Middle East Technical University

Department of Computer Engineering

CENG 351
Data Management & File Structures
Fall 2023–2024
In Class Activity 2 — RA

Student ID: . Name and Surname:

Background

Assume that you are in the theoretical exam stage of the hiring process for a position as lead database administrator in the IT department of a sports TV channel. The previous administrator created a database to manage the results and statistics of Volleyball Leagues worldwide. To be hired, you must respond to a few relational algebra questions on this database.

Question 1 – 5 points

Fill in the blanks with the correct numeric values. (Assume that all "xxxReferees" relations are union-compatible. Also, you should answer the questions by considering all possibilities.)
Imagine a league in which some teams compete with each other. Each match must be refereed by some referees. The role of each referee in a match (e.g. first referee, second referee, first line referee, second line referee, and VAR (Video Assistant Referee)) should be recorded. A referee may have refereed several matches, but some referees may have not refereed any matches yet. A referee might referee different matches with different roles. Assume that all the matches are played on different dates.
If any "role name" and "Referees" are concatenated as the string for a relation name, this relation refers to all the referees who refereed some matches as that role such as "firstReferees", "secondReferees, "firstLineReferees", and "VARReferees". For instance, the "firstReferees" relation refers to all the referees

who refereed some matches as the first referee role.

- The relation "firstReferees" contains 15 rows and the relation "secondReferees" contains 46 rows. The maximum number of rows in "firstReferees" - "secondReferees" can be _____ . Note that "-" denotes the set difference.
- The relation "firstLineReferees" contains 25 rows and the relation "VARReferees" contains 17 rows. Their intersection contains maximum _____ rows, minimum _____ rows.
- Assume that each match must be refereed by exactly 5 referees with diverse roles. The table "Referees" contains 22 rows and the table "Matches" contains 13 rows. The relation "RefereesInMatches" which is the relation between "Matches" and "Referees" must contain _____ rows.

Question 2 – 10 points

Considering the following schemas and tables:

- Referee1(PIN, matchID)
- Referee2(PIN, role, matchID)
- Referee3(PIN, matchID)

Referee1

PIN	matchID
101	100021
201	100022

Referee2

PIN	role	matchID
101	first referee	100021
203	VAR referee	100023

Referee3

PIN	matchID
205	100021
209	100027

Which of the expressions below is not valid (legal) under the above three RefereeX relation instances?

a) $\rho(Temp1, Referee1 \times Referee2)$
$\rho(Result, Referee1 \bowtie Temp1)$

b) $\rho(Temp1, Referee2 \times Referee2)$
$\rho(Result, Temp1 \times Referee1)$

c) $\rho(Temp1, Referee2 - Referee3)$
$\rho(Result, Temp1 \times Referee1)$

d) $\rho(Temp1, Referee1–Referee3)$
$\rho(Result, Referee1 \bowtie Temp1)$

e) $\rho(Temp1, Referee1 \bowtie Referee2)$
$\rho(Result, \ Referee1 \ \bowtie \ Temp1)$

Question 3 – 10 points

Consider the following small part of the schemata developed for this database.

- **Teams** (teamID, team_name, foundation_year, number_of_championships)
- **Players** (PIN, player_name, gender, date_of_birth, height, weight)
- **Referees** (PIN, referee_name, gender, date_of_birth, level)
- **Jersey Numbers** (jersey_no, position)
- **Matches** (matchID, home_teamID, away_teamID, date, duration, sets_home, sets_away) home_teamID REFERENCES **Teams** (teamID), away_teamID REFERENCES **Teams** (teamID)
- **Play_for** (PIN, teamID, jersey_no, start_date, end_date) PIN REFERENCES **Players** (PIN), teamID REFERENCES **Teams** (teamID), jersey_no REFERENCES **Jersey Numbers** (jersey_no)
- **RefereesInMatches** (PIN, matchID, role) PIN REFERENCES **Referees** (PIN), matchID REFERENCES **Matches** (matchID)

Write a relational algebra expression that lists **all matches** (only the **matchIDs) of the "Fenerbahçe"** team either at **home** or **away**.

Question 4 – 10 points

Consider the following small part of the schemata developed for this database.

- **Teams** (teamID, team_name, foundation_year, number_of_championships)
- **Players** (PIN, player_name, gender, date_of_birth, height, weight)
- **Referees** (PIN, referee_name, gender, date_of_birth, level)
- **Jersey Numbers** (jersey_no, position)
- **Matches** (matchID, home_teamID, away_teamID, date, duration, sets_home, sets_away) home_teamID REFERENCES **Teams** (teamID), away_teamID REFERENCES **Teams** (teamID)
- **Play_for** (PIN, teamID, jersey_no, start_date, end_date) PIN REFERENCES **Players** (PIN), teamID REFERENCES **Teams** (teamID), jersey_no REFERENCES **Jersey Numbers** (jersey_no)
- **RefereesInMatches** (PIN, matchID, role) PIN REFERENCES **Referees** (PIN), matchID REFERENCES **Matches** (matchID)

Write a relational algebra expression that lists **all players (PIN, player name, jersey no)** who **play for the "Galatasaray"** team.

Question 5 – 10 points

Consider the following small part of the schemata developed for this database.

