Exercise 3

SQL queries

Relational algebra

Example

Courses

Code	Name	nrStudents
AB123	Math	19
CD234	Physics	24
EF345	Karaoke	23
GH456	PERL	20

- SELECT name
  FROM Courses
  WHERE nrStudents > 20

- $\pi_{name}(\sigma_{nrStudents>20}(\text{Courses}))$

Physics
Karaoke

The most common question

What the hell is relational algebra good for?

The DBMS uses relational algebra

- A DBMS may have many different ways of implementing the relational algebra operations.
- The aim of query optimization is to choose the most efficient one.
- To do this, it uses formulae that estimate the costs for a number of options and selects the one with the lowest cost.

Projection

- Which attributes
- Which columns
- SELECT A,C,E
- $\pi_{A,C,E}$

A	B	C	D	E	F

Selection

- Which tuples
- Which rows
- WHERE E > 5
- $\sigma_{E > 5}$
- DON'T confuse SELECT (projection) for selection!!

A	B	C	D	E	F
				4	
				7	
				5	
				8	
				3	

SQL vs. Relational Algebra

SELECT	X	
FROM	T	
WHERE	C	
GROUP BY	Y	$\tau_Z(\pi_X(\sigma_D(\gamma_Y(\sigma_C(T)))))$
HAVING	D	
ORDER BY	Z	

Combining tables

- Set (actually bag) operations
- Cartesian product
- Joins

Set (actually bag) operations

R=

A	B
1	2
3	4

S=

A	B
1	2
5	6

R UNION ALL S
$R \cup S$

A	B
1	2
3	4
1	2
5	6

R INTERSECT S
$R \cap S$

A	B
1	2

R MINUS S
R - S

A	B
3	4

Joins

- 3 ”basic” joins:
  - Cartesian product
  - Conditional join
    - Theta join, Inner join, Equi join, Nonequi join, Natural join
  - Outer join

Cartesian product

R=

A	B
1	2
3	4

S=

C	D
5	6
7	8

R.A	R.B	S.C	S.D
1	2	5	6
1	2	7	8
3	4	5	6
3	4	7	8

R , S

$R \times S$

Conditional join = Inner join

R=

A	B
1	2
3	4

S=

B	C
3	4
4	5

R,S
WHERE C
or
R JOIN S ON C
$R \bowtie_{C} S$

If C is ”R.A = S.B”

R.A	R.B	S.B	S.C
3	4	3	4

If C is equality – Equi join

If C is inequality – Nonequi join

Beware of NULL!

Special case – Natural join

R=

A	B
1	2
3	4

S=

B	C
4	5
6	7

R NATURAL JOIN S

$R \bowtie S$

R.A	B	S.C
3	4	5

Outer join

R=

A	B
1	2
3	4

S=

C	D
4	5
6	7

R FULL OUTER JOIN S
ON R.B = S.C

R.A	R.B	S.C	S.D
1	2	NULL	NULL
3	4	4	5
NULL	NULL	6	7

R FULL/LEFT/RIGHT
OUTER JOIN S ON C

All rows in both/left/right table(s) will appear, and the rest will be filled with null if C does not match.

There is more…

Grouping
Renaming
Sorting

Also note that the terminology regarding joins is confused.
Inner join = Equi join (orafaq.com)
Inner join = Conditional join (Wikipedia)
Theta join = Conditional join (Course book)