

For this exam, consider the following schema of a simple university database. It includes information about instructors, students, and the courses offered. Feel free to remove this page from the exam.

- The **Students** table contains the id of the student (**sid**), his/her name (**sname**), age (in years), and gpa.

```
Students(sid: integer, sname: string,  
        age: integer, gpa: real)
```

– Key: **sid**

- The **Instructors** table contains information about instructors of the courses: their id (**iid**), name (**iname**) and department they belong to (**dept**). An instruct can teach many different courses.

```
Instructors(iid: string, iname: string, dept: string)
```

– Key: **iid**

- The **Courses** table contains information about courses: their id (**cid**), their name (**cname**), the department that offers it (**dept**), the id of its instructor (**iid**), and the maximum number of students who can take it (**maxenrol**). Every **iid** in this table is also found in the table **Instructors**.

```
Courses(cid: string, cname: string,  
        dept: string, iid: string,  
        maxenrol: integer  
)
```

– Key: **cid**

- The table **Enrolled** contains what students are registered to which courses, and the grade they receive (NULL if they have not received one yet). A student can only register once to any given course, but he/she can register to as many courses as necessary. Neither **sid** nor **cid** can be NULL. Every **sid** in this table is also found in the table **Students**, and every **cid** in this table is also found in the table **Courses**.

```
Enrolled(sid: integer, cid: string,  
        grade: integer)
```

– Key: (**sid**,**cid**)

3. Relational Algebra and SQL

For each of the following questions, provide a relational algebra expression to answer them, and its equivalent SQL query:

- (a) [2] What is the average **age** of the students who are taking at least one course? Result should have only one column (and one tuple). Hint. Make sure you average each student's age only once.

$D = \pi_{sid} E$ WITH D AS (SELECT sid FROM E)
 $\gamma_{avg(age)} \sigma_{sid \in D} S$ SELECT avg(age) FROM S WHERE sid IN (TABLE D)

- (b) [4] For every instructor that is teaching exactly two courses, list the **iid** of the instructor, their name **iname** and the course **cid** they are teaching. There are going to be two tuples for each instructor, one for each course they teach. For instance, your result should look something like this (three columns).

iid	iname	cid
342	M. Zastre	Seng 365
342	M. Zastre	CSC 360
123	D. German	CSC 370
456	D. German	CSC 225

$T = \sigma_{c=2} \gamma_{iid}^{count(x) \rightarrow c} C$

$\pi_{iid, iname, cid} (T \bowtie C \bowtie I)$

WITH T AS (
 SELECT iid, count(x) AS C
 FROM C
 GROUP BY iid
 HAVING count(x) = 2)
 SELECT iid, iname, cid FROM
 T NATURAL JOIN C
 NATURAL JOIN I;

- (c) [4] List the **sid** and **sname** of the students who are enrolled in the fewest courses. Your result should include three columns: **sid**, **sname** and total number of courses. Make sure you consider students who might not be taking any course (in that case they are enrolled to zero courses).

count of courses per student \rightarrow $D = \gamma_{sid}^{count(cid) \rightarrow c} (S \bowtie E)$
 minimum count. \rightarrow $M = \gamma_{min(c) \rightarrow c} D$
 $\pi_{sid, sname} (D \bowtie M \bowtie S)$

WITH D AS (
 SELECT sid, count(cid) AS C
 FROM S NATURAL JOIN E
 GROUP BY sid)
 M AS (
 SELECT min(c) AS C FROM D)
 SELECT sid, sname FROM
 D NATURAL JOIN M
 NATURAL JOIN S;

End of examination
 Total pages: 5
 Total marks: 30