Démonstrations des lois algébriques utilisées en C2QL

Santiago Bautista

Juin 2017

Structure des démonstrations

Puisque dans toutes les démonstrations qui suivent le but est de prouver, sous certaines conditions, l'égalité de deux fonctions f_1 et f_2 sur R (ou sur R^2 ou R^3 selon le cas), la structure de toutes les démonstrations sera la même : on considérera r une relation (ou une paire ou un triplet de relations, selon le cas), on commencera par montrer que $f_1(r)$ et $f_2(r)$ ont le même schéma relationnel, puis, on montrera que $f_1(r) \subset f_2(r)$ et ensuite que $f_2(r) \subset f_1(r)$.

On aura ainsi démontré par double inclusion que $f_1(r) = f_2(r)$.

Dans toutes les démonstrations qui suivent, quand on dit de deux fonctions f et g qu'elles coïncident sur un ensemble d, on entend par là qu'elles coïncident sur $D_f \cap D_g \cap d$.

Lois de projection

Projection et projection

$$\pi_{\delta_1} \circ \cdots \circ \pi_{\delta_n} = \pi_{\delta_1 \cap \cdots \cap \delta_n} \tag{1}$$

Soit r une relation. On pose $res_1 = \pi_{\delta_1} \circ \cdots \circ \pi_{\delta_n}(r)$ et $res_2 = \pi_{\delta_1 \cap \cdots \cap \delta_n}(r)$

Schéma relationnel

On peut démontrer par récurrence sur n que le schéma relationnel de res_1 est

$$\mathrm{sch}(res_1) = \mathrm{sch}(r) \cap \bigcap_{i \in \{1,\ldots,n\}} \delta_i$$

De même, par définition de la projection, on a

$$\mathrm{sch}(res_2) = \mathrm{sch}(r) \cap \bigcap_{i \in \{1,\ldots,n\}} \delta_i$$

Donc $\mathrm{sch}(res_1) = \mathrm{sch}(res_2)$

Première inclusion

Soit l une ligne de res_1.

Il existe l' une ligne de r telle que $l = \big((l'|_{\delta_n \cup \{id\}})|_{\ldots}\big)|_{\delta_1 \cup \{id\}} = l'|_{(\delta_1 \cap \cdots \cap \delta_n) \cup \{id\}}$. Or, par définition de la projection $\pi_{\delta_1 \cap \cdots \cap \delta_n}$, on a $l'|_{(\delta_1 \cap \cdots \cap \delta_n) \cup \{id\}} \in res_2$. Donc $l \in res_2$.

Ainsi, $res_1 \subset res_2$.

Deuxième inclusion

De même, si l est un élément de res_2, alors il existe une ligne l' de r telle que $l = l'|_{(\delta_1 \cap \cdots \cap \delta_n) \cup \{id\}} = \big((l'|_{\delta_n \cup \{id\}})|_{\ldots}\big)|_{\delta_1 \cup \{id\}}$ et, par définition de $\pi_{\delta_1} \circ \cdots \circ \pi_{\delta_n}$, on a $\big((l'|_{\delta_n \cup \{id\}})|_{\ldots}\big)|_{\delta_1 \cup \{id\}} \in res_1$, d'où $l \in res_1$ et $res_2 \subset res_1$.

Projection et sélection

$$\pi_\delta \circ \sigma_p = \sigma_p \circ \pi_\delta \qquad \text{si } \mathrm{dom}(p) \subset \delta \tag{2}$$

Soit δ un ensemble de noms d'attributs et p un prédicat sur les lignes tel que $\mathrm{dom}(p) \subset \delta$.
Soit r une relation. On pose $res_1 = (\pi_\delta \circ \sigma_p)(r)$ et $res_2 = (\sigma_p \circ \pi_\delta)(r)$

Schéma relationnel

Une sélection ne modifiant jamais le schéma relationnel d'une relation, la schéma relation de res_1 et de res_2 est $\mathrm{sch}(r) \cap \delta$.

Première inclusion

Soit l une ligne de res_1.
Il existe une ligne l' de $\sigma_p(res_1)$ telle que $l = l'|_{(\mathrm{sch}(r) \cap \delta) \cup \{id\}}$.
Puisque l et l' coïncident sur δ et que $\mathrm{dom}(p) \subset \delta$, on a $p(l) = p(l') = \mathit{true}$.
Or, par définition de π_δ, $l' \in \pi_\delta(r)$, donc $l' \in \sigma_p(\pi_\delta(r)) = res_2$.
Ainsi, $res_1 \subset res_2$.

Deuxième inclusion

De même, si l est un élément de res_2,alors $p(l) = \mathit{true}$ et $l \in \pi_\delta(r)$ donc il existe une ligne l' dans r telle que $l = l'|_{(\mathrm{sch}(r) \cap \delta) \cup \{id\}}$. l et l' coïncidant sur δ qui contient le domaine de p, l' vérifie le prédicat p donc l' $\in \sigma_p(r)$.
On en déduit par définition de π_δ que $l \in res_1$.
Ainsi, $res_2 \subset res_1$.

Projection et défragmentation (verticale)

En appelant δ_1 le schéma relationnel du premier argument et δ_2 le schéma relationnel du deuxième argument, on a :

$$\pi_\delta \circ \mathrm{defrag} = \mathrm{defrag} \circ (\pi_\delta, \pi_\delta) \qquad \text{si } \delta_1 \cap \delta_2 = \emptyset \tag{3}$$

Soit δ un ensemble de noms d'attributs. Soient r_1 et r_2 deux relations unifiables.
On pose $res_1 = (\pi_\delta \circ \mathrm{defrag})(r_1, r_2)$ et $res_2 = \mathrm{defrag} \circ (\pi_\delta, \pi_\delta)(r_1, r_2)$.

