# Section 10

*Priya Rao Chagaleti*

# XML / XQuery / XPath

# Problem a 1

Book(bid, title, publisher, price, year)
Author(aid, last, first)
Editor(eid, last, first, affiliation)
Writes(aid, bid)
Edits(edi, bid)

# Problem a 2 i

```
select Book.title
from book, edits, editor
where book.bid = edits.bid and edits.eid = editor.eid
 and book.year < 2005 and book.price < 99
 and editor.last = 'Samet'
```

# Problem a 2 ii

```
select x.title
from book x, writes y, author z, writes u, author v
where x.bid = y.bid and y.aid = z.aid
  and x.bid = y.bid and y.aid = z.aid
```

and z.last = “Hull” and v.last “Vianu”

# Problem b i

Let E stand for Element

```
select distinct xtitle.child
from E xbib, E xbook, E xeditor, E xlast, E xprice, E xtitle
where xbib.id = 0 and xbib.tag='bib'
  and xbook.id=xbib.child and xbook.tag='book'
  and xeditor.id=xbook.child and xeditor.tag='editor' and xeditor.child='Samet'
  and xprice.id=xbook.child and xprice.tag='price' and xprice.child <'99'
  and xtitle.id=xbook.child and xtitle.tag='title'
```

# Problem b ii

```
select distinct xtitle.child
from E xbib, E xbook, E xa1, E xl1, E xa2, E xl2, E xtitle
where xbib.id = 0 and xbib.tag='bib'
  and xbook.id=xbib.child and xbook.tag='book'
  and xa1.id = xbook.child and xa1.tag = 'author'
  and xl1.id = xa1.child and xl1.tag = 'last' and xl1.child = 'Hull'
  and xa2.id = xbook.child and xa2.tag = 'author'
  and xl2.id = xa2.child and xl2.tag = 'last' and xl2.child = 'Vianu'
  and xtitle.id = xbook.id and xtitle.tag = 'title'
```

# Conceptual design/ Constraints/ Views

# Problem 1 a - Solution 1

Solution 1:

| Table | $X^+ = ?$ | New table 1 | New table 2 |
|---|---|---|---|
| $R(A, B, C, D, E)$ | $ABC+ = ABCD$ | $R_1(A, B, C, D)$ | $R_2(A, B, C, E)$ |
| $R_1(A, B, C, D)$ | $AD+ = ACD$ | $R_3(A, C, D)$ | $R_4(A, B, D)$ |
| $R_2(A, B, C, E)$ | $E+ = BE$ | $R_5(B, E)$ | $R_6(A, C, E)$ |

Answer: $R_3(A, C, D), R_4(A, B, D), R_5(B, E), R_6(A, C, E)$.

# Problem 1 a - Solution 2

Solution 2:

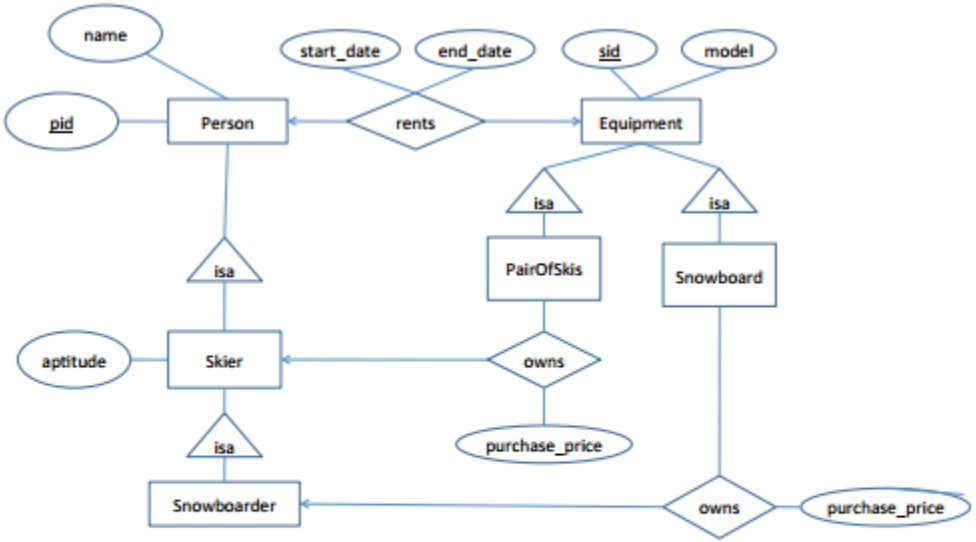
| Table | $X^+ = ?$ | New table 1 | New table 2 |
|---|---|---|---|
| $R(A, B, C, D, E)$ | $E+ = BE$ | $R_1(B, E)$ | $R_2(A, C, D, E)$ |
| $R_2(A, C, D, E)$ | $AD+ = ACD$ | $R_3(A, C, D)$ | $R_4(A, D, E)$ |

