Echanges

En réponse à l'Avis de Recherche n°12 de M. BOUTEILLER (Brive) : «Qui pourrait et voudrait bien expliquer, avec exemples, ce qu'est une catégorie opposée ou duale ?», Anne MICHEL-PAJUS (Paris) nous propose une explication très documentée et extrêmement complète.

CATÉGORIES et TOPOS : des univers à explorer

Anne MICHEL-PAJUS

Saint Cyr l'Ecole

Dans le *Bulletin* de décembre 1993, un lecteur lançait un «Avis de recherche» sur les catégories. Ce n'est peut-être pas un hasard, on en reparle beaucoup ces temps-ci. Il y a une vingtaine d'années, comme en témoignent des articles de ce même *Bulletin* [1], la théorie des catégories était avant tout perçue comme un outil permettant de dégager un certain nombre de notions communes à toutes les structures usuelles (groupes, espaces vectoriels, etc.), et donc de préciser la notion même de structure. Pour Bernard Charles, par exemple, elle représentait «un couronnement tout naturel de l'axiomatisation des mathématiques». Cependant, malgré le goût prononcé pour le formalisme qui marquait cette époque, ou peut-être à cause des excès commis en son nom, l'utilité de ce niveau d'abstraction supplémentaire était très contesté, même parmi les chercheurs.

Il me semble que dans l'approche actuelle, l'intérêt du cadre catégorique, par rapport au cadre ensembliste, réside plutôt dans sa capacité à

1 J.Fourastié. «Dans dix ans, la théorie des catégories remplacera-t-elle celle des ensembles ?» - *Bulletin 297, 1975*

B.Charles, A.Bouvier, J.Fourastié «Echanges : les catégories» *Bulletin 302, 1976.*

accepter et engendrer beaucoup plus d'objets. Par exemple, l'ensemble de tous les ensembles est exclu de la théorie des ensembles, mais une catégorie peut avoir comme collection d'objets celle de tous les ensembles (moyennant quelques précautions, quand même!). Une catégorie peut aussi recevoir des ensembles "forcés" comme ceux utilisés par Paul Cohen dans la théorie du forcing, qui a permis d'établir, en 1963 l'indépendance de l'hypothèse du continu [2]. Elle peut aussi rassembler des espèces d'ensembles "un peu mous" (qui n'ont rien à voir avec les ensembles flous), ce qui permet, entre autres, de définir des logiques "à plusieurs niveaux". Ces logiques, où la propriété de vérité dépend du lieu où l'on se place, peuvent fournir des modèles dans des domaines aussi divers que la psychologie, l'intelligence artificielle, ou le droit (avec la superposition des droits nationaux, européens, internationaux, etc.) [3]

Le volet géométrique est tout aussi intéressant, bien que plus difficile à saisir. Un topos peut être considéré comme un espace généralisé au sens de la topologie et de la géométrie algébrique (théorie des sites, des faisceaux) [4].

Pierre Cartier, lors des journées APMEP de 93 exprimait son bonheur devant le foisonnement d'objets qui ouvre à la recherche mathématique de nouvelles perspectives, en contraste avec la situation d'il y a une vingtaine d'années où tout semblait se clore, s'achever. Par exemple, la classification des groupes finis est terminée, mais on s'aperçoit que la structure de groupe n'épuise pas l'idée de symétrie. En soulignant le rôle des schémas de pensée catégoriciens, il comparait les mathématiques actuelles à une mare boueuse, d'où l'on ressort un poisson doré chaque fois que l'on plonge sa ligne...

C'est ainsi qu'en rentrant des journées de Poitiers, j'ai eu envie de donner

2) Rappelons que Gödel a montré en 1983 que l'hypothèse du continu n'est pas *contradictoire* avec l'axiomatique Zermelo-Frankel de la théorie des ensembles, et que Cohen a montré qu'elle n'est pas une *conséquence* de ces axiomes. La démonstration par les topos n'est pas dûe à Cohen, mais à Lawvere et Tierney. On la trouve au Chapitre VI de l'ouvrage cité en note 5.

3) Voici ce qu'on peut lire dans *Libération*, le 28 octobre 1993 :«[la convention de Schengen] impose aux Etats signataires de s'aligner sur la décision du premier Etat auquel le demandeur d'asile s'est adressé... Cependant une clause de cet accord laisse chaque Etat *libre de se conduire autrement* s'il le juge opportun pour des raisons de droit *interne*» (c'est moi qui souligne).

4) Les travaux de Grothendieck dans ce domaine en font le "père" de la théorie des topos. Il leur a donné ce nom qui signifie "lieu" en grec et s'écrit parfois "topoi" quand on utilise le pluriel grec. L'utilisation en logique a été imaginée par Lawvere et Tierney. Ces résultats ont émergé vers 1963. La notion de catégorie remonte à 1945, avec les travaux de Eilenberg et Mac Lane.

un aperçu intuitif, à titre culturel, de quelques concepts fondamentaux et de leurs applications possibles, pour ceux qui, comme moi, ignoraient cette théorie. [5]

Premières notions et exemples de base

Un **topos** est une catégorie qui possède certaines "**propriétés universelles**". On en trouvera quelques-unes décrites ci-dessous.
Par définition, une **catégorie** $\mathbb{C}$ est formée de deux collections : une collection d'**objets** et une collection de **flèches** (ou morphismes) qui relient certains objets. Sur ces collections sont définies 4 opérations vérifiant 4 axiomes.

Les opérations :

La première et la deuxième associent respectivement à chaque flèche f sa **source** $d_0(f)$ et son **but** $d_1(f)$, qui sont des objets de $\mathbb{C}$ (on écrit : $f : C \to D$ pour indiquer que $C = d_0(f)$ et $D = d_1(f)$).

La troisième opération associe à chaque objet C de $\mathbb{C}$ une flèche 1_C (ou id_C) qui est la flèche **identité** de C.

Enfin, la quatrième opération est la **composition** qui associe à tout couple de flèches (f,g) vérifiant $d_0(f) = d_1(g)$ leur composée $f \circ g$.

Les axiomes :

(i) pour tout objet C, $d_0(1_C) = C = d_1(1_C)$.

(ii) pour tout couple de flèches (f,g) vérifiant $d_0(f) = d_1(g)$, on a $d_0(f \circ g) = d_0(g)$ et $d_1(f \circ g) = d_1(f)$.

(iii) pour toute flèche f de source C et but D, $1_D \circ f = f$ et $f \circ 1_C = f$.

(iv) pour tout triplet de flèches telles que les deux membres de l'expression en cause soient définis, on a $(f \circ g) \circ h = f \circ (g \circ h)$.

Définition : la flèche $f : C \to D$ est appelée **isomorphisme** s'il existe une flèche $g : D \to C$ tel que $f \circ g = 1_D$ et $g \circ f = 1_C$.
Pour une catégorie donnée, on définit la catégorie **duale** en retournant les flèches (en échangeant source et but).

5) J'utilise essentiellement les notes prises aux cours donnés par René Guitart au Collège International de Philosophie en 92-93 et l'ouvrage de Saunders Mac Lane et Ieke Merdjik, *Sheaves in Geometry and Logic, a first introduction to Topos Theory*, Springer-Verlag, 1992.

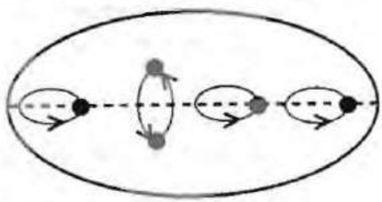
Exemples :

Tout ensemble muni d'un ordre partiel est une catégorie dont les objets sont les éléments, et les flèches les couples (p,q) : $p \to q$ ssi $p \leq q$, munis de la loi de composition définie par $(p,q) \circ (q,f) = (p,r)$.

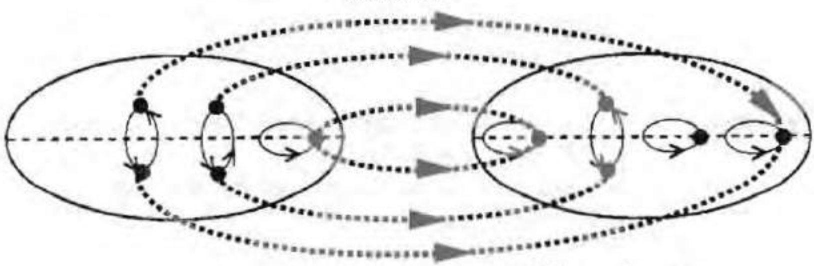
Moins évident : un groupe peut être considéré comme une catégorie à un seul objet et dont les flèches sont les éléments du groupe, munis comme loi de composition de la loi multiplicative du groupe.

Les objets de la catégorie $\mathbb{E}$ns sont les ensembles et ses flèches sont des applications munies de la composition usuelle. Si l'on considère des ensembles munis d'une structure (groupes, anneaux, corps,...), les flèches sont les morphismes (de groupes , d'anneaux, etc...)

Moins familier : une catégorie dont les objets ne sont pas des ensembles, mais des "miroirs". Les miroirs sont des collections d'éléments tous distincts sauf peut-être certains couples, dits en reflet. Plus mathématiquement, ce sont des couples (E,R), où E est un ensemble et R une involution sur E. Par cette involution, chaque élément de E est couplé avec un autre, éventuellement lui-même, appelé son reflet. Les flèches entre les deux miroirs (E,R) et (E',R') sont des applications de E dans E' compatibles avec le calcul des reflets. Ce qui se traduit, pour tout x élément de E, par $R'[f(x)] = f(R[x])$.



un miroir



l'ensemble des flèches en pointillé représente
une flèche de la catégorie des miroirs

On peut déployer sur les miroirs toutes les structures mathématiques. Une arithmétique, par exemple : On y distingue deux nombres correspondant à 2.

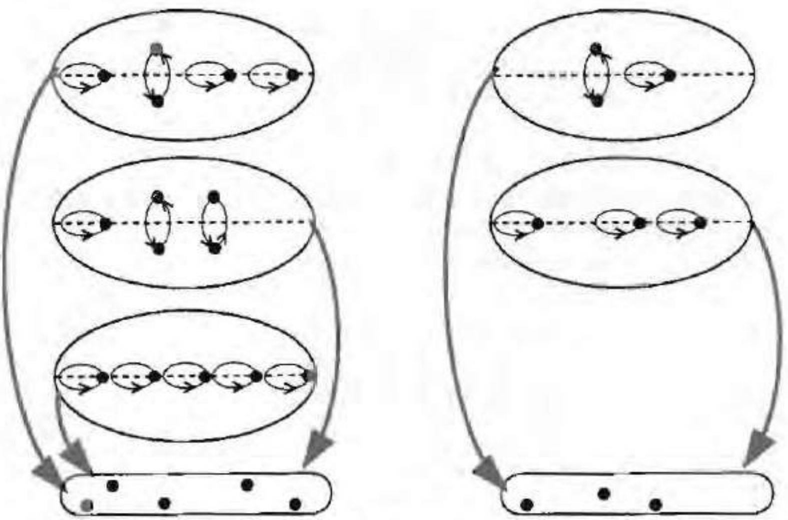
L'un, pour un couple de discernables, est désigné par 2, l'autre, pour la dyade d'éléments reflets l'un de l'autre, est désigné par D. Il existe alors deux nombres correspondant à 3 : 1 + 2 et 1 + D, etc... cette arithmétique est plus riche que l'arithmétique classique, dont elle redonne d'ailleurs les résultats quand on remplace D par 2.

Notion de foncteur

Bien entendu, il existe des "relations" entre les catégories : on les appelle des foncteurs. Puisque les catégories sont définies par deux collections, un **foncteur** F de $\mathbb{C}$ dans $\mathbb{D}$ est défini par un couple d'applications, l'une entre les objets des deux catégories, et l'autre entre les flèches, tout cela de façon à respecter source, but, identité et composition. C'est-à-dire :

- Si f est une flèche $A \to B$ de $\mathbb{C}$, $F(f)$ est une flèche $F(A) \to F(B)$ dans $\mathbb{D}$.
- Pour tout objet A de $\mathbb{C}$, $F(1_A) = 1_{F(A)}$
- Pour tout couple de flèches composable (f,g), $F(f \circ g) = F(f) \circ F(g)$.

Prenons l'exemple des foncteurs de la catégorie des miroirs vers celle des ensembles : le foncteur C dit "de coupure" associe au miroir (E,R) l'ensemble E de ses éléments et à une flèche de (E,R) dans (E',R') la même flèche de E dans E'.



Le foncteur C n'est évidemment pas injectif. L'image réciproque par C d'un ensemble E s'appelle sa fibre.

Cet exemple montre un peu comment une catégorie permet de prendre en compte la différence entre **distinct** et **discernable**, au sein d'une structure raisonnablement **cohérente**. Cette idée est développée dans l'exemple suivant.

La catégorie des ensembles empiriques [6]

Imaginons la situation suivante : j'observe les étoiles avec un instrument à réglage continu. Appelons X la collection des étoiles. Elle est assez mal déterminée : il y a celles que l'on ne voit pas, celles que l'on confond, et la liste d'étoiles dépend de l'instrument d'observation, de l'état du ciel qui se modifie avec le temps, des conditions atmosphériques, etc... Supposons que l'état de réglage de l'instrument soit caractérisé par la valeur d'un paramètre m décrivant $\mathbb{R}$ (on peut évidemment choisir plusieurs paramètres, l'état sera défini par un vecteur de $\mathbb{R}^r$). Nous pouvons rendre compte d'une observation en remplissant un grand tableau carré, où sont répertoriées en abscisse et ordonnée toutes les étoiles connues, et en indiquant dans la case correspondant à un couple (x,x') d'étoiles l'ensemble $V(x,x')$ des valeurs du paramètre pour lesquelles on voit x et x' confondues. Sur la diagonale, $V(x,x)$ est l'ensemble des valeurs de m pour lesquelles x est simplement vue.

