Fondements des mathématiques

Sylvain Poirier
sylvainpoirier.fr/logique

2. Constructions élémentaires

2.1. Uplets, familles

On appelle uplet (ou n-uplet) toute méta-fonction de domaine un méta-ensemble d'un nombre fini (n) de symboles de variables, à valeurs parmi les objets. C'est donc un système d'interprétations de ces variables. On peut les ajouter à toute théorie comme nouveau type d'objets, servant à abréger n variables de types préexistants en une seule du nouveau type. En pratique, pour chaque entier n on ne considèrera que les n-uplets de domaine le méta-ensemble $\mathcal{A}_n$ des n chiffres de 0 à $n-1$. Un 2-uplet s'appelle un *couple*, un 3-uplet est un *triplet*, un 4-uplet est un *quadruplet*...

La fixation du méta-argument curryfie l'évaluateur en n foncteurs appelés *projections*: les n-uplets sont séparément évalués en chaque $i \in \mathcal{A}_n$, par la i-ième projection π_i qui restitue la valeur de la i-ième variable. Les projections des couples seront notées π et π'.

Le définisseur de n-uplet est un opérateur n-aire (non liant), présentant ses n arguments dans une parenthèse et séparés par des virgules: $(\ ,\cdots, \ )$. Ces opérateurs sont reliés par les axiomes suivants dont le premier résume les $n+1$ autres : pour tous $x_0, \cdots, x_{n-1}$ et tout n-uplet x,

$$\begin{aligned} x = (x_0, \cdots, x_{n-1}) &\Leftrightarrow (\pi_0(x) = x_0 \wedge \cdots \wedge \pi_{n-1}(x) = x_{n-1}) \\ x_i &= \pi_i((x_0, \cdots, x_{n-1})) \quad \text{pour chaque } i \in \mathcal{A}_n \\ x &= (\pi_0(x), \cdots, \pi_{n-1}(x)) \end{aligned}$$

Les couples suffisent à construire les n-uplets pour tout $n > 2$. Par exemple les triplets $t = (x, y, z)$ peuvent se définir par $t = ((x, y), z))$ évalué par $x = \pi(\pi(t)), y = \pi'(\pi(t)), z = \pi'(t)$.

Connecteur conditionnel

Ce connecteur d'arité 3, noté $(\ \to \ | \ )$, qu'on lit "*Si A alors B sinon C*", se définit par

$$\begin{aligned} (A \to B|C) &\Leftrightarrow (\neg C \Rightarrow A \Rightarrow B) \Leftrightarrow ((A \Rightarrow B) \wedge (\neg A \Rightarrow C)) \Leftrightarrow (\neg A \to C|B) \Leftrightarrow (C, B)(A) \\ &\Leftrightarrow ((A \wedge B) \vee (\neg A \wedge C)) \Leftrightarrow ((C \Rightarrow A) \Rightarrow (A \wedge B)) \not\Leftrightarrow (A \to \neg B|\neg C) \end{aligned}$$

où $(C, B)(A)$ évalue le couple (C, B) en $A \in \mathbf{B} = \mathcal{A}_2 = \{0, 1\}$.

Il permet de réduire tout connecteur K d'arité $n+1$ en 2 connecteurs d'arité n:

$$K(A) \Leftrightarrow (A \to K(1)|K(0)).$$

Ainsi, $\neg A \Leftrightarrow (A \to 0|1)$; $(A \Rightarrow B) \Leftrightarrow (A \to B|1)$; $(A \vee B) \Leftrightarrow (A \to 1|B)$; $(A \Leftrightarrow B) \Leftrightarrow (A \to B|\neg B)$.

Familles

La théorie des ensembles peut représenter les n-uplets comme fonctions (réutilisant les outils des fonctions), en figurant $\mathcal{A}_n$ comme ensemble d'objets chacun désigné par une constante.

Une *famille* est une fonction interprétée comme généralisant la notion d'uplet: son argument est une variable, de domaine un ensemble d'objets éventuellement infini et/ou indéterminé (variable libre), mais vu comme un ensemble de méta-objets (simple, fixe, et extérieur au principal système étudié, comme dans la forme ensembliste d'une théorie quelconque). On parlera de "famille de..." pour préciser une classe d'arrivée, donc dire que son image est "un ensemble de...".

Le formalisme des familles est celui des fonctions déguisé à la ressemblance de celui des uplets (inapplicable faute de disposer d'une infinité de symboles). Ainsi la notation u_i, semblant celle d'un symbole variable de variable, sert d'évaluation de u en i (abrégeant $u(i)$ ou $\pi_i(u)$). Une famille définie par un terme t, se note $(t(i))_{i \in I}$ au lieu de $(I \ni i \mapsto t(i))$ ou de $(t(0), \cdots, t(n-1))$. L'argument i est appelé *indice*, et la famille est dite *indexée par I*. Une famille indexée par $\mathbb{N}$ est appelée une *suite*.

Structures et symboles liant une variable

Les prédicats n-aires $\mathcal{R}$ sont interprétables comme prédicats unaires $\mathcal{R}_1$ sur la classe des n-uplets $x = (x_0, \cdots, x_{n-1})$, par $\mathcal{R}_1(x) \Leftrightarrow \mathcal{R}(x_0, \cdots, x_{n-1})$. Ainsi les classes n-aires sont des classes de n-uplets. Tout opérateur n-aire T se traduit en foncteur T_1 sur une classe de n-uplets: $T_1(x) = T(x_0, \cdots, x_{n-1})$.

Remplaçant ces uplets par des familles, les symboles liants sont la généralisation des structures. Ainsi $\forall$ et $\exists$ généralisent les chaînes de $\wedge$ et de $\vee$: $(B_0 \wedge \cdots \wedge B_{n-1}) \Leftrightarrow (\forall i \in \mathcal{A}_n, B_i)$, rendant la condition d'égalité des couples $(x, y) = (z, t) \Leftrightarrow (x = z \wedge y = t)$ un cas particulier de celle des fonctions.

Soient $\mathcal{R}$ un prédicat unaire valide sur E, et C un booléen. On a les distributivités

$$\begin{aligned}(C \wedge \exists x \in E, \mathcal{R}(x)) &\Leftrightarrow (\exists x \in E, C \wedge \mathcal{R}(x))\\ (C \vee \forall x \in E, \mathcal{R}(x)) &\Leftrightarrow (\forall x \in E, C \vee \mathcal{R}(x))\\ (C \Rightarrow \forall x \in E, \mathcal{R}(x)) &\Leftrightarrow (\forall x \in E, C \Rightarrow \mathcal{R}(x))\\ ((\exists x \in E, \mathcal{R}(x)) \Rightarrow C) &\Leftrightarrow (\forall x \in E, \mathcal{R}(x) \Rightarrow C)\end{aligned}$$
$$(\exists x \in E, C) \Leftrightarrow (C \wedge E \neq \emptyset) \Rightarrow C \Rightarrow (C \vee E = \emptyset) \Leftrightarrow (\forall x \in E, C)$$

Ensembles finis, écriture en extension

Le foncteur Im, de domaine la classe des fonctions, définit un symbole liant ressemblant au symbole de compréhension et mixable avec celui-ci:

$$\{T(x)|x \in E\} = \mathrm{Im}(E \ni x \mapsto T(x))$$
$$\{T(x)|x \in E \wedge \mathcal{R}(x)\} = \{T(x)|x \in \{y \in E|\mathcal{R}(y)\}\}$$

L'image d'un uplet $(a, b, \cdots)$ se note $\{a, b, \cdots\}$, structure appelée l'*écriture en extension* de cet ensemble (qui énumère ses éléments). Ainsi l'ensemble $\mathcal{A}_n$ s'écrit $\{0, \cdots, n-1\}$. Ces images d'uplets sont des ensembles finis (la finitude sera formellement définie dans le texte 3).

On a déjà vu les cas d'arité 0 ($\emptyset$), 1 (singleton) et 2 (paire), dont les définitions (comme classe et comme domaine de quantificateur) traduisent celles de Im par traduction de $\forall$ et $\exists$ en $\wedge$ et $\vee$.

2.2. Autres operateurs sur les ensembles

Algèbre des parties, union et intersection d'une famille d'ensembles

Les connecteurs et les quantificateurs opérant entre booléens, définissent des opérations entre classes lorsque ces booléens dépendent d'une même variable libre. Et donc des opérations entre ensembles si ces classes sont des ensembles. Le résultat définit un ensemble s'il vaut faux pour un élément extérieur à tous les ensembles donnés. Lorsque ces ensembles sont des parties d'un ensemble E, le résultat peut se définir par compréhension dans E, mais reste indépendant de E.

