Name

UIN

Homework Assignment 1

Due Date: Thursday, Feb 08, 2024

CS480 - Database Systems 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 names of students majoring in `CS` or `PHYS`.

Solution

$$\pi_{name}(\sigma_{dept=CS \vee dept=PHYS}(Student))$$

Question 1.1.2 (6 Points)

Write a relational algebra expression that returns the title and credits of courses that have CS425 as a prerequisite. Note that the `Prereq` table stores pairs of courses (`cid`) and prerequisites (`pid`).

Solution

$$\pi_{title,credits}(Course \bowtie \sigma_{pid=CS425}(Prereq))$$

Question 1.1.3 (8 Points)

Write a relational algebra expression that returns the title of the course(s) with the largest number of credits.

Solution

$$Q_{maxc} \leftarrow \rho_{maxcred}(\gamma_{max(credits);}(Course))$$
$$Q \leftarrow \pi_{title}(Course \bowtie_{credits=maxcred} Q_{maxc})$$

Question 1.1.4 (8 Points)

Write a relational algebra expression that returns the name of student(s) with the largest number of credits (total credits that students has registered for) among all students.

Solution

$$Q_{sumc} \leftarrow \rho_{totalc,sid}(\gamma_{sum(credits);sid}(Student \bowtie Enroll \bowtie \rho_{d \leftarrow dept}(Course)))$$
$$Q_{maxc} \leftarrow \rho_{maxtot}(\gamma_{max(totalc);}(Q_{sumc}))$$
$$Q \leftarrow \pi_{name}(Q_{maxc} \bowtie_{maxtot=totalc} Q_{sumc} \bowtie Student)$$

Question 1.1.5 (8 Points)

Write a relational algebra expression that returns the sids of students under probation. A student is under probation if their GPA is less than 3.0. The GPA is calculated by multiplying the grade points the student has received in the course with the number of credits for the course, then summing up these numbers and dividing by the total number of credits the student has taken.

Solution

$$
\begin{aligned}
Q_{gps} &\leftarrow \rho_{gps,credits,sid}(\pi_{credits*gradepoint,credits,sid}(Enroll \bowtie Student \bowtie \rho_{d \leftarrow dept}(Course))) \\
Q_{gpa} &\leftarrow \rho_{gpa,sid}(\pi_{totalgps/sumcred,sid}(\rho_{totalgps,sumcred,sid}(\gamma_{sum(gps),sum(credits);sid}(Q_{gps})))) \\
Q &\leftarrow \pi_{sid}(\sigma_{gpa<3.0}(Q_{gpa}))
\end{aligned}
$$

Question 1.1.6 (8 Points)

Write a relational algebra expression that returns students (`sid`) that are missing a prerequisite, i.e., that have taken a course, but have not taken all the prerequisites for the course.

Solution

$$
\begin{aligned}
Q_{shouldtake} &\leftarrow \pi_{sid,pid}(\pi_{sid,cid}(Enroll) \bowtie Prereq) \\
Q_{missing} &\leftarrow \pi_{sid}(Q_{shouldtake} - \pi_{sid,cid}(Enroll))
\end{aligned}
$$

Question 1.1.7 (8 Points)

Write a relational algebra expression that returns for each course the number of students that have taken the course. Note: also return courses not taken by any students with a 0 count. For each course, return the `cid`, `title`, `department`, and `number of students` that have taken the course.

Solution

$$
\begin{aligned}
Q_{nottaken} &\leftarrow \pi_{title,dept,0}(Course \bowtie (\pi_{cid}(Course) - \pi_{cid}(Enroll))) \\
Q_{numtaken} &\leftarrow \rho_{title,dept,numtaken}(\gamma_{count(sid);cid,title,dept}(Course \bowtie Enroll)) \\
Q &\leftarrow Q_{numtaken} \cup Q_{nottaken}
\end{aligned}
$$

Question 1.1.8 (10 Points)

Write a relational algebra expression that returns for each major the student (their sid, name, and GPA) with the highest GPA and the student with the lowest GPA for that major. GPAs are calculated as in the question above. You may reuse queries you defined in the previous question for calculating GPAs.

Solution

$$
\begin{aligned}
Q_{gps} &\leftarrow \rho_{gps,credits,sid,name,dept}(\pi_{credits*gradepoint,credits,sid,name,dept}(Enroll \bowtie Student \bowtie \rho_{d \leftarrow dept}(Course))) \\
Q_{gpa} &\leftarrow \rho_{gpa,sid,name,dept}(\pi_{totalgps/sumcred,sid}(\rho_{totalgps,sumcred,sid,name,dept}(\gamma_{sum(gps),sum(credits);sid,name,dept}(Q_{gps})))) \\
Q_{lowestgpa} &\leftarrow \pi_{gpa,sid,name,dept}(\rho_{gpa,dept}(\gamma_{min(gpa);dept}(Q_{gpa})) \bowtie Student) \\
Q_{higehstgpa} &\leftarrow \pi_{gpa,sid,name,dept}(\rho_{gpa,dept}(\gamma_{max(gpa);dept}(Q_{gpa})) \bowtie Student) \\
Q &\leftarrow Q_{lowestgpa} \cup Q_{highestgpa}
\end{aligned}
$$

Question 1.1.9 (10 Points)

Write a relational algebra expression that calculates the credit distribution of students as follows. For each student calculate the number of credits the student has taken. Then for each number of credits return the number of students that have this many credits.

