

**luping xue**

Name

**A20453695**

CWID

# Homework Assignment 1

**Due Date: Friday, Sept 13, 2019  
(11:59 PM CDT)**

**CS425 - Database Organization**

---

*Please leave this empty!*

1.1

1.2

1.3

1.4

1.5

1.6

1.7

1.8

1.9

1.10

1.11

1.12

Sum

# Instructions

- Try to answer all the questions using what you have learned in the class.
- Please write queries in *Relational Algebra*. Do not write SQL queries in this assignment!
- When writing a query, write the query in a way that it would work over all possible database instances (schema) and not just for the given example instance! In other words, queries can be created without the given example instance.

Consider the following database schema and example instance:

Athlete			
<u>aid</u>	name	weight	age
01	Phil	250	39
02	Jay	240	46
03	Simeon	190	33
04	Jeremy	220	28
05	Jeff	195	25
...	...	...	...

cid	title	region	credit
MOL-US	Mr. Olympia	US	3
MOL-IL	Mr. Olympia	Illinois	2
MOL-CHI	Mr. Olympia	Chicago	1
MOL-NY	Mr. Olympia	New York	2
MOL-NYC	Mr. Olympia	New York City	1
ARC-US	Arnold Classic	US	3
ARC-IL	Arnold Classic	Illinois	2
ARC-CHI	Arnold Classic	Chicago	1
ARC-NY	Arnold Classic	New York	2
ARC-NYC	Arnold Classic	New York City	1
...	...	...	...

## Result

<u>cid</u>	<u>aid</u>	rank	score
MOL-US	01	1st	20
MOL-US	02	2nd	10
MOL-US	04	3rd	5
MOL-IL	02	1st	20
MOL-IL	07	2nd	10
MOL-IL	09	3rd	5
ARC-US	01	1st	20
ARC-US	04	2nd	10
ARC-US	08	3rd	5
ARC-IL	02	1st	20
...	...	...	...

## Subcompetition

cid	sid
MOL-US	MOL-IL
MOL-US	MOL-NY
MOL-IL	MOL-CHI
MOL-NY	MOL-NYC
...	...

## Hints:

- This database is designed for storing the information of bodybuilding competitions. The data types of attributes *weight (lbs)*, *age*, *credit*, and *score* are defined as *Number*. The data types of the remaining attributes are defined as *String*. Attribute *title* refers to the title of the championship in the competition.
- Each underlined attribute is the primary key of a relation.
- The attributes *cid* and *aid* of relation *Result* is a foreign key to relations *Competition* and *Athlete*, respectively. Both attributes *cid* and *sid* are foreign keys (two different foreign keys) to relation *Competition* (which records the information of *direct sub-competitions*). For instance, ‘MOL-CHI’ is the direct sub-competition of ‘MOL-IL’ and ‘MOL-IL’ is the direct subcompetition of ‘MOL-US’.
- Attribute *score* corresponds to attribute *rank* (e.g., an athlete receives score 20 if he is ranked as ‘1st’ in the competition). There exists some redundancy in relation *Result* (please skip this in the queries).

## Part 1.1 Relational Algebra (Total: 100 Points)

### Question 1.1.1 (6 Points)

List the names of all the athletes who participate in the competition ‘MOL-IL’.

$$\Pi \text{ name}(\sigma \text{ cid}='MOL-IL'(\text{Result} \bowtie \text{Athlete}))$$

### Question 1.1.2 (6 Points)

List the query result as “title, region, name, rank” where the athlete weight is greater than 230lbs.

$$\Pi \text{ name, title, region, rank } (\sigma \text{ weight} > 230 (\text{Athlete} \bowtie \text{Result} \bowtie \text{Competition}))$$

### Question 1.1.3 (8 Points)

List the athletes (result: aid, name) who participate in the direct subcompetitions of ‘ARC-IL’ and got a rank ‘2nd’. (Hint: only the direct subcompetition of ‘ARC-IL’)

$$\Pi \text{ a.aid , a.name } (\sigma \text{ s.cid}='ARC\_IL' \wedge \text{r.rank}='2nd' ( \rho \text{ s(Subcompetition)} \bowtie \text{s.sid}=\text{r.cid} \rho \text{ r(Result)} \bowtie \text{pa(Athlete)} ) )$$

#### Question 1.1.4 (10 Points)

List the athletes (result: aid, name) who have participated in all the competitions in the ‘US’. (Hint: only national competitions with region ‘US’, not including subcompetitions)

$$E1 \leftarrow \prod cid, aid (\sigma region='US' (Competition \bowtie Result))$$
$$E2 \leftarrow \prod cid (\sigma region='US' (Competition))$$
$$E3 \leftarrow E1 \div E2$$
$$\prod aid, name (Athlete \bowtie E3)$$