C'est le cas de la constante 0 traduite par $\emptyset$, mais non du connecteur $\neg$.

On définit la différence $E \backslash F = \{x \in E | x \notin F\}$ de sorte que $x \in E \backslash F \Leftrightarrow x \in E \wedge x \notin F$.

Lorsqu'on étudie les parties d'un ensemble fixé E, la négation se traduit pour une partie F de E par $\complement_E F = E \backslash F$, appelé le *complémentaire de F dans E* : $\forall x \in E, x \in F \not\Leftrightarrow x \in \complement_E F$.

Les foncteurs sur les ensembles s'étendent aux familles grâce au foncteur Im. Ainsi se définissent l'union d'une famille d'ensembles, et l'intersection d'une famille non vide ($I \neq \emptyset$) d'ensembles:

$$\bigcup_{i \in I} F_i = \bigcup\{F_i | i \in I\}$$
$$x \in \bigcup_{i \in I} F_i \Leftrightarrow \exists i \in I, x \in F_i$$
$$(\forall x \in \bigcup_{i \in I} F_i, \mathcal{R}(x)) \Leftrightarrow \forall i \in I, \forall x \in F_i, \mathcal{R}(x)$$
$$x \in A \cup B \Leftrightarrow (x \in A \vee x \in B)$$
$$A \subset A \cup B = B \cup A$$
$$A = A \cup B \Leftrightarrow B \subset A \Leftrightarrow \ B = A \cap B$$

$$\forall j \in I, \quad \bigcap_{i \in I} F_i = \{x \in F_j | \forall i \in I, x \in F_i\}$$
$$x \in \bigcap_{i \in I} F_i \Leftrightarrow \forall i \in I, x \in F_i$$
$$A \cap B = \{x \in A | x \in B\}$$
$$(\forall x \in A \cap B, \mathcal{R}(x)) \Leftrightarrow (\forall x \in A, x \in B \Rightarrow \mathcal{R}(x))$$
$$x \in A \cap B \Leftrightarrow (x \in A \wedge x \in B)$$
$$A \cap B = B \cap A \subset A$$
$$(\exists x \in A, x \in B) \Leftrightarrow (\exists x \in B, x \in A) \Leftrightarrow A \cap B \neq \emptyset$$

Pour une famille de parties F_i de E, $\bigcap_{i \in I} F_i = \bigcap\{F_i | i \in I\} = \{x \in E | \forall i \in I, x \in F_i\} = \complement_E \bigcup_{i \in I} \complement_E F_i$.

Fixant le choix de E, l'intersection ainsi définie de la famille vide donne E.

Deux ensembles A et B sont dits *disjoints* lorsque $A \cap B = \emptyset$.

La paire et l'union binaire suffisent à définir toute écriture en extension de plus grande arité.

Enfin le connecteur $\not\Leftrightarrow$ se traduit par la différence symétrique: $E \Delta F = (E \cup F) \backslash (E \cap F)$.

L'union et l'intersection ont les mêmes propriétés d'associativité et de distributivité que $\wedge$ et $\vee$:

$$A \cup B \cup C = (A \cup B) \cup C = A \cup (B \cup C) = \bigcup\{A, B, C\}$$
$$(\bigcup_{i \in I} A_i) \cap C = \bigcup_{i \in I}(A_i \cap C) \qquad (\bigcap_{i \in I} A_i) \cup C = \bigcap_{i \in I}(A_i \cup C)$$
$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C) \qquad (A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

Produit fini

On appelle *graphe* tout ensemble de couples. On définit le produit (ou produit cartésien) $E \times F$ de deux ensembles E et F, comme l'ensemble des (x, y) où $x \in E$ et $y \in F$. De même, le produit de n ensembles $E_0 \times \cdots \times E_{n-1}$ est l'ensemble des n-uplets $(x_0, \cdots, x_{n-1})$ où $\forall i \in \mathcal{A}_n, x_i \in E_i$.

Un symbole L liant une variable x à un graphe G, joue le rôle de symbole liant deux variables y, z sur une structure $S(y, z)$, sous la forme $(Lx \in G, S(x_0, x_1))$ (intérprétant S comme structure à un argument $x = (x_0, x_1) \in G$), qu'on notera par un couple de variables: $L(y, z) \in G, S(y, z)$.

L'existence du produit (en toute arité) se justifie par le principe de génération des ensembles:

$$(\exists (x, y) \in E \times F, \mathcal{R}(x, y)) \Leftrightarrow (\exists x \in E, \exists y \in F, \mathcal{R}(x, y)) \Leftrightarrow (\exists y \in F, \exists x \in E, \mathcal{R}(x, y))$$
$$(\forall (x, y) \in E \times F, \mathcal{R}(x, y)) \Leftrightarrow (\forall x \in E, \forall y \in F, \mathcal{R}(x, y)) \Leftrightarrow (\forall y \in F, \forall x \in E, \mathcal{R}(x, y))$$
$$((\exists x \in E, \mathcal{A}(x)) \vee (\exists x \in E, \mathcal{B}(x)) \;\Leftrightarrow\; (\exists x \in E, \mathcal{A}(x) \vee \mathcal{B}(x))$$
$$((\forall x \in E, \mathcal{A}(x)) \wedge (\forall x \in E, \mathcal{B}(x)) \;\Leftrightarrow\; (\forall x \in E, \mathcal{A}(x) \wedge \mathcal{B}(x))$$
$$(\exists x \in E, \forall y \in F, \mathcal{R}(x, y)) \Rightarrow (\forall y \in F, \exists x \in E, \mathcal{R}(x, y))$$

Si $E \neq \emptyset$ alors

$$(\exists x \in E, C \vee \mathcal{A}(x)) \Leftrightarrow (C \vee \exists x \in E, \mathcal{A}(x))$$
$$(C \wedge \forall x \in E, \mathcal{A}(x)) \Leftrightarrow (\forall x \in E, C \wedge \mathcal{A}(x))$$

On abrègera $\forall x \in E, \forall y \in E, \mathcal{R}(x, y)$ en $\forall x, y \in E, \mathcal{R}(x, y)$, et de même pour $\exists$.

Une opération n-aire est une fonction sur un produit de n ensembles. De même les ensembles de n-uplets serviront de relations n-aires: une relation entre E et F sera un graphe $G \subset E \times F$. (On peut rappeler au besoin la donnée des domaines E et F en prenant le triplet (E, F, G).)

Somme ou union disjointe

On définit le transposé d'un couple ${}^t(x, y) = (y, x)$; celui d'un graphe ${}^tR = \{(y, x) | (x, y) \in R\}$.

La curryfication convertit tout graphe R en foncteur $\overrightarrow{R}$ arrivant parmi les ensembles. Inversement, la *somme* (ou *union disjointe*) d'une famille d'ensembles $(E_i)_{I \in I}$, la reconvertit en graphe (décurryfie):

$$\overrightarrow{R}(x) = \{y | (x', y) \in R \wedge x' = x\} \qquad (x, y) \in R \Leftrightarrow y \in \overrightarrow{R}(x) \Leftrightarrow x \in \overleftarrow{R}(y) = \overrightarrow{{}^tR}(y)$$
$$\coprod_{i \in I} E_i = \bigcup_{i \in I} \{(i, x) | x \in E_i\} \subset I \times \bigcup_{i \in I} E_i \qquad (i \in I \wedge x \in E_i) \Leftrightarrow (i, x) \in \coprod_{i \in I} E_i$$

