ALGÈBRE RELATIONNELLE

Algèbre relationnelle *in a nutshell*

- Inventée par Edgar Frank Ted Codd en 1970
- Langage de requête dans les BDR
  - Expressions algébriques en notation fonctionnelle
    - Variables représentent des tables, instances de relations
    - Opérateurs sur ces tables
  - Il existe un interprète de ce langage, qui permet de calculer toute expression portant sur des tables.

Algèbre relationnelle *in a nutshell*

- Opérations sur des relations
- Dont le résultat est une nouvelle relation qui peut à son tour être manipulée
- L'algèbre relationnelle permet de faire des recherches dans les relations

Algèbre relationnelle *in a nutshell*

- Opérations ensemblistes
  de la théorie des ensembles
  - Union, Intersection, Différence
  - Produit cartésien
- Opérations relationnelles
  - Projection
  - Sélection
  - Jointure naturelle
  - Division
  - Renommage

Union, Intersection, Différence

Soient r et s deux instances d'un même schéma de relation R,

Les trois opérations suivantes définissent chacune une instance du même schéma de relation R :

$r \cup s = \{ t \mid t \in r \text{ ou } t \in s \}$

$r \cap s = \{ t \mid t \in r \text{ et } t \in s \}$

$r - s = \{ t \mid t \in r \text{ et } t \notin s \}$

Ces opérations sont impossible entre deux instances qui n'ont pas le même schéma de relation

Union, Intersection, Différence

A

Id	Nom	Prop
122	r21	renault
145	sparc	sun
223	spike	sun
147	r19	renault

B

Id	Nom	Prop
122	r21	renault
145	sparc	sun
149	r18	renault

A ∪ B

Id	Nom	Prop
122	r21	renault
145	sparc	sun
223	spike	sun
147	r19	renault
149	r18	renault

A ∩ B

Id	Nom	Prop
122	r21	renault
145	sparc	sun

A − B

Id	Nom	Prop
223	spike	sun
147	r19	renault

Union, Intersection, Différence

A

Id	Nom	Prop
122	r21	renault
145	sparc	sun
223	spike	sun
147	r19	renault

B

Id	Nom	Ville	Prop
122	r21	Boulogne	renault
145	sparc	New York	sun
149	r18	Marseille	renault

A ∪ B n'est pas défini
A ∩ B n'est pas défini
A − B n'est pas défini

Produit cartésien

Soient r et s deux instances de schémas relations

$r \times s = \{ (t_1,t_2) \mid t_1 \in r \text{ et } t_2 \in s\}$

Le produit cartésien entre deux instances de schémas de relation n'est possible que si ces schémas de relation n'ont aucun attribut commun

Produit cartésien

A

Id	Nom	Prop
122	r21	renault
145	sparc	sun
223	spike	sun

B

Id1	Ville1	Prop1
122	Boulogne	renault
145	New York	sun

A × B

Id	Nom	Prop	Id1	Ville1	Prop1
122	r21	renault	122	Boulogne	renault
145	sparc	sun	122	Boulogne	renault
223	spike	sun	122	Boulogne	renault
122	r21	renault	145	New York	sun
145	sparc	sun	145	New York	sun
223	spike	sun	145	New York	sun

Projection

Soient R un schéma de relation et A = $\{A_1, \ldots, A_n\}$ un sous-ensemble d'attributs de R

- Projection sur A d'un tuple t défini sur R:
  $\prod_A(t) = (t.A_1{:}A_1, \ldots, t.A_n{:}A_n)$
- Projection sur A de l'instance de relation r du schéma R :
  $\prod_A(r) = \{ \prod_A(t) \,,\, t \in r \}$

Projection

$marque$

Id	Nom	Classe	Prop
122	r21	14	renault
145	sparc	12	sun
223	spike	12	sun
147	r19	13	renault

