

- Une relation peut être vue comme une table dont les colonnes sont les attributs. Les lignes de la table sont appelées  $n$ -uplets en français, tuples en anglais.
- Algèbre relationnelle : définir des opérations sur les relations.
- Résultats des opérations : des relations.
- Peut paraître assez abstrait (au premier abord), mais constituent la base des requêtes SQL (dans un cours prochain).

- Sélection : choisir un sous-ensemble des  $n$ -uplets (lignes) qui vérifient la condition de sélection. Notation :  $\sigma$
- Sélection peut être vue comme filtrage :
  - Séparation des  $n$ -uplets en 2 paquets : ceux qui vérifient la condition (conservés) et les autres (rejetés).
- Exemples :
  - Films dont le titre est *True Grit* :  $\sigma_{\text{Titre}=\text{'True Grit'}}(\text{FILM})$

FilmId	Titre	Année
1001	True Grit	1969
1020	True Grit	2010

$\sigma_{\text{Titre}=\text{'True Grit'}}(\text{FILM})$

- Films sortis avant 2000 :  $\sigma_{\text{Année} < 2000}(\text{FILM})$

FilmId	Titre	Année
1001	True Grit	1969
1005	The Big Lebowski	1998

$\sigma_{\text{Année} < 2000}(\text{FILM})$

- Notation : sélection sur une relation R :  
 $\sigma_{\langle \text{condition de sélection} \rangle}(R)$
- Condition formée d'un nombre quelconque de clauses de la forme  
    <nom attribut> <op. comparaison> <valeur constante>  
    ou <nom attribut> <op. comparaison> <nom attribut>  
    où <op. comparaison> est un des opérateurs =, <, ≤, >, ≥, ≠.
- Les clauses peuvent être reliées par les opérateurs booléens et, ou, non (négation ¬).
- Exemple : sélectionner les personnes dont le numéro est 1 ou nées dans les années 1950 :

No	Nom	AnnéeNaiss
1	Jeff Bridges	1949
2	Ethan Coen	1954
3	Joel Coen	1957
4	John Goodman	1952

$\sigma_{\text{No} = 1 \text{ ou } (\text{AnnéeNaiss} \geq 1950 \text{ et } \text{AnnéeNaiss} \leq 1959)}(\text{PERSONNE})$

## Sélection – propriétés

### Propriétés

- Le résultat d'une sélection à partir d'une relation  $R$  est une relation de même degré (nombre d'attributs) que  $R$ .
- Le cardinal (nombres de  $n$ -uplets ou lignes) d'une sélection à partir d'une relation  $R$  est inférieur ou égal au cardinal de  $R$ .
- L'opération de sélection est commutative :

$$\sigma_{\langle \text{cond1} \rangle} (\sigma_{\langle \text{cond2} \rangle} (R)) = \sigma_{\langle \text{cond2} \rangle} (\sigma_{\langle \text{cond1} \rangle} (R)).$$

En effet,  $\sigma_{\langle \text{cond1} \rangle} (\sigma_{\langle \text{cond2} \rangle} (R)) = \sigma_{\langle \text{cond1} \rangle}$  et  $\langle \text{cond2} \rangle (R)$ .

- Projection : sélectionner certaines colonnes de la table (et jeter les autres). Notation :  $\pi$ .
- Projection utilisée si seuls certains attributs (colonnes) nous intéressent.
- Exemples : titres, années des films, puis titres des films

Titre	Année
True Grit	1969
The Big Lebowski	1998
True Grit	2010
Parasite	2019

$\pi_{\text{Titre, Année}}(\text{FILM})$

Titre
True Grit
The Big Lebowski
Parasite

$\pi_{\text{Titre}}(\text{FILM})$

- Les répétitions sont éliminées.

## Projection – cas général

- Notation : projection pour une relation  $R$  :  $\pi_{\langle \text{liste d'attributs} \rangle}(R)$
- Il est possible que des  $n$ -uplets (lignes) identiques apparaissent. Mais la projection supprime les  $n$ -uplets identiques.
- Le résultat de l'opération de projection est donc un ensemble d'uplets distincts deux à deux (et donc une relation valide).