On définit le graphe d'une fonction f par $\operatorname{Gr} f = \{(x, f(x)) | x \in \operatorname{Dom} f\} = \coprod_{x \in \operatorname{Dom} f} \{f(x)\}$.
Les notions de domaine et d'image s'étendent aux graphes ($\operatorname{Dom} f = \operatorname{Dom}(\operatorname{Gr} f)$ et $\operatorname{Im} f = \operatorname{Im}(\operatorname{Gr} f)$):

$$\overrightarrow{R}(x) \neq \emptyset \Leftrightarrow x \in \operatorname{Dom} R = \{x | (x, y) \in R\} = \operatorname{Im} {}^tR$$
$$R = \coprod_{i \in I} E_i \Leftrightarrow (\operatorname{Dom} R \subset I \wedge \forall i \in I, \overrightarrow{R}(i) = E_i) \Rightarrow ((\forall x \in R, A(x)) \Leftrightarrow (\forall i \in I, \forall y \in E_i, A(i, y)))$$
$$R_{|A} = \{(x, y) \in R | x \in A\} = \coprod_{x \in A} \overrightarrow{R}(x) \Rightarrow R_*(A) = \operatorname{Im} R_{|A} = \bigcup_{x \in A} \overrightarrow{R}(x)$$
$$\operatorname{Dom} R \subset A \Leftrightarrow R_{|A} = R \Rightarrow R_*(A) = \operatorname{Im} R = \{y | (x, y) \in R\}$$
$$R_*(\bigcup_{i \in I} A_i) = \bigcup_{i \in I} R_*(A_i) \qquad R_*(\bigcap_{i \in I} A_i) \subset \bigcap_{i \in I} R_*(A_i)$$
$$A \subset B \Rightarrow R_*(A) \subset R_*(B)$$
$$E \times F = \coprod_{x \in E} F$$
$$R \subset E \times F \Leftrightarrow (\operatorname{Dom} R \subset E \wedge \operatorname{Im} R \subset F)$$
$$E_0 \sqcup \cdots \sqcup E_{n-1} = \coprod_{i \in \mathcal{A}_n} E_i$$
$$E \times \emptyset = \emptyset = \emptyset \times E$$
$$(E \subset E' \wedge F \subset F') \Rightarrow E \times F \subset E' \times F'$$
$$(\forall i \in I, E_i \subset E'_i) \Rightarrow \coprod_{i \in I} E_i \subset \coprod_{i \in I} E'_i$$

2.3. Quantificateur d'unicité

Pour tous ensembles $F \subset E$, tout prédicat unaire $\mathcal{A}$ valide sur E, et tout $x \in E$,

$$\begin{aligned} x \in F &\Leftrightarrow \{x\} \subset F \Leftrightarrow (\exists y \in E, x = y \wedge y \in F) \Leftrightarrow (\forall y \in E, x = y \Rightarrow y \in F) \\ x \in F &\Rightarrow ((\forall y \in F, \mathcal{A}(y)) \Rightarrow \mathcal{A}(x) \Rightarrow \exists y \in F, \mathcal{A}(y)) \\ F \subset \{x\} &\Leftrightarrow (\forall y \in F, x = y) \Rightarrow ((\exists y \in F, \mathcal{A}(y)) \Rightarrow \mathcal{A}(x) \Rightarrow (\forall y \in F, \mathcal{A}(y))) \\ F = \{x\} &\Leftrightarrow (x \in F \wedge \forall y \in F, x = y) \Leftrightarrow (\forall y \in E, y \in F \Leftrightarrow x = y) \end{aligned}$$

Voici 3 nouveaux quantificateurs: $\exists 2$ (pluralité), ! (unicité), et $\exists!$ ("il existe un unique...”), qui appliqués à $\mathcal{R}$ dans E ne dépendent que de $F = \{x \in E, \mathcal{R}(x)\}$ (comme $\exists$ et contrairement à $\forall$) :

$$\begin{aligned} (\exists x \in E, \mathcal{R}(x)) &\Leftrightarrow (F \neq \emptyset) \Leftrightarrow (\exists x \in F, 1) \Leftrightarrow (\exists x \in E, \{x\} \subset F) \\ (\exists 2 x \in E, \mathcal{R}(x)) &\Leftrightarrow (\exists 2 : F) \Leftrightarrow (\exists x, y \in F, x \neq y) \Leftrightarrow (\exists x, y \in E, \mathcal{R}(x) \wedge \mathcal{R}(y) \wedge x \neq y) \\ (!x \in E, \mathcal{R}(x)) &\Leftrightarrow (!: F) \Leftrightarrow \neg(\exists 2 : F) \Leftrightarrow (\forall x, y \in F, x = y) \Leftrightarrow \forall x \in F, F \subset \{x\} \\ (\exists! x \in E, \mathcal{R}(x)) &\Leftrightarrow (\exists! : F) \Leftrightarrow (\exists x \in F, F \subset \{x\}) \Leftrightarrow (\exists x \in E, F = \{x\}) \\ F \subset \{x\} &\Rightarrow \forall y \in F, F \subset \{y\} \Leftrightarrow (!: F) \\ (\exists! : F) &\Leftrightarrow (F \neq \emptyset \wedge !: F) \\ F \neq \emptyset &\Rightarrow ((\forall y \in F, \mathcal{A}(y)) \Rightarrow (\exists y \in F, \mathcal{A}(y))) \\ (!: F) &\Rightarrow ((\exists y \in F, \mathcal{A}(y)) \Rightarrow (\forall y \in F, \mathcal{A}(y))) \\ F = \{x\} &\Rightarrow ((\exists y \in F, \mathcal{A}(y)) \Leftrightarrow \mathcal{A}(x) \Leftrightarrow \forall y \in F, \mathcal{A}(y)) \end{aligned}$$

Une fonction f est dite *constante* lorsque $!: \operatorname{Im} f$. La constance d'un uplet est la chaîne d'égalités: $x = y = z \Leftrightarrow !:\{x, y, z\} \Leftrightarrow ((x = y) \wedge (y = z)) \Rightarrow x = z$.

Traduction des opérateurs en prédicats

Dans une théorie générique, tout symbole de foncteur T est remplaçable par un symbole de prédicat R (où $xRy \Leftrightarrow (y = T(x))$) avec l'axiome $\forall x, \exists! y, xRy$, non dans les termes mais dans toute formule $A(T(x))$ (où x est un terme) par $(\exists y, xRy \wedge A(y))$, ou par $(\forall y, xRy \Rightarrow A(y))$.

De même, tout prédicat R tel que $\forall x, \exists! y, xRy$ définit implicitement un symbole d'opérateur T tel que $\forall x, y, T(x) = y \Leftrightarrow xRy$) (traduisible par R dans toute formule). Cela s'étend en arité supérieure.

En théorie des ensembles, ces règles briseraient la distinction entre énoncé et formule. A leur place, introduisons un nouvel opérateur ϵ, valide sur la classe $(\operatorname{Ens}(E) \wedge \exists! : E)$ des singletons, qui en extrait l'élément, suivant l'axiome $(\forall x, \epsilon\{x\} = x)$ qui s'écrit aussi $(\operatorname{Ens} E \wedge \exists! : E) \Rightarrow \epsilon E \in E$.

Alors pour tout prédicat unaire A, $A(\epsilon E) \Leftrightarrow (\exists x \in E, A(x)) \Leftrightarrow (\forall x \in E, A(x))$.

Opérateur conditionnel

Comme le connecteur de même nom, il choisit entre deux objets x, y suivant le booléen A:

$$(A \to x|y) = (y, x)_A = \epsilon\{z \in \{x, y\} | A \to z = x \,|\, z = y\}$$

de sorte que pour tout prédicat $\mathcal{R}, \mathcal{R}(A \to x|y) \Leftrightarrow (A \to \mathcal{R}(x)|\mathcal{R}(y))$. Combiné à des structures, il est le moyen naturel de définir tout autre para-opérateur structurant (comme celui traduisant les booléens en objets), rendant inutile l'inscription directe de ceux-ci au langage d'une théorie.

2.4. Propriétés des fonctions

On définit la *composition* des fonctions $f : E \to F$ et $g : F \to G$ par $g \circ f = (E \ni x \mapsto g(f(x)))$. De même avec $h : G \to H$, $h \circ g \circ f = (h \circ g) \circ f = h \circ (g \circ f) = (E \ni x \mapsto h(g(f(x))))$ et ainsi de suite.

La fonction *identité* sur un ensemble E se définit par $\operatorname{Id}_E = (E \ni x \mapsto x)$. La *restriction* d'une fonction f à $A \subset \operatorname{Dom} f$ est $f_{|A} = (A \ni x \mapsto f(x)) = f \circ \operatorname{Id}_A$ (donc $\operatorname{Gr}(f_{|A}) = (\operatorname{Gr} f)_{|A}$, et $f_{|\operatorname{Dom} f} = f$).

