Database management, spring 2002 **Exercise 1 (25.-27.3., 4.-5.4.)**

The minimum for active attendance: 3 tasks done
Obs! Neither lectures nor exercise sessions between 28.3. – 3.4.

The purpose of tasks 1-3 is simply to review some basics of the SQL language. For SQL, see the textbook (E&N, Ch. 8), Oracle manual etc. The SQL exercises concern the database described in E&N, page 205. The copy of the database is also in the course folder (room A412).

1. a) Write a SQL query to output the names of those employees who supervise some employee (a subordinate) working at the department other than their own department.
b) Modify the query so that it also outputs the names of the subordinates and the names of the departments (both for the supervisor and for the subordinate).

2. Write a SQL query to output the names of employees working in Houston. (Is the relation DEPT_LOCATIONS necessary at all in this database – or is it possible to find the locations of employees always using the project data?)
b) Write a SQL query to output the sum of the work hours for every employee. The result should be arranged alphabetically by the employee name.

3. Write a SQL query to output the names of the employees who have a spouse but no children.

4. In the Helsinki metropolitan area there are about 1 000 000 residents and (an inexact guess) about 2 000 associations. Let us assume that every resident is a member of 3 associations on the average. A database contains data for residents (at least ssn, name, city of residence, address), for associations (at least association_id, name, home city) and for memberships (resident, association, year_of_joining). Sketch the structure and contents of the database so that you can roughly estimate how much disk space is needed.

5. Let us implement the EMPLOYEE relation of the Company database as a heap file (E&N, Ch. 5.7). The size of a disk block is 6 KB, and there are 30 000 employees. Give a detailed representation for the tuple ('Franklin T Wong', 333445555, 08-DEC-45, '638 Voss, Houston, TX', 'M', 40000, 8886655555, 5) using a) fixed-length records, b) variable-length records according to Oracle format. The Oracle format is described on slide #27 in the (Finnish) lecture notes (Ch. 2) – and surely in some part of Oracle documentation which I, however, have forgotten just now!. (For every attribute, its identifier (1,2,3,…), length, and the value are stored.)
c) Give an estimate for the size of the file (in the number of blocks).

6. A file of 200 000 blocks is stored optimally on a disk pack having 20 disk surfaces, 4000 tracks per surface and 25 blocks per track.
a) How much time is spent in scanning through the file when the blocks are read in the order they are stored on the disk. The seek time is 1-12 ms (the average is 6 ms) and block transfer time 0.2 ms. (Gaps between sectors need not to be considered.)
b) What is the time needed to read all blocks separately (in an arbitrary order)?
c) Let us suppose that X arbitrary records are needed. With what number X it is more efficient to scan through the whole file as compared to the series of single requests?

7. What is the effect of the following changes to the components of disk access time (seek + latency + transfer; E&N, p. 121).
a) rotational speed is doubled,
b) packing density is doubled (i.e. the capacity of a track is doubled too),
c) every read/write arm is equipped with two read/write heads (instead of one),
d) block size of the file is doubled?