$\prod_{Classe,\,Prop}(marque)$

Classe	Prop
14	renault
12	sun
13	renault

Une instance de schéma de relation étant un ensemble de tuples, il n'y a jamais de doublons

Projection

Propriété

Soient r et s deux instances du schéma de relation R et A un sous-ensemble d'attributs de R

Distributivité par rapport à l'union:
$\prod_A(r \cup s) = \prod_A(r) \cup \prod_A(s)$

Projection

Distributivité par rapport à $\cup$

A

Id	Nom
122	r21
145	sparc
147	r19

B

Id	Nom
122	r21
149	r18

$A \cup B$

Id	Nom
122	r21
145	sparc
147	r19
149	r18

$\prod_{Nom}(A) \quad \cup \quad \prod_{Nom}(B) \quad = \quad \prod_{Nom}(A \cup B)$

Nom
r21
sparc
r19

Nom
r21
r18

Nom
r21
sparc
r19
r18

Projection

Propriétés

Soient r et s deux instances du schéma de relation R et A un sous-ensemble d'attributs de R

Distributivité par rapport à l'union:
$\prod_A(r \cup s) = \prod_A(r) \cup \prod_A(s)$

On n'a pas toujours :

$\prod_A(r \cap s) = \prod_A(r) \cap \prod_A(s)$

$\prod_A(r - s) = \prod_A(r) - \prod_A(s)$

Projection

Non distributivité par rapport à $\cap$

r

Id	Nom	Prop
122	r21	renault
145	sparc	sun
147	r19	renault

$\prod_{Prop}(r)$

Prop
renault
sun

s

Id	Nom	Prop
321	c1	citroen
223	spike	sun

$\prod_{Prop}(s)$

Prop
citroen
sun

$\prod_{Prop}(r) \cap \prod_{Prop}(r)$

Prop
sun

$\prod_{Prop}(r \cap s) = \varnothing$

car $r \cap s = \varnothing$

Sélection

Soient

- r une instance d'un schéma de relation R
- A un attribut de R
- a $\in$ dom(A) une constante

La sélection de r par filtrage de A sur a est une instance du même schéma de relation R définie par :

$$\sigma_{A=a}(r) = \{t \in r : t.A = a\}$$

Sélection

marque

Id	Nom	Classe	Prop
122	r21	14	renault
128	r30	14	renault
145	sparc	12	sun
223	spike	12	sun
147	r19	13	renault

$\sigma_{Classe=14}(marque)$

Id	Nom	Classe	Prop
122	r21	14	renault
128	r30	14	renault

Sélection

Propriétés

Commutativité

$$\sigma_{A=a}(\sigma_{B=b}(r)) = \sigma_{B=b}(\sigma_{A=a}(r)) = \sigma_{A=a,\ B=b}(r)$$

Sélection
Commutativité

$\sigma_{Prop='renault'}(\boldsymbol{marque})$

Id	Classe	Prop
122	14	renault
128	14	renault
147	13	renault

marque

Id	Classe	Prop
122	14	renault
128	14	renault
145	12	sun
111	14	citroen
147	13	renault

$\sigma_{Classe=14}(\boldsymbol{marque})$

Id	Classe	Prop
122	14	renault
128	14	renault
111	14	citroen

$\sigma_{Classe=14}(\sigma_{Prop='renault'}(\boldsymbol{marque}))$

$\sigma_{Prop='renault'}(\sigma_{Classe=14}(\boldsymbol{marque}))$

Id	Classe	Prop
122	14	renault
128	14	renault

Sélection

Propriétés

Commutativité

$\sigma_{A=a}(\sigma_{B=b}(r)) = \sigma_{B=b}(\sigma_{A=a}(r)) = \sigma_{A=a,\ B=b}(r)$

Distributivité sur les opérations ensemblistes

$\forall \text{Op} \in \{\cap, \cup, -\},\ \sigma_{A=a}(r \text{ Op } s) = \sigma_{A=a}(r) \text{ Op } \sigma_{A=a}(s)$