Remarque : L'hypothèse « r_1 et r_2 unifiables » garantit que les res_1 et res_2 sont bien définies. En effet, non seulement elle garantit que $\mathrm{defrag}(r_1, r_2)$ existe et donc que res_1 existe (la projection a été définie sur R tout entier), mais elle garantit également que $(\delta_1 \cap \delta) \cap (\delta_2 \cap \delta) = \emptyset$ et donc (vu que les projections conservent les identifiants) que $\pi_\delta(r_1)$ et $\pi_\delta(r_2)$ sont unifiables, donc que res_2 existe.

Schémas relationnels

Le schéma relationnel de $\mathrm{defrag}(r_1, r_2)$ est $\delta_1 \cup \delta_2$, donc celui de res_1 est $\delta \cap (\delta_1 \cup \delta_2)$.

Les schémas relationnels de $\pi_\delta(r_1)$ et de $\pi_\delta(r_2)$ sont respectivement $\delta \cap \delta_1$ et $\delta \cap \delta_2$, donc le schéma relationnel de res_2 est $(\delta \cap \delta_1) \cup (\delta \cap \delta_2) = \delta \cap (\delta_1 \cup \delta_2)$

Première inclusion

Soit l une ligne de res_1.

Il existe l_0 une ligne de $\mathrm{defrag}(r_1, r_2)$ de schéma relationnel $\delta_1 \cup \delta_2$ telle que $l = l_0|_{\delta \cup \{id\}}$. Il existe donc deux lignes l_1 et l_2 appartenant respectivement à r_1 et r_2 telles que $l_1 = l_0|_{\delta_1 \cup \{id\}}$ $l_2 = l_0|_{\delta_2 \cup \{id\}}$

Puisque l_1 appartient à r_1, il existe une ligne l'_1 dans $\pi_\delta(r_1)$ telle que $l'_1 = l_1|_{\delta \cup \{id\}} = l_0|_{(\delta \cap \delta_1) \cup \{id\}}$. De même, il existe une ligne l'_2 dans $\pi_\delta(r_2)$ telle que $l'_2 = l_2|_{\delta \cup \{id\}} = l_0|_{(\delta \cap \delta_2) \cup \{id\}}$.

De l'existence de l'_1 et l'_2 qui partagent même identifiant (et portent sur des schémas relationnels disjoints) on en déduit que $l'_1.l'_2$ appartient à res_2.

Or,

$$
\begin{aligned}
l'_1.l'_2 &= l_0|_{((\delta \cap \delta_1) \cup \{id\}) \cup ((\delta \cap \delta_2) \cup \{id\})} \\
&= l_0|_{(\delta \cap (\delta_1 \cup \delta_2)) \cup \{id\}} \\
&= \left(l_0|_{\delta_1 \cup \delta_2 \cup \{id\}}\right)|_{\delta \cup \{id\}} \\
&= l_0|_{\delta \cup \{id\}} = l
\end{aligned}
$$

Donc : $l \in res_2$.

Deuxième inclusion

Soit l une ligne de res_2.

Il existe des lignes l'_1 et l'_2 appartenant respectivement à $\pi_\delta(r_1)$ et $\pi_\delta(r_2)$ telles que $l = l'_1.l'_2$.

On en déduit qu'il existe deux lignes l_1 et l_2 appartenant respectivement à r_1 et r_2 telles que $l'_1 = l_1|_{\delta \cup \{id\}}$ et $l'_2 = l_2|_{\delta \cup \{id\}}$.

r_1 et r_2 étant unifiables, et l_1 et l_2 ayant même identifiant, l_1 et $_2$ sont des lignes correspondantes et on peut donc considérer $l_1.l_2$.

On a d'ailleurs $l = l'_1.l'_2 = (l_1|_{\delta \cup \{id\}}).(l_2|_{\delta \cup \{id\}}) = (l_1.l_2)|_{\delta \cup \{id\}}$.

Or, $l_1.l_2$ appartient à $\mathrm{defrag}(r_1, r_2)$ donc $l = (l_1.l_2)|_{\delta \cup \{id\}}$ appartient à res_1

Projection et déchiffrement d'un attribut projeté ou non

$$\pi_\delta \circ \mathrm{decrypt}_{\alpha,\mathsf{c}} \equiv \mathrm{decrypt}_{\alpha,\mathsf{c}} \circ \pi_\delta \tag{4}$$

Soit δ un ensemble de noms d'attributs et α un attribut (appartenant à δ ou pas). Soit r une relation. On pose $res_1 = (\pi_\delta \circ \mathrm{decrypt}_{\alpha,\mathsf{c}})(r)$ et $res_2 = (\mathrm{decrypt}_{\alpha,\mathsf{c}} \circ \pi_\delta)(r)$.

Schémas relationnels

Le déchiffrement ne changeant pas le schéma relationnel d'une relation,le schéma relationnel de res_1 et res_2 est $\mathrm{sch}(r) \cap \delta$.

Première inclusion

Soit l une ligne de res_1.

Il existe l' une ligne de $\mathrm{decrypt}_{\alpha,\mathtt{c}}(r)$ telle que $l = l'|_{\delta\cup\{id\}}$. l' étant un élément de $\mathrm{decrypt}_{\alpha,\mathtt{c}}(r)$, il existe une ligne l_0 de r telle que $l' = \mathtt{c}^{-1}(l_0)_\alpha$ et donc $l = \mathtt{c}^{-1}(l_0)_\alpha|_{\delta\cup\{id\}}$.

Puisque l_0 appartient à r, $l_0|_{\delta\cup\{id\}}$ appartient à $\pi_\delta(r)$ et donc $\mathtt{c}^{-1}(l_0|_{\delta\cup\{id\}})_\alpha$ appartient à res_2.

Montrons que $\mathtt{c}^{-1}(l_0)_\alpha|_{\delta\cup\{id\}} = \mathtt{c}^{-1}(l_0|_{\delta\cup\{id\}})_\alpha$. Les deux fonctions en question sont définies sur $(\mathrm{sch}(r)\cap\delta)\cup\{id\}$.

