BMEG3120: Midterm Exam

Please write all your solutions in the answer book.

Problem 1.(10%) Consider a table $T(A, B, C)$, namely, the table's name is T, and its attributes are A, B, and C. It has 2 candidate keys: $\{A, B\}$ and $\{B, C\}$. Can the following pairs of tuples co-exist in T, respectively?

(i) $(a1, b1, c1)$ and $(a1, b2, c2)$.
(ii) $(a1, b1, c1)$ and $(a1, b2, c1)$.
(iii) $(a1, b1, c1)$ and $(a2, b1, c1)$.

Answer. (i) yes (ii) yes (iii) no

Problem 2.(40%) Consider these tables:

- ACTOR(aid, name, country): *aid* is an actor's id, while the other attributes are self-explanatory. The candidate key is *aid*.
- MOVIE(mid, title, year): *mid* is a movie's id, *title* is the movie's title, and *year* is its production year. The candidate key is *mid*.
- DIRECTOR(did, dname, age): *did* is a director's id, *dname* is the director's name, while *age* is self-explanatory. The candidate key is *did*.
- PLAY(aid, mid, pay): Each tuple records that an actor played in a movie. Specifically, *aid* (*mid*) is the actor's (movie's) id, and *pay* gives how much money the actor made from the movie. The candidate key is (*aid*, *mid*).
- PRODUCE(did, mid): Each tuple records that a director produced a movie. Specifically, *did* (*mid*) is the director's (movie's) id. The candidate key is (*did*, *mid*).

Write relational algebra queries for the following tasks:
(i) Find the names of all actors from HK.
(ii) Find the titles of all movies directed by "James Cameron".
(iii) Find the highest amount of money an actor has ever made from a single movie.
(iv) If a director produced a move in which an actor played, we say that the director has *worked with* the actor. Find the aids of all the actors that "James Cameron" has ever worked with.
(v) Find the dids of the directors that have worked with all the actors.

Answer.
(i) $\Pi_{\text{name}}(\sigma_{\text{country}=\text{"HK"}}(\text{ACTOR}))$

(ii) $\Pi_{\text{title}}(\sigma_{\text{dname}=\text{"James Cameron"}}(\text{MOVIE} \bowtie \text{PRODUCE} \bowtie \text{DIRECTOR}))$

(iii) $T_1 \leftarrow \text{PLAY}$
$T_2 \leftarrow \text{PLAY}$
$\Pi_{\text{pay}}(\text{PLAY}) - \Pi_{\text{pay}}(\sigma_{T_1.\text{pay}<T2.\text{pay}}(T_1 \times T_2))$

(iv) $\Pi_{\text{aid}}(\sigma_{\text{dname}=\text{"James Cameron"}}(\text{PLAY} \bowtie \text{PRODUCE} \bowtie \text{DIRECTOR}))$

(v) $T_1 \leftarrow \Pi_{\text{did, aid}}(\text{PLAY} \bowtie \text{PRODUCE})$
$T_1 \div \Pi_{\text{aid}}(\text{ACTOR})$

Problem 3.(40%) Write SQL queries for the following tasks based on the tables in Problem 2.

(i) Find the names of all directors at least 50 years old.

(ii) For each actor, display her/his aid, and the total amount of money s/he has made from all movies.

(iii) For each director, display her/his name, and the number of distinct actors s/he has worked with.

(iv) For each actor that has played in at least 5 movies, display her/his name and country.

(v) Find the country with the largest number of actors.

Answer.

(i) select dname from DIRECTOR where age >= 50

(ii) select aid, sum(pay) from PLAY group by aid

(iii) select dname, count(distinct aid)
from PLAY PL, PRODUCE PR, DIRECTOR D
where PL.mid = PR.mid and PR.did = D.did
group by did, dname

(iv) select name, country from ACTOR A, PLAY P
where A.mid = P.mid
group by aid, name, country
having count(*) >= 5

(v) select country from ACTOR
group by country
having count(*) $\geq$ all (select count(*) from ACTOR group by country)

Problem 4.(10%) Consider the following table whose name is T:

A	B	C
1	10	100
2	10	10
3	40	100
4	30	200
5	25	90

Give the results of the following SQL queries:

(i)
select sum(C) from T
group by B
having count(*) >= 2

(ii)
select A from T as R
where not exists (
  select * from T
  where $T.B >= R.B$ and $T.C >= R.C$)

Answer.

(i) 110

(ii) The query returns an empty table.

A note on this question. Unfortunately the query has two typos – the two occurrences of ">=" should have been ">". The originally intended query is actually much more interesting, and returns the following answer:

A
3
4

The instructor has decided to regard both answers as being correct.