Fiche de TD no 1 : maths et algèbre relationnelle

Soit $f : A \to B$ une fonction entre deux ensembles. For $\alpha \subseteq A$ et $\beta \subseteq B$,

$$f^{-1}(\beta) := \{\, a \in A \mid f(a) \in \beta \,\}, \quad \text{(image inverse)},$$
$$\exists_f(\alpha) := \{\, b \in B \mid \exists a \in \alpha, f(a) = b \,\}, \quad \text{(image directe)}.$$

Les images directe et inverse des fonctions satisfont ces propriétés : pour $f : A \to B$ et $g : B \to C$ deux fonctions, on a :

1. $\exists_{g\circ f} = \exists_g \circ \exists_g$,
2. $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$,
3. $\exists_f$ et f^{-1} sont croissantes[1],
4. $\exists_f(\alpha) \subseteq \beta$ ssi $\alpha \subseteq f^{-1}(\beta)$, pour tout $\alpha \subseteq A$ et $\beta \subseteq B$,
5. $f^{-1}(\bigcup_{i\in I} \beta_i) = \bigcup_{i\in I} f^{-1}(\beta_i)$ et $f^{-1}(\bigcap_{i\in I} \beta_i) = \bigcap_{i\in I} f^{-1}(\beta_i)$, pour toute collection $\{\, \beta_i \subseteq B \mid i \in I \,\}$,
6. $\exists_f(\bigcup_{i\in I} \alpha_i) = \bigcup_{i\in I} \exists_f(\alpha_i)$ et $\exists_f(\bigcap_{i\in I} \alpha_i) \subseteq \bigcap_{i\in I} \exists_f(\alpha_i)$, pour toute collection $\{\, \alpha_i \subseteq A \mid i \in I \,\}$,
7. proposez A, B, $f : A \to B$ et une collection $\{\, \alpha_i \subseteq A \mid i \in I \,\}$ tels que $\bigcap_{i\in I} \exists_f(\alpha_i) \not\subseteq \exists_f(\bigcap_{i\in I} \alpha_i)$.

Exercice 1. (A faire chez vous.) Vérifiez les propriétés listées ci-dessus de l'image inverse et de l'image directe.

Dans la suite, nous allons considérer deux ensembles A et D. On pense à A comme un ensemble de noms de colonne et à D comme à un domaine (ensemble de valeurs). D sera un domaine unique, commun à toutes les relations/tables et à leurs colonnes.
Nous utiliserons la notation D^X pour l'ensemble des fonctions de X vers D. On pense à une tuple comme une fonction $t : \alpha \to D$, où $\alpha \subseteq A$. On a donc $t \in D^\alpha$.

Définition. Une *relation* (ou table, non-nommée) est un couple (α, R) où $\alpha \subseteq A$ et $R \subseteq D^\alpha$.

Dans une relation (α, R), α est le schéma—ou « header »—de la relation, c'est-à-dire, l'ensemble des noms de ses colonnes, et R est l'ensemble des ses tuples.

Question 2. Supposons $D = A = \mathbb{N}$, $\alpha = \{\, 0, 1, 2 \,\}$. Representez sous forme de table la relation $(\alpha, \{\, f, g \,\})$ où $f(x) = 5 + x \mod 2$ et $g(x) = 6 + x \mod 3$.

Solution.

0	1	2
1	0	1
0	1	2

□

Pour $\beta \subseteq \alpha$ et $t \in D^\alpha$, nous utiliserons $t\restriction_\beta$ pour denoter la *restriction*[2] de la fonction t au sous-ensemble β du domaine α de t. Donc, pour $t \in D^\alpha$ et $\beta \subseteq \alpha$, $t\restriction_\beta \in D^\beta$.

Question 3. Pour $\beta \subseteq \alpha$, quel est le codomaine de la correspondance $t \mapsto t\restriction_\beta$? Et son domaine ? Quels sont les domaines et codomaines des images directe et inverse de cette fonction ?

Solution. Pour $\beta \subseteq \alpha$, soit ψ^α_β cette correspondance. (on a donc, pour $t \in D^\alpha$, $\psi^\alpha_\beta(t) = t\restriction_\beta$). Le domaine (l'ensemble des argument de cette fonction) de ψ^α_β est l'ensemble de tuples dont le colonnes sont α ; nous avons

1. Une fonction est croissante—ou monotone, en anglais—si elle préserve les inclusions.
2. Attention : on utilise ici le jargon du matheux, non celui de l'informaticien. On parle ici de restriction d'une fonction à un sous-ensemble de son domaine, non pas de restriction d'une table à un sous-ensemble de tuples satisfaisants une condition logique (clause `WHERE` de sql).