#### Question 1.1.5 (10 Points)

List the athletes (result: aid, name) whose score in ‘MOL-US’ is higher than his score in ‘ARC-US’. (Hint: only count the athletes who have participated in both ‘MOL-US’ and ‘ARC-US’)

$$E1 \leftarrow \prod aid, cid (\sigma cid = 'MOL_US' (Result))$$
$$E2 \leftarrow \prod aid, cid (\sigma cid = 'ARC_US' (Result))$$
$$E3 \leftarrow \prod E1.aid (\sigma E1.score > E2.score (E1 \bowtie E2))$$
$$E \leftarrow \prod aid, name (Athlete \bowtie E3)$$

### Question 1.1.6 (8 Points)

List the athletes (result: aid, name) who never got a rank lower than ‘2nd’ in any competition he has participated in. (Hint: rank lower than ‘2nd’ is equivalent to “score less than 10”)

$$\Pi \text{ aid, name } (\text{Athlete}) - \Pi \text{ aid, name } ( \sigma r.\text{score} < 10 ( \rho a(\text{Athlete}) \bowtie \rho r(\text{Result}) ) )$$

### Question 1.1.7 (8 Points)

List all the ‘US’ competitions ‘Jay’ has not participated in (result: title, region). (Hint: only national competitions with region ‘US’, not including subcompetitions)

$$\begin{aligned} E1 &\leftarrow \Pi c\text{id} ( \sigma c.\text{region} = "US" (\text{Competition})) \\ E2 &\leftarrow \Pi c\text{id} ( \sigma a.\text{name} = "jay" ( \rho a(\text{Athlete}) \bowtie \rho r(\text{Result}) ) ) \\ E3 &\leftarrow E1 - E2 \\ \Pi c.\text{title}, c.\text{region} ( E3 \bowtie \rho c(\text{Competition}) ) \end{aligned}$$

### **Question 1.1.8 (8 Points)**

List the athletes and their weighted average scores (result: name, weighted) in all the competitions. (Hint: weighted = (score × credit)/total credits)

```
E1 ← Π cid,aid,(score* credit) as sc  (Competition ⋈ Result)
E2 ← aid G sum(sc) as tsc(E1)
E3 ← aid G sum(credit) as tc (Competition ⋈ Result)
E4← Π E2.aid, (tsc/tc) as weighted (E2 ⋈ E3)
result ← Π name , weighted ( Athlete ⋈ E4 )
```

### **Question 1.1.9 (8 Points)**

List all the competitions and the counts of their direct subcompetitions (result: cid, count) in the database. (Hint: if a competition does not have any direct subcompetition, return constant 0)

```
E1 ← Π cid (Competition) – Π cid (Subcompetition)
E2 ← Competition ⋈ Subcompetition
```

```
Competition.cid G count(*) (E2) U Π cid,0 (E1)
```

### Question 1.1.10 (6 Points)

List all the competitions where the average age of the athletes in the competition is lower than 30 (result: cid, avg\_age).

$E1 \leftarrow \prod r.cid, a.age (\rho r(Result) \bowtie r.aid=a.aid \rho a(Athlete))$

$E2 \leftarrow cid G \text{ avg}(age) \text{as avg\_age } (E1)$

$\text{result} \leftarrow \prod cid, avg\_age (\sigma (avg\_age < 30) (E2))$

### Question 1.1.11 (10 Points)

List all the competitions and the name(s) of the athlete(s) with the highest score (result: cid, name) in each competition. (Hint: please do not use *max* to query the highest score)

$E1 \leftarrow \prod r1.cid, r1.aid (\sigma r1.score < r2.score \wedge r1.cid=r2.cid (\rho r1(Result) \times \rho r2(Result)))$

$E2 \leftarrow \prod Result.cid, Result.aid (\prod cid, aid (Result) - E1)$

$\text{result} \leftarrow \prod cid, name (\sigma a.aid = e.aid (\rho a(Athlete) \times \rho e(E2)))$

### Question 1.1.12 (12 Points)

List the athletes (result: aid, name) who have participated in ‘ARC-US’ and all its subcompetitions. (Hint: all the subcompetitions of ‘ACR-US’ include not only the direct subcompetitions but also the subcompetitions of the subcompetition, … )

$R1 := \prod_{cid, aid, sid} (\sigma_{cid='ARC\_US'} (Result \bowtie Subcompetition))$

$R2 := R1 \cup (\prod_{cid, aid, sid} (\sigma_{R1.aid=r.aid \wedge r.cid = R1.sid} (\rho_{r(Result)} \times R1)))$

```
while change do
begin
T := R2
end
result ←  $\prod_{aid, name} (Athlete \bowtie T)$ 
```