L'*image directe* et l'*image réciproque* d'un ensemble par $f : E \to F$, se définissent par

$$\begin{aligned} A \subset E \Rightarrow f[A] = f_*(A) = \operatorname{Im}(f_{|A}) = (\operatorname{Gr} f)_*(A) = \{f(x) | x \in A\} &\subset \operatorname{Im} f \\ f^*(B) = ({}^t\operatorname{Gr} f)_*(B) = \{x \in E | f(x) \in B\} &= \bigcup_{y \in B} f^\bullet(y) \\ f^\bullet(y) = (\overleftarrow{\operatorname{Gr}} f)(y) = f^*(\{y\}) &= \{x \in E | f(x) = y\} \end{aligned}$$

Si $A \subset F$ alors $f^*(\complement_F A) = \complement_E f^*(A)$. Pour toute famille d'ensembles $(A_i)_{i\in I}$,

$$f^*(\bigcap_{i\in I} A_i) = \bigcap_{i\in I} f^*(A_i).$$

Un graphe $R \subset E \times F$ est dit *fonctionnel* s'il est le graphe d'une fonction:

$$R = \mathrm{Gr}\, f \Rightarrow (\forall x \in E, ! : \overrightarrow{R}(x)) \Leftrightarrow (\forall x, y \in G, x_0 = y_0 \Rightarrow x_1 = y_1) \Rightarrow R = \mathrm{Gr}(\epsilon \overrightarrow{R})$$
$$\epsilon \overrightarrow{R} = ((\mathrm{Dom}\, R) \ni x \mapsto \epsilon \overrightarrow{R}(x))$$
$$(x, y) \in \mathrm{Gr}\, f \Leftrightarrow (x \in \mathrm{Dom}\, f \wedge y = f(x))$$

Pour tout $f : E \to F$, tout $R \subset E \times F$ et toute fonction g,

$$\mathrm{Gr}\, f \subset R \Leftrightarrow \forall x \in E, f(x) \in \overrightarrow{R}(x)$$
$$\mathrm{Gr}\, f \subset \mathrm{Gr}\, g \Leftrightarrow ((E \subset \mathrm{Dom}\, g) \wedge f = g_{|E})$$

On dit que f est *injective* (ou : une *injection*) et on note $\mathrm{Inj}\, f$ si ${}^t\mathrm{Gr}\, f$ est un graphe fonctionnel:

$$(\forall y \in F, ! : f^\bullet(y)) \;\Leftrightarrow\; (\forall x, x' \in E, f(x) = f(x') \Rightarrow x = x') \;\Leftrightarrow\; (\forall x, x' \in E, x \neq x' \Rightarrow f(x) \neq f(x'))$$

On définit alors son inverse par $f^{-1} = \epsilon f^\bullet = (\mathrm{Im}\, f \ni y \mapsto \epsilon f^\bullet(y))$ de sorte que $\mathrm{Gr}(f^{-1}) = {}^t\mathrm{Gr}\, f$.

Une fonction $f : E \to F$ est dite *bijective* (ou une *bijection*) de E sur F, et on note $f : E \leftrightarrow F$, si elle est injective et surjective: $\forall y \in F, \exists! : f^\bullet(y)$, auquel cas $f^{-1} : F \leftrightarrow E$.

Une *permutation* (ou *transformation*) est une bijection $f : E \leftrightarrow E$.

Proposition. *Soient deux fonctions $f : E \to F$ et $g : F \to G$.*
1. Si f et g sont injectives alors $g \circ f$ est injective.
2. Si $g \circ f$ est injective alors f est injective.
3. $\mathrm{Im}(g \circ f) = g[\mathrm{Im}\, f] \subset \mathrm{Im}\, g$.
4. $\mathrm{Im}(g \circ f) = G \Rightarrow \mathrm{Im}\, g = G$.
5. $\mathrm{Im}\, f = F \Rightarrow \mathrm{Im}(g \circ f) = \mathrm{Im}\, g$, de sorte que $((f : E \twoheadrightarrow F) \wedge (g : F \twoheadrightarrow G)) \Rightarrow (g \circ f : E \twoheadrightarrow G)$.

Preuve:
1. $(\mathrm{Inj}\, f \wedge \mathrm{Inj}\, g) \Rightarrow \forall x, y \in E, g(f(x)) = g(f(y)) \Rightarrow f(x) = f(y) \Rightarrow x = y$.
2. $\forall x, y \in E, f(x) = f(y) \Rightarrow g(f(x)) = (g(f(y)) \Rightarrow x = y$.
3. $\forall z \in G, z \in \mathrm{Im}(g \circ f) \Leftrightarrow (\exists x \in E, g(f(x)) = z) \Leftrightarrow (\exists y \in \mathrm{Im}\, f, g(y) = z) \Leftrightarrow z \in g[\mathrm{Im}\, f]$.
Puis, 3. $\Rightarrow$ 4. puis 3. $\Rightarrow$ 5. □

Proposition. $\forall g : F \to G, \mathrm{Inj}\, g \Rightarrow (\forall f, f' : E \to F, g \circ f = g \circ f' \Rightarrow f = f') \Rightarrow (\mathrm{Inj}\, g \vee E = \emptyset)$.

Preuve: $g \circ f = g \circ f' \Leftrightarrow \forall x \in E, g(f(x)) = g(f'(x)) \Rightarrow \forall x \in E, f(x) = f'(x) \Rightarrow f = f'$.
$\forall y, y' \in F, g(y) = g(y') \Rightarrow g \circ (x \mapsto y) = g \circ (x \mapsto y') \Rightarrow (\forall x \in E, y = y') \Leftrightarrow (y = y' \vee E = \emptyset)$. □

Proposition. *Soient $f : E \to F$ et $\mathrm{Dom}\, g = F$. Alors*

$$\begin{aligned} g \circ f = \mathrm{Id}_E \Leftrightarrow\;& (\forall x \in E, \forall y \in F, f(x) = y \Rightarrow g(y) = x) \\ \Leftrightarrow\;& (\mathrm{Gr}\, f \subset {}^t\mathrm{Gr}\, g) \;\Leftrightarrow\; (\forall y \in F, f^\bullet(y) \subset \{g(y)\}) \\ \Leftrightarrow\;& (\forall x \in E, f(x) \in g^\bullet(x)) \;\Leftrightarrow\; (\mathrm{Inj}\, f \wedge g_{|\,\mathrm{Im}\, f} = f^{-1}) \\ \Rightarrow\;& E \subset \mathrm{Im}\, g \end{aligned}$$

$$(g : F \to E \wedge g \circ f = \mathrm{Id}_E) \Rightarrow (\mathrm{Im}\, f = F \Leftrightarrow \mathrm{Inj}\, g \Leftrightarrow f \circ g = \mathrm{Id}_F \Leftrightarrow g = f^{-1}).$$

Preuve de la dernière formule: $(\mathrm{Inj}\, g \wedge g \circ f \circ g = g \circ \mathrm{Id}_F) \Rightarrow f \circ g = \mathrm{Id}_F$. □

Proposition. *1) Si f, g sont injectives et $\mathrm{Im}\, f = \mathrm{Dom}\, g$, alors $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*
2) Si $f, h : E \to F$ et $g : F \to E$ alors $(g \circ f = \mathrm{Id}_E \wedge h \circ g = \mathrm{Id}_F) \Leftrightarrow ((g : F \leftrightarrow E) \wedge f = h = g^{-1})$.

Preuve:
1) $\forall x \in \mathrm{Dom}\, f, \forall y \in \mathrm{Im}\, g, (g \circ f)(x) = y \Leftrightarrow f(x) = g^{-1}(y) \Leftrightarrow x = (f^{-1} \circ g^{-1})(y)$.
Autre méthode: $(g \circ f)^{-1} = (g \circ f)^{-1} \circ g \circ g^{-1} = (g \circ f)^{-1} \circ g \circ f \circ f^{-1} \circ g^{-1} = f^{-1} \circ g^{-1}$.
2) On déduit $f = h$ de $f = h \circ g \circ f = h$, ou de $\mathrm{Gr}\, f \subset {}^t\mathrm{Gr}\, g \subset \mathrm{Gr}\, h$; le reste est évident. □

Points fixes; fonctions idempotentes

L'ensemble des points fixes d'une fonction f est $\mathrm{Fix}\, f = \{x \in \mathrm{Dom}\, f | f(x) = x\} \subset \mathrm{Im}\, f$. On a
$\mathrm{Im}\, f \subset \mathrm{Fix}\, g \Leftrightarrow (\mathrm{Im}\, f \subset \mathrm{Dom}\, g \wedge g \circ f = f)$
$\mathrm{Im}\, f = \mathrm{Fix}\, f \Leftrightarrow (\mathrm{Im}\, f \subset \mathrm{Dom}\, f \wedge f \circ f = f)$: une telle fonction f est dite *idempotente*.

2.5. L'axiome des parties; produit et puissance

Voici trois nouveaux opérateurs entre ensembles, désignant certaines classes $\mathcal{R}$ comme ensembles K avec l'axiome ($\forall$ paramètres), $\forall x, x \in K \Leftrightarrow \mathcal{R}(x)$, mais non justifiables par le principe de génération des ensembles (les $\forall_{\mathcal{R}}$ n'étant pas traduisibles en formules ensemblistes).

