Name

CWID

Homework Assignment 1

Due Date: Monday, Sep 26, 2022

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:

Student

sid	name	dept
001	Alice	CS
002	Bob	EE
003	Carol	CS
004	David	PHYS

Course

cid	title	dept	credits
CS425	Databases	CS	3
CS595	Database Security	CS	3
EE591	Microcomputers	EE	4
EE401	VLSI Design	EE	3
PHYS571	Radiation Physics	PHYS	3

Enroll

cid	sid	grade	gradepoint
CS425	001	A	4.0
CS595	001	B	3.0
CS595	002	A	4.0
EE401	001	A	4.0
EE401	002	B	3.0
EE401	004	A	4.0
PHYS571	002	C	2.0
PHYS571	004	A	4.0

Prereq

cid	pid
CS595	CS425
EE591	EE401
...	...

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)

Write a relational algebra expression that returns the titles of all the courses with more than 3 credits from the `EE` department.

Solution

$$\pi_{title}(\sigma_{credits>3 \wedge dept=EE}(Course))$$

Question 1.1.2 (6 Points)

Write a relational algebra expression that returns for each student the title of courses they have taken from their major, i.e., where the course's `dept` is equal to the student's `dept`.

Solution

$$\pi_{name,title}(Course \bowtie Enroll \bowtie Student)$$

Question 1.1.3 (8 Points)

Write a relational algebra expression that returns the `title` and `dept` of courses that do not have any prerequisites.

Solution

$$
\begin{aligned}
Q_{haspre} &= \pi_{cid}(Prereq) \\
Q_{nopre} &= \pi_{cid}(Course) - Q_{haspre} \\
Q &= \pi_{title,dept}(Course \bowtie Q_{nopre})
\end{aligned}
$$

Question 1.1.4 (8 Points)

Write a relational algebra expression that returns the names of students that have taken all courses in their department.

Solution

$$
\begin{aligned}
Q_{posscomb} &= \pi_{sid,cid}(Student \bowtie Course) \\
Q_{actual} &= \pi_{sid,cid}(Student \bowtie Enroll \bowtie Course) \\
Q_{all} &= \pi_{sid}(Student) - \pi_{sid}(Q_{posscomb} - Q_{actual}) \\
Q &= \pi_{name}(Student \bowtie Q_{all})
\end{aligned}
$$

Question 1.1.5 (8 Points)

Write a relational algebra expression that returns the names of students that have gotten a better grade in a course than in at least one of the prerequisites of the course.

Solution

$$Q_{better} = \pi_{sid}(\sigma_{grade>pgrade}(Enroll \bowtie Prereq \bowtie \rho_{pid,psid,pgrade}(\pi_{cid,sid,grade}(Enroll))))$$
$$Q = \pi_{name}(Student \bowtie Q_{better})$$

Question 1.1.6 (8 Points)

Write a relational algebra expression that returns the sid and name of students that are ready to graduate. A student is ready to graduate if they have taken at least 30 credits in total and at least 20 credits in their major.

Solution

$$Q_{majorcredit} = \pi_{sid}(\sigma_{totcred \geq 20}({}_{sid}\mathcal{G}_{sum(credit) \rightarrow totcred}(Student \bowtie Enroll \bowtie Course)))$$
$$Q_{totalcredit} = \pi_{sid}(\sigma_{totcred \geq 30}({}_{sid}\mathcal{G}_{sum(credit) \rightarrow totcred}(Student \bowtie Entroll \bowtie \pi_{cid,credits}(Course))))$$
$$Q = \pi_{sid,name}(Student) \bowtie Q_{majorcredit} \bowtie Q_{totalcredit}$$

Question 1.1.7 (8 Points)

Write a relational algebra expression that returns the titles of courses that all CS students have taken.

