Relational algebra - Part 2

Exercise

Consider the Supplier-Parts-Catalog schema from the previous question. State what the following queries compute:

$$\pi_{\texttt{sname}}((\sigma_{\texttt{colour}=\text{`}red\text{'}}P) \bowtie (\sigma_{\texttt{cost}<100}C) \bowtie S)$$

Solution

Find the Supplier names of the suppliers who supply a red part that costs less than 100 dollars

Exercise

$$\pi_{\texttt{sname}}(\pi_{\texttt{sid}}((\sigma_{\texttt{colour}=\text{`}red\text{'}}P) \bowtie (\sigma_{\texttt{cost}<100}C) \bowtie S))$$

Solution

Nothing. After the projection on `sid`, it is the only attribute left

Exercise

$$(\pi_{\texttt{sname}}((\sigma_{\texttt{colour}=\text{`}red\text{'}}P) \bowtie (\sigma_{\texttt{cost}<100}C) \bowtie S))$$

$$\cap(\pi_{\texttt{sname}}((\sigma_{\texttt{colour}=\text{`}green\text{'}}P) \bowtie (\sigma_{\texttt{cost}<100}C) \bowtie S))$$

Solution

Find the Supplier names of the suppliers who supply a red part that costs less than 100 dollars and a green part that costs less than 100 dollars

Exercise

$$(\pi_{\texttt{sid}}((\sigma_{\texttt{colour}=\text{`}red\text{'}}P) \bowtie (\sigma_{\texttt{cost}<100}C) \bowtie S))$$

$$\cap(\pi_{\texttt{sid}}((\sigma_{\texttt{colour}=\text{`}green\text{'}}P) \bowtie (\sigma_{\texttt{cost}<100}C) \bowtie S))$$

Solution

Find the Supplier ids of the suppliers who supply a red part that costs less than 100 dollars and a green part that costs less than 100 dollars.

Exercise

$$\pi_{\texttt{sname}}((\pi_{\texttt{sid},\texttt{sname}}((\sigma_{\texttt{colour}=\text{`}red\text{'}}P) \bowtie (\sigma_{\texttt{cost}<100}C) \bowtie S))$$

$$\cap(\pi_{\texttt{sid},\texttt{sname}}((\sigma_{\texttt{colour}=\text{`}green\text{'}}P) \bowtie (\sigma_{\texttt{cost}<100}C) \bowtie S)))$$

Solution

Find the Supplier names of the suppliers who supply a red part that costs less than 100 dollars and a green part that costs less than 100 dollars.

Exercise

Consider the following schema

```
Frequents(Drinker, Bar)
Serves (Bar, Beer)
Likes (Drinker, Beer)
```

You can abbreviate them as F, S, L

Print the bars that serve a beer that Joe likes

Solution

$$\pi_{\texttt{Bar}}(\sigma_{\texttt{Drinker}=\texttt{'Joe'}}(S \bowtie L))$$

Exercise

```
Frequents(Drinker, Bar)
Serves (Bar, Beer)
Likes (Drinker, Beer)
```

Print the drinkers that frequent at least one bar that serves a beer that they like

Solution

$$\pi_{\texttt{Drinker}}(S \bowtie F \bowtie L)$$

Exercise

```
Frequents(Drinker, Bar)
Serves (Bar, Beer)
Likes (Drinker, Beer)
```

Print the drinkers that frequent only bars that serve some beer that they like (assume each drinker likes at least one beer and frequents at least one bar)

Solution

1. Drinkers that frequent a bar that serves no beer that that they like

$$R = \pi_{\texttt{Drinker}}(F - \pi_{\texttt{Drinker},\texttt{Bar}}(S \bowtie L))$$

2. Solution

$$\pi_{\texttt{Drinker}}(F) - R$$

Exercise

```
Frequents(Drinker, Bar)
Serves (Bar, Beer)
Likes (Drinker, Beer)
```

Print the drinkers that frequent no bar that serves a beer that they like

Solution

$$\pi_{\texttt{Drinker}}(F) - \pi_{\texttt{Drinker}}(S \bowtie F \bowtie L)$$

Exercise

Consider the following relations containing airline flight information:

```
Flights(flno: integer, from: string, to: string, distance: integer,
        departs: time, arrives: time)
Aircraft( aid: integer, aname: string, cruisingrange: integer)
Certified(eid: integer, aid: integer)
Employees(eid: integer, ename: string, salary: integer)
```

Abbreviate the relations as F, A, C, E. Note that the `Employees` relation describes pilots and other kinds of employees as well; every pilot is certified for some aircraft (otherwise, he or she would not qualify as a pilot), and only pilots are certified to fly.