Le formalisme traditionnel de la théorie ZF garde ces K hors du langage, liés par $(\exists K)$, laissant les $\forall_K$ comme quantificateurs ouverts $\forall_{\mathcal{R}}$, et les autres usages de K abréger $(\exists K, (\forall x, x \in K \Leftrightarrow \mathcal{R}(x)) \wedge \ldots)$ contenant aussi $\forall_{\mathcal{R}}$. Notre approche distinguant énoncés et formules nécessite l'ajout de ces opérateurs au langage ensembliste fondamental.

Axiome des parties. $\forall_{\mathrm{Ens}} E, \forall F, F \in \mathcal{P}(E) \Leftrightarrow (\mathrm{Ens}(F) \wedge F \subset E)$.

On abrégera aussi $\in\mathcal{P}$ par $\subset$ dans les symboles liants: $(\forall A \subset E, \cdots) \Leftrightarrow (\forall A \in \mathcal{P}(E), \cdots)$.

Théorème de Cantor. *Aucune fonction $f : E \to \mathcal{P}(E)$ ne peut être surjective.*

Preuve: $F = \{x \in E | x \notin f(x)\} \Rightarrow (\forall x \in E, x \in F \not\Leftrightarrow x \in f(x)) \Rightarrow (\forall x \in E, F \neq f(x)) \Rightarrow F \notin \mathrm{Im}\, f$. □

(Le paradoxe de Russell peut se voir comme un cas particulier).

L'axiome des parties identifiant $\mathcal{P}(E)$ à la classe des parties de E, veut signifier que cette classe est un ensemble, donc fixe si l'univers croît. Donc supposer l'univers assez grand pour contenir toutes les parties possibles de E, et donc le "vrai" $\mathcal{P}(E)$. Mais formellement, sa détermination de $\mathcal{P}(E)$ dépend de l'univers. Aucun axiome ne peut exclure l'existence de parties de $\mathbb{N}$ (ou autre E infini) qui échappent à l'univers étudié (pouvant appartenir à un univers plus large avec son foncteur $\mathcal{P}$ de valeur différente sur le même E). En effet, la preuve du théorème de complétude donne à toute théorie axiomatique consistante un modèle fait d'objets extérieurement numérotables (une construction plus fine peut en plus lui donner le "vrai" $\mathbb{N}$). Cette méta-numérotation épuise le $\mathcal{P}(\mathbb{N})$ interne mais ne peut pas épuiser le $\mathcal{P}(\mathbb{N})$ externe d'après le théorème de Cantor, ce qui prouve leur différence.

Ensemble puissance. *Pour tous ensembles E et F, la classe des fonctions de E dans F est un ensemble noté F^E: pour tout f, $f \in F^E \Leftrightarrow f : E \to F$.*

Produit d'une famille d'ensembles. *On systématise les produits finis en un symbole liant unique applicable à toute famille d'ensembles même infinie:*

$$\forall x, x \in \prod_{i \in I} E_i \Leftrightarrow (\mathrm{App}(x) \wedge \mathrm{Dom}\, x = I \wedge \forall i \in I, x_i \in E_i).$$

Pour tout $i \in I$ on appelle *i-ième projection*, la fonction π_i de $\prod_{i \in I} E_i$ dans E_i qui évalue toute famille x en i : $\pi_i(x) = x_i$. C'est l'évaluateur de fonction vu comme curryfié dans l'ordre inhabituel.

Ces trois opérateurs sont "équivalents" (définissables les uns par les autres): notant $\mathbf{B} = \{0, 1\}$,

$$\mathcal{P}(E) = \{\{x \in E | f(x) = 1\} | f \in \mathbf{B}^E\}$$
$$F^E = \prod_{x \in E} F = \{\epsilon \overrightarrow{R} | R \subset E \times F \wedge \forall x \in E, \exists! : \overrightarrow{R}(x)\}$$
$$\prod_{i \in I} E_i = \{x \in (\bigcup_{i \in I} E_i)^I | \forall i \in I, x_i \in E_i\} = \{\epsilon \overrightarrow{R} | R \subset \prod_{i \in I} E_i \wedge \forall x \in E, \exists! : \overrightarrow{R}(x)\}$$

Même certains cas sont exprimables au moyen des outils précédents:

$$F^{\{a\}} = \{\{a\} \ni x \mapsto y | y \in F\} \qquad F^{\emptyset} = \{\emptyset\} \qquad \mathcal{P}(\{a\}) = \{\emptyset, \{a\}\}$$
$$F^{E \cup E'} = \{(E \cup E') \ni x \mapsto (x \in E \to f(x) | g(x)) | (f, g) \in F^E \times F^{E'}\}$$
$$(\exists i \in I, E_i = \emptyset) \Rightarrow \prod_{i \in I} E_i = \emptyset \qquad (\forall i \in I, \exists! : E_i) \Rightarrow \prod_{i \in I} E_i = \{(\epsilon E_i)_{i \in I}\}$$

et de même on peut formuler $\prod_{i \in I \cup J} E_i$.

Si $F \subset F'$ alors $\mathcal{P}(F) \subset \mathcal{P}(F')$, $F^E \subset {F'}^E$, et $(\forall i \in I, E_i \subset E'_i) \Rightarrow \prod_{i \in I} E_i \subset \prod_{i \in I} E'_i$.

Proposition. *Pour tout* $f : E \to F$,

$$\begin{aligned}\operatorname{Im} f = F &\Rightarrow \operatorname{Inj}(G^F \ni g \mapsto g \circ f) \Rightarrow (\operatorname{Im} f = F \vee !\colon G)\\ (\operatorname{Inj} f \wedge G \neq \emptyset) &\Rightarrow \{g \circ f | g \in G^F\} = G^E \Rightarrow (\operatorname{Inj} f \vee !\colon G)\\ (E \subset F \wedge G \neq \emptyset) &\Rightarrow \{g_{|E} | g \in G^F\} = G^E\\ (\operatorname{Inj} f \wedge E \neq \emptyset) &\Rightarrow \exists g \in E^F, g \circ f = \operatorname{Id}_E\end{aligned}$$

Preuves :
$\operatorname{Im} f = F \Rightarrow (g \circ f = g' \circ f \Leftrightarrow (\forall x \in E, g(f(x)) = g'(f(x))) \Leftrightarrow (\forall y \in F, g(y) = g'(y) \Leftrightarrow g = g'))$.
$\forall z, z' \in G, (y \mapsto z) \circ f = (y \mapsto (y \in \operatorname{Im} f \to z|z')) \circ f$ donc $\operatorname{Inj}(g \mapsto g \circ f) \Rightarrow (z = z' \vee \operatorname{Im} f = F)$.
$\forall z \in G, \forall h \in G^E, \operatorname{Inj} f \Rightarrow (F \ni y \mapsto (y \in \operatorname{Im} f \to h \circ f^{-1}(y)|z)) \circ f = h$.
$\forall x \in E, \forall z, z' \in G, \exists g \in G^F, g \circ f = (y \mapsto (y = x \to z|z'))$
$\Rightarrow \forall y \in E, (f(y) = f(x) \Rightarrow g(f(y)) = z \Rightarrow (y = x \vee z = z'))$
Les dernières formules se déduisent de la deuxième comme cas particuliers. □

Pour tout $f : E \to F$, notons $f^*_F : \mathcal{P}(F) \to \mathcal{P}(E)$ la fonction définie par f^*. Alors (par $G = \mathbf{B}$)

$$\begin{aligned}\operatorname{Inj} f &\Leftrightarrow \operatorname{Im} f^*_F = \mathcal{P}(E))\\ \operatorname{Im} f = F &\Leftrightarrow \operatorname{Inj} f^*_F\end{aligned}$$

La fonction $f_* = (\mathcal{P}(E) \ni A \mapsto f[A])$ a pour image $\mathcal{P}(\operatorname{Im} f)$ car $\forall B \subset \operatorname{Im} f, f[f^*(B)] = B$.

