PROJECT REPORT

Group members:

Name- Shubham Saoji
Ufid- 26364957

How to run project:

make clean //to remove compiled files

make test.out
./test.out //Run project

make gtest //to compile google test cases
./gtest //to run test cases

Youtube Video Link - https://www.youtube.com/watch?v=waclgQiPbeg&feature=youtu.be

yyparse()- used to parse CNF

Join_func- Prints corresponding function in case of Join.
Sum_func- Corresponding function is printed in case of sum.
Groupby_func- Prints function in case of Groupby.
InitPlotSchema- Plots schema for all tables.
InitStatistics- Initialize statistics obj.

Statistics Class and Methods:

The following are the container-oriented operations having to do with storing simple data about the relations and attributes that the Statistics object will operate over.

AddRel:
This operation adds another base relation into the structure. Name of Relation and number of tuples are provided as input parameters. The parameter set tells the statistics object what the name and size of the new relation is (size is given in terms of the number of tuples).

AddAtt:
This operation adds an attribute to one of the base relations in the structure. The parameter set tells the Statistics object what the name of the attribute is, what relation the attribute is attached to, and the number of distinct values that the relation has for that particular attribute. If numDistincts is initially passed in as a –1, then the number of distincts is assumed to be equal to the number of tuples in the associated relation.

Read:

Statistics object also has the ability to read from text file. File is loaded from filepath given as input parameter. Data is read from statistics.txt

Write:
The Statistics object also has the ability to write itself to a text file. So file at filepath provided as input parameter is loaded. If file is not found, error is not thrown as per requirement. In case file is found, object can write data into this text file. We are writing output in output41.txt file.

Apply:
The Apply operation uses the statistics stored by the Statistics class to simulate a join of all of the relations listed in the relNames parameter. This join is performed using the predicates listed in the parameter parseTree. Join can be understood as when two or more relations are within the same subset, it means that they have been “joined” and they do not exist independently anymore. This function doesn’t actually perform join as join has been implemented in relOps.

Estimate:
This operation is exactly like Apply, except that it does not actually change the state of the Statistics object. Instead, it computes the number of tuples that would result from a join over the relations in relNames, and returns this to the caller.

Screenshot of gtest: