Seonhong Hwang

CS430/630 – Homework 1

Released Feb 02, Due Feb 15 50 points (5/100 of final grade)

Instructions: The homework is due BEFORE CLASS on Thu Feb 15. Please hand in paper copies (either typeset or hand-written copies are fine, as long as the hand writing is clear).

Question 1 (25 points)

Consider a database schema with three relations:

```
Students (sid:integer, sname:string, age:integer
Enrolled (sid:integer, cid:integer, grade:integer)
Courses(cid:integer, cname:string, credits:integer)
```

The keys are underlined in each relation. Students are identified uniquely by `sid`, and courses by `cid`. Students enroll to take courses, and for each course they obtain a `grade` which is an integer. `sname` is the student name (string), `age` represents the student age and is an integer. `cname` is the course name (string), and `credits` is the number of credits for a particular course (integer).

Write **relational algebra** expressions for the following queries:

. (a) Find the names of students who got grade 10 in some course.

$\pi_{sname}((\sigma_{grade=10}\text{Enrolled})\bowtie\text{Students})$

. (b) Find the ages of students who take some course with 3 credits.

$\pi_{age}((\sigma_{credits=3}\text{Courses})\bowtie\text{Students})$

. (c) Find the names of students who take a course named ‘Calculus’.

$\pi_{sname}(\pi_{sid}(\pi_{cid}(\sigma_{cname='Calculus'}Courses)\bowtie Enrolled)\bowtie Students)$

. (d) Find the names of students who obtained grade at least 8 in some course that has less than 4 credits.

$p(TempGrade, \pi_{sid}((\sigma_{grade>=8}Enrolled)\bowtie Enrolled))$

$p(TempCredits, \pi_{sid}(\pi_{cid}(\sigma_{credits<4}Courses)\bowtie Enrolled))$

$\pi_{sname}((Tempgrade\cap Tempcredits)\bowtie Students)$

. (e) Find the names of students who obtained only grades of 10 (implies that they took at least one course).

$Students - \pi_{sname}(Students - \pi_{grades<>10}Enrolled)$

. (f) Find the names of students who took a course with three credits or who obtained grade 10 in some course.

$p(TempCredits, \pi_{sname}(\sigma_{credits=3}Courses)\bowtie Students)$

$p(TempGrade, Students - \pi_{sname}(Students-\pi_{grades<>10}Enrolled))$

$\pi_{sname}(TempCredits\vee TempGrade)\bowtie Students)$

. (g) Find the ages of students who attend ‘Calculus’ but never took any 4 credit course (assume there is a course ‘Calculus’ with 3 credits).

$p(\text{TempCalculus}, \pi_{sid}(\pi_{cid}(\sigma_{cname='Calculus'}\text{Courses})\bowtie\text{Enrolled})\bowtie\text{Students}))$

$p(\text{TempCredits}, \pi_{sid}(\pi_{cid}(\sigma_{credits<>4}\text{Courses})\bowtie\text{Enrolled})\bowtie\text{Students}))$

$\pi_{age}(\text{TempCalculus}\cap(\text{Students}-\text{TempCredits})\bowtie\text{Students})$

. (h) Find the names of students who have the lowest age.

p(S1, Students) p(S2, Students)

$p(\text{TempJoin}(1 \text{ -> } f1, 2 \text{ -> } f2, 3 \text{ -> } f3), S1\bowtie_{S1.age<S2.age}S2)$

$p(\text{LeftHalf}, \pi_{f1, f2, f3}\text{TempJoin})$

Question 2 (25 points)

Consider a database schema with three relations:

```
Movies (movie_id:integer, title:string, year:integer,
studio:string)
Actors (actor_id:integer, name:string, nationality:string)
StarsIn(actor_id:integer, movie_id:integer,
character:string)
```

The keys are underlined in each relation. Relation `Movies` stores information such as unique movie identifier, title, year and producing studio. `Actors` contains unique actor identifier, actor name and nationality. Relation `StarsIn` tracks which actor starred in which movie, and the name of the character interpreted in that movie. Assume that one actor plays at most one character in the same movie.

Write **relational algebra** expressions for the following queries:

. (a) Find the titles of movies produced by 'Universal' studio.

$\pi_{title}(\sigma_{studio='Universal'}Movies)$

. (b) Find the names of actors that played a character named 'Forrest Gump' in some movie.

$\pi_{name}(\pi_{actor\_id}((\sigma_{character='Forrest\ Gump'}StarsIn)\bowtie Actors))$

. (c) Find the names of actors of nationality 'German'

$\pi_{name}(\sigma_{nationality='German'}Actors)$

. (d) Find the nationality of actors who played a character named 'Forrest Gump' or who starred in a movie in year 1980

$p(\text{TempActors}, \pi_{\text{actor\_id}}((\sigma_{\text{character='Forrest Gump'}}\text{StarsIn}) \bowtie \text{Actors}))$

$p(\text{TempNationality}, \pi_{\text{actor\_id}}(\sigma_{\text{nationality='German'}}\text{Actors}))$

$p(\text{TempStarsIn}, \pi_{\text{actor\_id}}((\sigma_{\text{year='1980'}}\text{Movies}) \bowtie \text{StarsIn}))$

$\pi_{\text{name}}(((\text{TempActors} \cap \text{TempNationality}) \vee \text{TempStarsIn}) \bowtie \text{Actors})$

. (e) Find the names of actors that star in **exactly one** movie

?