Soit : $\beta \in (\mathrm{sch}(r)\cap\delta)\cup\{id\}$.

Si $\beta \neq \alpha$, on a :

$$\left\{\begin{array}{ll} \mathtt{c}^{-1}(l_0)_\alpha|_{\delta\cup\{id\}}(\beta) & = \mathtt{c}^{-1}(l_0)_\alpha(\beta) = l_0(\beta) \\ \mathtt{c}^{-1}(l_0|_{\delta\cup\{id\}})_\alpha(\beta) & = l_0|_{\delta\cup\{id\}}(\beta) = l_0(\beta) \end{array}\right.$$

Si $\alpha \in \mathrm{sch}(r)\cap\delta$, on a :

$$\left\{\begin{array}{ll} \mathtt{c}^{-1}(l_0)_\alpha|_{\delta\cup\{id\}}(\alpha) & = \mathtt{c}^{-1}(l_0)_\alpha(\alpha) = \mathtt{c}^{-1}(l_0(\alpha) \\ \mathtt{c}^{-1}(l_0|_{\delta\cup\{id\}})_\alpha(\alpha) & = \mathtt{c}^{-1}(l_0|_{\delta\cup\{id\}}(\alpha)) = \mathtt{c}^{-1}(l_0(\alpha)) \end{array}\right.$$

Ainsi, $\mathtt{c}^{-1}(l_0)_\alpha|_{\delta\cup\{id\}} = \mathtt{c}^{-1}(l_0|_{\delta\cup\{id\}})_\alpha$ donc l appartient à res_2.

Deuxième inclusion

Soit l une ligne de res_2.

Il existe une ligne l' de $\pi_\delta(r)$ telle que $l = \mathtt{c}^{-1}(l')_\alpha$.

Puisque l' appartient à $\pi_\delta(r)$, il existe l_0 dans r telle que $l' = l_0|_\delta$ et donc telle que $l = \mathtt{c}^{-1}(l_0|_{\delta\cup\{id\}})_\alpha$.

Vu que l_0 appartient à r, $\mathtt{c}^{-1}(l_0)_\alpha$ appartient à $\mathrm{decrypt}_{\alpha,\mathtt{c}}(r)$ et $\mathtt{c}^{-1}(l_0)_\alpha|_{\delta\cup\{id\}}$ appartient à res_1.

Or, l_0 étant une ligne de r, d'après la démonstration faite pour la première inclusion, on a : $\mathtt{c}^{-1}(l_0)_\alpha|_{\delta\cup\{id\}} = \mathtt{c}^{-1}(l_0|_{\delta\cup\{id\}})_\alpha$.

On en déduit que l appartient à res_1.

Projection et déchiffrement d'un attribut non projeté

$$\pi_\delta \circ \mathrm{decrypt}_{\alpha,\mathtt{c}} \equiv \pi_\delta \qquad \text{si } \alpha \notin \delta \tag{5}$$

Soit δ un ensemble de noms d'attributs et α un attribut n'appartenant pas à δ. Soit r une relation. On pose $res_1 = (\pi_\delta \circ \mathrm{decrypt}_{\alpha,\mathtt{c}})(r)$ et $res_2 = (\mathrm{decrypt}_{\alpha,\mathtt{c}} \circ \pi_\delta)(r)$.

Schémas relationnels

Le déchiffrement ne changeant pas le schéma relationnel d'une relation,le schéma relationnel de res_1 et res_2 est $\mathrm{sch}(r)\cap\delta$.

Inclusions

La seule chose qui change est la démonstration du fait que pour toute ligne l_0 de r $\mathtt{c}^{-1}(l_0)_\alpha|_{\delta\cup\{id\}} = \mathtt{c}^{-1}(l_0|_{\delta\cup\{id\}})_\alpha$.

En effet, si on suppose $\alpha \notin \delta$, un seul cas se présente, à savoir $\beta \in (\mathrm{sch}(r) \cap \delta) \cup \{id\} \wedge \beta \neq \alpha$, et on a alors

$$\begin{cases} \mathtt{c}^{-1}(l_0)_\alpha|_{\delta\cup\{id\}}(\beta) & = \mathtt{c}^{-1}(l_0)_\alpha(\beta) = l_0(\beta) \\ \mathtt{c}^{-1}(l_0|_{\delta\cup\{id\}})_\alpha(\beta) & = l_0|_{\delta\cup\{id\}}(\beta) = l_0(\beta) \end{cases}$$

d'où l'égalité voulue.

À partir de là, si l est une ligne de res_1, elle s'écrit $\mathtt{c}^{-1}(l_0)_\alpha|_{\delta\cup\{id\}}$ avec $l_0 \in r$ et $\mathtt{c}^{-1}(l_0|_{\delta\cup\{id\}})_\alpha$ appartient à res_2 donc l appartient à res_2.

Inversement, si l est une ligne de res_2, elle s'écrit $\mathtt{c}^{-1}(l_0|_{\delta\cup\{id\}})_\alpha$ avec $l_0 \in r$ et $\mathtt{c}^{-1}(l_0)_\alpha|_{\delta\cup\{id\}}$ appartient à res_1 donc l appartient à res_1.

Projection et jointure

En appelant δ_1 le schéma relationnel du premier argument et δ_2 le schéma relationnel du deuxième argument, on a :

$$\pi_\delta \circ \bowtie = \bowtie \circ (\pi_\delta, \pi_\delta) \qquad \text{si } \delta_1 \cap \delta_2 \subset \delta \tag{6}$$

Soit δ un ensemble de noms d'attributs, et r_1 et r_2 des relations. On pose $res_1 = (\pi_\delta \circ \bowtie)(r_1, r_2)$ et $res_1 = (\bowtie \circ (\pi_\delta, \pi_\delta))(r_1, r_2)$.

Schémas relationnels

Le schéma relationnel de $r_1 \bowtie r_2$ est $\mathrm{sch}(r_1) \cup \mathrm{sch}(r_2)$ donc celui de res_1 est $(\mathrm{sch}(r_1) \cup \mathrm{sch}(r_2)) \cap \delta$.

