CSC 365 Homework 1
Solution

May 10, 2023

Problem 1

1.a

$R \cap S$

R∩S:

A	B	C
b	c	3
c	c	3
c	a	3
b	b	1

1.b

$S - R$

S−R:

A	B	C
c	c	2
a	b	3
a	a	1

1.c

$R - S$

R−S:

A	B	C
c	a	4
b	a	2

1.d

$\pi_{A,B}(S)$

$\pi_{A,B}$(S):

A	B
c	c
c	a
b	b
a	b
b	c
a	a

1.e

$\pi_{B,C}(R) \cup \pi_{B,C}(S)$

$\pi_{B,C}$(R)$\cup\pi_{B,C}(S)$:

B	C
c	3
a	3
b	1
a	4
a	2
c	2
b	3
a	1

1.f

$\pi_{A,B}(R) - \pi_{A,B}(S)$

$\pi_{A,B}$(R)$-\pi_{A,B}$(S):

A	B
b	a

1.g

$\pi_C(W) \times \pi_A(S) \times \pi_B(T)$

$\{1,2,3,4\} \times \{a,b,c\} \times \{a,b,c,d\}$: 48 tuples total.

1.h

$\sigma_{E>F}(T)$

$\sigma_{E>F}$(T):

B	D	E	F
b	b	3	2
d	b	3	2
b	c	4	1

1.i

$\sigma_{A\neq B}(R)$

$\sigma_{A \neq B}$(R):

A	B	C
b	c	3
c	a	3
c	a	4
b	a	2

1.j

$\pi_{B,F}(\sigma_{F \geq E}(T))$

$\pi_{B,F}(\sigma_{F \geq E}$(T)):

B	F
a	2
c	4
a	3
d	4

1.k

$\sigma_{A=D \vee B=D}(\pi_{A,B}(R) \times W)$

A	B	C	D
b	c	2	b
b	c	3	c
c	c	3	c
c	a	3	c
c	a	1	a
b	b	2	b
b	a	1	a
b	a	2	b

(essentially, this is a join $\pi_{A,B}(R) \bowtie_{A=D \vee B=D} W$)

1.l

$\sigma_{A=b \wedge C>1}(R) \cup \sigma_{B=b \vee C \neq 3}(S))$

A	B	C
b	c	3
b	a	2
c	c	2
b	b	1
a	b	3
a	a	1

1.m

$\sigma_{\neg B=d}(T)$

B	D	E	F
a	a	1	2
c	d	2	4
b	b	3	2
a	a	2	3
b	c	4	1

1.n

$\pi_{A,B,R.C,D}(\sigma_{R.C=W.C}(R \times W))$

A	B	R.C	D
b	c	3	c
c	c	3	c
c	a	3	c
b	b	1	a
c	a	4	d
b	a	2	b

(This is basically a natural join $R \bowtie W$)

1.o

$W \bowtie R$

W.C	D	A	B
3	c	b	c
3	c	c	c
3	c	c	a
1	a	b	b
4	d	c	a
2	b	b	a

(same as 1.n, except for the order of attributes)

1.p

$W \bowtie_{W.C=R.C} R$

W.C	D	A	B	R.C
3	c	b	c	3
3	c	c	c	3
3	c	c	a	3
1	a	b	b	1
4	d	c	a	4
2	b	b	a	2

(This is an equijoin, so there is no projection, i.e., both $R.C$ and $W.C$ are kept.)

1.q

$T \bowtie_{F>C} W$

B	T.D	E	F	C	W.D
a	a	1	2	1	a
c	d	2	4	2	b
c	d	2	4	3	c
c	d	2	4	4	d
b	b	3	2	1	a
d	b	3	2	1	a
a	a	2	3	1	a
a	a	2	3	2	b
d	a	1	4	1	a
d	a	1	4	2	b
d	a	1	4	3	d

1.r

$R \bowtie S$

R⋈S:

A	B	C
b	c	3
c	c	3
c	a	3
b	b	1

(Notice anything?)

1.s

$R \bowtie_{R.B=S.A} S$

R:

R.A	R.B	R.C	S.A	S.B	S.C
b	c	3	c	c	2
b	c	3	c	a	3
b	c	3	c	c	3
c	c	3	c	c	2
c	c	3	c	a	3
c	c	3	c	c	3
c	a	3	a	b	3
c	a	3	a	a	1
b	b	1	b	b	1
b	b	1	b	c	3
c	a	4	a	b	3
c	a	4	a	a	1
b	a	2	a	b	3
b	a	2	a	a	1

1.t

$(R \bowtie T) \bowtie \pi_{A,C,D}(S \bowtie W)$

$\emptyset$
(after all the complex computations)

1.u

$\pi_{T1.D,T2.B}(\rho_{T1}(T) \bowtie_{T1.D=T2.B} \rho_{T2}(T))$

T1.D	T2.B
a	a
b	b

1.v

$\pi_{B,D,E}(\sigma_{F \leq C}(T \bowtie W))$

B	D	E
c	d	*2*
b	b	*3*
d	b	*3*
b	c	*4*

1.w

$\pi_{R.A,R.B}(R \bowtie_{R.C \neq S.C} S) \bowtie \sigma_{D=a}(T)$
(to disambiguiate things, we use $R.A$ and $R.B$ in the projection)

R.A	R.B	D	E	F
c	a	a	*1*	*2*
b	a	a	*1*	*2*
c	a	a	*2*	*3*
b	a	a	*2*	*3*

1.x

$\pi_A(\pi_B(\pi_C(R \cup S)))$
$\emptyset$
(Why?)

