Query Optimization

3. Exercise
Due 10.11.2014, 9 AM
submit via email (`Andrey.Gubichev@in.tum.de`)

Exercise 1

The selectivity estimations (from the previous homework) are far from perfect.

- Construct two SQL query examples for IMDB data, where our estimations are very "bad", i.e. the logical plan will be suboptimal (w.r.t. C_{out}), if we use these estimations. Give the optimal query plan.

- Construct a SQL query for IMDB where the PostgreSQL cardinality estimates are wrong and lead to suboptimal PostgreSQL plan. Provide the optimal plan and measure its performance.

Use our IMDB dump to load data into PostgreSQL, and have a look at two example SQL queries to get familiar with the schema.

Exercise 2

Give an example query instance where the optimal join tree (using C_{out}) is bushy and includes a cross product. Note: the query graph should be connected!