Write the following queries in relational algebra, or specify when this is impossible.

Find the eids of pilots certified for some Boeing aircraft

Solution

$$\pi_{\texttt{eid}}(\sigma_{\texttt{aname}=\text{`}Boeing\text{'}}(A \bowtie C))$$

Exercise

```
Flights(flno: integer, from: string, to: string, distance: integer,
        departs: time, arrives: time)
Aircraft( aid: integer, aname: string, cruisingrange: integer)
Certified(eid: integer, aid: integer)
Employees(eid: integer, ename: string, salary: integer)
```

Find the names of pilots certified for some Boeing aircraft.

Solution

$$\pi_{\texttt{ename}}(\sigma_{\texttt{aname}=\text{`}Boeing\text{'}}(A \bowtie C \bowtie E)$$

Exercise

```
Flights(flno: integer, from: string, to: string, distance: integer,
        departs: time, arrives: time)
Aircraft( aid: integer, aname: string, cruisingrange: integer)
Certified(eid: integer, aid: integer)
Employees(eid: integer, ename: string, salary: integer)
```

Find the aids of all aircraft that can be used on non-stop flights from Rome to Mumbai.

Solution

1. $R = \sigma_{\texttt{from}=\text{`}Rome\text{'} \wedge \texttt{to}=\text{`}Mumbai\text{'}}(F)$
2. Solution:

$$\pi_{\texttt{aid}}(\sigma_{\texttt{cruisingrange}>=\texttt{distance}}(A \times R))$$

Exercise

```
Flights(flno: integer, from: string, to: string, distance: integer,
        departs: time, arrives: time)
Aircraft( aid: integer, aname: string, cruisingrange: integer)
Certified(eid: integer, aid: integer)
Employees(eid: integer, ename: string, salary: integer)
```

Identify the flights that can be piloted by a pilot whose salary is more than $100,000.

Solution

$$\pi_{\texttt{flno}}(\sigma_{\texttt{cruisingrange}>=\texttt{distance}\wedge\texttt{salary}>10000}(A \bowtie F \bowtie C \bowtie E))$$

Exercise

```
Flights(flno: integer, from: string, to: string, distance: integer,
        departs: time, arrives: time)
Aircraft( aid: integer, aname: string, cruisingrange: integer)
Certified(eid: integer, aid: integer)
Employees(eid: integer, ename: string, salary: integer)
```

Find the names of pilots who can operate planes with a range greater than
3,000 miles but are not certified on any Boeing aircraft.

Solution

1. $R = \pi_{\texttt{eid}}(\sigma_{\texttt{cruisingrange}>=3000}(A \bowtie C))$
2. Solution

$$\pi_{\texttt{ename}}(E \bowtie (R - \pi_{\texttt{eid}}(\sigma_{\texttt{aname}=\text{`}Boeing\text{'}}(A \bowtie C))$$

Exercise

```
Flights(flno: integer, from: string, to: string, distance: integer,
        departs: time, arrives: time)
Aircraft( aid: integer, aname: string, cruisingrange: integer)
Certified(eid: integer, aid: integer)
Employees(eid: integer, ename: string, salary: integer)
```

Find the eids of employees who make the highest salary.