1.y

$\sigma_{A \neq a}(S) \bowtie \sigma_{D \neq c}(W)$

A	B	S.C	D
c	c	*2*	b
b	b	*1*	a

1.z

$\sigma_{C=1}(R) \bowtie \sigma_{C=2}(S)$
$\emptyset$
(Quite obvious, isn't it?)

Problem 2

1. Find all musicians who played with their band in year 1970. Output the names of the musicians.

$$\pi_{\mathsf{Name}}\left(\sigma_{\mathsf{From}\leq 1970 \wedge \mathsf{To}\geq 1970}(\mathsf{M})\right)$$

2. Find all musicians who played with their band in year 1970. Output the names of the musicians and the name of the band they played in.

$$\pi_{\mathsf{M.Name},\mathsf{B.Name}}\left(\sigma_{\mathsf{From}\leq 1970 \wedge \mathsf{To}\geq 1970}(\mathsf{M}) \bowtie_{\mathsf{BandId}=\mathsf{Id}} B\right)$$

3. Find all musicians who played for `‘‘Gong’’`. Output the names of the musicians and the years they played in the band.

$$\pi_{\mathsf{M.Name},\mathsf{From},\mathsf{To}}\left(\sigma_{\mathsf{Name}='\mathsf{Gong}'}(B) \bowtie_{\mathsf{Id}=\mathsf{BandId}} M\right)$$

4. Find all musicians who played in the band `‘‘King Crimson’’` in the year 1974. Output the names of the musicians.

$$\pi_{\mathsf{M.Name}}\left(\sigma_{\mathsf{Name}='\mathsf{KingCrimson}'}(B) \bowtie_{\mathsf{Id}=\mathsf{BandId}} \sigma_{\mathsf{From}\leq 1974 \wedge \mathsf{To}\geq 1974}(M)\right)$$

5. Find all musicians who played in their band in the band’s year of inception. Output the names of the musicians and the names of the bands.

$$\pi_{\mathsf{M.Name},\mathsf{B.Name}}\left(B \bowtie_{\mathsf{Id}=\mathsf{BandId}\wedge \mathsf{Formed\_in}=\mathsf{From}} M\right)$$

6. Find the band that released the album `’’Loaded’’`. Report the name of the band.

$$\pi_{\mathsf{Name}}\left(\sigma_{\mathsf{Title}='\mathsf{Loaded}'}(A) \bowtie_{\mathsf{BandId}=\mathsf{Id}} B\right)$$

7. Find the band in which `‘‘Michael Karoli’’` played. Report the name of the band.

$$\pi_{\mathsf{B.Name}}\left(\sigma_{\mathsf{Name}='\mathsf{JimMorrison}'}(M) \bowtie_{\mathsf{BandId}=\mathsf{Id}} B\right)$$

8. Find all albums recoreded by bands from UK. For each record output its name, year and the name of the band that recorded it.

$$\pi_{\mathsf{Title},\mathsf{Year},\mathsf{Name}}\left(\sigma_{\mathsf{Country}='\mathsf{UK}'}(B) \bowtie_{\mathsf{Id}=\mathsf{BandId}} A\right)$$

9. Find all musicians who participated in recording of the album `‘‘Fragile’’`. Output the names of the musicians.

$$\pi_{\textsf{Name}}\left(\sigma_{\textsf{Title}='\textsf{Fragile}'}(A) \bowtie_{\textsf{A.BandId=M.BandId}\wedge\textsf{Year}\geq\textsf{From}\wedge\textsf{Year}\leq\textsf{To}} M\right)$$

10. Find all `‘‘Pink Floyd’’` band members who did NOT participate in the recording of the album `‘‘Meddle’’`. (note: a musician did not participate in a recording of an album if he did not play in the band that year).

$$\pi_{\textsf{M.Name}}\left(M \bowtie_{\textsf{BandId=Id}} \sigma_{\textsf{Name}='\textsf{PinkFloyd}'}(B)\right) -$$
$$\pi_{\textsf{Name}}\left(\sigma_{\textsf{Title}='\textsf{Meddle}'}(A) \bowtie_{\textsf{A.BandId=M.BandId}\wedge\textsf{Year}\geq\textsf{From}\wedge\textsf{Year}\leq\textsf{To}} M\right)$$