La plupart des mathématiques courantes (dont la physique) peuvent se fonder sur l'*Arithmétique du second ordre* n'accepte que $\mathbb{N}$ et $\mathcal{P}(\mathbb{N})$, refusant $\mathcal{P}(\mathcal{P}(\mathbb{N}))$. En effet $\mathcal{P}(\mathbb{N})$ permet de spécifier $\mathbb{N}$ (par la propriété "toute partie non vide a un plus petit élément") et de construire $\mathbb{R}$. Les suites récurrentes sur E peuvent se définir au moyen de $\mathcal{P}(\mathbb{N} \times E)$ ou de $E^{\mathbb{N}}$, or $\mathcal{P}(\mathbb{N})^{\mathbb{N}}$ est représentable par $\mathcal{P}(\mathbb{N})$. Puis l'utilité de $\mathcal{P}$ s'estompe après quelques $\forall X \subset \mathcal{P}(\mathbb{N})$ (d'usage souvent contournable); seulement pour interpréter une théorie dans un univers $\mathcal{U}$, une méta-théorie utilise en gros $\mathcal{P}(\mathcal{U}^{\mathbb{N}})$; sans $\mathbb{N}$ on a besoin de $\mathcal{P}(\mathcal{P}(E))$ ou de E^E pour définir la finitude d'un ensemble E et reconstruire $\mathbb{N}$ si E est infini.

La relativité du sens de $\mathcal{P}$ qui rapporte en formules ensemblistes sur des objets donnés, des indécidabilités qui sinon ne seraient que celles d'énoncés (liées au flou sur l'étendue de l'univers), est la clé de bien des paradoxes des fondements des mathématiques. Ainsi s'explique comment la prouvabilité peut être indécidable: elle dépend de $\mathbb{N}$ (ou du concept de finitude) qui peut peut varier suivant le modèle car la caractérisation de $\mathbb{N}$ dépend de $\mathcal{P}(\mathbb{N})$ qui n'est pas définissable dans l'absolu. D'où l'intérêt de systèmes axiomatiques forts comme ZF enrichissant certains ensembles, de parties qui pourraient être ignorées d'univers plus petits; et donc résolvant certaines de ces indécidabilités.

Pour simplifier, notre théorie des ensembles fondatrice acceptera l'axiome des parties dans tous les cas, avec l'axiome de l'infini (existence de $\mathbb{N}$ ou d'un ensemble infini).

2.6. Propriétés des relations binaires sur un ensemble.

Une *relation binaire sur* E est une relation $R \subset E \times E$. Notant $x\ R\ y \Leftrightarrow (x, y) \in R$ et sous-entendant le domaine E des quantificateurs, elle sera dite

réflexive	$\Leftrightarrow$	$\forall x,\ x\ R\ x$
antiréflexive	$\Leftrightarrow$	$\forall x,\ \neg(x\ R\ x)$
symétrique	$\Leftrightarrow$	$R = {}^tR$
antisymétrique	$\Leftrightarrow$	$\forall x, y,\ (x\ R\ y \wedge y\ R\ x) \Rightarrow x = y$
transitive	$\Leftrightarrow$	$\forall x, y, z,\ (x\ R\ y \wedge y\ R\ z) \Rightarrow x\ R\ z$

Toute relation binaire transitive et antiréflexive est antisymétrique.

Un **préordre** est une relation binaire réflexive et transitive. C'est un **ordre** s'il est antisymétrique. Un ensemble E muni d'un préordre (resp. ordre) R (implicitement un couple (E, R)) est un *ensemble préordonné* (resp. *ordonné*), et alors

$$\forall x, y,\ x\ R\ y \Leftrightarrow \forall z, (z\ R\ x \Rightarrow z\ R\ y)$$

i.e. $x\ R\ y \Leftrightarrow \overleftarrow{R}(x) \subset \overleftarrow{R}(y)$.
Preuve: La transitivité se réécrit $\forall x, y,\ x\ R\ y \Rightarrow \forall z, (z\ R\ x \Rightarrow z\ R\ y)$. Puis, R étant réflexive, $\forall x, y,\ (\forall z, z\ R\ x \Rightarrow z\ R\ y) \Rightarrow (x\ R\ x \Rightarrow x\ R\ y) \Rightarrow x\ R\ y$. □

Une **relation d'équivalence** est un préordre symétrique. Alors $\forall x, y,\ x\ R\ y \Leftrightarrow \forall z, (z\ R\ x \Leftrightarrow z\ R\ y)$, i.e. $x\ R\ y \Leftrightarrow \overleftarrow{R}(x) = \overleftarrow{R}(y)$.

Preuve : $\forall x, y, x\ R\ y \Leftrightarrow (x\ R\ y \wedge y\ R\ x) \Leftrightarrow (\overleftarrow{R}(x) \subset \overleftarrow{R}(y) \wedge \overleftarrow{R}(y) \subset \overleftarrow{R}(x)) \Leftrightarrow (\overleftarrow{R}(x) = \overleftarrow{R}(y))$. □

Note. *Si R est réflexive et $\forall x, y, z, (x\ R\ y \wedge z\ R\ y) \Rightarrow z\ R\ x$ alors R est une relation d'équivalence.*

Preuve : on vérifie la symétrie: $\forall x, y, (x\ R\ y \wedge y\ R\ y) \Rightarrow y\ R\ x$. La transitivité en découle. □

Exemple. Soit $A \subset E^E$ et $R = \bigcup_{f\in A} \operatorname{Gr} f$. Alors

$$\begin{aligned}(\mathrm{Id}_E \in A) &\Rightarrow R \text{ réflexive}\\ (\forall f, g \in A, g\circ f \in A) &\Rightarrow R \text{ transitive}\\ (\forall f \in A, f : E \leftrightarrow E \wedge f^{-1} \in A) &\Rightarrow R \text{ symétrique}\end{aligned}$$

2.7. Bijections canoniques

Pour tous objets x, y on dira que x détermine y s'il existe un foncteur invariant T tel que $T(x) = y$. Alors on peut utiliser x dans le rôle de y, en remplaçant y par $T(x)$ dans chaque expression. Ce prédicat est un méta-préordre sur l'univers.

On appelle *bijection canonique* entre deux ensembles E et F une bijection définie par un foncteur invariant. Elle sera dite *bicanonique* si son inverse est canonique. On notera $E \cong F$ (resp. $E \equiv F$) s'il existe une bijection canonique (resp. bicanonique) de E sur F, pour sous-entendre le terme qui la définit. C'est un méta-préordre entre ensembles, préservé par constructions: par exemple si $E \cong E'$ et $F \cong F'$ alors $E \times F \cong E' \times F'$ et $F^E \cong F'^{E'}$. Il aura souvent une allure d'identité remarquable car l'existence d'une bijection entre ensembles finis implique l'égalité de leurs nombres d'éléments.

La transposition des couples ($E \times F \equiv F \times E$) s'étend aux graphes ($\mathcal{P}(E\times F) \equiv \mathcal{P}(F \times E)$) et aux opérations: $G^{E\times F} \equiv G^{F\times E}$ où on transpose $f \in G^{E\times F}$ suivant ${}^t f(x, y) = f(y, x)$.

Les bijections canoniques non bicanoniques viennent généralement de foncteurs invariants non injectifs: $\mathbf{B}^E \cong \mathcal{P}(E)$, $\{x\}^E \cong \{x\}$ et $E \times \{x\} \cong E$, tandis que $\{x\} \times E \equiv E^{\{x\}}$ et $E \equiv \{0\} \times E$.

Somme d'ensembles, somme de fonctions

Le foncteur $\coprod$, convertissant $(E_i)_{i\in I}$ en graphe $S = \coprod_{i\in I} E_i$, définit $\prod_{i\in I} \mathcal{P}(E_i) \cong \mathcal{P}(S)$. Son inverse $S \supset R \mapsto \overrightarrow{R}_I = (I \ni i \mapsto \overrightarrow{R}(i))$ dépend de I.

En particulier pour deux ensembles E et F on a $(\mathcal{P}(F))^E \cong \mathcal{P}(E \times F)$.

On définit la somme sur I des fonctions f_i où $\forall i \in I, E_i = \operatorname{Dom} f_i$ par

$$\coprod_{i\in I} f_i = (S \ni (i, x) \mapsto f_i(x))$$

$$f = \coprod_{i\in I} f_i \Leftrightarrow (\operatorname{Dom} f = S \wedge \forall i \in I, f_i = \overrightarrow{f}(i) = f \circ j_i) \quad \text{où } j_i = (E_i \ni x \mapsto (i, x))$$

D'où les bijections canoniques (bicanoniques si $I = \operatorname{Dom} S$, donc si $E \neq \emptyset$ dans le cas $I \times E$)

$$\prod_{i\in I} F^{E_i} \cong F^S \qquad \prod_{i\in I}\prod_{x\in E_i} F_{(i,x)} \cong \prod_{c\in S} F_c$$

$$(F^E)^I \cong F^{I\times E} \qquad (E\times F)\times G \equiv E \times F \times G$$

$$(f_i)_{i\in I} \in (F^E)^I \Rightarrow \coprod_{i\in I} \operatorname{Gr} f_i \subset \mathcal{P}(I \times (E \times F)) \equiv \mathcal{P}((I\times E)\times F) \supset \operatorname{Gr} \coprod_{i\in I} f_i$$

Produit de fonctions ou recurryfication

La transposition échange les deux manières $\overrightarrow{R}$ et $\overleftarrow{R}$ currifier une relation, donnant une bijection $(\mathcal{P}(F))^E \leftrightarrow (\mathcal{P}(E))^F$ de seul paramètre F (et E pour l'inverse, pour retrouver le domaine).