Solution

$$
\begin{aligned}
Q_{allcomb} &= \pi_{sid}(Student) \times \pi_{cid}(Course) \\
Q_{actual} &= \pi_{sid,cid}(Student \bowtie Enroll) \\
Q_{missing} &= \pi_{cid}(Q_{allcomb} - Q_{actual}) \\
Q_{all} &= \pi_{cid}(Course) - Q_{missing} \\
Q &= \pi_{title}(Course \bowtie Q_{all})
\end{aligned}
$$

Question 1.1.8 (10 Points)

Write a relational algebra expression that returns for each major, the name and GPA of the student with the highest GPA among all students of that major. 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

$$
\begin{aligned}
Q_{gp} &\leftarrow \pi_{Enroll.cid,sid,(gradepoint*credits) \rightarrow gp}(Enroll \bowtie Course) \\
Q_{totalgp} &\leftarrow_{sid} \mathcal{G}_{sum(gp) \rightarrow tgp}(Q_{gp}) \\
Q_{totalcre} &\leftarrow_{sid} \mathcal{G}_{sum(credits) \rightarrow tcr}(Enroll \bowtie Course) \\
Q_{GPA} &\leftarrow \pi_{sid,name,dept,(tgp/tcr) \rightarrow GPA}(totalgp \bowtie totalcre \bowtie Student) \\
Q_{maxgpa} &\leftarrow_{dept} \mathcal{G}_{max(GPA) \rightarrow GPA}(Q_{GPA}) \\
Q &\leftarrow \pi_{dept,name,GPA}(Q_{GPA} \bowtie Q_{maxgpa})
\end{aligned}
$$

Question 1.1.9 (8 Points)

Write a relational algebra expression that returns the titles of courses that have more than one prerequisite.

Solution

$$\pi_{title}(\sigma_{npre>1}({}_{cid,title}\mathcal{G}_{count(1)\rightarrow npre}(Course \bowtie Prereq)))$$

Question 1.1.10 (8 Points)

Write a relational algebra expression that returns for each student the number of failed credits. The failed credits for a student is the number of credits of courses they have taken and have scored a `gradepoint` of less than 2.0.

Solution

$${}_{sid}\mathcal{G}_{sum(credits)\rightarrow failedcred}(\sigma_{gradepoint<2.0}(Enroll) \bowtie Course)$$

Question 1.1.11 (10 Points)

Write a relational algebra expression that returns courses that have more than 5 direct or indirect prerequisites. An indirect prerequisite is defined recursively as follows:

- Every direct prerequisite of a course is also an indirect prerequisite.
- If p is an indirect prerequisite of a course c, then any direct prerequisite of p is also an indirect prerequisite of c.

You only have to consider prerequisites with up to two levels of indirection, i.e., the prerequisite of a prerequisite of a prerequisite.

Solution

$$
\begin{aligned}
Q_{level1} &\leftarrow \pi_{cid,pid}(\rho_{cid,icid}(Prereq) \bowtie \rho_{icid,pid}(Prereq)) \\
Q_{level2} &\leftarrow \pi_{cid,pid}(\rho_{cid,icid}(Q_{level1}) \bowtie \rho_{icid,pid}(Prereq)) \\
Q_{iprereq} &\leftarrow Prereq \cup Q_{level1} \cup Q_{level2} \\
Q &\leftarrow \pi_{cid}(\sigma_{pcount>5}({}_{cid}\mathcal{G}_{count(1)\rightarrow pcount}(Q_{iprereq})))
\end{aligned}
$$

Question 1.1.12 (12 Points)

List all the students (sid and name) that are missing one or more prerequisites for some course they have taken.

Solution

$$
\begin{aligned}
Q_{studentpre} &\leftarrow \pi_{sid,pid}(Enroll \bowtie Prereq) \\
Q_{takenpre} &\leftarrow \pi_{sid,pid}(Q_{studentpre} \bowtie_{pid=cid \wedge sid=sid} Enroll) \\
Q_{missedpre} &\leftarrow \pi_{sid}(Q_{studentpre} - Q_{takenpre}) \\
Q &\leftarrow \pi_{sid,name}(Q_{missedpre} \bowtie Student)
\end{aligned}
$$

Question 1.1.13 (BONUS QUESTION) (10 Points)

Write a relational algebra expression that returns for the two students with sid `001` and `002` the `cid` of courses that only one of them has taken.

Solution

$$
\begin{aligned}
Q_{001} &\leftarrow \pi_{cid}(\sigma_{sid=001}(Enroll)) \\
Q_{002} &\leftarrow \pi_{cid}(\sigma_{sid=002}(Enroll)) \\
Q &\leftarrow (Q_{001} - Q_{002}) \cup (Q_{002} - Q_{001})
\end{aligned}
$$