Topics

- Relational Model

- Relational Algebra (Foundations of SQL)
  - Select, Project, Set Operators, Rename
  - Natural Join, Theta Join
  - Outer Join
  - Rename
  - Grouping, Aggregation
  - “Write” Operations: Insert, Delete, Update
Get Some Data Out of this DB!

- **Relational Algebra** (Chap 6.1 in the book)
  - Defines the syntax and semantics for operations performed over relations
  - Provides the theoretical foundation for SQL
  - *A "closed" language: Result of every expression is a valid relation*

- For this lecture, assume that $R(A_1, \ldots, A_n)$ is a relation containing $n$ attributes, $n \geq 1$
Select (Primitive)

- Syntax: $\sigma_c(R)$

- Given relation $R$ and boolean condition $c$
  - Returns an unnamed relation $R'$ that contains only the tuples in $R$ that satisfy $c$
  - The condition can be a combination of predicates using:
    - AND $\land$
    - OR $\lor$
    - NOT $\neg$

- Processing: For each tuple in $R$, evaluate $c$.
  - Retain tuple if true, eliminate if false.
Return planes that hold fewer than 400 people and achieve 600+ MPH

• One solution:

\[ \sigma_{\text{mph}>600}(\sigma_{\text{capacity}<400}(\text{plane})) \]

• Another:

\[ \sigma_{\text{capacity}<400}(\sigma_{\text{mph}>600}(\text{plane})) \]
Honing in on Results

- To reduce data, I don't always need all the attributes.

  - For example, \( \sigma_{\text{capacity} < 400 \land \text{mph} > 600}(\text{plane}) \) returns:

    | \text{tail no} | \text{make}            | \text{model} | \text{capacity} | \text{mph} |
    |----------------|------------------------|--------------|----------------|----------|
    | 3              | McDonnel Douglas       | DC10        | 380            | 610      |

  - But what if the query only asked for the Make and Model of such a jet?

    | \text{make}          | \text{model} |
    |----------------------|--------------|
    | McDonnel Douglas     | DC10         |
Project (Primitive)

- Syntax: $\Pi_{A_1,\ldots,A_k}(R)$

- Meaning: Given a relation $R$:
  - Returns an unnamed relation $R'$ containing only attributes $A_1,\ldots,A_k$
    - Eliminates unspecified columns from result
  - The order of the attributes matters!

- Example: Return the last names of all passengers.

\[
\pi_{\text{name}}(\text{passenger}) \quad //\text{follow up: what gets returned?}
\]
Return the make and model of all jets that holds less than 400 passengers.

\[ \pi_{\text{make, model}}(\sigma_{\text{capacity}<400 \land \text{mph}>600}(\text{plane})) \]

<table>
<thead>
<tr>
<th>make</th>
<th>model</th>
</tr>
</thead>
<tbody>
<tr>
<td>McDonnel Douglas</td>
<td>DC10</td>
</tr>
</tbody>
</table>

Flip the operators. Is this expression equivalent to the above?

- Why (or why not)?

\[ \sigma_{\text{capacity}<400 \land \text{mph}>600}(\pi_{\text{make, model}}(\text{plane})) \]
Relational Algebra Expression Trees

- In regular algebra, \(7/(1-2) + 9\) can be represented a binary tree:
  - Nodes are either operands or operators
  - Bottom-up execution
- Useful to visualizing and determining order of a large expression
Expression Trees (Cont.)

- Same with relational algebra:
  - Also two types of nodes in a binary tree:
    - Relational operators and relations as operands

- Example: $\pi_{make,model}(\sigma_{capacity<400 \land mph>600}(plane))$

- Is represented by:
Set Operators: Union, Difference, Intersection

- Union of relations $R_1 \cup R_2$
- Intersection of relations $R_1 \cap R_2$
- Difference of relations $R_1 \setminus R_2$ (or $R_1 - R_2$)

Example: $R(W, X)$ and $S(Y, Z)$

- $R = \{ ('a',4), ('b',5) \}$ and $S = \{ ('b',5), ('c',6), ('d',7) \}$
- $R \setminus S = ?$
- $R \setminus \{ ('a',2,2) \} = ???$
- $R \cup \{ (3, 'f') \} = ???$
Compatibility of Set Operators

- **Caveat:** To carry out any of these, \( R_1 \) and \( R_2 \) must be *compatible*.

\[
R_1 \cup R_2 \\
R_1 \cap R_2 \\
R_1 \setminus R_2
\]

- **Given two relations** \( R_1(A_1, \ldots, A_n) \) and \( R_2(B_1, \ldots, B_n) \), \( R_1 \) and \( R_2 \) are *compatible* if both hold:

1. \( R_1 \) and \( R_2 \) have the same number of attributes, and
2. for all \( i (1 \leq i \leq n) \): \( \text{domain}(A_i) = \text{domain}(B_i) \)
Administrivia 9/11

- Hwk 1 due Friday
- Project 1 posted (9/20)

Important Team information:

- For all teams:
  - Username: dbteam, Password: qwe123 ←—you should change this after first login

- Teams and IP addresses to servers:
  - Brody, Lukas, Aaron (129.114.104.171)
  - Maddy, Ethan, Ricardo (149.165.157.244)
  - Sarah M, Sarah W, Lia (149.165.168.159)
  - Katrina, Daniel, Matt C. (149.165.168.195)
  - Montana, Ashton, Jiman (149.165.168.9)
  - Spencer, Skye, Olivia, Ana-Lea (129.114.104.163)