De même pour les opérations : si $I \neq \emptyset$, $(F^E)^I \cong (F^I)^E$; laissant F dépendre de $i \in I$,

$$\prod_{i\in I} F_i^E \cong (\prod_{i\in I} F_i)^E$$

Cette bijection, arrivant dans le produit des ensembles d'arrivée comme une somme de fonctions partait de la somme des domaines, se nomme *produit des fonctions $f_i \in F_i^E$*:

$$\prod_{i\in i} f_i = (E \ni x \mapsto (f_i(x))_{i\in I})$$

$$\forall f, \quad f = \prod_{i\in i} f_i \Leftrightarrow (\operatorname{Dom} f = E \wedge \forall i \in I, f_i = \pi_i \circ f)$$

$$\operatorname{Dom} f = \operatorname{Dom} g = E \Rightarrow \quad f \times g = (E \ni x \mapsto (f(x), g(x)))$$

$$I^E \times F^E \simeq (I \times F)^E$$

$$\coprod_{\phi\in I^E}\prod_{x\in E} F_{\phi(x)} \simeq (\coprod_{i\in I} F_i)^E.$$

2.8. Relations d'équivalence et partitions

Familles-partitions

Pour toute fonction $f : E \to F$, le foncteur $f^\bullet$ définit une famille $f^\bullet_F$ de parties de E indexée par F. En particulier on notera encore $f^\bullet$ abusivement pour désigner la famille de parties non vides $f^\bullet_{\mathrm{Im}\, f}$, appelée la *famille-partition de E associée à f*. De façon équivalente, une *famille-partition de E* est une famille de parties $A_i \subset E$ pour $i \in I$, non vides, deux à deux disjoints et dont l'union est E.

Preuve. Les A_i sont deux à deux disjoints ssi le graphe ${}^t\coprod_{i\in I} A_i$ est fonctionnel:

$$\begin{aligned}(\forall i, j \in I, i \neq j \Rightarrow A_i \cap A_j = \emptyset) &\Leftrightarrow \forall i, j \in I, \forall x \in E, (i \neq j \Rightarrow \neg(x \in A_i \wedge x \in A_j))\\ &\Leftrightarrow \forall x \in E, !i \in I, x \in A_i\end{aligned}$$

(On peut aussi noter que $x \in f^\bullet(y) \cap f^\bullet(z) \Leftrightarrow y = f(x) = z$).

Puis $(f : E \to I) \Rightarrow ((E = \bigcup_{i\in I} f^\bullet(i)) \wedge (\forall i \in I, i \in \mathrm{Im}\, f \Leftrightarrow f^\bullet(i) \neq \emptyset))$. □

Toute famille-partition est injective : $i \neq j \Rightarrow (A_i \cap A_j = \emptyset \neq A_i) \Rightarrow A_i \neq A_j$.

Relation d'équivalence associée à une fonction

Tout $f : E \to F$ définit une relation d'équivalence sur E par $\forall x, y \in E, x \underset{f}{\sim} y \Leftrightarrow f(x) = f(y)$, autrement dit $\underset{f}{\overleftarrow{\sim}} = f^\bullet \circ f = \underset{f}{\overrightarrow{\sim}}$. Ses propriétés de réflexivité, symétrie et transitivité sont évidentes; l'autre formule s'obtient ainsi: $(\mathrm{Dom}\, g = \mathrm{Im}\, f \wedge \mathrm{Inj}\, g) \Rightarrow (\forall x, y \in E, f(x) = f(y) \Leftrightarrow g(f(x)) = g(f(y)))$. Or $f^\bullet$ est injective. Donc $x \underset{f}{\sim} y \Leftrightarrow f^\bullet(f(x)) = f^\bullet(f(y)) \Leftrightarrow \underset{f}{\overleftarrow{\sim}}(x) = \underset{f}{\overleftarrow{\sim}}(y)$.

Théorème. $\forall f : E \twoheadrightarrow F, \forall g \in G^E, (\forall x, x' \in E, f(x) = f(x') \Rightarrow g(x) = g(x')) \Rightarrow \exists! h \in G^F, g = h \circ f$.

Preuve. Le graphe $H = \mathrm{Im}(f \times g)$ est fonctionnel : $\forall (y, z), (y', z') \in H, y = y' \Rightarrow z = z$.
Puis, $\mathrm{Dom}\, H = \mathrm{Im}\, f = F$. Enfin, $g = h \circ f \Leftrightarrow (\forall (y, z) \in H, z = h(y)) \Leftrightarrow H \subset \mathrm{Gr}\, h$. □

Finalement, $h \mapsto h \circ f$ est bijective de G^F vers $\{g \in G^E | \underset{f}{\sim} \subset \underset{g}{\sim}\}$, avec $\mathrm{Inj}\, h \Leftrightarrow (\underset{f}{\sim} = \underset{g}{\sim})$.

Partition, surjection canonique

Une *partition de E* est un ensemble de parties de E non vides, deux à deux disjoints et dont l'union est E. C'est l'image de toute famille-partition: $P = \mathrm{Im}\, f^\bullet = \mathrm{Im}(f^\bullet \circ f)$.

Pour toute relation binaire R sur E, si $P = \mathrm{Im}\, \overleftarrow{R}$ alors

$$(\forall x, y \in E, x \in \overleftarrow{R}(y) \Leftrightarrow \overleftarrow{R}(x) = \overleftarrow{R}(y)) \Leftrightarrow (\forall x \in E, \forall A \in P, x \in A \Leftrightarrow \overleftarrow{R}(x) = A) \Leftrightarrow \mathrm{Id}_P = \overleftarrow{R}^\bullet$$

donc si R est une relation d'équivalence alors P est une partition.

Réciproquement pour toute partition P de E, $\exists! R \subset E \times E, \mathrm{Id}_P = \overleftarrow{R}^\bullet$ donc $P = \mathrm{Dom}\, \overleftarrow{R}^\bullet = \mathrm{Im}\, \overleftarrow{R}$ donc R est une relation d'équivalence (qu'on peut définir par $x\, R\, y \Leftrightarrow (\exists A \in P, x \in A \wedge y \in A)$).

Il y a ainsi bijection canonique entre partitions de E et relations d'équivalence sur E.

La partition $\mathrm{Im}\, \overleftarrow{R}$ associée à une relation d'équivalence R sur E est appelée le *quotient de E par R* et notée E/R; et la fonction $\overleftarrow{R}$ est appelée *surjection canonique* de E sur E/R. Pour tout $x \in E$, l'élément $\overleftarrow{R}(x)$, unique élément de E/R contenant x, est appelé la *classe de x par R*.

Pour toute relation d'équivalence $R \subset \underset{g}{\sim}$ sur $\mathrm{Dom}\, g$, on note $g/R = \epsilon\, \mathrm{Im}(\overleftarrow{R} \times g) : E/R \twoheadrightarrow \mathrm{Im}\, g$, de sorte que $g = (g/R) \circ \overleftarrow{R}$.

Plus généralement si $R = \underset{f}{\sim}$ alors $f^\bullet : \mathrm{Im}\, f \leftrightarrow E/R$, peut faire jouer à $\mathrm{Im}\, f$ le rôle de E/R, avec f en guise de surjection canonique. Le nom de "quotient" vient de son rapport au produit, voyant la projection $E \times F \to E$ dans le rôle de surjection canonique si $F \neq \emptyset$.

2.9. Axiome du choix

Axiome du choix (AC). *Il s'énonce $\forall_{\mathrm{Ens}} X, \mathrm{AC}_X$, où AC_X désigne les énoncés équivalents*
1) Tout produit sur X d'ensembles non vides est non vide : $(\forall x \in X, A_x \neq \emptyset) \Rightarrow \prod_{x\in X} A_x \neq \emptyset$.
2) $\forall_{\mathrm{Ens}} E, \forall R \subset X \times E$, $(\forall x \in X, \exists y \in E, xRy) \Rightarrow (\exists f \in E^X, \forall x \in X, xRf(x))$.
3) Pour toute fonction $g : E \twoheadrightarrow X, \exists f \in E^X, g \circ f = \mathrm{Id}_X$.

