

Database Management Systems

SQL (IV)

Sample Database Schema

- Departments(did, dname, managerId)
- Emps(eid, firstname, lastname, salary, workdept, hireDate)
- Projects(pno, title, respEid, respDid)
- WorksOn(eid, pno, sdate, edate)

Universal Qualification

- List the name of each employee who worked on EVERY project.
- Interpretation: For each employee, if there doesn't exist a project such that this employee didn't work on it, then this employee's name should be in the result.
- Mathematically, universal quantification can be translated to existential quantification plus double (table level) negations:
 $(\forall x, x \text{ is a project} \Rightarrow \text{emp } y \text{ worked on } x)$
 \equiv (is equivalent to)
 $(\nexists x, x \text{ is a project and emp } y \text{ NEVER worked on } x)$

Solution

- SQL:

```
select firstname, lastname
from Emps E
where not exists (select *
                  from Projects P
                  where P.pno not in (select pno
                                      from WorksOn W
                                      where W.eid = E.eid));
```

- Datalog:

```
Result(fn, ln) ::= Emps(eid, fn, ln, _, _, _)
                and ( $\forall x$ , Projects(x, _, _, _)  $\Rightarrow$  WorksOn(eid, x, _, _))
```

```
Result(fn, ln) ::= Emps(eid, fn, ln, _, _, _)
                and not (Projects(x, _, _, _) and not WorksOn(eid, x, _, _))
```

Aggregation in SQL

- Clause “group by” list divide tuples in the result table into groups according to the value(s) of the column(s) listed in the group by clause.
- aggregation functions:
max, min, count, sum, avg
- How is NULL handled? – It is ignored.
- If a column appears in the select list but not in the group by list, it must be inside an aggregation function
- having condition is applied on each group formed by group by clause
- usually aggregation functions are used in having condition

Example

- For each employee, list his/her name and count how many (distinct) projects he/she has/had worked on. (What about those employees who never worked on any project?)
- For each department, list the department's name and the highest salary earned by its employees.
- For each department, list the department's name, the number of employees working in it and their average and total salary.
- List the name(s) of the employee(s) who has the highest salary. (Note that query needs to return ALL qualified data.)
- List the name(s) of the departments that have more than 2 employees.

Sort the result

- using “order by” clause
- according to a column ascending (asc) or descending (desc), a secondary column asc or desc
- the default order is ascending
- null is treated as the largest value in Oracle

Set/Bag Based Operators

- What is “union compatible”?
- Union/Union All
- Intersect/Intersect All
- Minus/Minus All
- SQL_Q_1
union all — set/bag operators
SQL_Q_2;

Summary

- SELECT — decides the schema of the query result (relation/table)
- FROM — forms a big table by join/cross product all tables in this list, as the data source
- WHERE — applies to each tuple of the data source table formed in the FROM clause, screens out the ones that return false
- GROUP BY — divides the remaining tuples into groups with common values
- HAVING — applies to each group formed by the GROUP BY clause, screens out the whole groups that return false
- ORDER BY — sort the result table. It will only change the presentation of the result, but not the result itself.
- Sub queries, Set/Bag operators