

Prodotto cartesiano e theta join

L'operazione di **prodotto cartesiano** tra due relazioni R e S è la relazione ottenuta componendo tutte le righe di R con tutte le righe di S . Il prodotto cartesiano di R e S si indica con: $R \times S$. Per esempio le seguenti tabelle $Rep1$ e $Rep2$ rappresentano i dipendenti di due reparti con i rispettivi stipendi, mentre la terza tabella è il prodotto cartesiano $Rep1 \times Rep2$.

Rep1

Dip	Stip
Rossi	1800,00
Rossoni	1456,00
Vescovi	1678,00

Rep2

Dip	Stip
Nassini	1734,00
Peroni	1876,00
Mangini	1765,00
Proietti	1546,00
Sicurella	1678,00

Rep1 × Rep2

Rep1.Dip	Rep1.Stip	Rep2.Dip	Rep2.Stip
Rossi	1800,00	Nassini	1734,00
Rossoni	1456,00	Nassini	1734,00
Vescovi	1678,00	Nassini	1734,00
Rossi	1800,00	Peroni	1876,00
Rossoni	1456,00	Peroni	1876,00
Vescovi	1678,00	Peroni	1876,00
Rossi	1800,00	Mangini	1765,00
Rossoni	1456,00	Mangini	1765,00
Vescovi	1678,00	Mangini	1765,00
Rossi	1800,00	Proietti	1546,00
Rossoni	1456,00	Proietti	1546,00
Vescovi	1678,00	Proietti	1546,00
Rossi	1800,00	Sicurella	1678,00
Rossoni	1456,00	Sicurella	1678,00
Vescovi	1678,00	Sicurella	1678,00

Nella quasi totalità dei casi il prodotto cartesiano di tabelle produce una relazione priva di ogni valore informativo. Potrebbe invece essere utile una forma di congiunzione che permetta di scegliere tra le righe prodotte dal prodotto cartesiano. Questa forma di congiunzione prende il nome di *theta join*.

Il **theta join** tra due relazioni R e S è la relazione ottenuta componendo tutte le righe di R con tutte le righe di S e includendo nella congiunzione le sole righe che soddisfano a una determinata condizione. Il *theta join* di R e S si indica con: $R \bowtie_P S$, dove P rappresenta la condizione.

È utile osservare che il *theta join* non è altro che il prodotto cartesiano seguito da una selezione:

$$R \bowtie_P S \equiv \sigma_P (R \times S)$$

Per esempio, la tabella seguente è ottenuta impostando la condizione che lo stipendio dei dipendenti del primo reparto sia maggiore dello stipendio dei dipendenti del secondo reparto.

Join con Rep1.Stip>Rep2.Stip

Rep1.Dip	Rep1.Stip	Rep2.Dip	Rep2.Stip
Rossi	1800,00	Sicurella	1678,00
Rossi	1800,00	Proietti	1546,00
Rossi	1800,00	Mangini	1765,00
Rossi	1800,00	Nassini	1734,00
Vescovi	1678,00	Proietti	1546,00

Proviamo ora a porre la condizione $Rep1.Stip = Rep2.Stip$, in sostanza vogliamo ottenere l'elenco dei dipendenti che hanno lo stesso stipendio nei due reparti. La tabella che si ottiene è la seguente:

Join con Rep1.Stip=Rep2.Stip

Rep1.Dip	Rep1.Stip	Rep2.Dip	Rep2.Stip
Vescovi	1678,00	Sicurella	1678,00

Per altra strada siamo arrivati a costruire l'**equi-join**. A questo punto, pensare all'eliminazione delle colonne duplicate, e stabilire un qualche automatismo perché i campi da congiungere in due tabelle siano quelli con il medesimo nome, è ovvio, e così "naturale", da indicare con il nome di *join naturale* l'operazione.

La discussione fatta non è così astratta come potrebbe sembrare, in quanto il modo con il quale si costruiscono le diverse congiunzioni nel linguaggio SQL segue proprio questa strada. Si effettua il prodotto cartesiano delle tabelle da congiungere e, sulle righe così ottenute, si operano opportune selezioni e proiezioni.