dénoté cet ensemble par D^α. Le codomaine (l'ensemble des valeur retournés par la fonction) est l'ensemble de tuples dont le colonnes sont β, c'est-à-dire D^β.
On a donc

$$\begin{aligned}\psi^\alpha_\beta &: D^\alpha \to D^\beta\\ \exists_{\psi^\alpha_\beta} &: P(D^\alpha) \to P(D^\beta)\\ (\psi^\alpha_\beta)^{-1} &: P(D^\beta) \to P(D^\alpha)\,.\end{aligned}$$

□

Considérons une table (α, R) et soit $\beta \subseteq \alpha$. La projection peut se définir ainsi :

$$\Pi_\beta(\alpha, R) = (\beta, \{\, t\restriction_\beta \mid t \in R \,\})\,.$$

Question 4. Quel est la relation entre la définition donnée ci-dessus et les opérateurs $\exists_f, f^{-1}$?

Solution. Si on regarde bien la definition, on a

$$\Pi_\beta(\alpha, R) = (\beta, \exists_{\psi^\alpha_\beta}(R))\,.$$

□

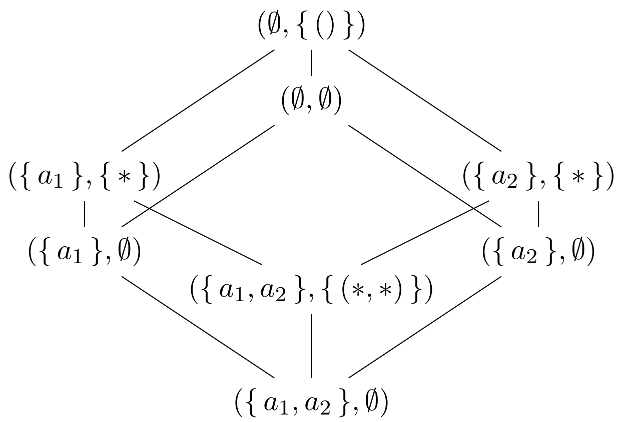
Posons, par la suite,

$$(\alpha_1, R_1) \leq (\alpha_2, R_2) \text{ ssi } \alpha_2 \subseteq \alpha_1 \text{ et } \{\, t\restriction_{\alpha_2} \mid t \in R_1 \,\} \subseteq R_2\,. \tag{1}$$

Exercice 5. Argumentez que $\leq$ ainsi défini est une relation d'ordre (réflexive, transitive, antisymétrique) sur les relations. Est ce qu'il existe une plus petite relation ? Et une plus grande relation ? Decrivez-les.

Exercice 6. Dessinez le diagramme de Hasse de cet ordre avec $A = \{\, a, b \,\}$ et $D = \{\, * \,\}$.

Solution. Voici :



On remarque que l'on retrouve un ensemble ordonné isomorphe à une algèbre de Boole avec trois atomes (ensemble des parties de $\{\, 1, 2, 3 \,\}$). □

Définition. La *réunion interne* de deux relations (α_i, R_i), $i = 1, 2$, se définit ainsi :

$$(\alpha_1, R_1) \vee (\alpha_2, R_2) := \Pi_{\alpha_1 \cap \alpha_2}(\alpha_1, R_1) \cup \Pi_{\alpha_1 \cap \alpha_2}(\alpha_2, R_2)\,.$$

Exercice 7. Montrez que $(\alpha_1, R_1) \vee (\alpha_2, R_2)$ est la borne sup de l'ensemble $\{\, (\alpha_1, R_1), (\alpha_2, R_2) \,\}$, par rapport à l'ordre sur les relations défini en (1).

Solution. Observons que :

$$(\alpha_1, R_1) \vee (\alpha_2, R_2) = (\, \alpha_1 \cap \alpha_2,\, \exists_{\psi^{\alpha_1}_{\alpha_1\cap\alpha_2}}(R_1) \cup \exists_{\psi^{\alpha_2}_{\alpha_1\cap\alpha_2}}(R_2)\,)\,.$$

Nous allons montrer que l'expression sur la droite donne une borne sup de l'ensemble $\{\, (\alpha_i, R_i) \mid i = 1, 2 \,\}$ par rapport à l'ordre défini en (1).
Observons d'abord que

$$\alpha_1 \cap \alpha_2 \subseteq \alpha_i$$

et

$$\{\, t\restriction_{\alpha_1\cap\alpha_2} \mid t \in R_i \,\} = \exists_{\psi^{\alpha_i}_{\alpha_1\cap\alpha_2}}(R_i) \subseteq \exists_{\psi^{\alpha_1}_{\alpha_1\cap\alpha_2}}(R_1) \cup \exists_{\psi^{\alpha_2}_{\alpha_1\cap\alpha_2}}(R_2)\,,$$

