Query Rewriting
(for Query "Optimization")

• Main Strategy: Make intermediate results small by applying selection and projection early.
• Additional Strategies: Remove unnecessary operations, execute common subexpressions only once, ...
Rewriting Rules 1 & 2

1. \( \pi_X e = e \)
   Note: the scheme of \( e \) is \( X \).
   
   \[ \Pi_{\text{RoomNr, Name, NrBeds, Cost}} r = r \]

2. \( \pi_X \sigma_f e = \pi_X \sigma_f \pi_XY e \)
   Note: \( f \) mentions the attrs. of \( Y \).
   
   \[ \Pi_{\text{RoomNr, Name}} \sigma_{\text{Cost > 75}} r \]
   \[ = \Pi_{\text{RoomNr, Name}} \sigma_{\text{Cost > 75}} \Pi_{\text{RoomNr, Name, Cost}} r \]
Rewriting Rules 3 & 4

3. \( \sigma_f (e_1 \mid \times \mid e_2) = \sigma_{f_1} (\sigma_{f_2} e_1 \mid \times \mid \sigma_{f_3} e_2) \)
Note: \( f = f_1 \land f_2 \land f_3 \); each pertains to its respective expression.

\[
\sigma_{\text{ArrivalDate} = 15 \text{ May} \land \text{City} = \text{Boston}} (g \mid \times \mid s) \\
= \sigma_{\text{City} = \text{Boston}} g \mid \times \mid \sigma_{\text{ArrivalDate} = 15 \text{ May}} s
\]

4. \( \pi_X (e_1 \mid \times \mid e_2) = \pi_X (\pi_{X_1 Y} e_1 \mid \times \mid \pi_{X_2 Y} e_2) \)
Note: \( X_1 = X \cap \text{scheme}(e_1), \)
\( X_2 = X \cap \text{scheme}(e_2), \) and
\( Y = (\text{scheme}(e_1) \cap \text{scheme}(e_2)) - X. \)

\[
\pi_{\text{RoomNr}} (r \mid \times \mid g) \\
= \pi_{\text{RoomNr}} (\pi_{\text{RoomNr}}, \text{Name} r \mid \times \mid \pi_{\text{Name}} g)
\]

QueryRewriting: 3
Rewriting Rules 5 & 6

5. $\pi_X \pi_Y e = \pi_X e$

$\Pi_{\text{Name}} \Pi_{\text{RoomNr, Name}} r = \Pi_{\text{Name}} r$

6. $\sigma_{f_1} \sigma_{f_2} e = \sigma_{f_1 \land f_2} e$

$\sigma_{\text{NrBeds} = 2} \sigma_{\text{Cost} = 80} r = \sigma_{\text{NrBeds} = 2 \land \text{Cost} = 80} r$
Rewriting Rules 7, 8 & 9

Natural join $\times$ is (7) commutative, (8) associative, and (9) idempotent.

\[
(r \times (g \times s)) \times r \\
=_{8} r \times g \times s \times r \\
=_{7} r \times r \times g \times s \\
=_{9} r \times g \times s
\]
Query Rewriting – Example

$$\Pi_{\text{Name}, \text{StreetNr}, \text{City}} \sigma_{\text{ArrivalDate} = 10 \text{ May}} (\sigma_{\text{NrBeds} = 2} (r \times g \times s))$$

$$=_{3,6,8} \Pi_{\text{Name}, \text{StreetNr}, \text{City}} (\sigma_{\text{NrBeds} = 2} r \times g \times \sigma_{\text{ArrivalDate} = 10 \text{ May}} s)$$

$$=_{4,8} \Pi_{\text{Name}, \text{StreetNr}, \text{City}} (\Pi_{\text{RoomNr}, \text{Name}} \sigma_{\text{NrBeds} = 2} r \times \Pi_{\text{GuestNr, Name, StreetNr, City}} g \times \Pi_{\text{GuestNr, RoomNr}} \sigma_{\text{ArrivalDate} = 10 \text{ May}} s)$$

$$=_{1} \Pi_{\text{Name}, \text{StreetNr}, \text{City}} (\Pi_{\text{RoomNr, Name}} \sigma_{\text{NrBeds} = 2} r \times g \times \Pi_{\text{GuestNr, RoomNr}} \sigma_{\text{ArrivalDate} = 10 \text{ May}} s)$$