Instructions: Print this assignment using single-side pages. Fill in your name above, and write in the solutions in the space provided below each question. You are allowed to use the back of each page. If you used any scratch paper to show your work, append those to the end. **Note:** It is important you use this format for gradescope.

Submission: After you've filled in the answers, scan all pages into a PDF, and submit to canvas.

Problems

1. Design a *schema diagram* for a company like Netflix that minimizes redundancy and NULL values. You should recall that an attribute cannot hold a list of values. Indicate all primary keys (underlined) and foreign key constraints (directed edges). You need to store data on the following:

   - A Movie has a title, description, country of origin, and a release-year. Assume that there may be two movies with the same name (e.g., Godzilla), but they are never released in the same year. You may assign each movie a unique identifier, but you shouldn't need to.

   - An Actor has a name and a biography. Assume that there may be two actors with the same name. You may assign each actor a unique identifier.

   - A User has a userid and password. The company does not allow two different users to share the same userid.

   - You need to keep track of what movies the users have seen.

   - Users can rate movies from 1 to 5 stars. Note that they need not have seen the movie to rate it! A user can only rate a movie once (though that rating can change later).

   - Users can also comment on each actor and/or movie. A comment has a date, time, and the comment itself. Unlike ratings, users can leave a multiple comments on the same movie or actor.

   - Actors play certain roles in certain movies. Actors may also play multiple roles in the same movie (see: Eddie Murphy and Mike Myers).

Space for your schema diagram is given on the next page.

2. Assume we have two relations $R(A, B)$ and $S(B, C)$. Assume that R $= \{(1,2), (2,3), (3,4)\}$, and $S = \{(2,2), (2,3), (5,1)\}$. How many tuples are in the result of $\pi_A(R \bowtie S)$?

3. Let's define some new relational operations! Assume we have relations $R(A_1, ..., A_n)$ and $S(B_1, ..., B_m)$.

   (a) The right-semi join $R \ltimes S$ retains all tuples in S participate in the natural join between R and S. For instance, the right-semi join between R and S as defined in the previous problem produces $\{(2,2), (2,3)\}$. (You may use $\bowtie$ operation in your expression.)

   $R \ltimes S = ?$

   (b) The right-anti join $R \triangleleft S$ retains all tuples in S that do not participate in the natural join between R and S. For instance, the right-anti join between R and S as defined in the previous problem produces $\{(5,1)\}$.

   $R \triangleleft S = ?$

For the remaining problems, consider the relation instances below that model competitions in the world of college sports. Assume that each city may have multiple schools (*e.g.*, both Columbia and NYU are in New York). Schools are split into two conferences: A and B. Within each conference, there are four divisions: North, South, East, and West. Games are played by "home" and "away" (*i.e.*, the traveling) schools, identified by the schoolIDs. For instance, game 7 was between GaTech (24 points) and ASU (6 points) on 9/8/1980.

School

schoolID	city	school	conference	division	spent
0	Seattle	Washington	B	West	500M
1	Cleveland	Case	A	North	100M
2	Pittsburgh	CMU	A	North	-50M
3	San Francisco	Berkeley	B	West	800M
4	Oakland	Mills	A	West	350M
5	Tempe	ASU	A	West	400M
6	Miami	Miami	A	East	200M
7	Houston	Rice	B	South	250M
8	New York	Columbia	A	East	800M
9	New York	NYU	A	East	400M
10	Buffalo	SUNY	B	East	100M
11	Atlanta	GaTech	A	South	200M
12	Baltimore	JHU	A	North	0M
13	Lafayette	Purdue	A	South	90M

Game

gameID	Away	Home	date	year	awayScore	homeScore
0	0	11	9/2	2012	0	3
1	3	13	10/17	2009	23	0
2	5	10	10/10	2012	10	10
3	4	9	11/20	2015	17	7
4	2	7	9/27	2014	7	14
5	9	8	10/30	1990	14	15
6	8	3	8/30	1980	21	9
7	11	5	9/8	1980	24	6
8	4	2	10/28	1981	35	15
9	12	10	11/27	2012	3	40
...	...	...	...	...	...	...

Coach

name	schoolID	title
Carroll	0	Head Coach
Jackson	1	Head Coach
Tomlin	2	Head Coach
Kelly	3	Assistant Coach
Brown	1	Head Coach
Day	0	Assistant Coach
…	...	...

4. Write a relational-algebra expression that lists all schools (city and school) in the Eastern and Northern divisions of Conference "A."

5. List all 2015 games (gameID, awayScore, homeScore) in which the home team lost.

6. List all games (gameID, date, year) in which 'Columbia' participated.

7. List all the schools (city, school) for which 'Kelly' has coached as an assistant coach.