1) $\Rightarrow$ 2) avec $A_x = \overrightarrow{R}(x)$; 2) $\Rightarrow$ 1) avec $R = \coprod_{x\in X} A_x$ et $E = \mathrm{Im}\, R$.
2) $\Rightarrow$ 3) par $R = {}^t\mathrm{Gr}\, g$; ou 1) $\Rightarrow$ 3) par $A_x = g^\bullet(x)$.
3) $\Rightarrow$ 2) : $\mathrm{Im}\, \pi_{|R} = \mathrm{Dom}\, R = X \Rightarrow \exists f \in R^X, \pi_{|R} \circ f = \mathrm{Id}_X$ donc $\forall x \in X, f(x) = (x, \pi'(f(x))) \in R$. □

Théorème. *Les énoncés suivants sont équivalents à l'axiome du choix:*
4) Pour tout ensemble E d'ensembles, $\emptyset \notin E \Rightarrow (\prod_{A\in E} A) \neq \emptyset$.
5) Pour toute partition P d'un ensemble E, $\exists K \subset E, \forall A \in P, \exists! : K \cap A$
6) Pour tous ensembles E, F, G et tout $g : F \twoheadrightarrow G, \{g \circ f | f \in F^E\} = G^E$.

Preuves: 1) $\Rightarrow$ 4) est évident ;
4) $\Rightarrow$ 5) : $(h \in \prod_{A\in P} A \wedge K = \operatorname{Im} h) \Rightarrow (K \subset E \wedge \forall A \in P, \{h(A)\} = K \cap A)$
car $x \in K \cap A \Rightarrow \exists B \in P, x = h(B) \in A \cap B$ donc $A = B$.
5) $\Rightarrow$ 3) : Soit $P = \operatorname{Im} g^\bullet$. Alors $f = (X \ni x \mapsto \epsilon(K \cap g^\bullet(x))) = g^{-1}_{|K} \Rightarrow g \circ f = \mathrm{Id}_X$.
AC_E2) $\Rightarrow$ 6) : $\forall h \in G^E, (\forall x \in E, \exists y \in F, g(y) = h(x)) \Rightarrow (\exists f \in F^E, \forall x \in E, g(f(x)) = h(x))$
AC_G3) $\Rightarrow$ 6) : $\exists i \in F^G, g \circ i = \mathrm{Id}_G \wedge \forall h \in G^E, i \circ h \in F^E \wedge g \circ i \circ h = h$.
6) $\Rightarrow$ 3) : $E = G \Rightarrow \mathrm{Id}_E \in \{g \circ f | f \in F^E\}$. □

Notes: 4) $\Rightarrow$ 1) est aussi facile : $\emptyset \notin \{A_i | i \in I\} = E$, puis $f \in \prod_{A\in E} A \Rightarrow (f(A_i))_{i\in I} \in \prod_{i\in I} A_i$.
Le 6) a une réciproque : $(\operatorname{Dom} g = F \wedge E \neq \emptyset \wedge \{g \circ f | f \in F^E\} = G^E) \Rightarrow \operatorname{Im} g = G$.
On verra que AC_X est démontrable si X est fini.

Un grand succès de la logique mathématique est la preuve (inabordable ici) de l'indépendance de l'axiome du choix : chaque univers où il est vrai en contient un autre où il est faux, et inversement. Dans tout univers, le sous-univers des objets “constructibles” (en un sens subtil formant un univers) ayant dans chaque ensemble non vide un “premier objet construit”, satisfait l'axiome du choix.

Les contre-exemples possibles à l'axiome du choix sont des familles d'ensembles dont une infinité sont dépourvus d'outil de choix. Ces derniers sont principalement les ensembles à plusieurs éléments purs et les ensembles infinis de parties inconstructibles de $\mathbb{N}$. Ainsi AC peut être faux sur la partition de $\mathcal{P}(\mathbb{N})$ formée des $\{B \subset \mathbb{N} | A \Delta B \text{ est fini}\}$ pour tous $A \subset \mathbb{N}$, même si $\mathrm{AC}_\mathbb{N}$ est vrai.

AC étant intuitif et aux conséquences généralement plus utiles que son contraire (à spécifier), est habituellement admis pour les questions qui en dépendent. Mais bien des questions n'en dépendent pas, ou peuvent se résoudre par un axiome plus faible comme $\mathrm{AC}_\mathbb{N}$. Contrairement à l'axiome des parties, l'axiome du choix ne sera pas nécessaire au présent exposé des fondements des mathématiques.

2.10. Notions sur les ensembles ordonnés

Pour toute relation binaire transitive R on note $x\ R\ y\ R\ z \Leftrightarrow ((x\ R\ y) \wedge (y\ R\ z)) \Rightarrow x\ R\ z$. Une relation d'ordre $x < y$ peut se lire “x est plus petit que y”, “x est inférieur à y”, “y est supérieur à x” ou encore “y est plus grand que x”.

Deux éléments x et y sont dits *incomparables* lorsque $\neg(x < y \vee y < x)$. (Ceci implique que $x \neq y$).

Pour tout préordre (resp. ordre) $R \subset E \times E$, sa restriction $R \cap (F \times F)$ à une partie $F \subset E$ est un préordre (resp. ordre) sur F. Ainsi toute partie d'un ensemble ordonné est ordonné.

Tout ensemble d'ensembles est naturellement ordonné par la relation d'inclusion.

Ordre quotient d'un préordre

La propriété caractérisant un ensemble préodonné (E, R), $\forall x, y \in E, x\ R\ y \Leftrightarrow \overleftarrow{R}(x) \subset \overleftarrow{R}(y)$ indique qu'il est modélisé à travers $\overleftarrow{R}$ par l'ensemble $\operatorname{Im} \overleftarrow{R}$ ordonné par $\subset$, avec

$$x \underset{\overleftarrow{R}}{\sim} y \Leftrightarrow (\overleftarrow{R}(x) \subset \overleftarrow{R}(y) \wedge \overleftarrow{R}(y) \subset \overleftarrow{R}(x)) \Leftrightarrow (x\ R\ y \wedge y\ R\ x)$$

de sorte que $\overleftarrow{R}$ est injective si et seulement si R est un ordre; et la relation $\subset$ dans $\operatorname{Im} \overleftarrow{R}$ est une copie de *l'ordre quotient du préordre* R sur l'ensemble $E/(R \cap {}^tR)$.

Sur chaque ensemble (ordonné) E, on n'utilisera généralement qu'un seul ordre noté $<_E$, ou abusivement $<$, ce qui serait justifiable en définissant les ensembles ordonnés comme ensembles d'ensembles, ordonnés par $\subset$, ignorant leurs éléments et autres parties.

fonctions croissantes, décroissantes, strictement croissantes

Entre deux ensembles ordonnés E et F, une fonction $f : E \to F$ sera dite :
— *croissante* ssi $\forall x, y \in E, x < y \Rightarrow f(x) < f(y)$
— *décroissante* ssi $\forall x, y \in E, x < y \Rightarrow f(y) < f(x)$
— *strictement croissante* ssi $\forall x, y \in E, x < y \Leftrightarrow f(x) < f(y)$
— *strictement décroissante* ssi $\forall x, y \in E, x < y \Leftrightarrow f(y) < f(x)$.

Toute composée de fonctions croissantes ou décroissantes, est croissante ou décroissante suivant que le nombre de fonctions décroissantes est pair ou impair.

Toute fonction strictement croissante ou strictement décroissante est injective.

Si $f \in F^E$ et $g \in E^F$ sont toutes deux croissantes (resp. toutes deux décroissantes) et $g \circ f = \mathrm{Id}_E$, alors f est strictement croissante (resp. strictement décroissante).

Ordre sur les ensembles de fonctions

Pour tous ensembles E, F où F est ordonné, l'ensemble F^E est ordonné par

$$f < g \Leftrightarrow (\forall x \in E, f(x) < g(x))$$

Alors, $\forall h \in E^G, f < g \Rightarrow f \circ h < g \circ h$, autrement dit $f \mapsto f \circ h$ est toujours croissante.

Si F et G sont ordonnés et $u \in G^F$ est croissante (resp. décroissante) alors $F^E \ni f \mapsto u \circ f \in G^E$ est croissante (resp. décroissante).

Dans tout ensemble ordonné E, une fonction $f \in E^E$ est dite *extensive* ssi $\forall x \in E, x < f(x)$, i.e. $\mathrm{Id}_E < f$. La composée de deux fonctions extensives est extensive.