Sélection
Distributivité sur $\cup$

$\boldsymbol{m1}$

Id	Classe	Prop
122	14	renault
145	12	sun

$\boldsymbol{m2}$

Id	Classe	Prop
145	12	sun
147	13	renault

$\boldsymbol{m1 \cup m2}$

Id	Classe	Prop
122	14	renault
145	12	sun
147	13	renault

$\sigma_{Prop='renault'}(\boldsymbol{m1})$

Id	Classe	Prop
122	14	renault

$\sigma_{Prop='renault'}(\boldsymbol{m2})$

Id	Classe	Prop
147	13	renault

$\sigma_{Prop='renault'}(\boldsymbol{m1 \cup m2}) =$

$\sigma_{Prop='renault'}(\boldsymbol{m1}) \cup \sigma_{Prop='renault'}(\boldsymbol{m1})$

Id	Classe	Prop
122	14	renault
147	13	renault

Sélection

Propriétés

Commutativité

$\sigma_{A=a}(\sigma_{B=b}(r)) = \sigma_{B=b}(\sigma_{A=a}(r)) = \sigma_{A=a,\ B=b}(r)$

Distributivité sur les opérations ensemblistes

$\forall \text{Op} \in \{\cap, \cup, -\},\ \sigma_{A=a}(r \text{ Op } s) = \sigma_{A=a}(r) \text{ Op } \sigma_{A=a}(s)$

Commutativité avec la projection *sur même attribut considéré*

$\prod_A (\sigma_{A=a}(r)) = \sigma_{A=a}(\prod_A (r))$

Sélection
Commutativité avec la projection

marque

Id	Nom	Classe
122	r21	14
128	r30	14
145	sparc	12
223	spike	12
147	r19	13

$\pi_{Classe}(\boldsymbol{marque})$

Classe
14
12
13

$\sigma_{Classe=14}(\boldsymbol{marque})$

Id	Nom	Classe
122	r21	14
128	r30	14

$\sigma_{Classe=14}(\pi_{Nom}(\boldsymbol{marque}))$

$\pi_{Nom}(\sigma_{Classe=14}(\boldsymbol{marque}))$

Nom
14

Sélection étendue

Soient

- r une instance de relation de schéma R,
- $\{A_1,\ldots,A_n\}$ un sous-ensemble d'attributs de R,
- f une fonction booléenne définie sur $dom(A_1)\times\ldots\times dom(A_n)$

La sélection de r par $f(A_1,\ldots,A_n)$ est une instance de relation de schéma R définie par :

$$\sigma_{f(A1,\ldots,An)}(r) = \{t \in r \mid f(t.A_1,\ldots,t.A_n)\}$$

Sélection étendue

Exemples:
Soient r une instance de relation de schéma R et A et B deux attributs de R

$\sigma_{A<13}(r)$

$\sigma_{A\neq13}(r)$

$\sigma_{A<13 \vee A>67}(r)$

$\sigma_{A=13 \wedge B>12}(r)$

Jointure naturelle

Contrairement aux autres opérations, elle est définie pour les tuples avant d'être définie pour les relations.

Jointure naturelle de tuples

Si t_1 et t_2 sont deux tuples n'ayant aucun attribut de même nom, alors on peut les joindre et le résultat est la concaténation de t_1 et t_2.

Sinon, on ne peut les joindre que s'ils coïncident sur leurs attributs de même nom.

Jointure naturelle de tuples

Quand on peut joindre deux tuples, l'ensemble des attributs du tuple résultat est l'union des ensembles d'attributs des deux tuples joints.

On ne duplique pas les attributs de même nom.

