COSC 4820

Relational Algebra

Kim Buckner

University of Wyoming

Jan. 20, 2023

Algebraic Query Language

Consists of:

- Operators
- Atomic operands
- Can build complex expressions
- Use parentheses for disambiguation
- This gets us thinking in the right direction

Operands & Operations

- Variables that stand for relations
- Constants
- Operations
  - Set
  - Removal of part of a relation
  - Combine relations
  - Renaming

Set operations (1)

Some conditions on set operations are:

- Sets must have matching schema
  - Have the same number of attributes
  - Same types
- Columns must be ordered the same way

Set operations (2)

- UNION
  - Like 'normal' set Union
  - Result is "distinct" (no duplicates in result)
- INTERSECTION
  - Like 'normal' set Intersection
  - Result only contains elements that exist it both sets.

Set operations (3)

- DIFFERENCE
  - $A - B$
  - Result is all elements of A not contained in B

Remove parts (1)

- Projection
  - Display a subset of columns from a relation
  - No change in the number of rows
- Selection
  - Display a subset of rows from a relation
  - No change in the number of columns
  - Selection based on a conditional expression

Remove parts (2)

- Conditional expressions
  - Standard operators: $\geq, =, \leq, <, >, \cdots$
  - Expressions combined using AND, OR, NOT

Combine operations (1)

- CARTESIAN PRODUCT
  - $A \times B$
  - Each tuple of A matched with each tuple of B
  - The relations do **not** have to have to matching schema
  - Common names have to be disambiguated.

Combine operations (2)

- NATURAL JOIN
  - $A \bowtie B$
  - Rows selected if attributes in common match
    - Matched rows sometimes called *joined tuple*
    - Unmatched rows are then DANGLING TUPLES

Combine operations (3)

- THETA JOIN
  - $A \bowtie_C B$
  - This is CARTESIAN PRODUCT with SELECTION
  - The C in the notation represents the condition
- A tutorial for joins.

Renaming, etc

- Can change the name of a relation
- Can change the name of some/all attributes
- Makes it easier to write complex expressions
- Renaming can be for the duration of the particular expression
- Or it can result in the creation of a 'temporary' table.

Constraints

- Every database has some constraints
- These support consistency
- Can be as simple as "the expression results in an empty set"
- Can be to support *referential integrity*

Referential integrity

- Referential integrity refers to the accuracy and consistency of data within a relationship.
- In relationships, data is linked between two or more tables.
- This is achieved by having the foreign key (in the associated table) reference a primary key value (in the primary - or parent - table).

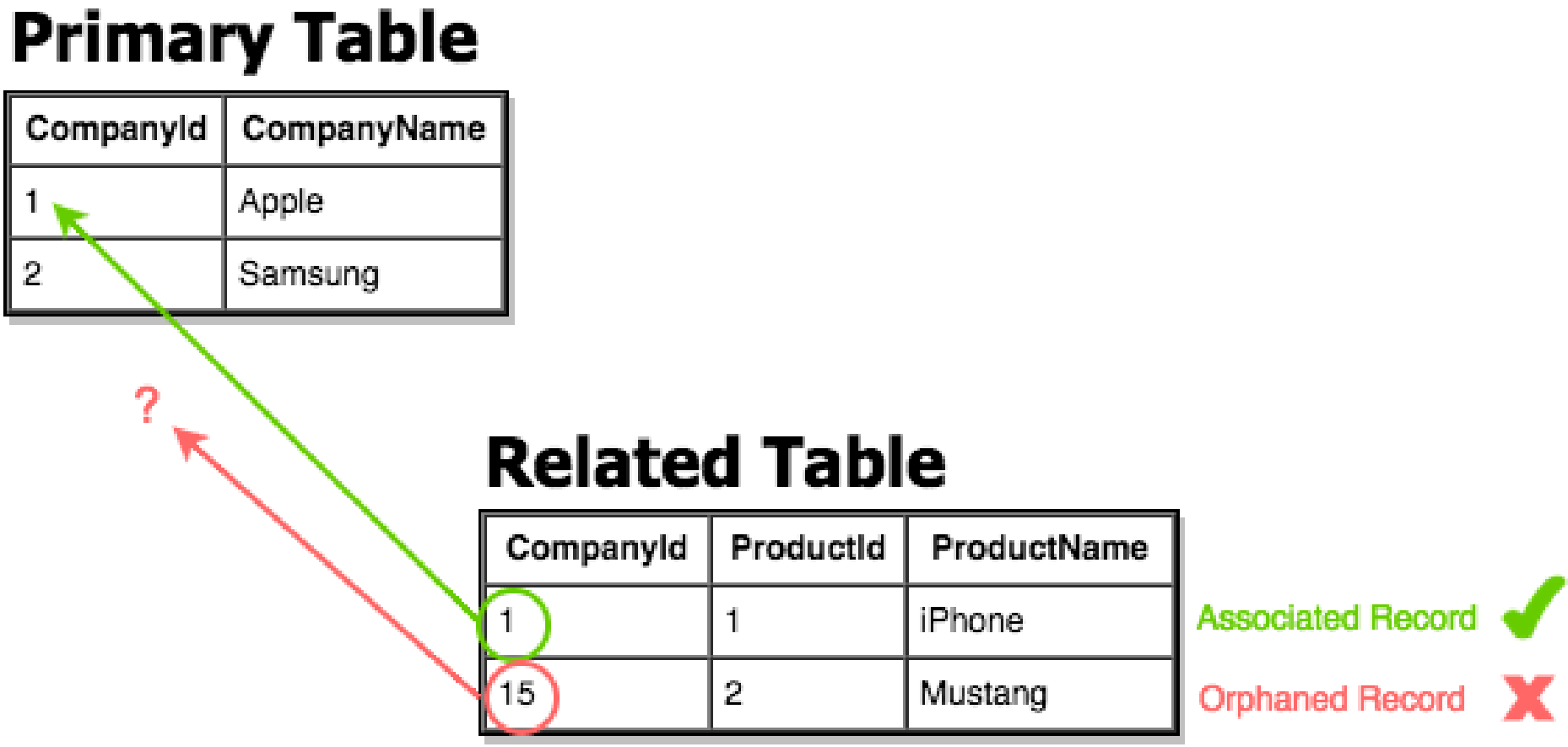
(more . . .)

- Because of this, we need to ensure that data on both sides of the relationship remain intact.
- So, referential integrity requires that, whenever a foreign key value is used it must reference a valid, existing primary key in the parent table.

This means

- Referential integrity will prevent users from:
  - Adding rows to a related table if there is no associated row in the primary table.
  - Changing values in a primary table that result in orphaned records in a related table.
  - Deleting rows from a primary table if there are matching related rows.

(more . . .)


Primary Table
CompanyId
CompanyName
1
Apple
2
Samsung
?
Related Table
CompanyId
ProductId
ProductName
1
1
iPhone
15
2
Mustang
Associated Record
Orphaned Record

Assignment

- This is easy, just read all of Chapter 2 in the text if you have not already done so.