{\neg(branch\_city=\text{"Brooklyn"})}(branch)$
c. $\sigma_{\neg(branch\_city<\text{"Brooklyn" v assets}<5000)}(branch)$

8.6. Query plans

Execute the following query on the univerity database from the second exercise:

SELECT student.ID, takes.grade
FROM student INNER JOIN takes ON(student.ID=takes.ID)
WHERE student.name = "Zhang" AND takes.year=2009

To do that

1. Open your browser at http://localhost/phpmyadmin/
2. Click on the university database and open the SQL tab and execute the query
3. Click on “explain SQL”

a. Find out what the columns contain
b. Formulate the query plan in text