

Simulating a Simple Query Processor that evaluates a SQL Query in Select-From-Where

Q:

Select Fname, Lname, SSN, Dno, Dname
From Employee, Department
Where Dno = 5 and Dno = DNumber;

There are multiple different ways to evaluate the query Q. Let's assume that the query Q is transformed by Optimizer to the following Query Execution Plan (Steps) in a query tree structure (as shown in Relational Algebra here):

QA:

STEP1: EMPS_DNO5 $\leftarrow \sigma_{dno=5}$ (EMPLOYEE)

STEP2: EMP_DEPT_DNO5 \leftarrow (EMPS_DNO5 $\bowtie_{DNO=DNUMBER}$ DEPARTMENT)

STEP3: EMP_DEPT_MGR_DEPENDENT
 $\leftarrow \pi_{FNAME, LNAME, SSN, DNO, DNAME}$ (EMP_DEPT_DNO5)

The query string Q is parsed, built and transformed to a query execution steps like QA in a tree structure as explained in class.

However, to avoid the complexity of the tree operations in your lab2, **we assume** that our query execution steps are written in a text file (input file) in the following operator and operand structure in one execution step per line instead of a query tree structure:

Selection InTable_Name Selection_Conditions OutTab_Name1

Nested Loop Join LTable RTable Join_Conditions OutTab_Name2

Projection InTable_Name2 ProjectColumn_List OutTab_Name3

Then our Query Execution Steps QA would look like as follow in a file which is an input to your query evaluator:

Input file:

Selection EMPLOYEE DNO=5 EMPS_DNO5

Join EMPS_DNO5 DEPARTMENT DNO=DNUMBER EMP_DEPT_DNO5

Projection EMP_DEPT_DNO5 Fname, Lname, SSN, Dno, Dname EMP_DEPT_MGR_DEPENDENT

Implement a query evaluator that takes a query execution steps (Query execution plan) in an **input file** and performs the three relational operators in the sequence (from top to bottom) as written in the input file to evaluate the query Q.

Implement three operators that read the operands (input tables) and perform each operation with given the condition list/column list in the input file, then generates a New Table with the given output file name to write the output of each operation.

Selection

Nested Loop Join

Projection

1. Create the **input file** that has the execution steps in a text before start.
2. We assume that the Optimizer finished building the query execution plan in the **input file** (instead of in a query tree structure) that is passed to your query evaluator. We also assume that your query evaluator finished replacing it with the physical execution plan in the **input file**.
3. Your program reads the query execution steps of Q from the **input file** and executes each step in that order.
4. Base tables are EMPLOYEE and DEPARTMENT that is stored in the file system. of the database server.
5. For each operator, create a new table with the given output table name specified in the execution steps to write the output of the operation.
6. Note that the input relation (table) of the next step is the output table that was created from the previous step. The last table generated in the last step is your final query result.

7. There are two ways to read your base tables Employee and Department in your program:
 1. You can get Employee and Department table from the SQL Server from your program with JDBC/ODBC call
 2. Save each table in CSV file then read directly from each CSV file for Employee and Department
8. You can assume anything else to reduce the complexity.

Document your Lab2 in doc file by explaining all the assumptions and the design of your program. Add all the execution outputs with the output of each New Table of each operation step and the final result of the query. Submit your documentation, your source codes, and the executables with all the input and output files.

COMPANY DATABASE

EMPLOYEE

FNAME	MINIT	LNAME	SSN	BDATE	ADDRESS	SEX	SALARY	SUPERSSN	DNO
John	B	Smith	123456789	09-Jan-55	731 Fondren, Houston, TX	M	30000	987654321	5
Franklin	T	Wong	333445555	08-Dec-45	638 Voss, Houston, TX	M	40000	888665555	5
Joyce	A	English	453453453	31-Jul-62	5631 Rice, Houston, TX	F	25000	333445555	5
Ramesh	K	Narayan	666884444	15-Sep-52	975 Fire Oak, Humble, TX	M	38000	333445555	5
James	E	Borg	888665555	10-Nov-27	450 Stone, Houston, TX	M	55000		1
Jennifer	S	Wallace	987654321	20-Jun-31	291 Berry, Bellaire, TX	F	43000	888665555	4
Ahmad	V	Jabbar	987987987	29-Mar-59	980 Dallas, Houston, TX	M	25000	987654321	4
Alicia	J	Zelaya	999887777	19-Jul-58	3321 Castle, SPring, TX	F	25000	987654321	4

DEPARTMENT

DNAME	DNUMBER	MGRSSN	MGRSTARTDATE
Headquarters	1	888665555	19-Jun-71
Administration	4	987654321	01-Jan-85
Research	5	333445555	22-May-78
Automation	7	123456789	06-Oct-05

DEPENDENT

ESSN	DEPENDENT_NAME	SEX	BDATE	RELATIONSHIP
123456789	Alice	F	31-Dec-78	Daughter
123456789	Elizabeth	F	05-May-57	Spouse
123456789	Michael	M	01-Jan-78	Son
333445555	Alice	F	05-Apr-76	Daughter
333445555	Joy	F	03-May-48	Spouse
333445555	Theodore	M	25-Oct-73	Son
987654321	Abner	M	29-Feb-32	Spouse

DEPT_LOCATIONS

DNUMBER	DLOCATION
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

PROJECT

PNAME	PNUMBER	PLOCATION	DNUM
ProductX	1	Bellaire	5
ProductY	2	Sugarland	5
ProductZ	3	Houston	5
Computerization	10	Stafford	4
Reorganization	20	Houston	1
Newbenefits	30	Stafford	4

WORKS_ON

ESSN	PNO	Hours
123456789	1	32.5
123456789	2	7.5
333445555	2	10
333445555	3	10
333445555	10	10
333445555	20	10
453453453	1	20
453453453	2	20
666884444	3	40
888665555	20	
987654321	20	15
987654321	30	20
987987987	10	35
987987987	30	5
999887777	10	10
999887777	30	30