11. Find all bands that recorded two albums in the same year. For each band, output its name, titles of both albums and the year of their release.

$$\pi_{\textsf{Name,A.Title,A2.Title}}\left(\left(A \bowtie_{\textsf{A.Year=A1.Year}\wedge\textsf{A.Aid<A2.Aid}\wedge\textsf{A.BandId=A2.BandId}} \rho_{A2}(A)\right) \bowtie_{A.BandId=Id} B\right)$$

12. Find all bands in which `"Lou Reed"` DID NOT play. Output the names of the bands.

$$\pi_{\textsf{Name}}(B) - \pi_{\textsf{Name}}\left(\sigma_{\textsf{Name}='\textsf{LouReed}'}(M) \bowtie_{BandId=Id} B\right)$$

13. Find all albums recorded by the bands who had at least one musician leave befire 1972. Output the name of the album and the year of release.

$$\pi_{\textsf{A.Title,A.Year}}\left(\left(\sigma_{\textsf{To}<1972}(M) \bowtie_{\textsf{M.BandId=B.Id}} B\right) \bowtie_{\textsf{B.Id=A.BandId}} (A)\right)$$

14. Find all albums recorded by `‘‘King Crimson’’` before `‘‘Adrian Belew’’` joined the band. Report the album titles and years.

$$\pi_{\textsf{A.title,A.Year}}\left(\sigma_{\textsf{Name}='\textsf{AdrianBelew}'}(M) \bowtie_{\textsf{M.Year>A.Year}} \left(\sigma_{\textsf{Name}='\textsf{KingCrimson}'}(B) \bowtie_{\textsf{B.Id=A.BandId}} (A)\right)\right)$$

15. Find all musicians who played in at least two different bands. Report their names.

$$\pi_{\textsf{M1.Name}}\left(\left(\rho_{B1}(B) \bowtie_{\textsf{B1.Id=M1.BandId}} \rho_{M1}(M)\right) \bowtie_{\textsf{M1.Name=M2.Name}} \left(\rho_{B2}(B) \bowtie_{\textsf{B2.Id=M2.BandId}} \rho_{M2}(M)\right)\right)$$

Problem 3

1. One possible candidate key is hinted at directly in the problem statement: `(FirstName, LastName, TeamName, Season)`. A row of the table reports the stats for a given NBA player for a specific season. `(FirstName, LastName, TeamName)` identifies the player, and `Season` identifies the season for the player's stats. Please note here, that in NBA team names (`Lakers`, `Jazz`, `Raptors` and so on) are unique, while team cities are not - e.g., there are two teams from Los Angeles. As a result, the following: `(FirstName, LastName, TeamCity, Season)` is **not a candidate key** - there is a possibility that a person with the same first name and last name plays for each of two teams from the same city in a given season.

2. `(TeamName, Shots, Rebounds)` is not a valid candidate key, because (a) it does not contain `Season`, and (b) there is no reason to assume that each team would have unique pairs of number of shots/number of rebounds as statistics for their players. In fact, at the beginning of a given season, all players on the team will have 0 shots and 0 rebounds.

3. `(LastName, Season)` is not a valid candidate key because it is too restrictive. If this was a candidate key no two different players with the same last name could play in the same season. Over its history NBA has consistently featured multiple players with the same last name in most of its seasons.

4. `(FirstName, LastName, TeamCity, TeamName, Height, Position, Season, Games)` is not a candidate key because it contains `(FirstName, LastName, TeamName, Season)` - which according to Part 1 of this question is a valid candidate key. By definition, no candidate key can contain another candidate key.

Problem 4

1. From the description of the attribute, it is clear that `RouteNumber` is a candidate key. However, it is not the only one.

   We can make a reasonable assumption that there is no reason to have two trains connecting the same cities that have the same departure and arrival times. In fact, we can make a somewhat stronger assumption that no two trains leave the same origin point for the same destination point at the same time. At the same time we also notice, that two cities **can** be connected by multiple train routes. This gives rise to the following candidate key: `(Origin, Destination, Departure)`.

   Symmetrically, we can assume that no two trains arrive to the same city at the same time if they originated from the same point. This would lead to another candidate key: `(Origin, Destination, Arrival)`.

If you wanted to make a more conservative assumption, `(Origin, Departure, Destination, Arrival)` could be considered a valid candidate key instead of the two candidate keys described above.

2. Of the three candidate keys described, one, `RouteNumber` is the simplest, and therefore it fits better as a primary key.

3. `(RouteNumber, Origin, NDays, NStops)` is not a candidate key, because it contains `RouteNumber`, which itself is a candidate key.

4. `(Origin, Destination)` is not a candidate key because (see Part 1) is should be possible for two cities to be connected by more than one train in a given day. In fact, many cities in Europe are connected by trains that leave every hour or so.

   `(Origin, TrainType, NStops)` is not a candidate key because two trains of the same type (e.g., "regular") can depart from a given point of origin at different times (for example), and follow the same route with the same number of stops.