### Propriétés

- Le cardinal (nombres de  $n$ -uplets ou lignes) d'une projection à partir d'une relation  $R$  est inférieur ou égal au cardinal de  $R$ .
- La projection est idempotente :

$$\pi_{\langle \text{cond} \rangle}(\pi_{\langle \text{cond} \rangle}(R)) = \pi_{\langle \text{cond} \rangle}(R).$$

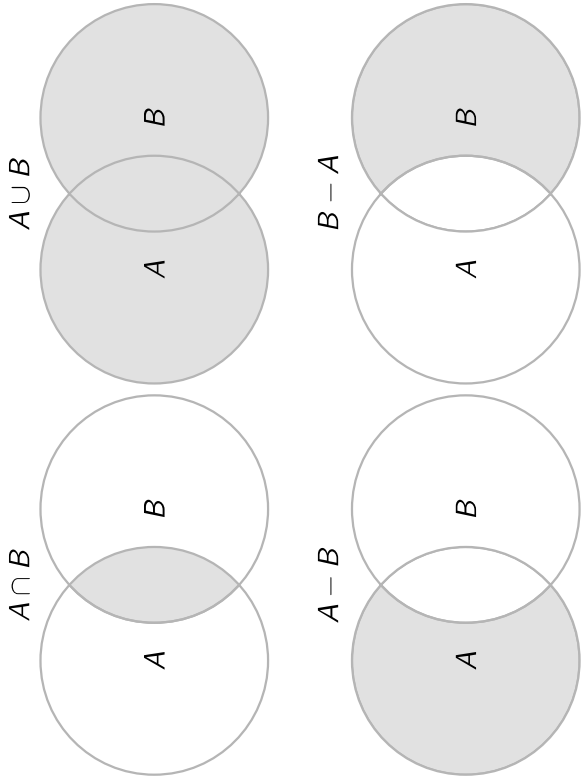
# Renommage – définition

- Le renommage consiste à changer le nom d'une colonne (un attribut).
- Exemple : renommer l'attribut Année en Sortie

FilmNo	Titre	Sortie
1001	True Grit	1969
1005	The Big Lebowski	1998
1020	True Grit	2010
1030	Parasite	2019

/Année/Sortie (FILM)

- On peut utiliser les opérations standard de la théorie des ensembles pour manipuler les relations.





## Union, intersection, différence

- Les relations  $R(A_1, \dots, A_n)$  et  $S(B_1, \dots, B_m)$  sont dites compatibles si les deux conditions suivantes sont vérifiées :
  - elles ont même degré (même nombre d'attributs, ou encore de colonnes, ou encore  $n = m$ )
  - pour tout  $i$ ,  $1 \leq i \leq n$ ,  $A_i$  et  $B_i$  ont même domaine.
- Pour deux relations  $R$  et  $S$  compatibles,
  - l'union de  $R$  et  $S$ , notée  $R \cup S$ , est la relation qui contient les  $n$ -uplets qui sont dans  $R$ , dans  $S$ , et dans les deux à la fois. Les répétitions sont éliminées.
  - L'intersection de  $R$  et  $S$ , notée  $R \cap S$ , est la relation qui contient les  $n$ -uplets qui sont dans  $R$  et dans  $S$ .
  - La différence de  $R$  et  $S$ , notée  $R - S$  est la relation qui contient les  $n$ -uplets qui sont dans  $R$  et pas dans  $S$ .
- Par convention, les noms des attributs sont ceux de la 1<sup>ère</sup> relation. On peut toujours faire un renommage au besoin.

FilmNo	Titre	Année
1001	True Grit	1969
1005	The Big Lebowski	1998
1020	True Grit	2010
1030	Parasite	2019

FILM1

FilmNo	Titre	Année
1001	True Grit	1969
1011	The African Queen	1951
1025	Birdman	2014

FILM2

Les relations FILM1 et FILM2 sont compatibles.

FilmNo	Titre	Année
1001	True Grit	1969

FILM1 ∩ FILM2

FilmNo	Titre	Année
1005	The Big Lebowski	1998
1020	True Grit	2010
1030	Parasite	2019

FILM1 – FILM2

FilmNo	Titre	Année
1001	True Grit	1969
1005	The Big Lebowski	1998
1020	True Grit	2010
1030	Parasite	2019
1011	The African Queen	1951
1025	Birdman	2014

FILM1 ∪ FILM2

FilmNo	Titre	Année
1011	The African Queen	1951
1025	Birdman	2014

FILM2 – FILM1

FilmNo	Titre	Année
1001	True Grit	1969
1005	The Big Lebowski	1998
1020	True Grit	2010
1030	Parasite	2019

No	Nom	AnnéeNaiss
1	Jeff Bridges	1949
2	Ethan Coen	1954
3	Joel Coen	1957
4	John Goodman	1952

Les relations FILM et PERSONNE sont compatibles.

