

**Helsinki university, Department of Computer Science
Introduction to Databases, 17.4.2001, H.Laine**

Write your name, date of birth (personal number if you have not registered), name and date of the course and your signature on each answer paper.

1. Let's consider the relations $A(a,b,c,d,e)$, cardinality $n > 0$ tuples, and $B(a \rightarrow A, f, e, d, g)$, cardinality $m > 0$ tuples. Notations: π projection, \times Cartesian product, $| \times |$ join, $*$ natural join, σ selection.
 - a) List the columns of $A * B$.
 - b) What is the join condition in $A * B$.
 - c) What is the cardinality (number of tuples) of $A | \times |_{A.a=B.a} B$?
 - d) Compare the cardinalities of $\pi_a(A)$ and $\pi_a(B)$?
 - e) In the operation $A \cup B$ allowed and if it is what are the prerequisites ?
 - f) Can there be null values in column f of relation B? (6 p)

Lets consider the following tables dealing with thesis administration

```
student(identifier, name, address, speciality)
thesis(number, title, date_specified, prepared_by ->instructor,
       reserved_for ->student, date_reserved, date_started, date_finished,
       state, gredit, description)
    State: under_preparation, started, finished
instructor(name, degree)
    Name in form last name space first name
instruction(work->thesis, instructor->instructor, first_date, last_date, role)
    Role: 1 = primary instructor, 2 = assisting instructor
evaluation(work->thesis, phase, date_submitted, date_evaluated, statement)
    Phase: plan, final
evaluator((work,phase,date_submitted)->evaluation, inspector->instructor).
```

2. Express the following queries in SQL. Determine the proper order for the results.
 - a) Titles, and student's names and dates started for theses started in year 2001..
 - b) Titles and the role of instructor for theses instructed by Timo Alanko and not finished yet, primary instructions first.
 - c) Theses (title, name of student) that have been started but have not been assigned instructors.
 - d) The average duration (days) of thesis writing for theses finished in year 2000.
 - e) The yearly amounts of finished theses per grades. The function yearOf(Date) is assumed to give the year of the parameter Date.
 - f) Top ten evaluators in year 2000, i.e. a list of instructors and their number of evaluations in year 2000, the one with most evaluations first and only ten topmost listed. (12p)
3. The instructors of the thesis are usually assigned as its evaluators. Let's assume that this is also true for the final evaluation of the thesis whose title is 'MYCT-architecture'. Give the operations for registering the submission of this thesis for its final evaluation and assigning the evaluators. You may assume that the title uniquely identifies the thesis. (4p)
4. Explain the role and use of the Java class ResultSet in a JDBC database query. (3p)
5. You may substitute the exercise points (early Spring) by taking the following task. *If you want to include the exercise points just reply: "EXERCISE POINTS". A list of exercise points achieved will be available in the exam.*

Explain briefly the components of a database management system (5p)