Les schémas relationnels respectifs de $\pi_\delta(r_1)$ et $\pi_\delta(r_2)$ sont $\mathrm{sch}(r_1) \cap \delta$ et $\mathrm{sch}(r_2) \cap \delta$ donc celui de res_2 est $(\mathrm{sch}(r_1) \cap \delta) \cup (\mathrm{sch}(r_2) \cap \delta) = (\mathrm{sch}(r_1) \cup \mathrm{sch}(r_2)) \cap \delta$.

Première inclusion

Soit l une ligne de res_1.

Il existe une ligne l' de $r_1 \bowtie r_2$ telle que $l = l'|_{\delta\cup\{id\}}$. Puisque l' appartient à $r_1 \bowtie r_2$, il existe deux lignes l_1 et l_2 appartenant respectivement à r_1 et r_2 telles que $l' = l_1.l_2$. Ainsi, $l = (l_1.l_2)|_{\delta\cup\{id\}}$.

Puisque l_1 et l_2 se correspondent et que $\delta_1 \cap \delta_2 \subset \delta$, $l_1|_{\delta\cup\{id\}}$ et $l_2|_{\delta\cup\{id\}}$ se correspondent aussi. Or, $l_1|_{\delta\cup\{id\}}$ (respectivement $l_2|_{\delta\cup\{id\}}$) appartient à $\pi_\delta(l_1)$ (resp. $\pi_\delta(r_2)$), donc $l_1|_{\delta\cup\{id\}}.l_2|_{\delta\cup\{id\}}$ appartient à res_2.

Montrons que $(l_1.l_2)|_{\delta\cup\{id\}} = l_1|_{\delta\cup\{id\}}.l_2|_{\delta\cup\{id\}}$. Ces deux fonctions sont définies sur $((\delta_1 \cup \delta_2) \cap \delta) \cup \{id\}$. Soit β un élément de $(\delta_1 \cup \delta_2) \cap \delta$.

$$(l_1.l_2)|_{\delta\cup\{id\}}(\beta) = l_1.l_2(\beta) = \begin{cases} l_1(\beta) & \text{si } \beta \in \delta_1 \\ l_2(\beta) & \text{si } \beta \in \delta_2 \end{cases}$$

$$l_1|_{\delta\cup\{id\}}.l_2|_{\delta\cup\{id\}}(\beta) = \begin{cases} l_1|_{\delta\cup\{id\}}(\beta) = l_1(\beta) & \text{si } \beta \in \delta_1 \\ l_2|_{\delta\cup\{id\}}(\beta) = l_2(\beta) & \text{si } \beta \in \delta_2 \end{cases}$$

De plus, $(l_1.l_2)|_{\delta\cup\{id\}}(id) = l_1|_{\delta\cup\{id\}}.l_2|_{\delta\cup\{id\}}(id) = l_1(id).l_2(id)$. Donc on a bien l'égalité souhaitée et on en déduit que l appartient à res_2.

Deuxième inclusion

Soit l une ligne de res_2.

Il existe deux lignes l'_1 et l'_2 de $\pi_\delta(r_1)$ et $\pi_\delta(r_2)$ respectivement telles que $l = l'_1.l'_2$. Or, il existe deux lignes l_1 et l_2 appartenant respectivement à r_1 et r_2 telles que $l'_1 = l_1|_{\delta\cup\{id\}}$ et $l'_2 = l_2|_{\delta\cup\{id\}}$.

Donc $l = l_1|_{\delta\cup\{id\}}.l_2|_{\delta\cup\{id\}}$.

D'autre part, vu que $l_1|_{\delta\cup\{id\}}$ et $l_2|_{\delta\cup\{id\}}$ se correspondent, $l_1|_{\delta\cup\{id\}}$ et $l_2|_{\delta\cup\{id\}}$ coïncident sur $((\delta_1 \cap \delta) \cap (\delta_2 \cap \delta)$. Or, $\delta_1 \cap \delta_2 \subset \delta$, donc l_1 et l_2 coïncident sur $\delta_1 \cap \delta_2$ donc l_1 et l_2 se correspondent et donc $l_1.l_2$ appartient à $r_1 \bowtie r_2$.

On en déduit que $(l_1.l_2)\,|_{\delta\cup\{id\}}$ appartient à res_1.

Grâce à l'égalité prouvée lors de la preuve de l'autre inclusion, on en déduit que l appartient à res_1.

Projection et agrégation

$$\text{group}_\delta \circ \pi_{\delta'} \equiv \pi_{\delta'} \circ \text{group}_\delta \qquad\qquad \text{si } \delta \subset \delta'$$

Soient δ et δ' deux ensembles de noms d'attributs.

Soit δ_1 le schéma relationnel de l'argument.

Soit r une relation. On pose $res_1 = (\text{group}_\delta \circ \pi_{\delta'})(r)$ et $res_2 = (\pi_{\delta'} \circ \text{group}_\delta)(r)$.

Schémas relationnels

La fonction group préserve les schémas relationnels, donc res_1 et res_2 ont tous deux pour schémas relationnels $\delta_1 \cap \delta'$.

Premier cas : si δ est vide

Dans ce cas-là, pour montrer que $res_1 = res_2$, on va directement calculer res_1 et res_2.

Calcul de res_1 :

Le seul nom de groupe minimal de $\pi_{\delta'}(r)$ pour $\delta = \emptyset$ est l'application vide qu'on notera également $\emptyset$.

Donc, res_1 a une seule ligne, à savoir $\text{lg}_{\pi_{\delta'}(r),\emptyset}$, qu'on appellera, pour simplifier les notations, l_1.

l_1 est définie sur $(\delta_1 \cap \delta') \cup \{id\}$ et est entièrement déterminée par

$$\forall \alpha \in (\delta_1 \cap \delta') \cup \{id\}, l_1(\alpha) = r_\emptyset(\alpha)$$

Calcul de res_2 :

De même, le seul nom de groupe minimal de r pour $\emptyset$ est $\emptyset$ donc $\text{group}_\delta(r)$ a un seul élément, que nous appelleront l'_2 qui est défini sur $\delta_1 \cup \{id\}$ par $\forall \alpha \in \delta_1 \cup \{id\}, l'_2(\alpha) = r_\emptyset(\alpha)$.