FilmNo	Titre	Année
FILM∩PERSONNE		

FilmNo	Titre	Année
1001	True Grit	1969
1005	The Big Lebowski	1998
1020	True Grit	2010
1030	Parasite	2019
1	Jeff Bridges	1949
2	Ethan Coen	1954
3	Joel Coen	1957
4	John Goodman	1952

FILM ∩ PERSONNE

FilmNo	Titre	Année
1001	True Grit	1969
1005	The Big Lebowski	1998
1020	True Grit	2010
1030	Parasite	2019

FILM – PERSONNE

No	Nom	AnnéeNaiss
1	Jeff Bridges	1949
2	Ethan Coen	1954
3	Joel Coen	1957
4	John Goodman	1952

PERSONNE – FILM

## Produit cartésien

- R et S deux relations (pas de compatibilité requise).
- Le produit cartésien de R et S, noté  $R \times S$  est la relation dont les n-uplets (lignes) sont obtenus en combinant tous les n-uplets de R avec ceux de S.
- Exemple :  $FILM \times RÔLE$

FILM.FilmNo	Titre	Année	Rôle	ActeurNo	RÔLE.FilmNo
1001	True Grit	1969	Walter Sobchak	4	1005
1001	True Grit	1969	The Dude	1	1005
1001	True Grit	1969	Rooster Cogburn	8	1001
1001	True Grit	1969	Rooster Cogburn	1	1020
1001	True Grit	1969	Mattie Ross	7	1020
1005	The Big Lebowski	1998	Walter Sobchak	4	1005
1005	The Big Lebowski	1998	The Dude	1	1005
1005	The Big Lebowski	1998	Rooster Cogburn	8	1001
1005	The Big Lebowski	1998	Rooster Cogburn	1	1020
1005	The Big Lebowski	1998	Mattie Ross	7	1020
1020	True Grit	2010	Walter Sobchak	4	1005
1020	True Grit	2010	The Dude	1	1005
1020	True Grit	2010	Rooster Cogburn	8	1001
1020	True Grit	2010	Rooster Cogburn	1	1020
1020	True Grit	2010	Mattie Ross	7	1020
1030	Parasite	2019	Walter Sobchak	4	1005
1030	Parasite	2019	The Dude	1	1005
1030	Parasite	2019	Rooster Cogburn	8	1001
1030	Parasite	2019	Rooster Cogburn	1	1020
1030	Parasite	2019	Mattie Ross	7	1020

## Produit cartésien – cas général

- $R(A_1, A_2, \dots, A_n)$  et  $S(B_1, B_2, \dots, B_m)$  deux relations.
- Alors  $R \times S$  relation de degré  $n + m$  et d'attributs

$$(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m).$$

- $\text{Cardinal}(R \times S) = \text{Cardinal}(R) \times \text{Cardinal}(S)$
- Quel intérêt ? Seul, peu pertinent. Plus utile suivi d'une sélection pour faire correspondre les attributs.

FILM.FilmNo	Titre	Année	Rôle	ActeurNo	RÔLE.FilmNo
1001	True Grit	1969	Rooster Cogburn	8	1001
1005	The Big Lebowski	1998	Walter Sobchak	4	1005
1005	The Big Lebowski	1998	The Dude	1	1005
1020	True Grit	2010	Rooster Cogburn	1	1020
1020	True Grit	2010	Mattie Ross	7	1020

$$\sigma_{\text{FILM.NoFilm} = \text{RÔLE.NoFilm}}(\text{FILM} \times \text{RÔLE})$$

- $\rightsquigarrow$  Films avec les rôles correspondants.
- Produit cartésien puis sélection : opération très courante, peut être vue comme une jointure.

- Jointure, dénotée par  $\bowtie$ , utilisée pour combiner des  $n$ -uplets apparentés de deux relations en un  $n$ -uplet « plus long ».
- Pour obtenir l'ensemble des films avec les rôles.

FILM.FilmId	Titre	Année	Rôle	ActeurId	RÔLE.FilmId
1001	True Grit	1969	Rooster Cogburn	8	1001
1005	The Big Lebowski	1998	Walter Sobchak	4	1005
1005	The Big Lebowski	1998	The Dude	1	1005
1020	True Grit	2010	Rooster Cogburn	1	1020
1020	True Grit	2010	Mattie Ross	7	1020

FILM $\bowtie$ <sub>FILM</sub> FILMId=RÔLE.FILMIdRÔLE



## Jointure : cas général

- Forme générale d'une jointure sur deux relations  $R(A_1, A_2, \dots, A_n)$  et  $S(B_1, B_2, \dots, B_m)$  :  

$$R \bowtie_{\langle \text{condition jointure} \rangle} S$$
- Résultat de cette opération : relation à  $n + m$  attributs  

$$(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m).$$
- Les n-uplets de la jointure sont formés d'uplets de  $R$  concaténés avec des n-uplets de  $S$  pourvu qu'ils vérifient la condition de jointure.
- La condition de jointure est spécifiée sur les attributs de  $R$  et de  $S$ .