Jointure naturelle de tuples

Deux tuples $t_r(R)$ et $t_s(S)$ sont **joignables** ssi il existe un tuple $t(R\cup S)$ tel que :

$$\prod_R(t) = t_r \quad \text{et} \quad \prod_S(t) = t_s$$

Ce tuple unique est noté :

$$t = t_r \bowtie t_s$$

Jointure naturelle de tuples

$(t_r \bowtie t_s)$ existe $\Leftrightarrow$ $\prod_{R\cap S}(t_r) = \prod_{R\cap S}(t_s)$

Si $(t_r \bowtie t_s)$ existe, alors
$\prod_R((t_r \bowtie t_s) = t_r$ et $\prod_S((t_r \bowtie t_s) = t_s$

Jointure naturelle de tuples

('Chocolat':NomProd, 302:CodeProd, 4.5:Prix)

⋈

('StéX':NomDistrib, 'Chocolat':NomProd)

=

('Chocolat':NomProd, 302:CodeProd, 4.5:Prix, 'StéX':NomDistrib)

Jointure naturelle de tuples

('Chocolat':NomProd, 302:CodeProd,4.5:Prix)

⋈

('StéX':NomDistrib, 'CremeDeMarron':NomProd)

n'existe pas

Jointure naturelle d'instances de relation

- Toujours possible
- Ensemble éventuellement vide de toutes les jointures de tuples joignables

Jointure naturelle d'instances de relation

marque

idM	NomM	Classe	IdProp
145245	sparc	27	sun
223423	spike	27	sun
147064	renegade	24	renault

prop

IdProp	NomProp	Ville
renault	renaultSA	Boulogne
sun	sun micro	Edmonton
jeep	Jeep inc.	Detroit

prop ⋈ marque

IdProp	NomProp	Ville	IdM	NomM	Classe
renault	renaultSA	Boulogne	147064	renegade	24
sun	sun micro	Edmonton	145245	sparc	27
sun	sun micro	Edmonton	223423	spike	27

Jointure naturelle d'instances de relation

marque

idM	NomM	Classe	IdProp
145245	sparc	27	sun
223423	spike	27	sun
147064	renegade	24	renault

prop

IdProp	NomProp	Ville
jeep	Jeep inc.	Detroit

prop ⋈ marque = ∅

IdProp	NomProp	Ville	IdM	NomM	Classe

Jointure naturelle d'instances de relation

Propriétés

- Associativité
  r ⋈ (s ⋈ t) = (r ⋈ s) ⋈ t
- Distributivité vis-à-vis de l'union
  r ⋈ (s ∪ t) = (r ⋈ s) ∪ (r ⋈ t)
- Jointure et produit cartésien
  Si R∩S = ∅ alors r ⋈ s est isomorphe à r ✕ s

Jointure naturelle d'instances de relation

Jointure et produit cartésien

marque

idM	NomM	Classe	IdProp
145245	sparc	27	sun
223423	spike	27	sun
147064	renegade	24	renault

pays

Code	Nom
FR	France
US	Etats Unis

pays $\bowtie$ marque = pays $\times$ marque car schéma(pays) $\cap$ shéma(marque) = $\varnothing$

Code	Nom	IdM	NomM	Classe	idProp
FR	France	145245	sparc	27	sun
FR	France	223423	spike	27	sun
FR	France	147064	renegade	24	renault
US	Etats Unis	145245	sparc	27	sun
US	Etats Unis	223423	spike	27	sun
US	Etats Unis	147064	renegade	24	renault

Division

Soient r(R) et s(S) deux instances de relation, avec $S \subset R$,

r $\div$ s est la relation de schéma Q=R$-$S définie par:

$$r \div s = \{ t_q \in \pi_Q(r) \mid \forall t_s \in s,\ \ t_q \bowtie t_s \in r\}$$

r $\div$ s regroupe tous les éléments qui dans r sont reliés à tous les éléments de s :

r $\div$ s est le plus grand ensemble q de π_Q(r) tel que (q $\bowtie$ s) $\subset$ r

Division

Quels sont les athlètes qui ont participé à toutes les épreuves?