On en déduit que res_2 a une seule ligne, que nous appellerons l_2.

l_2 est définie sur $(\delta_1 \cap \delta) \cup \{id\}$ par

$$\forall \alpha \in (\delta_1 \cap \delta) \cup \{id\}, l_2(\alpha) = r_\emptyset(\alpha)$$

Donc on a $l_1 = l_2$ et on en déduit $res_1 = res_2$.

Deuxième cas : si δ est non vide

Première inclusion :

Soit l une ligne de res_1.

Soit n un nom de groupe sur δ associé (i. e. $n = l|_\delta$).

Pour simplifier les notations, on pose $r' = \pi_{\delta'}(r)$.

Il existe $l'_1, \ldots, l'_m$ des lignes distinctes de r' telles que $r'_n = \{l'_1, \ldots, l'_m\}$.

Les l'_i appartenant à r', il existe des lignes $l_1, \ldots, l_m$ de r telles que

$$\forall i \in \{1, \ldots n\}, l'_i = l_i|_{\delta' \cup \{id\}}$$

Montrons par double inclusion que $\{l_1, \ldots, l_m\} = r_n$.

Les $l'_1, \ldots, l'_m$ sont les restriction des $l_1, \ldots, l_m$ à δ' et elles coïncident entre elles sur δ. Or $\delta \subset \delta'$ donc les $l_1, \ldots, l_m$ coïncident sur δ. On en déduit que $\{l_1, \ldots, l_m\} \subset r_n$.

Soit maintenant l_0 un élément de r_n. Puisque l_1 appartient à r_n, l_0 coïncide avec l_1 sur δ ; donc $l_0|_{\delta' \cup \{id\}}$ coïncide sur δ avec $l_1|_{\delta' \cup \{id\}}$, et par conséquent avec n donc $l_0|_{\delta' \cup \{id\}} \in r'_n$

On en déduit qu'il existe $i \in \{1, \ldots, n\}$ tel que $l_0|_{\delta' \cup \{id\}} = l'_i$.

l_0 coïncide donc avec l_i sur $\delta' \cup \{id\}$ donc en particulier $l_0(id) = l_i(id)$ et, comme l'identifiant de chaque ligne dans une relation est supposé unique et l_0 et l_1 appartiennent tous les deux à la relation r, on a : $l_0 = l_i$.

Ainsi, $l_0 \in \{l_1, \ldots, l_m\}$.

On en déduit que $r_n \subset \{l_1, \ldots, l_m\}$ et on a donc l'égalité.

On pose $l' = (\text{lg}_{r,n})|_{\delta' \cup \{id\}}$, qui est donc un élément de res_2.

Montrons que $l' = l$.

Ces deux fonctions sont définies sur $\delta' \cup \{id\}$, et pour $\alpha \in \delta' \cup \{id\}$ on a :

$$\begin{cases} l'(\alpha) = n(\alpha) = l(\alpha) & \text{si } \alpha \in \delta \\ l'(\alpha) = r_n(\alpha) = r'_n(\alpha) = l(\alpha) & \text{si } \alpha \notin \delta \end{cases}$$

d'où l'égalité.

On en déduit que l appartient à res_2, d'où la première inclusion.

Deuxième inclusion :

Soit l une ligne de res_2.

Soit n un nom de groupe pour δ tel que $l = \text{lg}_{r,n}|_{\delta' \cup \{id\}}$.

Il existe des lignes $l_1, \ldots, l_m$ telles que $r_n = \{l_1, \ldots, l_m\}$.

En appelant r' la relation $\pi_{\delta'}(r)$ et en appelant, pour i dans $\{1, \ldots, m\}$, $l'_i = l_i|_{\delta' \cup \{id\}}$ montrons par double inclusion que $r'_n = \{l'_1, \ldots, l'_m\}$.

Puisque les $l_1, \ldots, l_m$ et n coïncident sur δ, leurs restrictions $l'_1, \ldots, l'_m$ coïncident sur δ également et coïncident sur δ avec n, donc $\{l'_1, \ldots, l'_m\} \subset r'_n$.

Dans l'autre sens, soit l'_0 un élément de r'_n. Il existe l_0 élément de r tel que $l'_0 = l_0|_{\delta' \cup \{id\}}$. Puisque l'_0 coïncide avec n sur δ, que $\delta \subset \delta'$ et que l_0 coïncide avec l'_0 sur δ', l_0 coïncide avec n sur δ, d'où $l_0 \in r_n$.

On en déduit qu'il existe i dans $\{1, \ldots, m\}$ tel que $l_0 = l_i$. Par définition de l_0 et des l'_i, on en déduit que $l'_0 = l'_i$, donc que $l'_0 \in \{l'_1, \ldots, l'_m\}$.

Ainsi, $r'_n \subset \{l'_1, \ldots, l'_m\}$, d'où l'égalité.

Projection et réduction d'un attribut projeté ou non

$$\text{fold}_{\alpha,f,z} \circ \pi_\delta \equiv \pi_\delta \circ \text{fold}_{\alpha,f,z} \tag{7}$$

Soit δ un ensemble de noms d'attributs, α un attribut, f une fonction de $\mathcal{V}$ dans val et z un élément de $\mathcal{V}$.

Soit r une relation. On pose $res_1 = (\text{fold}_{\alpha,f,z} \circ \pi_\delta)(r)$ et $res_2 = (\pi_\delta \circ \text{fold}_{\alpha,f,z})(r)$.

Schémas relationnels

Ni la projection ni la réduction ne changent les schémas relationnels, donc res_1, r et res_2 ont tous les trois le même schéma relationnel.

