Score: 96/100

Name: Divya Thota

Student ID: 01994101

Question 1 (30 points)

Consider a database schema with three relations:
Books(bid:integer, bname:string, author:string, pubyear:integer, pubcompany:string)
Students(sid:integer,sname:string, age:real, state:string)
Reads(sid:integer,bid:integer,year:integer)

(The primary keys are underlined in each relation. A book is uniquely identified by bid. A student is uniquely identified by sid. If a student reads a book, a record will be present in the Reads relation, with that sid and bid and the year the book was read.)

and the following relations instances: Using the relation instances from above, show the resulted relation for each of the following relational algebra expressions:

Question 1: (30/30)

a) $\sigma_{\text{sname='mary'}} Students$

5/5

sid	sname	age	state
20	mary	21	MA
40	mary	21	VT

b) $\pi_{\text{sname,age}}(\sigma_{\text{sname='mary'}} Students)$

5/5

sname	age
mary	21

c) $(\sigma_{\text{state='MA'}} Students) \bowtie Reads$

5/5

sid	sname	age	state	bid	year
20	mary	21	MA	101	2020
20	mary	21	MA	102	2021
30	joe	21	MA	103	2020

d) $(\sigma_{\text{state='MA'}} Students) \times (\sigma_{\text{year=2020}} Reads)$

5/5

(sid)	sname	age	state	(sid)	bid	year
20	mary	21	MA	20	101	2020
30	joe	21	MA	20	101	2020
60	linda	23	MA	20	101	2020
20	mary	21	MA	30	103	2020
30	joe	21	MA	30	103	2020
60	linda	23	MA	30	103	2020

e) $\rho(A(bname \rightarrow name), \sigma_{\text{(bname='ulysses')v(banme='other book'}} Books)$ 5/5

A

bid	name	author	pubyear	pubcompany
102	ulysses	joyce	1920	simon
103	Other book	joyce	1920	penguin

5/5

f) [CS630 only] $Students \bowtie Reads \bowtie Books$

sid	sname	age	state	bid	year	bname	author	pubyear	pubcompany
20	mary	21	MA	101	2020	Lord of the rings	tolkien	1954	allen
20	mary	21	MA	102	2021	ulysses	joyce	1920	simon
30	joe	21	MA	103	2020	Other book	joyce	1920	penguin

Note: for each resulted relation, you must show both the header with the column names, in the expected order, as well as the records.

Question 2 (40 points) (36/40)

Consider a database schema with three relations:
Actors (aid:integer, aname:string, age:real, city:string, state:string)
Playsin(aid:integer, mid:integer, character:string)
Movies(mid:integer, mname:string, year:integer, studio:string)

Primary keys are underlined in each relation. Attribute aid uniquely identifies an actor in Actors relation. An actor has a an id (aid), a name(attr. aname), an age (attr. age), and a city and state(attributes city and state). Attribute mid uniquely identifies a movie in relation Movies. A movie has an id(mid), a name(attr. mname), a year (attr. year) and a studio that

produced it (attr. studio). Relation Playsin contains information about actors who played in movies. Attribute character = the character name of the actor with aid when playing in mid.

Write relational algebra queries for the following queries:

a) **Find the names of actors who live in ‘MA’ state.** 4/4
$\pi_{\text{aname}}(\sigma_{\text{state=’MA’}}\text{Actors})$

b) **Find the names and the ages of actors who are either from ‘Boston’ or ‘New York’.** 4/4
$\pi_{\text{aname,age}}(\sigma_{\text{(city=’Boston’)v(city=’New York’)}}\text{Actors})$

c) **Find the names of actors who are younger than 25 and are from ‘Los Angeles’.** 4/4
$\pi_{\text{aname}}(\sigma_{\text{(age<25)\textasciicircum(city=’Los Angeles’)}}\text{Actors})$

d) **Find the names, cities and states of actors who played as character ‘Peter Pan’.** 4/4
$\pi_{\text{aname,city,state}}(\ (\sigma_{\text{character=’Peter Pan’}}\ \text{Playsin}) \bowtie\ \text{Actors})$

e) **Find the names of the actors who played in a movie named ‘Titanic’.** 4/4
$\pi_{\text{aname}}(\ ((\sigma_{\text{mname=’Titanic’}}\ \text{Movies}) \bowtie\ \text{Playsin}) \bowtie \text{Actors})$

f) **Find the name of actors who played in some movie in year 2010, and did not play in any movie in 2020.** 3/4, set diff needs to include aid
$\pi_{\text{aname}}(\ ((\sigma_{\text{(year=2010)}}\ \text{Movies}) \bowtie\ \text{Playsin}) \bowtie\ \text{Actors}) - \pi_{\text{aname}}(\ ((\sigma_{\text{(year=2020)}}\ \text{Movies}) \bowtie\ \text{Playsin}) \bowtie\ \text{Actors})$

g) **Find the names of the actors who played in movies only in 2000.** 3/4, set diff needs to include aid
$\pi_{\text{aname}}(\ ((\sigma_{\text{(year=2000)}}\ \text{Movies}) \bowtie\ \text{Playsin}) \bowtie\ \text{Actors}) - \pi_{\text{aname}}(\ (\sigma_{\text{(year!=2000)}}\ \text{Movies}) \bowtie\ \text{Playsin} \bowtie\ \text{Actors})$