r

Athlète	Epreuve
Pierre	200 m
Pierre	400 m
Pierre	800 m
Paul	400 m
Jacques	200 m

$s = \prod_{\text{Epreuve}}(r)$

Epreuve
200 m
400 m
800 m

q

Athlète
Pierre

Division

Propriété

La division s'exprime en fonction des opérateurs précédents :

$$r \div s = \pi_Q(r) - \pi_Q(\, (\pi_Q(r) \bowtie s\,) - r)$$

$$r \div s = \pi_Q(r) - \pi_Q(\, (\pi_Q(r) \times s\,) - r)$$

Division

Quels sont les athlètes qui ont participé à toutes les épreuves?

r

Athlète	Epreuve
Pierre	200 m
Pierre	400 m
Pierre	800 m
Paul	400 m
Jacques	200 m

$s = \prod_{\text{Epreuve}}(r)$

Epreuve
200 m
400 m
800 m

$\pi_Q(r)$

Athlète
Pierre
Paul
Jacques

$s \times \pi_Q(r)$

Athlète	Epreuve
Pierre	200 m
Pierre	400 m
Pierre	800 m
Paul	200 m
Paul	400 m
Paul	800 m
Jacques	200 m
Jacques	400 m
Jacques	800 m

Division

Quels sont les athlètes qui ont participé à toutes les épreuves?

$s \times \pi_Q(r)$

Athlète	Epreuve
~~Pierre~~	~~200 m~~
~~Pierre~~	~~400 m~~
~~Pierre~~	~~800 m~~
Paul	200 m
~~Paul~~	~~400 m~~
Paul	800 m
~~Jacques~~	~~200 m~~
Jacques	400 m
Jacques	800 m

—

r

Athlète	Epreuve
Pierre	200 m
Pierre	400 m
Pierre	800 m
Paul	400 m
Jacques	200 m

temp

Athlète	Epreuve
Paul	200 m
Paul	200 m
Jacques	400 m
Jacques	800 m

$\pi_Q(\text{temp})$

Athlète
Paul
Jacques

Division

Quels sont les athlètes qui ont participé à toutes les épreuves?

$\pi_Q(r)$

Athlète
Pierre
Paul
Jacques

—

$\pi_Q(temp)$

Athlète
Paul
Jacques

q

Athlète
Pierre

Division

Existe-t-il une société qui possède *toutes* les marques de la classe 14?

$\Pi_{IdProp,\ IdM}(\text{marque}) \div \Pi_{IdM}(\sigma_{Classe=14}(\text{marque})) =$

$\Pi_{IdProp}(\text{marque})\ -$

$\Pi_{IdProp}($

$\Pi_{IdProp}(\text{marque}) \bowtie \Pi_{IdM}(\sigma_{Classe=14}(\text{marque}))$

$- \Pi_{IdProp,IdM}(\sigma_{Classe=14}(\text{marque}))$

$)$

Renommage

Le renommage de A en B dans une instance r(R) est une instance du schéma R'= R — {A} ∪ {B} notée :

$\delta_{A \leftarrow B}(r)$

On peut étendre le renommage à plusieurs attributs, sous réserve de ne pas créer de collision de noms :

$\delta_{A1,\ldots An \leftarrow B1,\ldots.Bn}(r)$

Renommage

Pourquoi renommer?

Parce qu'on ne dispose que de la jointure naturelle, qui dépend de l'homonymie des attributs

prop = {IdProp, NomProp, Pays, Ville}
enreg = {NumEnr, IdM, Date, Deposant}

L'attribut *Deposant* de *enreg* et l'attribut *IdProp* de *prop* correspondent à la même notion, mais ne portent pas le même nom.

Renommage

prop = {IdProp, NomProp, Pays, Ville}
enreg = {NumEnr, IdM, Date, Deposant}

Quels sont les noms de propriétaires ayant déposé au moins une marque avant le 15 janvier 91?