Première inclusion

Soit l une ligne de res_1.

Il existe l' dans $\pi_\delta(r)$ telle que $l = \text{red}_{\alpha,f,z,l'}$ et l'' dans r telle que $l' = l''|_{\delta\cup\{id\}}$.

l'' appartient à r donc $\text{red}_{\alpha,f,z,l''}$ appartient à $\text{fold}_{\alpha,f,z}(r)$ et, en posant $\tilde{l} = \text{red}_{\alpha,f,z,l''}|_{\delta\cup\{id\}}$, $\tilde{l}$ appartient à res_2.

Montrons que $l = \tilde{l}$.

Soit $\beta \in (\text{sch}(r) \cap \delta) \cup \{id\}$.

Si $\beta \neq \alpha$, on a :

$$\tilde{l}(\beta) = \text{red}_{\alpha,f,z,l''}(\beta) = l''(\beta) = l'(\beta) = \text{red}_{\alpha,f,z,l'}(\beta) = l(\beta)$$

Si $\beta = \alpha$, on a :

$$\tilde{l}(\alpha) = \text{red}_{\alpha,f,z,l''}(\alpha) = \text{red}_{f,z}(l''(\alpha)) = \text{red}_{f,z}(l'(\alpha)) = \text{red}_{\alpha,f,z,l'}(\alpha) = l(\alpha)$$

Puisqu'on a l'égalité souhaitée, on peut en déduire que l appartient à res_2, d'où la première inclusion.

Deuxième inclusion

Puisque la projection comme la réduction préservent le nombre de lignes dans la relation, res_1 et res_2 sont des ensembles finis ayant tous les deux le même cardinal (à savoir le cardinal de r). Donc la première inclusion implique la deuxième.

Projection et réduction d'un attribut non projeté

$$\text{fold}_{\alpha,f,z} \circ \pi_\delta \equiv \pi_\delta \qquad \text{si } \alpha \notin \delta \cup \{id\} \tag{8}$$

Soit δ un ensemble de noms d'attributs, α un attribut, f une fonction de $\mathcal{V}$ dans val et z un élément de $\mathcal{V}$.

Soit r une relation. On pose $res_1 = (\text{fold}_{\alpha,f,z} \circ \pi_\delta)(r)$ et $res_2 = \pi_\delta(r)$

Schéma et cardinal

Pour les mêmes raisons que lors de la démonstration précédente, res_1 et res_2 ont tous les deux le même schéma relationnel et le même cardinal que r.

Inclusion de res_2 dans res_1

Soit l une ligne de res_2.
On pose $l' = \mathrm{red}_{\alpha,f,z,l}$, qui est un élément de res_1.
Montrons que $l = l'$.
Ces deux lignes là sont définies sur $(\mathrm{sch}(r) \cap \delta) \cup \{id\}$.
Pour $\beta \in (\mathrm{sch}(r) \cap \delta) \cup \{id\}$, on a forcément $\beta \neq \alpha$ car $\alpha \notin \delta \cup \{id\}$, donc

$$l'(\beta) = l(\beta)$$

d'où l'égalité entre l et l' et l'appartenance de l à res_1.

Ainsi, $res_2 \subset res_1$. Vu que de plus les deux ensembles sont finis de même cardinal, on en déduit qu'ils sont égaux.

Lois de sélection

Sélection et sélection

$$\sigma_{p_1} \circ \cdots \circ \sigma_{p_n} \equiv \sigma_{p_1 \wedge \cdots \wedge p_n} \tag{9}$$

Vu que le ET logique est associatif, pour démontrer cette loi, il suffit de le démontrer pour deux sélections.
Soient p_1 et p_2 deux prédicats et r une relation.
On pose $res_1 = (\sigma_{p_1} \circ \sigma_{p_2})(r)$ et $res_2 = \sigma_{p_1 \wedge p_2}(r)$.

Schémas relationnels

La sélection préserve le schéma relationnel, donc res_1 et res_2 ont tous les deux le même schéma relationnel que r.

Première inclusion

Soit l une ligne de res_1.
Puisque l appartient à res_1, l appartient à $\sigma_{p_2}(r)$ et l vérifie p_1.
Or, si l appartient à $\sigma_{p_2}(r)$, l appartient à r et vérifie p_2.
Donc l appartient à r et vérifie à la fois p_1 et p_2, donc vérifie $p_1 \wedge p_2$.
Par conséquent, l appartient à res_2.

On en déduit que $res_1 \subset res_2$.

Deuxième inclusion

Soit l un élément de res_2.
Puisque l appartient à res_2, l appartient à r et vérifie $p_1 \wedge p_2$.
On en déduit que l vérifie et p_1 et p_2.
Comme l appartient à r et vérifie p_2, l appartient aussi à $\sigma_{p_2}(r)$. Or l vérifie p_1, donc l appartient à $\sigma_{p_1}(\sigma_{p_2}(r))$, autrement connue sous le nom de res_1.

On en déduit que $res_2 \subset res_2$.

Sélection et défragmentation

En appelant δ_1 le schéma relationnel du premier argument, et δ_2 le schéma relationnel du deuxième arguemnt,

$$\sigma_p \circ \text{defrag} \equiv \text{defrag} \circ (\sigma_p, \text{id}) \qquad \text{si } dom(p) \subset \delta_1 \tag{10}$$

$$\sigma_p \circ \text{defrag} \equiv \text{defrag} \circ (\text{id}, \sigma_p) \qquad \text{si } dom(p) \subset \delta_2 \tag{11}$$

Les démonstrations des deux lois étant tout à fait analogues, je ne démontrerai que la loi (10).
Soit p un prédicat sur les lignes. Soient r_1 et r_2 deux relations unifiables.
On pose $res_1 = \sigma_p(\text{defrag}(r_1, r_2))$ et $res_2 = \text{defrag}(\sigma_p(r_1), r_2)$.

