

Esempio 3: Ricerca del Max/Min in una colonna, in Algebra Relazionale

Trovare il valore Massimo fra i valori dell'attributo Numero della relazione NUMERI.

NUMERI	
Numero	
3	
1	
10	
5	

Da una prima analisi visiva, è immediato individuare il valore 10, contenuto nella tupla in posizione 3, come valore MAX per l'attributo Numero. Abbiamo scelto il 10 perché, per definizione di MAX, per il 10 valgono:

- $10 \geq 3$; $10 \geq 1$; $10 \geq 10$; $10 \geq 5$

Abbiamo cioè confrontato il 10 con ogni altro valore nella colonna e ci siamo accorti che è sempre \geq di tutti gli altri. Come esprimere questa condizione in Algebra Relazionale? Sarebbe necessario definire un operatore universale che imponga “Per ogni valore **a** dell'attributo Numero nella relazione NUMERI verifica che sia $\mathbf{a} \geq \mathbf{b}$, per ogni (\forall) valore **b** dell'attributo Numero”. Tuttavia, abbiamo detto che un operatore universale non è definibile in Algebra Relazionale.

Possiamo allora notare che affinché la condizione

“Per ogni valore **a** dell'attributo Numero nella relazione NUMERI verifica che sia $\mathbf{a} \geq \mathbf{b}$, per ogni (\forall) valore **b** dell'attributo Numero”

sia **falsa** basta che

“Esiste (\exists) almeno un valore **b** dell'attributo Numero tale che $\mathbf{a} \geq \mathbf{b}$ è falsa (cioè, $\mathbf{a} < \mathbf{b}$ è vera)”.

Ragionando quindi al contrario, proviamo a cercare tutti quei numeri che hanno almeno un numero più grande di loro, e che non possono pertanto essere considerati il MAX. Giungiamo allora ad un altro problema: come confrontare fra loro in Algebra Relazionale gli elementi di una colonna?

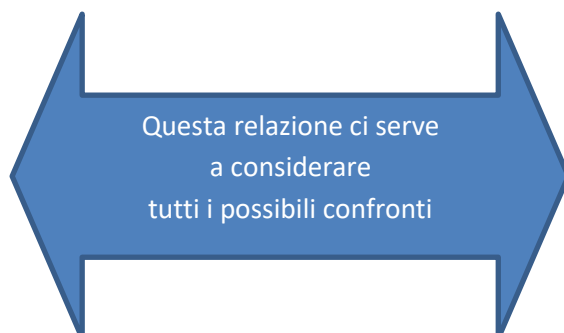
Ricordandoci che in Algebra Relazionale possiamo applicare condizioni sulle righe tramite l'operatore SELEZIONE, abbiamo bisogno di creare una tabella in cui, sulle stesse righe, avremo il primo elemento confrontato col secondo, poi il primo confrontato col terzo, poi il primo confrontato col quarto, ... poi il secondo confrontato col terzo, poi il secondo confrontato col quarto, ... e così via.

Dobbiamo cioè generare, su due colonne, tutte le possibili combinazioni degli elementi di una colonna per se stessa. Sappiamo che questa operazione equivale al prodotto cartesiano della colonna con se stessa.

E' importante notare che la JOIN di una relazione per se stessa, nel nostro caso $\text{NUMERI} \bowtie \text{NUMERI}$, restituirebbe in output la relazione NUMERI stessa. Per ottenere il prodotto cartesiano è necessario che gli schemi siano totalmente diversi (l'intersezione sia vuota), pertanto è necessario ridenominare gli attributi su cui vogliamo effettuare i confronti.

NUMERI \bowtie ($\rho_{\text{Numero2} \leftarrow \text{Numero}}$ NUMERI)

Numero	Numero2
3	3
3	1
3	10
3	5
1	3
1	1
1	10
1	5
10	3
10	1
10	10
10	5
5	3
5	1
5	10
5	5



3 < 3?
3 < 1?
3 < 10?
3 < 5?
1 < 3?
1 < 1?
1 < 10?
1 < 5?
10 < 3?
10 < 1?
10 < 10?
10 < 5?
5 < 3?
5 < 1?
5 < 10?
5 < 5?

Possiamo quindi procedere selezionando tutte le tuple in cui Numero < Numero2 e, poiché siamo interessati solo a Numero e non ci servono doppioni, procedere con una PROIEZIONE su Numero:

$\pi_{\text{Numero}}(\sigma_{\text{Numero} < \text{Numero2}}(\text{NUMERI} \bowtie (\rho_{\text{Numero2} \leftarrow \text{Numero}} \text{NUMERI})))$

Numero
3
1
5

Questi sono tutti i valori che NON sono MAX. L'unico valore escluso quale sarà? Proprio il MAX. Procediamo quindi applicando una differenza fra i valori della relazione NUMERO e quelli ottenuti nella precedente query. Notiamo inoltre che la differenza è possibile perché hanno lo stesso schema.

$\text{NUMERI} - \pi_{\text{Numero}}(\sigma_{\text{Numero} < \text{Numero2}}(\text{NUMERI} \bowtie (\rho_{\text{Numero2} \leftarrow \text{Numero}} \text{NUMERI})))$

Numero
10

Notiamo inoltre che $\sigma_{\text{Numero} < \text{Numero2}}(\text{NUMERI} \bowtie (\rho_{\text{Numero2} \leftarrow \text{Numero}} \text{NUMERI}))$ è una query equivalente a $\text{NUMERI} \bowtie_{\text{Numero} < \text{Numero2}} (\rho_{\text{Numero2} \leftarrow \text{Numero}} \text{NUMERI})$, dove impieghiamo il THETA-JOIN.

Per quanto riguarda la ricerca del MIN, si procede in maniera analoga, prestando attenzione a modificare la formule proposizionale utilizzata nella SELEZIONE in modo appropriato. Si lascia allo studente la dimostrazione della ricerca del MIN.

Nella pratica, difficilmente si avranno relazioni con un solo attributo. Una possibile soluzione potrebbe essere quella di effettuare una PROIEZIONE sull'attributo di interesse, ricercare l'elemento MAX o MIN, e sfruttare il risultato ottenuto tramite un'ulteriore JOIN sulla relazione di partenza.