- **Teams** (teamID, team_name, foundation_year, number_of_championships)
- **Players** (PIN, player_name, gender, date_of_birth, height, weight)
- **Referees** (PIN, referee_name, gender, date_of_birth, level)
- **Jersey Numbers** (jersey_no, position)
- **Matches** (matchID, home_teamID, away_teamID, date, duration, sets_home, sets_away) home_teamID REFERENCES **Teams** (teamID), away_teamID REFERENCES **Teams** (teamID)
- **Play_for** (PIN, teamID, jersey_no, start_date, end_date) PIN REFERENCES **Players** (PIN), teamID REFERENCES **Teams** (teamID), jersey_no REFERENCES **Jersey Numbers** (jersey_no)
- **RefereesInMatches** (PIN, matchID, role) PIN REFERENCES **Referees** (PIN), matchID REFERENCES **Matches** (matchID)

Write a relational algebra expression that lists **all matches** (only the **matchIDs**) of team **"Beşiktaş" after 01.11.2022** either **away** or **home** with **result of "3 -1"** either **won** or **lost**. (Note: $x > y$ when comparing dates means x is a date later than y).

Question 6 – 20 points

Consider the following small part of the schemata developed for this database.

- **Teams** (teamID, team_name, foundation_year, number_of_championships)
- **Players** (PIN, player_name, gender, date_of_birth, height, weight)
- **Referees** (PIN, referee_name, gender, date_of_birth, level)
- **Jersey Numbers** (jersey_no, position)
- **Matches** (matchID, home_teamID, away_teamID, date, duration, sets_home, sets_away) home_teamID REFERENCES **Teams** (teamID), away_teamID REFERENCES **Teams** (teamID)
- **Play_for** (PIN, teamID, jersey_no, start_date, end_date) PIN REFERENCES **Players** (PIN), teamID REFERENCES **Teams** (teamID), jersey_no REFERENCES **Jersey Numbers** (jersey_no)
- **RefereesInMatches** (PIN, matchID, role) PIN REFERENCES **Referees** (PIN), matchID REFERENCES **Matches** (matchID)

The **referee roles** that are taken by **both of the referees "Nurper Özbar" and "Tuncay Sevim" are called as PopularRoles.** By using the schemata in above, write a relational algebra expression that **lists referees (PIN and referee name)** who refereed matches by taking **all of the PopularRoles.**

Question 7 – 20 points

Consider the following small part of the schemata developed for this database.

- **Teams** (teamID, team_name, foundation_year, number_of_championships)
- **Players** (PIN, player_name, gender, date_of_birth, height, weight)
- **Referees** (PIN, referee_name, gender, date_of_birth, level)
- **Jersey Numbers** (jersey_no, position)
- **Matches** (matchID, home_teamID, away_teamID, date, duration, sets_home, sets_away) home_teamID REFERENCES **Teams** (teamID), away_teamID REFERENCES **Teams** (teamID)
- **Play_for** (PIN, teamID, jersey_no, start_date, end_date) PIN REFERENCES **Players** (PIN), teamID REFERENCES **Teams** (teamID), jersey_no REFERENCES **Jersey Numbers** (jersey_no)
- **RefereesInMatches** (PIN, matchID, role) PIN REFERENCES **Referees** (PIN), matchID REFERENCES **Matches** (matchID)

Write a relational algebra expression that lists **all referees (PIN and referee name)** who refereed matches with the roles of **both the "first referee" and "second referee"** but **without** the role of **"VAR referee"**.

Question 8 – 15 points

Considering the following tables:

Teams

teamID	team_name	foundation_year	number_of_championships
1001	METU	1956	25
1002	Hacettepe	1967	24
1003	Fenerbahçe	1907	19
1004	Galatasaray	1905	23
1005	Beşiktaş	1903	16

Matches

matchID	home_teamID	away_teamID	date	duration	sets_home	sets_away
20210113	1003	1004	13.01.2021	97	3	1
20210201	1002	1005	1.02.2021	79	3	0
20210610	1001	1002	10.06.2021	76	3	0
20220303	1002	1001	3.03.2022	93	1	3
20220419	1002	1001	19.04.2022	81	0	3
20220508	1001	1002	8.05.2022	121	3	2
20230221	1005	1004	21.02.2023	85	3	0
20230702	1001	1003	2.07.2023	102	3	1
20230827	1002	1001	27.08.2023	100	1	3
20230901	1001	1002	1.09.2023	125	2	3

a) By using the schemata in Question 7, write a relational algebra expression that **lists all distinct match scores (i.e., pairs of sets_home and sets_away)** that occurred in the matches. **Draw** the output relation when your RA expression is applied to the instance of Matches relation provided above.

b) $\rho(Metu, \Pi_{teamID}(\sigma_{team\_name='METU'}(Teams)))$

$\rho(Tmp, \Pi_{Matches.matchID}(Metu \bowtie_{Metu.teamID=Matches.home\_teamID}$
$(\sigma_{date>30.03.2022 \wedge date<30.07.2023}(Matches))))$

$\rho(Tmp2, (Tmp \bowtie Matches))$

$\rho(Result, \Pi_{Tmp2.date,Teams.team\_name,Tmp2.sets\_home,Tmp2.sets\_away}$
$(Teams \bowtie_{Teams.teamID=Tmp2.away\_teamID} Tmp2))$

What is the output of this relational algebra expression? **Draw** the relation table for the result. (Note: $x > y$ when comparing dates means x is a date later than y).