car l'image directe préserve les inclusions.
Soit (γ, S) un majorant de l'ensemble $\{\, (\alpha_i, R_i) \mid i = 1, 2 \,\}$, montrons que $(\alpha_1, R_1) \vee (\alpha_2, R_2) \leq (\gamma, S)$.
On a donc (pour $i = 1, 2$) $\gamma \subseteq \alpha_i$ et $\exists_{\psi^{\alpha_i}_{\gamma}}(R_i) \subseteq S$. Par conséquent, $\gamma \subseteq \alpha_1 \cap \alpha_2$ et $\exists_{\psi^{\alpha_1}_{\gamma}}(R_1) \cup \exists_{\psi^{\alpha_2}_{\gamma}}(R_2) \subseteq S$.
Nous avons ainsi

$$\begin{aligned}
\exists_{\psi^{\alpha_1\cap\alpha_2}_{\gamma}}(\, \exists_{\psi^{\alpha_1}_{\alpha_1\cap\alpha_2}}(R_1) \cup \exists_{\psi^{\alpha_2}_{\alpha_1\cap\alpha_2}}(R_2)\,) &= \exists_{\psi^{\alpha_1\cap\alpha_2}_{\gamma}}(\, \exists_{\psi^{\alpha_1}_{\alpha_1\cap\alpha_2}}(R_1)\,) \cup \exists_{\psi^{\alpha_1\cap\alpha_2}_{\gamma}}(\, \exists_{\psi^{\alpha_2}_{\alpha_1\cap\alpha_2}}(R_2)\,) \\
&= \exists_{\psi^{\alpha_1}_{\gamma}}(R_1) \cup \exists_{\psi^{\alpha_2}_{\gamma}}(R_2) \subseteq S\,,
\end{aligned}$$

car, pour $i = 1, 2$,

$$(\, \exists_{\psi^{\alpha_1\cap\alpha_2}_{\gamma}}\,) \circ (\, \exists_{\psi^{\alpha_1}_{\alpha_1\cap\alpha_2}}\,) = \exists_{(\,\psi^{\alpha_1\cap\alpha_2}_{\gamma}\,)\circ(\,\psi^{\alpha_1}_{\alpha_1\cap\alpha_2}\,)} = \exists_{\psi^{\alpha_i}_{\gamma}}\,.$$

□

Exercice 8. Donnez une définition formelle de la jointure naturelle en utilisant l'image inverse de la restriction et l'intersection :

$$(\alpha_1, R_1) \bowtie (\alpha_2, R_2) := \ldots$$

Solution.

$$\begin{aligned}
(\alpha_1, R_1) \bowtie (\alpha_2, R_2) &= (\alpha_1 \cup \alpha_2, \{\, t \mid t\restriction_{\alpha_1} \in R_1 \,\} \cap \{\, t \mid t\restriction_{\alpha_2} \in R_2 \,\}) \\
&= (\alpha_1 \cup \alpha_2, (\psi^{\alpha_1\cup\alpha_2}_{\alpha_1})^{-1}(R_1) \cap (\psi^{\alpha_1\cup\alpha_2}_{\alpha_2})^{-1}(R_2))\,.
\end{aligned}$$

□

Exercice 9. Montrez que $(\alpha_1, R_1) \bowtie (\alpha_2, R_2)$ est la borne inf de l'ensemble $\{\, (\alpha_1, R_1), (\alpha_2, R_2) \,\}$, par rapport à l'ordre sur les relations défini en (1).

Un treillis est un ensemble ordonné P tel que, pour tout couple $x, y \in P$ la borne inf $x \wedge y$ et la borne sup $x \vee y$ existent dans P. On a donc vu que l'ensemble des relations sur A et D (avec l'ordre défini en (1)) est un treillis.

Exercice 10. Montrez que le treillis des relations sur A et D ne satisfait pas la loi distributive

$$x \wedge (y_1 \vee y_2) = (x \wedge y_1) \wedge (x \wedge y_2)\,.$$

Idée : avec avec $A = \{\, a, b, c \,\}$ et $D = \{\, 0, 1 \,\}$, considérez

x	a	b	c
	1	1	1

y_1	a	b
	0	1

y_2	b	c
	1	0

Solution.

$x \wedge y_i$	a	b	c

$(x \wedge y_1) \vee (x \wedge y_2)$	a	b	c

$y_1 \vee y_2$	b
	1

$x \wedge (y_1 \vee y_2)$	a	b	c
	1	1	1

$\square$

Exercice 11. Considérez un ensemble partiellement ordonné $(P, \leq)$ tel que, pour tout sous-ensemble $X \subseteq P$, la borne inf $\bigwedge X$ existe. Montrez que

$$\bigwedge \{\, y \in P \mid \forall x \in X,\ x \leq y \,\}$$

est la borne sup de $X \subseteq P$.