Homework Assignment

1

Due Date: Tuesday, Sept 26, 2017

CS425 - Database Organization

Results

Please leave this empty!

1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
1.9
1.10
1.11
1.12

Sum
Instructions

- Try to answer all the questions using what you have learned in class
- When writing a query, write the query in a way that it would work over all possible database instances and not just for the given example instance!

Consider the following database schema and example instance:

<table>
<thead>
<tr>
<th>Student</th>
<th>Course</th>
</tr>
</thead>
<tbody>
<tr>
<td>sid name dept</td>
<td>cid title dept credits</td>
</tr>
<tr>
<td>001 Alice CS</td>
<td>CS425 Databases CS 3</td>
</tr>
<tr>
<td>002 Bob EE</td>
<td>CS595 Database Security CS 3</td>
</tr>
<tr>
<td>003 Carol CS</td>
<td>EE591 Microcomputers EE 4</td>
</tr>
<tr>
<td>004 David PHYS</td>
<td>EE401 VLSI Design EE 3</td>
</tr>
<tr>
<td>004</td>
<td>PHYS571 Radiation Physics PHYS 3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Enroll</th>
<th>Prereq</th>
</tr>
</thead>
<tbody>
<tr>
<td>cid sid grade gradepoint</td>
<td>cid pid</td>
</tr>
<tr>
<td>CS425 001 A 4.0</td>
<td>CS595 CS425</td>
</tr>
<tr>
<td>CS595 001 B 3.0</td>
<td>EE591 EE401</td>
</tr>
<tr>
<td>CS595 002 A 4.0</td>
<td>... ...</td>
</tr>
<tr>
<td>EE401 001 A 4.0</td>
<td></td>
</tr>
<tr>
<td>EE401 002 B 3.0</td>
<td></td>
</tr>
<tr>
<td>EE401 004 A 4.0</td>
<td></td>
</tr>
<tr>
<td>PHYS571 002 C 2.0</td>
<td></td>
</tr>
<tr>
<td>PHYS571 004 A 4.0</td>
<td></td>
</tr>
</tbody>
</table>

Hints:

- Attributes shown with grey background form the primary key of a relation.
- The attribute cid and sid of relation Enroll is a foreign key to relations Course and Student, respectively. All the attributes cid and pid (except for the one in Course) are a foreign key to relation Course.
- Attribute gradepoint is converted from the letter grade (4.0 scale).
Part 1.1 Relational Algebra (Total: 100 + 10 bonus Points)

Question 1.1.1 (6 Points)
Find the names of all the students enrolled in course ‘EE401’.

Solution

\[ \pi_{name}(\sigma_{cid='EE401'}(Enroll \bowtie Student)) \]

Question 1.1.2 (6 Points)
Write an relational algebra expression that for each student the title of the courses he/she has received an ‘A’ grade in. Return this information as tuples \((name, title)\) where title and name represent the course title and student name, respectively.

Solution

\[ \pi_{name,title}(\sigma_{grade='A'}(\pi_{cid,title}(Course) \bowtie Enroll \bowtie Student)) \]
Question 1.1.3  (8 Points)

Find the students (sid and name) that have taken at least one of the prerequisite(s) of course ‘CS595’ and got an ‘A’ grade in this prerequisite course.

Solution

$$\pi_{\text{sid, name}}(\sigma_{\text{Enroll.cid='CS595'} \land \text{grade='A'} \land \text{Enroll.\bigtriangleup Enroll.cid=Prereq.pid Prereq.\bigtriangleup Student}})$$

Question 1.1.4  (8 Points)

Find all the ‘EE’ students (sid and name) that have taken all the courses offered by the ‘CS’ department.

Solution

$$E_1 \leftarrow \pi_{\text{sid, name, cid}}(\sigma_{\text{dept='EE'} \land \text{Enroll.\bigtriangleup Student}})$$
$$E_2 \leftarrow \pi_{\text{cid}}(\sigma_{\text{dept='CS'} \land \text{Course}})$$
$$E \leftarrow E_1 \div E_2$$
Question 1.1.5  (8 Points)

Find the IDs of all the students (sid), whose grade in ‘EE401’ is lower than their grade in ‘CS595’.

