

Calcul relationnel

Le calcul relationnel à variable n-uplet repose sur une idée simple : une requête correspond à la définition d'un *ensemble* de n-uplets satisfaisant une condition logique. Une expression typique s'écrit :

$$\{ t.A \mid P(t) \}$$

Cette écriture représente l'ensemble de toutes les valeurs $t.A$ telles que la variable d'uplet t vérifie le prédicat $P(t)$.

Ensemble

Un ensemble désigne une collection d'éléments sans ordre particulier et sans doublon. Dans le cadre du modèle relationnel, ces éléments sont des n-uplets : chaque n-uplet correspond à un enregistrement complet dans une relation.

Ainsi, lorsqu'une requête retourne un résultat, ce résultat est toujours un ensemble d'éléments vérifiant une condition, jamais une liste ordonnée ni indexée.

Projection

La projection consiste à extraire seulement certains attributs d'un n-uplet. Si t représente un n-uplet d'une relation, alors $t.A$ désigne la valeur de l'attribut ou du groupe d'attributs A dans ce n-uplet.

La projection permet donc de « réduire » les informations retournées : on garde uniquement les colonnes souhaitées tout en respectant le principe d'ensemble (pas de doublon, absence d'ordre).

Exemple de relations

ID	Nom	Type
1	Triple Karmeliet	Blonde
2	Guinness	Stout
3	Chimay Bleue	Trappiste

Table 1: Bières

ID	Nom	Ville
1	Le Houblon d'Or	Lille
2	La Chope Sacrée	Paris
3	The Black Barrel	Dublin

Table 2: Bars

ID_Bar	ID_Biere
1	1
1	3
2	1
2	2
3	2

Table 3: Relation Bar–Bière

Interprétation générale

Dans le calcul relationnel, une variable d'uplet représente un enregistrement complet d'une relation. Cette variable peut correspondre :

- à un n-uplet d'une seule relation (ex. : un seul bar),
- ou à la concaténation de plusieurs n-uplets issus de relations différentes (ex. : un bar combiné avec une bière donnée).

Lorsque plusieurs relations interviennent, les variables sont posées en parallèle. Cela revient, sur le plan mathématique, à considérer le *produit cartésien* des relations impliquées. Chaque combinaison possible d'uplets constitue alors un élément du produit cartésien, et donc une affectation potentielle pour les variables.

Le prédicat $P(t)$ sert ensuite à filtrer ce produit cartésien : on élimine toutes les combinaisons incohérentes pour ne conserver que celles satisfaisant les contraintes logiques.

Exemple d'interprétation

Pour comprendre progressivement le mécanisme du calcul relationnel, commençons par des requêtes ne portant que sur une seule relation.

Exemple 1 : une seule relation

On veut obtenir les noms des bières de type *Blonde*. On ne travaille qu'avec la relation **Bieres**. On introduit une seule variable d'uplet :

- t pour la relation **Bieres**.

La requête s'écrit :

$$\left\{ t.Nom \mid \begin{array}{l} t \in \text{Bieres} \\ \wedge t.Type = \text{Blonde} \end{array} \right\}$$

Ici, aucune jointure : on parcourt simplement tous les n-uplets de **Bieres** et on filtre sur un attribut.

Résultat :

$$\{\text{Triple Karmeliet}\}$$

Exemple 2 : introduction d'une deuxième relation

On souhaite maintenant obtenir les bières servies par un bar donné. Il faut alors combiner deux relations : **Bars** et **BarBiere**. On introduit deux variables :

- b pour **Bars**,
- x pour **BarBiere**.

Objectif : récupérer les identifiants des bières servies par le bar *Le Houblon d'Or*.

$$\left\{ x.ID_Biere \mid \begin{array}{l} b \in \text{Bars} \\ \wedge x \in \text{BarBiere} \\ \wedge b.ID = x.ID_Bar \\ \wedge b.Nom = \text{Le Houblon d'Or} \end{array} \right\}$$

Résultat :

$$\{1, 3\}$$

Exemple 3 : jointure complète avec trois relations

On veut maintenant les *noms* des bières, ce qui nécessite la troisième relation **Bieres**. On ajoute une troisième variable t .

- b pour **Bars**,
- x pour **BarBiere**,
- t pour **Bieres**.

$$\left\{ t.Nom \left| \begin{array}{l} b \in \text{Bars} \\ \wedge x \in \text{BarBiere} \\ \wedge t \in \text{Bieres} \\ \wedge b.ID = x.ID_Bar \\ \wedge t.ID = x.ID_Biere \\ \wedge b.Nom = \text{Le Houblon d'Or} \end{array} \right. \right\}$$

Résultat :

$$\{\text{Triple Karmeliet, Chimay Bleue}\}$$

Fonctions d'agrégat

Dans le calcul relationnel pur (au sens théorique du modèle relationnel), il n'existe pas d'opérations d'agrégation comme la somme, la moyenne ou le comptage. Le formalisme ne définit que des ensembles et des filtrages logiques.

