CS 430 Database Systems
Solution for Assignment 2

1. Given two relations $R1$ and $R2$, where $R1$ contains $N1$ tuples and $R2$ contains $N2$ tuples, and $N2 > N1 > 0$, give the maximum and minimum possible sizes (in tuples) for the result relation produced by each of the following relational algebra expressions. In each case, state any assumptions about the schemas for $R1$ and $R2$ that are needed to make the expression meaningful.

   (a) $R1 \cup R2$
   Assumption: $R1$ and $R2$ are union compatible
   Min: $N2$
   Max: $N1 + N2$

   (b) $R1 \cap R2$
   Assumption: $R1$ and $R2$ are union compatible
   Min: 0
   Max: $N1$

   (c) $R1 - R2$
   Assumption: $R1$ and $R2$ are union compatible
   Min: 0
   Max: $N1$

   (d) $R1 \times R2$
   Min: $N1 * N2$
   Max: $N1 * N2$

   (e) $\sigma_{a=5}(R1)$
   Assumption: $R1$ has an attribute named a
   Min: 0
   Max: $N1$

   (f) $\pi_a(R1)$
   Assumption: $R1$ has an attribute named a

Min: 1
Max: $N1$

(g) $R1/R2$
Assumption: The set of attributes of $R2$ is a subset of the attributes of $R1$.
Min: 0
Max: 0

2. Consider the following schema:

Suppliers(sid: *integer*, sname: *string*, address: *string*)
Parts(pid: *integer*, pname: *string*, color: *string*)
Catalog(sid: *integer*, pid: *integer*, cost: *real*)

The key fields are underlined, and the domain of each field is listed after the field name. Thus *sid* is the key for Suppliers, *pid* is the key for Parts, and *sid* and *pid* together form the key for Catalog. The Catalog relation lists the prices charged for parts by Suppliers. Write the following queries in relational algebra.

(a) Find the *names* of suppliers who supply some red part.
$\pi_{sname}(\pi_{sid}((\pi_{pid}\sigma_{color=\text{"red"}}Parts) \bowtie Catalog) \bowtie Suppliers)$

(b) Find the *sids* of suppliers who supply some red or green part.
$\pi_{sid}((\pi_{pid}\sigma_{color=\text{"red"} \vee color=\text{"green"}}Parts) \bowtie Catalog)$

(c) Find the *sids* of suppliers who supply some red part or are at 221 Packer Ave.

$$\pi_{sid}((\pi_{pid}\sigma_{color=\text{"red"}}Parts) \bowtie Catalog) \cup \pi_{sid}(\sigma_{address=\text{"221ParkerAve."}}Supplier)$$

(d) Find the *sids* of suppliers who supply some red part and some green part.

$$\pi_{sid}((\pi_{pid}\sigma_{color=\text{"red"}}Parts \cap (\pi_{pid}\sigma_{color=\text{"green"}}Parts)) \bowtie Catalog)$$

(e) Find the *sids* of suppliers who supply every part.

$$(\pi_{sid,pid}Catalog)/(\pi_{pid}Parts)$$

(f) Find the *sids* of suppliers who supply every red part.

$$(\pi_{sid,pid}Catalog)/(\pi_{pid}\sigma_{color=\text{"red"}}Parts)$$

(g) Find the *sids* of suppliers who supply every red or green part.

$$(\pi_{sid,pid}Catalog)/(\pi_{pid}\sigma_{color="red"\vee color="green"}Parts)$$

(h) Find the *sids* of suppliers who supply every red part or supply every green part.

$$(\pi_{sid,pid}Catalog)/(\pi_{pid}\sigma_{color="red"}Parts)\cup(\pi_{sid,pid}Catalog)/(\pi_{pid}\sigma_{color="green"}Parts)$$

(i) Find the pairs of *sids* such that the supplier with the first *sid* charges more for some part than the supplier with the second *sid*.

$$\pi_{R1.sid,R2.sid}(\sigma_{R1.pid=R2.pid\wedge R1.sid\neq R2.sid\wedge R1.cost>R2.cost}(\rho_{R1}Catalog\times\rho_{R2}Catalog))$$

(j) Find the *pids* of parts that are supplied by at least two different suppliers.

$$\pi_{R1.pid}(\sigma_{R1.pid=R2.pid\wedge R1.sid\neq R2.sid}(\rho_{R1}Catalog\times\rho_{R2}Catalog))$$

(k) Find the *pids* of the most expensive parts supplied by suppliers names Yosemite Sham.

$t1\leftarrow((\pi_{sid}\sigma_{sname="Yosemite\ Sham"}Suppliers)\bowtie Catalog)$
$t4\leftarrow\pi_{t2.sid,t2.pid,t2.cost}\sigma_{t2.cost<t3.cost}(\rho_{t2}(t1)\times\rho_{t3}(t1))$
$\pi_{pid}(\rho_{t5(sid,pid,cost)}t1-(\rho_{t6(sid,pid,cost)}t4)$

(l) Find the *pids* of parts supplied by every supplier at less than \$200. (If any supplier either does not supply the part or charges more than \$200 for it, the part is not selected.)

$$(\pi_{pid,sid}(\sigma_{cost<200}Catalog))/(\pi_{sid}Suppliers)$$