Solution

\[ E_1 \leftarrow \sigma_{\text{cid} = 'EE401'}(\text{Enroll}) \]
\[ E_2 \leftarrow \sigma_{\text{cid} = 'CS595'}(\text{Enroll}) \]
\[ E \leftarrow \pi_{E_1.\text{sid}}(\sigma(E_1.\text{sid} = E_2.\text{sid}) \land (E_1.\text{gradepoint} < E_2.\text{gradepoint})(E_1 \times E_2)) \]

Question 1.1.6  (8 Points)

List all the students (sid and name) who never got a grade lower than ‘B’ (grade point below 3.0).

Solution

\[ S' \leftarrow \pi_{\text{sid}}(\text{Student}) - \pi_{\text{sid}}(\sigma_{\text{gradepoint} < 3.0}(\text{Enroll})) \]
\[ S \leftarrow \pi_{\text{sid}, \text{name}}(S' \bowtie \text{Student}) \]
Question 1.1.7  (8 Points)

List the titles of all the courses that student ‘Alice’ has not taken.

Solution

\[ \pi_{\text{title}}(Course) - \pi_{\text{title}}(\sigma_{\text{name} = 'Alice'}(Student \bowtie Enroll \bowtie \pi_{\text{cid}, \text{title}}(Course))) \]

Question 1.1.8  (10 Points)

List all the students and their GPA (result schema: sid and GPA). The GPA is calculated by summing up the grade of each course multiplied the number of credits for the course and then dividing the result by the total number of credits the student has taken.

Solution

\[
\begin{align*}
gp & \leftarrow \pi_{\text{Enroll}.\text{cid}, \text{sid}.(\text{gradepoint} \times \text{credits})} \rightarrow gp(\text{Enroll} \bowtie \text{Course}) \\
\text{totalgp} & \leftarrow \text{sid} \ G_{\text{sum}(gp)} \rightarrow \text{gp}(\text{Enroll} \bowtie \text{Course}) \\
\text{totalcre} & \leftarrow \text{sid} \ G_{\text{sum}(\text{credits})} \rightarrow \text{tcr}(\text{Enroll} \bowtie \text{Course}) \\
\text{result} & \leftarrow \pi_{\text{totalgp}, \text{sid}.(\text{tgp} / \text{tcr})} \rightarrow \text{GPA} (\text{totalgp} \bowtie \text{totalcre})
\end{align*}
\]

or using one aggregation

\[
\begin{align*}
gp & \leftarrow \pi_{\text{Enroll}.\text{cid}, \text{sid}.(\text{gradepoint} \times \text{credits})} \rightarrow gp.\text{credits}(\text{Enroll} \bowtie \text{Course}) \\
\text{totalgpcre} & \leftarrow \text{sid} \ G_{\text{sum}(gp) \rightarrow \text{tgpp}, \text{sum}(\text{credits}) \rightarrow \text{tcr}(\text{gp})} \\
\text{result} & \leftarrow \pi_{\text{totalgp}, \text{sid}.(\text{tgpp} / \text{tcr})} \rightarrow \text{GPA} (\text{totalgpcre})
\end{align*}
\]
Question 1.1.9  (8 Points)

List all the courses (their id) and for each course return the number of courses it has as a prerequisite.

Solution

\[
\begin{align*}
\text{courseprereq} & \leftarrow Enroll \bowtie_{Enroll.cid=Prereq.cid} Prereq \\
noprereq & \leftarrow \pi_{cid}(\text{Course}) - \pi_{cid}(\text{Prereq}) \\
\text{numpre} & \leftarrow Enroll.cid \sigma_{\text{avg}(\text{gradepoint}) < 3.0}(\text{courseprereq}) \cup \pi_{cid,0}(\text{noprereq})
\end{align*}
\]

Question 1.1.10  (8 Points)

Return the number of courses for which the average grade of all students enrolled in the course is lower than ‘B’ (grade point below 3.0).

Solution

\[
\begin{align*}
\text{avgG} & \leftarrow cid \sigma_{\text{avg}(\text{gradepoint})}(\text{Enroll}) \\
\text{num} & \leftarrow \sigma_{\text{avg}(\text{gradepoint}) < 3.0}(\text{avgG})
\end{align*}
\]
Question 1.1.11  (10 Points)

For every course, return the names of the highest-scoring students (result schema: course title and student name).

