

IAT-1

Q3)

B)

Grouping and Aggregation : $\gamma_x(R)$

① Given a relation R , partition its tuples according to their values in on set of attributes G

② The set G is called the grouping attributes

③ Then, for each group, aggregate the values in certain other attributes, aggregation functions :
Sum, COUNT, AVG, MIN, MAX, ...

④ In the notation, x is a list of elements that can be :

- A grouping attribute

- An expression $\sigma(A)$, where σ is one of the five aggregation functions and A is an attribute NOT among the grouping attributes.

⑤ Map: For a tuple (a, b, c) emit a key/value pair (a, b)

Reduce: Each key a represents a group, with values $[b_1, b_2, \dots, b_n]$

Apply σ to the list $[b_1, b_2, \dots, b_n]$

Emit the key/value pair (a, x)

where $x = \sigma([b_1, b_2, \dots, b_n])$

⑥ Set operations include union, Intersection and difference.

Union

map: For each tuple t in R or S , emit a key/value pair (t, t)

Reduce: For each key t , emit a key/value pair (t, t)

⑦ Intersections

Map: For each tuple t in R, S emit a key/value pair (t, t)

Reduce: If key t has value list $[t, t]$
emit a key/value pair (t, t)
otherwise, emit a key/value pair $(t, NULL)$

⑧ Difference

Map: For a tuple t in R emit a key/value pair $(t, 'R')$

For a tuple t in S , emit a key/value pair $(t, 'S')$

Reduce: If key t has value list $[R]$,
emit a key/value pair (t, t)
otherwise, emit a key/value pair $(t, NULL)$
i.e. $[R, 'S']$ or $['S', 'R']$ or $['S']$.