Toutefois, dans la pratique, on introduit souvent une extension appelée *calcul relationnel avec agrégats*, qui autorise des constructions de la forme :

$$\{ \text{SUM}(t.Montant) \mid t \in \text{Relation} \wedge P(t) \}$$

Supposons maintenant que la relation **BarBiere** contienne un attribut **Prix**. On veut obtenir la somme des prix des bières servies au bar *Le Houblon d'Or*.

On introduit deux variables :

- $b \in \text{Bars}$,
- $x \in \text{BarBiere}$.

La requête devient alors, en utilisant l'extension avec agrégat :

$$\left\{ \text{SUM}(x.Prix} \left| \begin{array}{l} b \in \text{Bars} \\ \wedge x \in \text{BarBiere} \\ \wedge b.ID = x.ID_Bar \\ \wedge b.Nom = \text{Le Houblon d'Or} \end{array} \right. \right\}$$

Cette écriture exprime l'ensemble contenant un seul élément : la somme de tous les prix des bières du bar.

Exercices

Les relations ci-dessous serviront pour l'ensemble des exercices. Elles représentent un championnat sportif simple. Des exemples de tables sont fournis pour aider à visualiser les données.

Relations

- JOUEUR(NJO, EQ, TAILLE, AGE)
- EQUIPE(NEQ, VILLE, COULEUR, STP)
- SPONSORISE(NSP, NJO, SOMME)
- MATCH(EQ1, EQ2, DATE, ST)
- DISTANCE(ST1, ST2, NBKM)

JOUEUR

NJO	EQ	TAILLE	AGE
Messi	PIÉPLA	170	35
Manon	DIREKT	182	29
Ramos	PIÉPLA	185	36
Lia	TÈTOKARÉ	178	25

EQUIPE

NEQ	VILLE	COULEUR	STP
PIÉPLA	Paris	Rouge	ParcA
DIREKT	Lyon	Bleu	GrandL
TÈTOKARÉ	Marseille	Vert	SudM

SPONSORISE

NSP	NJO	SOMME
Adadis	Messi	8000
Adadis	Ramos	6000
Sponz	Manon	9000
Sponz	Lia	5000

MATCH

EQ1	EQ2	DATE	ST
PIÉPLA	DIREKT	2024-02-10	ParcA
DIREKT	TÈTOKARÉ	2024-02-15	GrandL
PIÉPLA	TÈTOKARÉ	2024-03-01	SudM

DISTANCE

ST1	ST2	NBKM
ParcA	GrandL	420
GrandL	SudM	310
ParcA	SudM	600

Sélections simples

1. Donner les joueurs mesurant plus de 180 cm.
2. Donner les joueurs âgés de plus de 30 ans.
3. Donner les noms des joueurs sponsorisés par *Adadis*.
4. Donner les équipes situées à Paris.
5. Donner les équipes dont la couleur est Bleu.
6. Donner les stades situés à plus de 300 km du stade GrandL.
7. Donner les dates des matchs joués dans le stade ParcA.
8. Donner les sponsors ayant versé plus de 7000.
9. Donner les joueurs mesurant plus de 175 cm et ayant moins de 30 ans.
10. Donner les joueurs sponsorisés par *Sponz* et appartenant à l'équipe DIREKT.
11. Donner les joueurs sponsorisés par au moins un sponsor et mesurant plus de 180 cm.
12. Donner les équipes ayant un stade préféré commençant par la lettre P.

Jointures

13. Donner les noms des joueurs ayant un sponsor.
14. Donner les noms des sponsors des joueurs de l'équipe PIÉPLA.
15. Donner les couples (joueur, équipe).
16. Donner les joueurs ayant participé à un match (via leur équipe).
17. Donner les équipes ayant joué un match dans leur propre stade.
18. Donner les joueurs de l'équipe qui a joué le 2024-02-15.
19. Donner les joueurs ayant joué dans le stade GrandL.
20. Donner les stades où ont joué les équipes sponsorisées par *Adadis*.

Auto-jointures

21. Donner les matchs ayant eu lieu le même jour dans des stades différents.

Différences

22. Donner les joueurs non sponsorisés.
23. Donner les équipes n'ayant jamais joué dans le stade SudM.
24. Donner les dates des matchs où PIÉPLA n'a pas joué.
25. Donner les équipes n'ayant rencontré ni PIÉPLA ni DIREKT.