La division d'algèbre relationnelle

La *division* (ou le quotient) d'une relation R de schéma $R(A_1, \ldots, A_k)$ par la relation S de schéma $S(A_{p+1}, \ldots, A_k)$ est la relation T de schéma $T(A_1, \ldots, A_p)$ formée de tous les tutles t tels que $\{t\} \times S \subseteq R$. C'est à dire de tous les tutles qui concaténés à chacun des tuples de S donnent un tuple de R. On note $T = R \div S$. Par exemple :

$Voyage$

num − client	ville − dep	ville − arr
$x231$	$Paris$	$Marseille$
$v485$	$Marseille$	$Lyon$
$r871$	$Paris$	$Nantes$
$f872$	$Marseille$	$Avignon$
$f872$	$Paris$	$Nantes$
$r871$	$Paris$	$Marseille$
$f872$	$Caen$	$LeMans$
$x231$	$Paris$	$Nantes$

$Ville$

ville − dep	ville − arr
$Paris$	$Marseille$
$Paris$	$Nantes$

$Voyage \div Ville$:

num − client
$x231$
$r871$

Vous pouvez aussi utiliser en algèbre relationnelle et en `SQL` les opérateurs suivants :

Nom	**Algèbre rel.**	SQL
$Union$	$R(A_1, \ldots, A_k) \cup S(A_1, \ldots, A_k)$	`UNION`
$Intersection$	$R(A_1, \ldots, A_k) \cap S(A_1, \ldots, A_k)$	`INTERSECT`
$Difference$	$R(A_1, \ldots, A_k) - S(A_1, \ldots, A_k)$	`EXCEPT`

Les questions

1. (a) Les identifiants des joueurs qui ont joué en double avec un italien (ils ont alors un même numéro d'enregistrement).

Solution: On commence par trouver tous les doubles en faisant une jointure de `Registration_big` avec elle-même. On pose $\mathtt{R_1} = \mathtt{Registration\_big}$ et $\mathtt{R_2} = \mathtt{Registration\_big}$.

$$A = \pi_{\mathtt{R_1.pid},\mathtt{R_2.pid}} \left(\mathtt{R_1} \bowtie_{\substack{(\mathtt{R_1.pid} \neq \mathtt{R_2.pid} \\ \mathtt{R_1.registrnum} = \mathtt{R_2.registrnum})}} \mathtt{R_2} \right)$$

Puis on fait une jointure avec `Player_big` avec une sélection sur les joueurs italiens.

$$Q_1 = \pi_{\mathtt{R_1.pid}} \left(\mathtt{A} \bowtie_{(\mathtt{R_2.pid=Player\_big.pid})} \sigma_{\mathtt{countrycode='ITA'}} (\mathtt{Player\_big}) \right)$$

(b) Les identifiants des joueurs qui ont joué en double avec tous les italiens.

Solution: On réutilise `A` sauf que cette fois-ci on va utiliser une division plutôt qu'une jointure.

$$Q_2 = A \div \pi_{\mathtt{pid}} \left(\sigma_{\mathtt{countrycode='ITA'}} (\mathtt{Player\_big}) \right)$$

(c) Les identifiants des joueurs qui n'ont jamais joué en double avec un italien.

Solution: Les joueurs qui n'ont jamais joué en double avec un italien sont le complémentaire des joueurs qui ont déjà joué avec un italien (question 1). D'où :

$$Q_3 = \pi_{\mathtt{pid}} (\mathtt{Player\_big}) - Q_1$$

2. (a) Le numéro d'enregistrement des joueurs qui ont participé à "Wimbledon" et à l'"US Open".

Solution: On commence par trouver tous les joueurs qui ont participé à une édition de Wimbledon en faisant une jointure des table `played_in_big` et `tournament_big` après avoir bien entendu sélectionné seulement les tournois qui nous intéressent :

$$W = \pi_{\mathtt{registrnum}} \left(\mathtt{played\_in\_big} \bowtie_{(\mathtt{tid})} \sigma_{\mathtt{name='Wimbledon'}} (\mathtt{tournament\_big}) \right)$$

On trouve même les joueurs qui ont participé à une édition de l'US Open.

$$O = \pi_{\mathtt{registrnum}} \left(\mathtt{played\_in\_big} \bowtie_{(\mathtt{tid})} \sigma_{\mathtt{name=USOpen'}} (\mathtt{tournament\_big}) \right)$$

Finalement, on ne garde que l'instersection de ces deux ensembles :

$$Q_4 = W \cap O$$

(b) Le numéro d'enregistrement des joueurs qui ont participé à "Wimbledon" ou à l'"US Open".

