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 $g(x) = 6 + x \mod 3$.

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 ?

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 $\exists_f, f^{-1}$?

Posons, par la suite,

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

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.

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).

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 6. 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).

Exercice 7. Donnez une définition formelle de la jointure naturelle en utilisant l'image inverse de la restriction et l'intersection :

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

Exercice 8. 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).

Question 9. Considérez un ensemble partiellement ordonné $(P, \leq)$ tel que, de 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 $X \subseteq P$.

Définissez une (sorte de) distance δ sur l'ensemble (de tuples sur toutes les colonnes) D^A de cette façon :

$$\delta(t, t') := \{\, a \in A \mid t(a) \neq t'(a) \,\}\,.$$

Exercice 10. Montrez que δ ainsi définie satisfait :

1. $\delta(t, t') = \delta(t', t)$,
2. $\delta(t, t) = \emptyset$,
3. $\delta(t, t') \subseteq \delta(t, s) \cup \delta(s, t')$.

pour tout $t, t', s \in D^A$.

Un sous-ensemble $X \subseteq A \uplus D^A$ (réunion disjointe de A et D^A) est *fermé* si $f \in X \cap D^A$ et $\delta(g, f) \subseteq X \cap A$ implique $g \in X$.

Exercice 11. Montrez que si, pour tout $i \in I$, $X_i \subseteq D^A$ est fermé, alors $\bigcap_{i \in I} X_i$ est aussi fermé. Concluez que, pour chaque collection $\{\, X_i \subseteq D^A \mid i \in I \,\}$ tel que tout X_i est fermé, il existe un plus petit ferme qui contient tous les X_i.

Exercice 12. Considérez la correspondance ψ définie par

$$(\alpha, R) \mapsto \psi(\alpha, R) := (A \setminus \alpha) \cup \{\, t \in D^A \mid t\restriction_\alpha \in R \,\}\,.$$

Montrez que

1. $\psi(\alpha, R)$ est fermé, pour toute relation (α, R),
2. $(\alpha_1, R_1) \leq (\alpha_2, R_2)$ ssi $\psi(\alpha_1, R_1) \subseteq \psi(\alpha_2, R_2)$ (en particulier, ψ est injective),
3. si $X \subseteq A \uplus D^A$ est fermé, alors il existe $\alpha \subseteq A$ et $R \subseteq D^\alpha$ tel que $\psi(\alpha, R) = X$.

Exercice 13. Montrez que :

1. pour $X_1, X_2 \subseteq A \uplus D^A$ fermés,
$$X_1 \cup X_2 \cup \{\, g \in D^A \mid \exists f \in X_1 \cup X_2, \delta(g, f) \subseteq X \cap A \,\}$$
est le plus petit fermé contenant X_1 et X_2.
2. le treillis des relations sur A et D ne satisfait pas la loi distributive
$$x \wedge (y_1 \vee y_2) = (x \wedge y_1) \wedge (x \wedge y_2)\,.$$