

Database Systems, ITU, Fall 2006

Rasmus Pagh

November 8, 2006

Group project – part 3

This is the third part of the **mandatory** group project. It should be handed in **by e-mail** to the teaching assistant associated with your group, no later than:

Tuesday November 28, 23.59 PM.

Learning outcomes

The third group hand-in develops several core competences mentioned in the course description. After working on this hand-in you should have improved your ability to implement databases based on the relational model, and program complex queries in SQL.

Case description – continued

The hand-in concerns the continued development of the new database of IT@. At this point you should have a final E/R diagram for your system, and have created the corresponding normalized relations in Oracle. You should populate your database with data that allows you to test your system, in particular the queries below. You may use the method described at www.itu.dk/people/pagh/DBSE06/generating-sample-data.txt to generate sample data. It is fine that groups help each other with this part (import each other's data).

To shed light on the future use of the system, the management at IT@ have described a number of “use cases”. To see if your database design supports these, without implementing a complete system, you are given the task to program the queries specified below. If you did not make a data model for the whole system, ignore the queries that do not make sense in your data model. Note that a number of the queries are quite challenging. If some query is difficult (or impossible) to write, maybe because of the way you represented data, you should explain the difficulty. You may want to add or change data in order to test your queries properly. Five of your queries should additionally be stated as *relational algebra expressions*.

Queries on the common part for all groups:

1. “How many students do we have of each nationality?”
2. “What is the average total number of ECTS for students in each study line?”
3. “How many persons are registered in each of the different groups of persons?” (this may require several queries).
4. “Who in the faculty lives on the same address as another faculty?”
5. “What was the total number of ECTS for courses in the fall of 2005, specified for each study line?”

6. “Which students have taken a course (s)he did not have the competences for?”
7. “How many students need more competences to get a degree from their study line?”
8. “What students did not pass any course in 2005?”
9. “Who in the faculty are also alumni, and what were their average grade?”
10. “How many of our students live in an area with postal code above 2999?”

Queries on the optional parts:

1. “What is the average grade in courses and projects, respectively, for each student?”
2. “Is it true that grades in projects are higher than grades in courses?”
3. “How many people were employed each year (give a list of years and numbers)?”
4. “What is the size of the different departments?”
5. “How much does each department cost in salary?”
6. “Who was in the various bodies at the beginning of 2004?”
7. “Which body has had the highest number of different members?”
8. “What students enrolled in 2005 are lacking a tutor?”
9. “What tutors had less than three students assigned?”
10. “How much time has been spent in meetings (past or planned)?”

Extra credit: For queries that you have written with a correlated subquery, try to come up with an alternative query that has no correlated subquery.

What must be handed in

You must hand in:

- **Oracle user name.** The user name of an Oracle account at ITU where the relations reside. You don’t have to provide a password. Instead, you should make the relations viewable by others by issuing the command `GRANT SELECT ON <MyRelation> TO public` a number of times, substituting `<MyRelation>` with the name of each of your relations.
- **Your SQL queries, and the corresponding query results.** You may use the `spool` feature of SQL*Plus, as follows: Enter the command `spool myfile.txt` in SQL*Plus, after which the terminal output is recorded in the file `myfile.txt` in the Unix file system (written when you exit SQL*Plus). To retrieve the file use an SSH client (with file transfer mode), or move it into your `public_html` directory (viewable with a web browser).

For queries that you could not complete, you should explain what the difficulty is, and perhaps give some partial result (e.g. a query that seems useful as a subquery, or computes part of the result).

- **Relational algebra expressions corresponding to 5 of your queries.**

On the first page, clearly specify the members of your group (if someone dropped the project, don’t include him/her). The project should be sent as a **single file in PDF format** to your group’s TA. Start the subject line of the e-mail with `DBS:`.