## $\theta$ -jointure

Une condition de jointure générale est de la forme  
<condition> et <condition> et ... et <condition>  
où <condition> est de la forme

$$A_i \theta B_j,$$

où  $A_i$  (attribut de  $R$ ) et  $B_j$  (attribut de  $S$ ) ont le même domaine, et  $\theta$  est une des comparaisons  $=, <, \leq, >, \geq, \neq$ .

Une jointure avec une telle condition générale est appelée  $\theta$ -jointure.

- Les jointures avec égalité (=) sont appelées équijointures.
  - Le résultat d'une équijointure comporte deux attributs identiques, l'un provenant de la relation R1, l'autre provenant de R2.
  - $\rightsquigarrow$  jointure naturelle, notation \*
    - il faut que les deux attributs aient le même nom.
    - Exemple : combiner FILM et RÔLE pour obtenir aussi le nom du film correspondant à chaque rôle.
- La jointure naturelle supprime l'un des deux attributs dupliqués.

Filmo	Titre	Année	Rôle	ActeurNo
1001	True Grit	1969	Rooster Cogburn	8
1005	The Big Lebowski	1998	Walter Sobchak	4
1005	The Big Lebowski	1998	The Dude	1
1020	True Grit	2010	Rooster Cogburn	1
1020	True Grit	2010	Mattie Ross	7

FILM \* RÔLE

- Dans la jointure naturelle précédente, le film *Parasite* n'apparaît pas, parce qu'aucun rôle de ce film n'est renseigné dans la relation RÔLE.
- Jointure externe gauche de R et S :
  - uplets de R et de S qui vérifient le critère de jointure sur leurs noms d'attributs communs (i.e. jointure naturelle),
  - ainsi que les uplets de R qui n'ont pas d'uplets correspondants dans S.
- Notation :  $R \bowtie S$ .
- Exemple :  $FILM \bowtie RÔLE$

FilmNo	Titre	Année	Rôle	ActeurNo
1001	True Grit	1969	Rooster Cogburn	8
1005	The Big Lebowski	1998	Walter Sobchak	4
1005	The Big Lebowski	1998	The Dude	1
1020	True Grit	2010	Rooster Cogburn	1
1020	True Grit	2010	Mattie Ross	7
1030	Parasite	2019	NULL	NULL

FILM  $\bowtie$  RÔLE

## Jointures externes droite et totale

- Jointure externe droite de R et S (notation  $R \bowtie S$ ) :
  - uplets de R et de S qui vérifient le critère de jointure sur leurs noms d'attributs communs (i.e. jointure naturelle),
  - ainsi que les uplets de S qui n'ont pas d'uplets correspondants dans R.
- Jointure externe totale de R et S (notation  $R \Join S$ ) :
  - uplets de R et de S qui vérifient le critère de jointure sur leurs noms d'attributs communs (i.e. jointure naturelle),
  - ainsi que les uplets de R qui n'ont pas d'uplets correspondants dans S.
  - ainsi que les uplets de S qui n'ont pas d'uplets correspondants dans R.

- Définissons une relation ANCIENS =  $\sigma_{\text{AnnéeNaiss} < 1955}(\text{PERSONNE})$  et renommons l'attribut RéalNo en No pour préparer la jointure.

No	Nom	AnnéeNaiss	No	FilmNo
1	Jeff Bridges	1949	3	1005
2	Ethan Coen	1954	2	1020
4	John Goodman	1952	3	1020
5	Henry Hathaway	1898	5	1001
8	John Wayne	1907	6	1030

ANCIENS =  $\sigma_{\text{AnnéeNaiss} < 1955}(\text{PERSONNE})$       RÉAL =  $\rho_{\text{RealNo/No}}(\text{RÉALISATION})$

- Jointure externe totale :

No	Nom	AnnéeNaiss	FilmNo
1	Jeff Bridges	1949	NULL
2	Ethan Coen	1954	1020
3	NULL	NULL	1005
3	NULL	NULL	1020
4	John Goodman	1952	NULL
5	Henry Hathaway	1898	1001
6	NULL	NULL	1030
8	John Wayne	1907	NULL

## Combiner toutes ces opérations

- On peut combiner toutes les opérations pour obtenir les informations qui nous intéressent.
- Tous les rôles de Jeff Bridges avec les titres correspondants :  
 $\pi_{\text{Titre}, \text{Rôle}} \left( (\text{RÔLE} \star \text{FILM}) \star \rho_{\text{No}/\text{ActeurNo}} \left( \sigma_{\text{Nom} = \text{'Jeff Bridges'}} (\text{PERSONNE}) \right) \right)$

Titre	Rôle
The Big Lebowski	The Dude
True Grit	Rooster Cogburn