MapReduce Exercises

October 31, 2017

Consider the document `mapreduce.pdf` in the Lecture Notes. In that document, you can find in Sections 2.3.3-2.3.7 how certain relational algebra operations can be implemented in MapReduce. In particular, look at the map and reduce functions for various RA operators.

In the following exercises, you are asked to write simulations of MapReduce in PostgreSQL.[1]

1. Write a simulation in PostgreSQL of a MapReduce program that implements $\pi_A(R)$ where $R(A, B)$ is a relation. You can assume that the domains of A and B are INTEGER.

2. Write a simulation in PostgreSQL of a MapReduce program that implements the set difference of two unary relations $R(A)$ and $S(A)$, i.e., the relation $R(A) - S(A)$. You can assume that the domain of A is INTEGER.

3. Write a simulation in Postgres of a MapReduce program that implements the natural join $R \bowtie S$ of two relations $R(A, B)$ and $S(B, C)$. You can assume that the domains of A, B, and C are INTEGER.

[1] To do this, you should look at the document `word_count_mapreduce.pdf` for an example of such a simulation. In particular, you should write your solutions in the style as done in that document.