CSE 444 - Section 2

E/R Diagrams, Group-by's, sub-queries and more...

09 April 2009
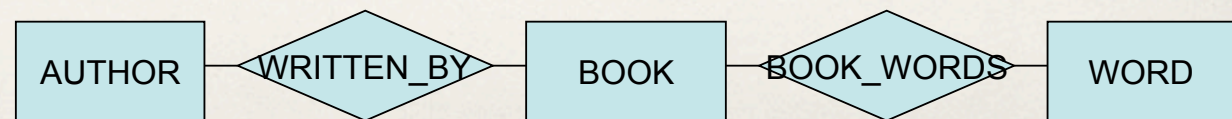
Today's Plan

* E/R Diagrams - Schema design

* SQL Exercises (to practice `group-by`)

* Discussion about sub-queries

* Questions?

E/R Diagrams - Schema Design

Our Schema

- AUTHOR(aid, name, age)

- WRITTEN_BY(bid, aid)

- BOOK (bid, title)

- BOOK_WORDS (bid, word)

- WORD(word) - we can do without this.

Constraints

* What are examples of ICs constraints that we might want?

    * Keys, foreign keys

    * Attribute-level constraints

    * Tuple-level

    * Global constraints

Inserts, Updates, Deletes

* `INSERT INTO AUTHOR VALUES(312,'Michael Chabon',45);`

* `UPDATE AUTHOR SET AGE=46 WHERE aid=312;`

* `DELETE FROM AUTHOR WHERE aid=312;`
[be careful! don't forget the WHERE condition!]

* But what happens to all the books he wrote?
    * under REJECT policy?
    * under CASCADE policy?
    * under SET-NULL policy?

Group-by Exercises - 1

Find names of authors who wrote more than 20 books.

Without group-by:
```
SELECT  name
FROM    AUTHOR a
WHERE   (SELECT COUNT(*)
          FROM    WRITTEN_BY wb
          WHERE   wb.aid = a.aid) > 20
```

With group-by:
```
SELECT    name
FROM      AUTHOR a, WRITTEN_BY wb
WHERE     a.aid = wb.aid
GROUP BY  a.aid, a.name --note that we must include a.name
HAVING    COUNT(*) > 20
```

The dreaded group-by error

* Column '___' is invalid in the select list because it is not contained in either an aggregate function or the GROUP BY clause.

Group-by Exercises - 2

Find authors who have a vocabulary of more than 10k words.

Without group-by:
```
SELECT name
FROM    AUTHOR A
WHERE   (SELECT COUNT(DISTINCT word)
         FROM    WRITTEN_BY wb, Book_Words bw
         WHERE   A.aid = wb.aid AND wb.bid = bw.bid) > 10000
```

With group-by:
```
SELECT    name
FROM      AUTHOR a, WRITTEN_BY wb, BOOK_WORDS bw
WHERE     a.aid = wb.aid AND wb.bid = bw.bid
GROUP BY  a.aid, a.name
HAVING    COUNT(DISTINCT word) > 10000
```

Group-by Exercises - 3

For each author, report average number of words per book.

<u>With group-by:</u>
```
SELECT    aid, AVG(num)
FROM      ( SELECT   aid, bid, COUNT(*) num
            FROM     AUTHOR a, WRITTEN_BY wb, BOOK_WORDS bw
            WHERE    …
            GROUP BY aid, bid) t
GROUP BY aid
```

Can we make this into a VIEW?

Group-by Exercises - 4

Find most frequently used word.

With group-by:
```
SELECT    word
FROM      BOOK_WORDS
GROUP BY  word
HAVING    count(*) >= ALL(SELECT    count(*)
                          FROM      BOOK_WORDS bw1
                          GROUP BY  bw1.word)
```

Discussion about Sub-queries

* Where can sub-queries occur?

    * SELECT / FROM / WHERE / HAVING

* If I make a subquery S in the ____ (one of the above) clause,

    * Where can I access S?

    * What properties must S have?