$\Pi_{NomProp}($

$\delta_{Deposant \leftarrow IdProp}($

$\sigma_{Date<910115}(\text{enreg})$

$)$

$\bowtie \text{prop}$

$)$

Exemples de requêtes

Schéma de la base de données :

employe (nomEmploye, ville, rue)

entreprise (nomEntreprise, ville, rue)

travaille (nomEmploye, nomEntreprise, salaire)

Trouver les noms de tous les employés qui travaillent pour 'Banquissimo'

Exemples de requêtes

Trouver les noms de tous les employés qui travaillent pour ‘Banquissimo’

$$\prod_{nomEmploye}\big(\sigma_{nomEntreprise=\text{‘Banquissimo’}}(travaille) \big)$$

Exemples de requêtes

Schéma de la base de données :

- employe (nomEmploye, ville, rue)
- entreprise (nomEntreprise, ville, rue)
- travaille (nomEmploye, nomEntreprise, salaire)

Trouver les villes dans lesquelles résident au moins un des employés qui travaillent pour ‘Banquissimo’

Exemples de requêtes

Trouver les villes dans lesquelles résident au moins un des employés qui travaillent pour ‘Banquissimo’

$$\prod_{ville}\big(\sigma_{nomEntreprise=\text{‘Banquissimo’}}(travaille) \bowtie employe \big)$$

Exemples de requêtes

Schéma de la base de données :

- employe (nomEmploye, ville, rue)
- entreprise (nomEntreprise, ville, rue)
- travaille (nomEmploye, nomEntreprise, salaire)

Trouver les noms des employés qui travaillent dans leur ville de résidence

Exemples de requêtes

Trouver les noms des employés qui travaillent dans leur ville de résidence

$$\prod_{ville}\big(travaille \bowtie employe \bowtie \Pi_{nomEntreprise,ville}(entreprise) \big)$$

Exemples de requêtes

Schéma de la base de données :

- employe (nomEmploye, ville, rue)
- entreprise (nomEntreprise, ville, rue)
- travaille (nomEmploye, nomEntreprise, salaire)

Trouver les noms des entreprises localisées dans au moins deux villes différentes

Exemples de requêtes

Trouver les noms des entreprises localisées dans au moins deux villes différentes

$\prod_{nomEntreprise}($
$\sigma_{ville \neq ville2}($
entreprise
$\bowtie$
$\delta_{ville \leftarrow ville2,\ rue \leftarrow rue2}$(entreprise)
)
)

Exemples de requêtes

Schéma de la base de données :
- employes (nomEmploye, ville, rue)
- entreprises (nomEntreprise, ville, rue)
- travaille (nomEmploye, nomEntreprise, salaire)

Trouver les noms des entreprises localisées dans au moins deux villes différentes

Exemples de requêtes

Trouver les noms des entreprises localisées dans au moins deux villes différentes

$\prod_{nomEntreprise}($
$\sigma_{ville \neq ville2}($
$\Pi_{NomEntreprise,ville}$(entreprise)
$\bowtie$
$\delta_{ville \leftarrow ville2}$ ($\Pi_{NomEntreprise,ville}$(entreprise)
)))

Exemples de requêtes

Schéma de la base de données :
- employe (nomEmploye, ville, rue)
- entreprise (nomEntreprise, ville, rue)
- travaille (nomEmploye, nomEntreprise, salaire)

Trouver les noms des villes dans lesquelles toutes les entreprises sont implantées

Exemples de requêtes

Trouver les noms des villes dans lesquelles toutes les entreprises sont implantées

$\Pi_{nomEntreprise,\ ville}$(entreprise)
$\div$
Π_{ville} (entreprise)

Pour en savoir plus

- https://fr.wikipedia.org/wiki/Algèbre_relationnelle
- https://en.wikipedia.org/wiki/Relational_algebra
- http://georges.gardarin.free.fr/Livre_BD_Contenu/XX-TotalBD.pdf (chapitre 7)