Schémas relationnels et unifiabilité

La sélection préserve le schéma relationnel, donc r_1 et r_2 sont unifiables si et seulement si $\sigma_p(r_1)$ et r_2 le sont.
De plus, le schéma relationnel après défragmentation est l'union des schémas relationnels initiaux, donc les schémas relationnels de res_1 et res_2 sont tous les deux $\delta_1 \cup \delta_2$.

Première inclusion

Soit l une ligne de res_1.
l vérifie la propriété p et appartient aussi à $\text{defrag}(r_1, r_2)$. Or, puisque $\text{dom}(p) \subset \delta_1$, toute ligne coïncidant avec l sur δ_1 vérifie la propriété p.
Il existe deux lignes l_1 et l_2 appartenant respectivement à r_1 et r_2 telles que $l = \text{Unif}(l_1, l_2)$. l_1 coïncide avec l sur $\delta_1 \cup \{id\}$ donc sur δ_1, donc l_1 vérifie la propriété p.

On en déduit que l_1 appartient à $\sigma_p(r_1)$. Or l_2 appartient à r_2 et, par définition, l_1 et l_2 sont unifiables, donc $l = \text{Unif}(l_1, l_2)$ appartient à $\text{defrag}(\sigma_p(r_1, r_2)) = res_2$.

Deuxième inclusion

Soit l une ligne de res_2.
Il existe l_1 et l_2 unifiables appartenant respectivement à $\sigma_p(r_1)$ et r_2 telles que $l = \text{Unif}(l_1, l_2)$.
Puisque l_1 appartient à $\sigma_p(r_1)$, l_1 vérifie la condition p et appartient à r_1.
Or l_2 appartient à r_2 et l_1 et l_2 sont unifiables, donc $l = \text{Unif}(l_1, l_2)$ appartient à $\text{defrag}(r_1;_2)$.
Puisque l_1 coïncide avec l sur son domaine de définition qui contient δ_1 qui lui même contient $\text{dom}(p)$ et que l_1 vérifie p, l vérifie p.
On en déduit que l appartient à $\sigma_p(\text{defrag}(r_1, r_2)) = res_2$.

Sélection et déchiffrement non sélectif

$$\sigma_p \circ \text{decrypt}_{\alpha,\mathtt{c}} = \text{decrypt}_{\alpha,\mathtt{c}} \circ \sigma_p \qquad \text{si } \alpha \notin \text{dom}(p) \tag{12}$$

Soit p un prédicat sur les lignes, $\mathtt{c}$ un chiffrement et α un attribut n'étant pas contenu dans le domaine de p.

Soit r une relation et δ_1 son schéma relationnel.

On pose $res_1 = \sigma_p(\text{decrypt}_{\alpha,\mathtt{c}}(r))$ et $res_2 = \text{decrypt}_{\alpha,\mathtt{c}}(\sigma_p(r))$.

Schémas relationnels

La sélection et le déchiffrement préservant tous deux les schémas relationnels, res_1 et res_2 ont tous deux pour schéma relationnel δ_1.

Première inclusion

Soit l une ligne de res_1.

l appartient à $\text{decrypt}_{\alpha,\mathtt{c}}(r)$ et l vérifie la propriété p.

Il existe une ligne l' de r telle que $l = \mathtt{c}^{-1}(l')_\alpha$.

Puisque l et l' (qui sont toutes deux définies sur $\delta_1 \cup \{id\}$) coïncident partout sauf éventuellement sur α mais que le domaine de p ne contient pas alpha, on a $l|_{\text{dom}(p)} = l'|_{\text{dom}(p)}$. De plus, l vérifie la propriété p donc l' vérifie la propriété p.

On en déduit que l' appartient à $\sigma_p(r)$.

Par conséquent, $l = \mathtt{c}^{-1}(l')_\alpha$ appartient à $\text{decrypt}_{\alpha,\mathtt{c}}(\sigma_p(r)) = res_2$.

Deuxième inclusion

Soit l une ligne de res_2.

Il existe l' dans $\sigma_p(r)$ telle que $l = \mathtt{c}^{-1}(l')_\alpha$.

l' appartient à r et vérifie p.

Pour les mêmes raisons que précédemment, l et l' coïncident sur le domaine de p et donc l vérifie la propriété p.

Or, puisque l' appartient à r, l appartient à $\text{decrypt}_{\alpha,\mathtt{c}}(r)$, donc l appartient à $\sigma_p(\text{decrypt}_{\alpha,\mathtt{c}}(r)) = res_1$.

Sélection et déchiffrement d'un attribut sélectif

$$\sigma_p \circ \text{decrypt}_{\alpha,\mathtt{c}} = \text{decrypt}_{\alpha,\mathtt{c}} \circ \sigma_{\mathtt{c}\Rightarrow p} \qquad \text{si } p \text{ est compatible avec } \mathtt{c} \tag{13}$$

Soit p un prédicat, soit α un nom d'attribut, soit $\mathtt{c}$ un chiffrement compatible avec p pour α, r une relation et δ_1 son schéma relationnel.

On pose $res_1 = \sigma_p(\text{decrypt}_{\alpha,\mathtt{c}}(r))$ et $res_2 = \text{decrypt}_{\alpha,\mathtt{c}}(\sigma_{\mathtt{c}_\alpha\Rightarrow p})$.

Schémas relationnels

Le chiffrement et la sélection préservent le schéma relationnel, donc res_1 et res_2 ont tous les deux pour schéma relationnel δ_1.

Première inclusion

Soit l un élément de res_1.

l vérifie le prédicat p et appartient à $\text{decrypt}_{\alpha,\mathtt{c}}(r)$. Il existe donc une ligne l_0 de r telle que $l = \mathtt{c}^{-1}(()l_0)_\alpha$, d'où $l_0 = \mathtt{c}(l)_\alpha$.