Solution

\[
\text{nonhigh} \leftarrow \pi_{Enroll.cid, Enroll.sid}((Enroll.gradepoint < x.gradepoint) \land Enroll.course = x.course (Enroll \times \rho_x(Enroll)))
\]

\[
cidsid \leftarrow \pi_{Enroll cid, Enroll sid} (\pi_{cid, sid}(Enroll) - \text{nonhigh})
\]

\[
result \leftarrow \pi_{title, name} (\pi_{cid, title}(Course) \bowtie cidsid \bowtie Student)
\]

Alternatively, you can use aggregation.

\[
h s \leftarrow \text{cid} \ G_{\text{max}(\text{gradepoint})} \rightarrow \text{maxgp} (Enroll)
\]

\[
result \leftarrow \pi_{hs . cid, name, title} (hs \bowtie \text{hs . cid} = Enroll . cid \land \text{maxgp} = \text{gradepoint} Enroll \bowtie Student \bowtie \pi_{cid, title}(Course))
\]

Question 1.1.12  (12 Points)

List all the students (sid and name) that are enrolled in courses for which they have taken the prerequisites.

Solution

\[
cidpre \leftarrow \pi_{Enroll.cid, Enroll.sid, Prereq . pid} (Enroll \bowtie Enroll . cid = Prereq . cid Prereq)
\]

\[
nopre \leftarrow \pi_{sid} ((\pi_{cid}(Course) - \pi_{cid}(Prereq)) \bowtie Enroll)
\]

\[
pretaken \leftarrow \pi_{cidpre . sid} ((\pi_{cidpre . pid} = x . cid) \land (cidpre . sid = x . sid)) (cidpre \times \rho_x(Enroll))
\]

\[
result \leftarrow \pi_{sid, name} ((pretaken \cup \text{nopre}) \bowtie Student)
\]
Question 1.1.13 (BONUS QUESTION) (10 Points)

Write a query that returns the names of all students that are ready to graduate. A student is ready to graduate if the student fulfills all the following requirements:

1. The student has a GPA of 3.0 or higher
2. The student has taken at least 30 credits of coursework
3. The student is not missing any prerequisite of the courses he/she has taken. That is, for every course the student has taken that has a prerequisite, the student has taken one or more of these prerequisites

Solution

\[
\begin{align*}
gp & \leftarrow \pi_{\text{Enroll.cid}, \text{sid}, (\text{gradepoint} \times \text{credits})} \rightarrow gp(\text{Enroll} \bowtie \text{Course}) \\
totalgp & \leftarrow \text{sid} \sum_{\text{gp}} \rightarrow \text{tgp}(gp) \\
totalcre & \leftarrow \text{sid} \sum_{\text{credits}} \rightarrow \text{tcr}(\text{Enroll} \bowtie \text{Course}) \\
highGPA & \leftarrow \pi_{\text{sid} \mid \sigma_{\text{GPA} > 3.0}}(\pi_{\text{sid}, (\text{tgp/totalcre})} \rightarrow \text{GPA}(\text{totalgp} \bowtie \text{totalcre})) \\
enoughCred & \leftarrow \pi_{\text{sid} \mid \sigma_{\text{tcr} \geq 30}}(\text{totalcre}) \\
takenPreq & \leftarrow \pi_{\text{sid, cid} \mid \text{Enroll} \bowtie \text{Prereq} \bowtie \text{pid} = \text{cid}' \land \text{sid} = \text{sid}' \land \text{cid} \rightarrow \text{cid}' \land \text{sid} \rightarrow \text{sid}'}(\text{Enroll}) \\
hasPreq & \leftarrow \pi_{\text{sid, cid} \mid \text{Enroll} \bowtie \text{Prereq}} \\
missingPreq & \leftarrow \pi_{\text{sid} \mid \text{hasPreq} \setminus \text{takenPreq}} \\
allPreq & \leftarrow \pi_{\text{sid} \mid \text{student}} \setminus \text{missingPreq} \\
allReqs & \leftarrow \text{highGPA} \cap \text{enoughCred} \cap \text{allPreq} \\
result & \leftarrow \pi_{\text{name}}(\text{students} \bowtie \text{allReqs})
\end{align*}
\]