Solution

$$Q_{totcred} \leftarrow \rho_{totc,sid}(\gamma_{sum(credits);sid}(Enroll \bowtie Course))$$
$$Q \leftarrow \rho_{totc,numst}(\gamma_{count(sid);totc}(Q_{totcred}))$$

Question 1.1.10 (10 Points)

Write a relational algebra expression that returns for each department the total number of credits taken by students from this department. Also return the number of credits students from the department have taken within their own department.

Solution

$$Q_{totcred} \leftarrow \rho_{totc,dept}(\gamma_{sum(credits);dept}(Enroll \bowtie Course))$$
$$Q_{totdcred} \leftarrow \rho_{totcin,dept}(\gamma_{sum(credits);dept}(Enroll \bowtie Course \bowtie Student))$$
$$Q \leftarrow Q_{totcred} \bowtie Q_{totdcred}$$

Question 1.1.11 (8 Points)

Write a relational algebra expression that returns for each student the department the student has taken courses with.

Solution

$$Q \leftarrow \pi_{sid,dept}(Enroll \bowtie Course)$$

Question 1.1.12 (10 Points)

Write a relational algebra expression that checks whether there are circular prerequisite dependencies. You have to only consider cycles of length up to 3 (e.g., CS100 $\rightarrow$ CS201 $\rightarrow$ CS100 would be a cycle of length 3). If there are circular dependencies, then the query should return $\{(1)\}$ and the empty set otherwise.

Solution

$$Q_{2hop} \leftarrow \pi_{cid,sid}(\rho_{cid,fpid}(Prereq) \bowtie \rho_{fpid,spid}(Prereq))$$
$$Q_{3hop} \leftarrow \pi_{cid,sid}(\rho_{cid,fpid}(Q_{2hop}) \bowtie \rho_{fpid,spid}(Prereq))$$
$$Q_{cycle} \leftarrow \pi_{1}(\sigma_{cid=sid}(Q_{2hop} \cup Q_{3hop}))$$

Question 1.1.13 (BONUS QUESTION) (10 Points)

Write a relational algebra expression that returns pairs of courses such that the first course is a direct or indirect prerequisite of the second. Indirect prerequisites are chains in the prereq table of a length less than or equal to 64. For instance, if CS480 is a prerequisite of CS580 and CS580 is a prerequisite of CS680, then CS480 is an indirect prerequisite of CS680 (chain of length 2).

Solution

$$
\begin{aligned}
Q_{2hop} &\leftarrow \pi_{cid,sid}(\rho_{cid,fpid}(Prereq) \bowtie \rho_{fpid,spid}(Prereq)) \\
Q_{4hop} &\leftarrow \pi_{cid,sid}(\rho_{cid,fpid}(Q_{2hop}) \bowtie \rho_{fpid,spid}(Q_{2hop})) \\
Q_{8hop} &\leftarrow \pi_{cid,sid}(\rho_{cid,fpid}(Q_{4hop}) \bowtie \rho_{fpid,spid}(Q_{4hop})) \\
Q_{16hop} &\leftarrow \pi_{cid,sid}(\rho_{cid,fpid}(Q_{8hop}) \bowtie \rho_{fpid,spid}(Q_{8hop})) \\
Q_{32hop} &\leftarrow \pi_{cid,sid}(\rho_{cid,fpid}(Q_{16hop}) \bowtie \rho_{fpid,spid}(Q_{16hop})) \\
Q_{up2hop} &\leftarrow Q_{2hop} \cup Prereq \\
Q_{up4hop} &\leftarrow Q_{up2hop} \cup \pi_{cid,sid}(\rho_{cid,fpid}(Q_{up2hop}) \bowtie \rho_{fpid,spid}(Q_{2hop})) \\
Q_{up8hop} &\leftarrow Q_{up4hop} \cup \pi_{cid,sid}(\rho_{cid,fpid}(Q_{up4hop}) \bowtie \rho_{fpid,spid}(Q_{4hop})) \\
Q_{up16hop} &\leftarrow Q_{up8hop} \cup \pi_{cid,sid}(\rho_{cid,fpid}(Q_{up8hop}) \bowtie \rho_{fpid,spid}(Q_{8hop})) \\
Q_{up32hop} &\leftarrow Q_{up16hop} \cup \pi_{cid,sid}(\rho_{cid,fpid}(Q_{up16hop}) \bowtie \rho_{fpid,spid}(Q_{16hop})) \\
Q_{up64hop} &\leftarrow Q_{up32hop} \cup \pi_{cid,sid}(\rho_{cid,fpid}(Q_{up32hop}) \bowtie \rho_{fpid,spid}(Q_{32hop}))
\end{aligned}
$$