Solution: Cette fois-ci, c'est l'union de W et de O qui nous intéresse :

$$Q_5 = W \cup O$$

(c) Le numéro d'enregistrement des joueurs qui ont participé à "Wimbledon" et jamais à l'"US Open".

Solution: Cette fois-ci, c'est la différence de W et de O qui nous intéresse :

$$Q_6 = W - O$$

3. Les numéros d'enregistrement des joueurs qui n'ont jamais gagné.

Solution: Il est facile de trouver les numéros d'identifications des joueurs qui on déjà gagné : c'est le champ `Winner` de la table `MatchResults`. Les joueurs qui n'ont jamais gagné sont donc ceux qui n'apparaissent pas dans ce champ. D'où :

$$Q_7 = \pi_{\texttt{registrnum}}(\texttt{Registration\_big}) - \pi_{\texttt{Winner}}(\texttt{MatchResults})$$

4. (a) Les numéros d'enregistrement des joueurs qui ont gagné un match à "Wimbledon".

Solution: Il faut encore utiliser le champ `Winner` de `MatchResults` mais cette fois-ci après avoir sélectionner seulement les matchs qui ont eu lieu à Wimbledon. Ceux-ci peuvent être calculés par la requête :

$$M_w = \pi_{\texttt{mid}}\left(\texttt{games\_big} \bowtie_{(\texttt{tid})} \sigma_{\texttt{name}='\texttt{Wimbledon}'}(\texttt{tournament\_big})\right)$$

Et donc :

$$Q_8 = \pi_{\texttt{Winner}}\left(\texttt{MatchResults} \bowtie_{(\texttt{mid})} M_w\right)$$

(b) Les numéros d'enregistrement des joueurs qui n'ont jamais gagné un match à "Wimbledon".

Solution: On a les joueurs qui ont gagné un match à Wimbledon, il est donc facile de trouver ceux qui n'en ont jamais gagné un avec une différence.

$$Q_9 = \pi_{\texttt{registrnum}}(\texttt{Registration\_big}) - M_w$$

(c) Les joueurs qui ont gagné tout les matchs qu'ils ont joués à "Wimbledon".

Solution: On commence par associer à chaque joueur ses numéros d'enregistrement et les identifiants des matchs de Wimbledon auquel il a participé. Pour cela, on utilise une grosse jointure :

$$R_w = \pi_{\texttt{pid,registrnum,mid}}(\texttt{player\_big} \bowtie \texttt{registration\_big}$$
$$\bowtie_{\substack{registrnum=registrnum1\\ \vee registrnum=registrnum2}} \texttt{game\_big} \bowtie \sigma_{\texttt{name}='\texttt{Wimbledon}'}(\texttt{tournament\_big}))$$

On effectue ensuite une division entre R_w et `MatchResults` :

$$Q_{10} = R_w \div \pi_{\texttt{winner,mid}}(\texttt{MatchResults})$$

(d) Les joueurs qui ont gagné "Wimbledon", c'est-à-dire qu'ils ont gagné tout les matchs qu'ils ont joués à partir du deuxième round, pour une édition donnée.

Solution: On adapte l'idée ci-dessus, mais il faut cette fois-ci conserver un peu plus d'information dans R_w. Tout d'abord, on ne veut garder que les matchs qui se déroulent à partir du deuxième round. De plus, dans la division que nous allons effectuer, on veut tenir compte de l'édition de Wimbledon. Il va donc falloir conserver `tid`. Soit :

$$R'_w = \pi_{\texttt{pid,registrnum,mid,tid}}(\texttt{player\_big} \bowtie \texttt{registration\_big} \bowtie_{\substack{registrnum=registrnum1 \\ \vee registrnum=registrnum2}} \sigma_{\texttt{round}>=2}\left(\texttt{game\_big}\right) \bowtie \sigma_{\texttt{name}='\texttt{Wimbledon}'}\left(\texttt{tournament\_big}\right))$$

On effectue encore la division de R'_w avec `MatchResults`, mais il faut à nouveau prendre `tid` en compte. D'où

$$Q_{11} = R'_w \div \pi_{\texttt{winner,mid,tid}}\left(\texttt{MatchResults} \bowtie_{(\texttt{mid})} \texttt{game}\right)$$