Database Management Systems: Practice Problem Set I: Relational Algebra

Q1. Suppose you are given the relational model below:

REGION(R_REGIONKEY, R_NAME, R_COMMENT)
PRIMARY KEY(R_REGIONKEY)

NATION(N_NATIONKEY, N_NAME, N_REGIONKEY, N_COMMENT)
PRIMARY KEY(N_NATIONKEY)
FOREIGN KEY(N_REGIONKEY) REFERENCES REGION(R_REGIONKEY)

PART(P_PARTKEY, P_NAME, P_MFGR, P_BRAND, P_TYPE, P_SIZE, P_CONTAINER, P_RETAILPRICE, P_COMMENT)
PRIMARY KEY(P_PARTKEY)

SUPPLIER(S_SUPPKEY, S_NAME, S_ADDRESS, S_NATIONKEY, S_PHONE, S_ACCTBAL, S_COMMENT)
PRIMARY KEY(S_SUPPKEY)
FOREIGN KEY(S_NATIONKEY) REFERENCES NATION(N_NATIONKEY)

PARTSUPP(PS_PARTKEY, PS_SUPPKEY, PS_AVAILQTY, PS_SUPPLYCOST, PS_COMMENT)
PRIMARY KEY(PS_PARTKEY, PS_SUPPKEY)
FOREIGN KEY PS_PARTKEY REFERENCES PART(P_PARTKEY)
FOREIGN KEY PS_SUPPKEY REFERENCES SUPPLIER(S_SUPPKEY)

Write down the following queries using relational algebra. You may use any valid algebra expression. Make sure you rename attributes whenever selection conditions are ambiguous.

1. Find all parts in the PARTS relation.

2. Find all parts with available quantity greater than 1000 or are supplied by a country in 'East Asia' geographic region. Return the name of the part.

3. Find all nations with at least two different suppliers. Return the name of the nation.

4. Find all suppliers who do not supply any parts with retail price greater than 1,000. Return the name of the suppliers.

5. Find pairs of suppliers who supply exactly the same parts, no more, no less. Return the supplier key pairs.

Answer.

1. $\sigma_{true} PART$.
2. Find parts with available quantity is greater than 1000.
   $Temp1 := \Pi_{P\_NAME}(PART \bowtie_{P\_PARTKEY=PS\_PARTKEY} (\sigma_{PS\_AVAILQTY>1000}(PARTSUPP)))$

   Find parts which are supplied by a country in 'East Asia' geographic region
   $Temp2 := \Pi_{P\_NAME}(PART \bowtie_{P\_PARTKEY=PS\_PARTKEY} (PARTSUPP \bowtie_{PS\_SUPPKEY=S\_SUPPKEY} (SUPPLIER \bowtie_{S\_NATIONKEY=N\_NATIONKEY} (NATION \bowtie_{N\_REGIONKEY=R\_REGIONKEY} (\sigma_{R\_NAME='EastAsia'}(REGION))))))$

   $Result := Temp1 \cup Temp2$.

3. $Temp1 := (\Pi_{S\_SUPPKEY,S\_NATIONKEY} SUPPLIER)[SUPPKEY1, NATIONKEY1]$

   Find the nations which have two different suppliers.
   $Temp2 := \Pi_{S\_NATIONKEY}(SUPPLIER \bowtie_{S\_NATIONKEY=NATIONKEY1 \text{ AND } SUPPKEY1<>S\_SUPPKEY} Temp1)$
   $Result := \Pi_{N\_NAME}(NATION \bowtie_{S\_NATIONKEY=N\_NATIONKEY} Temp2)$.

4. Find the suppliers one of whose parts's retail price is greater than 1000
   $Temp1 := \Pi_{PS\_SUPPKEY}(PARTSUPP \bowtie_{PS\_PARTKEY=P\_PARTKEY} (\sigma_{P\_RETAILPRICE>1000}(PART)))$

   From all suppliers, substract the suppliers in Temp1.
   $Temp2 := \Pi_{N\_NAME}(SUPPLIER \bowtie_{S\_SUPPKEY=S\_SUPPKEY} (\Pi_{S\_SUPPKEY} SUPPLIER - Temp1))$.

5. Find all pairs of partkey and suppkey
   $Temp1 := (\Pi_{PS\_PARTKEY,PS\_SUPPKEY}(PARTSUPP))[PKEY1, SKEY1]$
   Rename the partkey and suppkey
   $Temp2 := Temp2[PKEY2, SKEY2]$
   Find all pairs of suppliers who supply one common part
   $Temp3 := \Pi_{SKEY2,SKEY1,PKEY1}(Temp1 \bowtie_{SKEY1<>SKEY2 AND PKEY1=PKEY2} Temp2)$
   Find all pairs of suppliers who could possibly supply a common part
   $Temp4 := (\Pi_{SUPPKEY}(PARTSUPP))[SKEY2] \times Temp1$
   Find all pairs of suppliers who don't supply one common part
   $Temp5 := \Pi_{SKEY2,SKEY1}(Temp4 - Temp3)$
   $Result := \Pi_{SKEY1,SKEY2}(Temp3) - Temp5$

Q2. Consider the following schema of a *company database*:

Employees(eid: integer, ename: string, address: string, supereid: integer)
Departments(did: integer, dname: string)
Projects(pid: integer, pname: string, did: integer)
Works_on(eid: integer, pid: integer, hours: integer)

Each Employee has a supervisor (another Employee) referenced by his/her supereid. Projects are uniquely assigned to a Department. The Works_on relation records which Employee works on which Project for how many hours a week.

Formulate each of the following queries in relational algebra (RA).

1. For each Employee, find his / her name and the name of his / her supervisor.

$$\rho(Supervisors, (1 \to supeid, 2 \to supname), Employees)$$
$$\rho(EmpSup, (\sigma_{supereid=supeid}(Employees \times Supervisors)))$$
$$\pi_{ename,supname} EmpSup$$

2. Find the eids of Employees who work on a project of every Department, i.e. find the eids of Employees who work for (a project of) every Department.

$$(\pi_{eid,did}(Works\_on \infty Projects))/(\pi_{did} Departments)$$

3. Find the pid of Projects of Department with dname = "Toys" for which at least two different Employees work.

$$\rho(R1, \pi_{eid,pid}(Works\_on \infty Projects \infty (\sigma_{dname="Toys"} Departments)))$$
$$\rho(R2, (1 \to eid1, 2 \to pid1, 3 \to eid2, 4 \to pid2), R1 \times R1)$$
$$\rho(R3, \sigma_{pid1=pid2 \text{ AND } eid1 \neq eid2} R2)$$
$$\pi_{pid1} R3$$