

CSE 562: Project #3 (due 12/10/06)

Description

You are given a subset of SQL2 consisting of:

- relation definitions of the form:

```
CREATE TABLE Table(  
   $A_1$  VARCHAR2( $n_1$ )  
  ...  
   $A_k$  VARCHAR2( $n_k$ ))
```

Additionally, one attribute is designated as **PRIMARY KEY**. For this attribute a primary index should be built.

- DROP TABLE *Table*.
- index definitions of the form:

```
CREATE INDEX I ON Table(A).
```

- DROP INDEX *I*.
- SQL2 single-tuple INSERT commands.
- the @File command to execute an .sql file.
- queries of the following form:

```
SELECT List of Attributes  
FROM List of Range Variables  
WHERE Condition.
```

The **WHERE Condition** is a conjunction of atomic equality conditions of the form $R.A = c$ or $R.A = S.B$ where A and B are attributes, c a constant, and R and S different range variables. You can assume that there are at most *two* range variables in the query.

You are supposed to implement:

1. a parser for the input
2. a query evaluator for the above subset of SQL2, whose repertoire includes the following operations:
 - table scan
 - index scan

- filter
- cartesian product
- nested-loops join
- index join.

3. display of the selected evaluation plan.

The evaluator should pick an evaluation plan that uses the existing indexes as much as possible. The evaluator should use the implementation of the file and index operations developed in Project #2. The operations should be pipelined: no intermediate results should be created.

Extra credit

You may get up to 10% of the final grade for implementing one or more of the following:

- cost-based optimization in the presence of multiple range variables
- other join methods
- aggregation with GROUP BY
- views.