Answer: $R_1(B, E), R_3(A, C, D), R_4(A, D, E)$.

# Problem 1 b

| FD | Holds ? |
|---|---|
| $B \rightarrow A$ | holds |
| $C \rightarrow A$ | fails: tuples 1,3 |
| $A \rightarrow B$ | fails: tuples 1,2 |
| $C \rightarrow B$ | fails: tuples 1,3 |
| $A \rightarrow C$ | fails: tuples 1,2 |
| $B \rightarrow C$ | fails: tuples 3,4 |
| $BC \rightarrow A$ | holds |
| $AC \rightarrow B$ | holds |
| $AB \rightarrow C$ | fails: tuples 3,4 |

# Problem 2 a




Note: This solution assumes that when a person returns a piece of equipment, the corresponding tuple is deleted from the **rents** relation. If we wanted to keep the historical data, we would change the relationship multiplicity to many-many and we could enforce the constraint that dates cannot overlap using a check constraint or trigger. Additionally, we do not prevent a piece of equipment form being both a PairOfSkis and a Snowboard at the same time. Such a hybrid still counts as one piece of equipment to rent.

# Problem 2 b

```
CREATE TABLE owns(skier_id INT REFERENCES Skier,
                  ski_id   INT REFERENCES PairOfSkis,
                  purchase_price INT,
                  PRIMARY KEY(ski_id),
                  CHECK ( purchase_price > 0))
```

# Problem 3 a

A superkey is a set of attributes X s.t. $X^+ =$ all attributes.

From the FDs above, we can derive:

$$\{A, B, D\}^+ = \{A, B, C, D\}^+ = \{A, B, D, E\}^+ = \{A, B, C, D, E\}^+ = \{A, B, C, D, E\}$$

Hence,
$\{A, B, D\}, \{A, B, C, D\}, \{A, B, D, E\}$, and $\{A, B, C, D, E\}$ are all superkeys.

A key is a set of attributes which form a superkey and for which no subset is a superkey. In our example, $\{A, B, D\}$ is the only key.

# Problem 3 b

Both functional dependencies violate BCNF.

Try $\{A, B\}^+ = \{A, B, E\}$. Decompose into R1(A,B,E) and R2(A,B,C,D).

For R1, AB $\rightarrow$ E is the only FD and $\{A, B\}$ is a key, so R1 is in BCNF.

R2 is not in BCNF, since $\{D\}$ is not a key and we have D $\rightarrow$ C.

Try $\{D\}^+ = \{C, D\}$. Decompose into R3(C, D) and R4(A, B, D )

End result: R1(A,B,E), R3(C, D), and R4 (A, B, D )

# Parallel Data Processing

# Problem 4 a

| Job $J$ consists of: | Type of Parallelism | | |
|---|---|---|---|
| | Inter-query | Inter-operator | Intra-operator |
| 1 SQL query | 100 | 100 | 10 |
| 1000 SQL queries | 10 | 100 | 10 |

# Problem 4 b

i)16MB/128KB = 128

ii)16MB/64KB = 256

iii)365: # days/year

# Problem 4 c

i) 50 * 100 = 5000

ii) 5GB/50 = 100MB

iii) 100*1/10 + 50* 1/10 = 15 minutes

iv) 15 + 9 = 24 minutes