Puisque l vérifie p, l_0 vérifie $\mathtt{c}_\alpha \Rightarrow p$ et donc l_0 appartient à $\sigma_{\mathtt{c}_\alpha \Rightarrow p}(r)$. On en déduit que l appartient à $\text{decrypt}_{\alpha,\mathtt{c}}(\sigma_{\mathtt{c}_\alpha \Rightarrow p})(r) = res_2$.

Deuxième inclusion

Soit l une ligne de res_2.

Il existe l_0 dans $\sigma_{\mathtt{c}_\alpha \Rightarrow p}(r)$ telle que $l = \mathtt{c}^{-1}(l_0)_\alpha$ et donc telle que $l_0 = \mathtt{c}(l)_\alpha$.

Puisque l_0 appartient à $\sigma_{\mathtt{c}_\alpha \Rightarrow p}(r)$, l_0 appartient à r et vérifie $\mathtt{c}_\alpha \Rightarrow p$. Par conséquent, l appartient à $\text{decrypt}_{\alpha,\mathtt{c}}(r)$ et l vérifie p donc l appartient à $\sigma_p(\text{decrypt}_{\alpha,\mathtt{c}}(r)) = res_1$.

Sélection et jointure

Soit δ_1 le schéma relationnel du premier argument et δ_2 le schéma relationnel du deuxième argument.

$$\sigma_p \circ \bowtie = \bowtie \circ(\sigma_p, \text{id}) \qquad \text{si dom}(p) \subset \delta_1 \tag{14}$$
$$\sigma_p \circ \bowtie = \bowtie \circ(\text{id}, \sigma_p) \qquad \text{si dom}(p) \subset \delta_2 \tag{15}$$

Les deux lois se démontrant de façons tout à fait symétriques, je ne vais démontrer que la loi (14).

Soit p un prédicat et r_1 et r_2 deux relations.

On pose $res_1 = \sigma_p(r_1 \bowtie r_2)$ et $res_2 = \sigma_p(r_1) \bowtie r_2$.

Schémas relationnels

La sélection préservant les schémas relationnels, les schémas relationnels de res_1 et res_2 sont tous les deux $\delta_1 \cup \delta_2$.

Première inclusion

Soit l un élément de res_1.

l vérifie p et appartient à $r_1 \bowtie r_2$. Donc il existe deux lignes l_1 et l_2 se correspondant et appartenant respectivement à r_1 et r_2 telles que $l = l_1.l_2$.

Puisque l coïncide avec l_1 sur δ_1, que $\text{dom}(p) \subset \delta_1$ et que l vérifie p, l_1 vérifie p. Donc l_1 appartient à $\sigma_p(r_1)$ et, vu que l_2 appartient à r_2 et que l_1 et l_2 se correspondent, $l = l_1.l_2$ appartient à $\sigma_p(r_1) \bowtie r_2 = res_2$.

Deuxième inclusion

Soit l une ligne de res_2.

Il existe l_1 et l_2 appartenant respectivement à $\sigma_p(r_1)$ et r_2 telles que $l = l_1.l_2$.

Puisque l coïncide avec l_1 là où l_1 est définie donc en particulier sur δ_1 qui contient le domaine de p, et que l_1 vérifie p, l vérifie p.

Or l_1 appartient aussi à r_1 donc $l_1.l_2 = l$ appartient aussi à $r_1 \bowtie r_2$, donc à $\sigma_p(r_1 \bowtie r_2) = res_1$.

Sélection et agrégation

$$\text{group}_\delta \circ \sigma_p \equiv \sigma_p \circ \text{group}_\delta \qquad\qquad \text{si } \text{dom}(p) \subset \delta \tag{16}$$

Soit δ un ensemble d'attributs, p un prédicat et r une relation.
On pose $res_1 = \text{group}_\delta(\sigma_p(r))$ et $res_2 = \sigma_p(\text{group}_\delta(r))$.

Schémas relationnels

L'agrégation conserve les schémas relationnels, donc res_1 et res_2 ont tous les deux pour schéma δ_1.

Première inclusion

Soit l une ligne de res_1.
Pour simplifier les notations, on pose $r' = \sigma_p(r)$.
On pose $n = l|_\delta$ qui est donc le nom du groupe de r' pour δ auquel est associée l.
Il existe $l_1, \ldots, l_m$ des lignes de r' telles que $r'_n = \{l_1, \ldots, l_m\}$.

Toute ligne de r' appartient également à r donc (entre autres) $r'_n \subset r_n$. Montrons qu'on a également $r_n \subset r'_n$. Soit l_0 un élément de r_n. l_1 coïncide avec l_0 sur δ, donc en particulier sur $\text{dom}(p)$, et l_1 vérifie p, donc l_0 vérifie p et appartient à r'. Comme l_0 coïncide avec les l_i sur δ, l_0 appartient à r'_n. Ainsi, $r'_n \subset r_n$ et $r'_n = r_n$.

On en déduit que l appartient à $\text{group}_\delta(r)$. Or, l coïncide avec les l_i sur $\text{dom}(p) \subset \delta$, et ceux-ci vérifient p, donc l appartient à $\sigma_p(r') = res_2$.

Deuxième inclusion

Soit l une ligne de res_2.
l vérifie p et appartient à $\text{group}_\delta(r)$.
On pose $n = l|_\delta$. Il existe $l_1, \ldots, l_m$ des lignes de r telles que $r_n = \{l_1, \ldots, l_m\}$.

Les l_i coïncident avec l sur δ donc sur $\text{dom}(p)$ donc vérifient p.
Pour simplifier les notations, on pose $r' = \sigma_p(r)$. Les l_i vérifient p donc appartiennent tous à r', donc, vu qu'ils coïncident sur δ, $r_n = \{l_1, \ldots, l_m\} \subset r'_n$. Vu que toute relation de r' appartient aussi à r, on a aussi l'inclusion $r'_n \subset r_n$.

Ainsi $r_n = r'_n$ et on en déduit que $l = \text{lg}_{r,n} = \text{lg}_{r',n}$ appartient à $\text{group}_\delta(r') = res_1$.