h) **Find the names of actors who are older than 25 and played in some movie produced by ‘Universal’ studio.** 4/4
$\pi_{\text{aname}}(\ ((\sigma_{\text{(studio=’Universal’)}}\ \text{Movies}) \bowtie\ \text{Playsin}\ ) \bowtie\ (\sigma_{\text{(age>25)}}\text{Actors}))$

i) **[CS630 only] Find the names of the actors who played in a movie produced by ‘WB’ studio in 2018 and did not play in any movie produced by ‘Universal’ (in any year).** 3/4, set diff needs to include aid
$\pi_{\text{aname}}(\ ((\sigma_{\text{(studio=’WB’)\textasciicircum(year=2018)}}\ \text{Movies}) \bowtie\ \text{Playsin}) \bowtie\ \text{Actors}) - \pi_{\text{aname}}(\ ((\sigma_{\text{(studio=’Universal’)}}\ \text{Movies}) \bowtie\ \text{Playsin}) \bowtie\ \text{Actors})$

j) **[CS630 only] Find the names of actors who played in movies both in 2018 and 2020.**
$(\pi_{\text{aname}}(\ ((\sigma_{\text{(year=2018)}}\ \text{Movies}) \bowtie\ \text{Playsin}) \bowtie\ \text{Actors})) \cap (\pi_{\text{aname}}(\ ((\sigma_{\text{(year=2020)}}\ \text{Movies}) \bowtie\ \text{Playsin}) \bowtie\ \text{Actors}))$
3/4, INTERSECT need to be done on aid

Question 3 (30 points) (30/30)

Consider a database schema with three relations:

Books(<u>bid:integer</u>, bname:string, author:string, pubyear:integer,pubcompany:string)
Students(<u>sid:integer</u>,sname:string, age:real, state:string)
Reads(<u>sid:integer,bid:integer</u>,year:integer)

Primary keys are underlined in each relation. A book is uniquely identified by bid. A book has an id (bid), a name (bname), one author (attribute author), a publication year (pubyear), and a publishing company (pubcompany). A student is uniquely identified by sid. A student has an id (sid), a name (attr. sname), age (attr. age) and a state (attr. state). If a student reads a book, a record will be present in the Reads relation, with that sid and bid and the year the book was read.

Write the relational algebra expressions for the following queries:

a) Find the students who live in ‘MA’ and they are either younger than 25 or older than 35.

$\sigma_{(\text{state='MA'})\text{^}((\text{age}<25)\text{v}(\text{age}>35))}\text{Students})$ 5/5

b) Find the oldest books (hint: pubyear is min). 5/5

$\rho(\text{B1},Books)$

$\rho(\text{B2},Books)$

$\rho(\text{Temp}(1 \rightarrow \text{fbid}, 2 \rightarrow \text{fbname}, 3 \rightarrow \text{fauthor}, 4 \rightarrow \text{fpubyear}, 5 \rightarrow \text{fpubcompany}_{),}$

$\text{B1} \bowtie_{\text{B1.pubyear > B2.pubyear}} \text{B2})$

$\rho(\text{TempLeft}, \pi_{\text{fbid,fbname,fauthor,fpubyear,fpubcompany}}\ \text{Temp})$

Books - TempLeft

c) Find the names and authors of the newest books. 5/5

$\rho(\text{B1},Books)$

$\rho(\text{B2},Books)$

$\rho(\text{Temp}(1 \rightarrow \text{fbid}, 2 \rightarrow \text{fbname}, 3 \rightarrow \text{fauthor}, 4 \rightarrow \text{fpubyear}, 5 \rightarrow \text{fpubcompany}_{),}$

$\text{B1} \bowtie_{\text{B1.pubyear < B2.pubyear}} \text{B2})$

$\rho(\text{TempLeft}, \pi_{\text{fbid,fbname,fauthor,fpubyear,fpubcompany}}\ \text{Temp})$

$\pi_{\text{bname,author}}\ (\text{Books} - \text{TempLeft})$

d) Find the names of the students who read some book in the same year when that book was published. 5/5

$\pi_{\text{sname}}\ ((\text{Books} \bowtie_{\text{Books.pubyear=Reads.year}} \text{Reads}) \bowtie \text{Students})$

e) Find the names of the students who read all books.

$\rho(\text{TempIds}, \pi_{\text{sid,bid}}\text{Reads}/\pi_{\text{bid}}\text{Books})$ 5/5

$\pi_{\text{sname}}\ (\text{TempIds} \bowtie \text{Students})$

f) [CS630 only] Find the names and ages of the students who read all books published by ‘penguin’ publishing company in any year.

$\rho(\text{TempIds}, \pi_{\text{sid,bid}}\text{Reads}/\pi_{\text{bid}}(\sigma_{\text{pubcompay='penguin'}}\text{Books}))$ 5/5

$\pi_{\text{sname}}\ (\text{TempIds} \bowtie \text{Students})$