Solution

1. $R_1 = E$
2. $R_2 = E$
3. Find the employees who do not make the highest salary

$$R_3 = \pi_{R2.\texttt{eid}}(R_1 \bowtie_{R_1.\texttt{salary} > R_2.\texttt{salary}} R_2)$$

4. Solution

$$\pi_{\texttt{eid}} E - R_3$$

Exercise

```
Flights(flno: integer, from: string, to: string, distance: integer,
        departs: time, arrives: time)
Aircraft( aid: integer, aname: string, cruisingrange: integer)
Certified(eid: integer, aid: integer)
Employees(eid: integer, ename: string, salary: integer)
```

Find the eids of employees who make the second highest salary.

Solution

The approach taken is similar to the solution for the previous exercise. First find all the employees who do not have the highest salary. Remove these from the original list of employees and repeat.

1. $R_1 = E$
2. $R_2 = E$
3. Find employees who do not have the highest salary

$$R_3 = \pi_{R_2.\texttt{eid}}(R_1 \bowtie_{R_1.\texttt{salary} > R_2.\texttt{salary}} R_2)$$

4. $R_4 = E \bowtie R_3$
5. $R_5 = E \bowtie R_3$

$$R_6 = \pi_{R_5.\texttt{eid}}(R_4 \bowtie_{R_4.\texttt{salary} > R_5.\texttt{salary}} R_5)$$

6. Solution $\pi_{\texttt{eid}}(R_3) - R_6$

Exercise

```
Flights(flno: integer, from: string, to: string, distance: integer,
        departs: time, arrives: time)
Aircraft( aid: integer, aname: string, cruisingrange: integer)
Certified(eid: integer, aid: integer)
Employees(eid: integer, ename: string, salary: integer)
```

Find the eids of employees who are certified for the largest number of aircraft.

Solution

This cannot be expressed in the relational algebra without the ability to count (which we shall study later for SQL)

Exercise

```
Flights(flno: integer, from: string, to: string, distance: integer,
        departs: time, arrives: time)
Aircraft( aid: integer, aname: string, cruisingrange: integer)
Certified(eid: integer, aid: integer)
Employees(eid: integer, ename: string, salary: integer)
```

Find the eids of employees who are certified for exactly three aircraft.

Solution

1. Find the employees who are certified for at least three aircraft and at least four aircraft. Subtract the second from the first.
2. $R_1 = C$
3. $R_2 = C$
4. $R_3 = C$
5. $R_4 = C$
6. $$R_5 = \pi_{\texttt{eid}}(\sigma_{(R_1.\texttt{eid}=R_2.\texttt{eid}=R_3.\texttt{eid}))\wedge(R_1.\texttt{aid}\neq R_2.\texttt{aid}\neq R_3.\texttt{aid})}(R_1 \times R_2 \times R_3))$$
7. $$R_6 = \pi_{\texttt{eid}}(\sigma_{(R_1.\texttt{eid}=R_2.\texttt{eid}=R_3.\texttt{eid}=R_4.\texttt{eid})\wedge(R_1.\texttt{aid}\neq R_2.\texttt{aid}\neq R_3.\texttt{aid}\neq R_4.\texttt{aid})}(R_1 \times R_2 \times R_3 \times R_4))$$
8. Solution $R_5 - R_6$

Reminder: As explained in class, the inequalities are just an abbreviation for more complicated formulae

Exercise

```
Flights(flno: integer, from: string, to: string, distance: integer,
        departs: time, arrives: time)
Aircraft( aid: integer, aname: string, cruisingrange: integer)
Certified(eid: integer, aid: integer)
Employees(eid: integer, ename: string, salary: integer)
```

Find the total amount paid to employees as salaries.

Solution

Impossible, as there is no operator to sum values

Exercise

```
Flights(flno: integer, from: string, to: string, distance: integer,
        departs: time, arrives: time)
Aircraft( aid: integer, aname: string, cruisingrange: integer)
Certified(eid: integer, aid: integer)
Employees(eid: integer, ename: string, salary: integer)
```

Is there a sequence of flights from Madison to Timbuktu? Each flight in the sequence is required to depart from the city that is the destination of the previous flight; the first flight must leave Madison, the last flight must reach Timbuktu, and there is no restriction on the number of intermediate flights. Your query must determine whether a sequence of flights from Madison to Timbuktu exists for any input Flights relation instance.

Solution

Impossible