University of Helsinki, Department of Computer Science
Introduction to databases, course exam, 16.10.2006, H. Laine

Write the name of the course, the date of the exam, your name clearly, your date of birth, and your signature on each separate answer paper.

Tasks 1-3 deal with the following recipe database. Column *noOfServings* in table *Course* indicates the number of servings in the recipe. *CookingTime* is expressed in minutes. Categories of courses include, for example, *soup*, *salad*, *appetiser*, *dessert*, and *main course*. Column *type* in table *Material* contains values like '*fish*', '*pork*', and '*vegetable*'. *Amount* in *Ingredients* is expressed in units specified in table *Material* (for example *kg*, *table spoon*, *apiece*). It contains the amount of material needed for the whole recipe. Table *biggest_courseid* contains the biggest *courseID* value in use. All identifiers are integers. Notation *x -> y* indicates that *x* is a foreign key that refers to table *y*.

```
course (courseID, name, easeOfPreparation, noOfServings, cookingTime)
  [1000 rows]
categories (course->course, category) [3000 rows]
material (materialID, name, type, unit, unitPrice) [200 rows]
ingredients(course->course, materialID->material, amount) [20000 rows]
instruction(course->course, phaseNo, description) [10000 rows]
biggest_courseid(highvalue) [1 row]
```

1.

a) Is the natural join of tables *course* and *material* allowed? If it is, what is the join condition and what is expected as the result of this operation?
b) How many rows there are in the result of the projection $\pi_{\mathbf{materialID}}$ **(*ingredients*)**?
c) How are the cardinalities (number of rows) of projections $\pi_{\mathbf{materialId}}$ **(*material*)** and $\pi_{\mathbf{materialId,unit}}$ ***(material)*** related to each other ?
d) How many rows there are in the result of the join ***course*** $\rhd\!\lhd_{\mathbf{courseID=course}}$ **ingredients** ?
e) A row is deleted from table *ingredients*. What should be done to the other tables to preserve the referential integrity of the database?

If you are not able to deduce the exact number of rows, give an estimate and justify it. (10p)

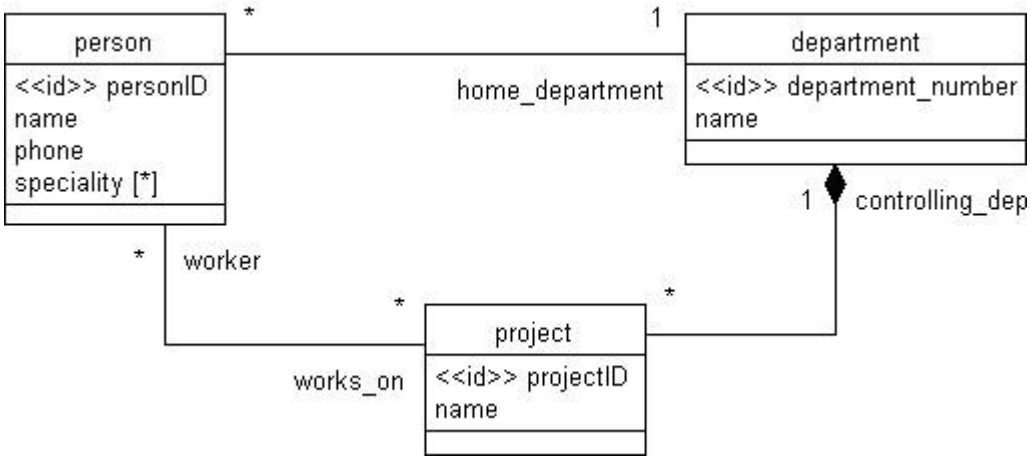
2. Consider the database defined above. Express the following queries in SQL. Define a proper order for the result tables.

a) Make a list of courses that are categorized as soups but not as appetisers.
b) What is the material cost for one portion of pea soup (course identifier is 100)?
c) List the names of courses that are defined to be *easy* to cook but contain more than 20 ingredients. (12p)

3. Consider the recipe database. A 'Tomato soup' is introduced as a variation of 'Tomato and onion soup' (courseID= 230). This variation does not contain onions (materialID=120). Otherwise the ingredients are the same. Two first cooking phases of the original recipe should be omitted. Explain what database operations are needed to register this variation in the database by making full use of the already existing recipe for the Tomato and onion soup. In addition to explaining the operations give them also in SQL. (8p)

Turn the paper for tasks 4 and 5.

4. The following schema defines a part of a project management database in the conceptual level. Stereotype <<id>> indicates an identifying attribute or relationship. Notation [*] indicates a multivalued attribute. How should this information be stored as tables in a relational database? Define the schema using the notation used in the beginning of this paper. (9p)



5. Let's consider a relation that contains information on purchase orders

```
purchase_order (personID, person_name, person_address, order_number,
        delivery_address, order_item_number, product_id, amount_ordered,
        dateOrdered).
```

a) Explain the meaning of the functional dependency *personID* à *order_number*

b) Express the rule 'A person can make only one order on the same day' as a functional dependency. (6p)

Turn the paper for tasks 1-3