Pour des raisons de continuité, on peut penser que si $m \in V(x,x')$ et si m' est assez voisin de m, on continuera à confondre x et x'. Il est donc raisonnable de prendre pour $V(x,x')$ un ouvert. Nous avons ainsi une "matrice d'observation" **V**, dont les coefficients sont des ouverts de $\mathbb{R}$, qui définit une application V de X × X dans Ouv($\mathbb{R}$), l'ensemble des ouverts de $\mathbb{R}$. Il est également raisonnable de poser en axiomes :
la symétrie : $\forall (x,x') \in X \times X,\ V(x,x') = V(x',x)$
la transitivité : $\forall (x,x',x'') \in X \times X \times X,\ V(x,x') \cap V(x',x'') \subset V(x,x'')$(encore que la transitivité trop répétée se discute, mais ne compliquons pas trop !).

Les deux axiomes se traduisent immédiatement sur **V**. La symétrie correspond à ${}^t\mathbf{V} = \mathbf{V}$ (${}^t\mathbf{V}$ représente la transposée de **V**). La transitivité équivaut à : $\forall (x,x') \in X \times X,\ V(x,x') = \cup \{V(x,x'') \cap V(x'',x')\ ;\ x'' \in X\}$ qui rappelle le produit de matrices, si l'on remplace $\cup$ par Σ et $\cap$ par $\times$. Nous adopterons désormais cette notation. Les relations ci-dessus se traduisent alors par $\mathbf{V}^2 = \mathbf{V}$. Tout cela conduit à la définition suivante :

On appelle ensemble empirique tout ensemble (X,V), où V est une application
$$V : X \times X \to \mathrm{Ouv}(\mathbb{R})$$
$$(x,x') \mapsto V(x,x')$$
symétrique et transitive au sens ci-dessus.

6) Tout ce paragraphe est inspiré de l'article de Jean Bénabou : «Rapports entre le Fini et le Continu», *Le labyrinthe du Continu*, Salanskis et Sinaceur, Springer-Verlag, 1992, p.178-189. On y trouve d'autres développements de cette théorie.

Une série d'observations conduit à une collection d'ensembles empiriques (X_i, V_i), caractérisés par leur matrice d'observation $\mathbf{V_i}$. Pour faire de notre collection d'ensembles empiriques une catégorie, il faut définir des flèches entre les (X_i, V_i). Une application quelconque f entre X_1 et X_2 ne conviendra pas toujours. Pour être cohérente avec les applications V_i, il est nécessaire et suffisant que $V_1(x_1,x'_1) \subset V_2(f(x_1),f(x'_1))$. Mais pour coller davantage à la situation de départ, il faut envisager la situation où $f(x_1)$ est confondue (et pas seulement égale) avec une étoile de X_2.

Soit $F(x_2,x_1) = V_2(x_2,f(x_1)) = \{m \in \mathbb{R}$ tels que $f(x_1)$ est vue confondue avec x_2 dans $X_2\}$. Posons d'autres axiomes raisonnables :

$\forall(x_1,x_2) \in X_1 \times X_2$, $F(x_2,x_1)$ est un ouvert (continuité)

$\forall(x_1,x_2) \in X_1 \times X_2$, $F(x_2,x_1) \cap V_1(x_1,x'_1) \subset F(x_2,x'_1)$ (compatibilité avec V_1)

$\forall(x_1,x_2) \in X_1 \times X_2$, $V_2(x'_2,x_2) \cap F(x_2,x_1) \subset F(x'_2,x_1)$ (compatibilité avec V_2)

$\forall(x_1,x_2) \in X_1 \times X_2$, $F(x_2,x_1) \cap F(x'_2,x_1) \subset V_2(x_2,x'_2)$ (f est une fonction)

$\forall x_1 \in X_1$, $V_1(x_1,x_1) \subset \cup\{F(x_2,x_1)\ ;\ x_2 \in X_2\}$ (f est partout définie sur X_1 : pour une valeur de m où x est vue dans X_1, il doit exister au moins un x_2 dans X_2 confondu avec $f(x_1)$.

Si l'on interprète F comme une matrice rectangulaire **F**, avec des éléments de X_1 en abscisse et ceux de X_2 en ordonnée en utilisant le "produit" de matrices défini ci-dessus, les quatre derniers axiomes sont respectivement équivalents à :

$$\boxed{\begin{array}{l} \mathbf{FV_1 = F} \\ \mathbf{V_2F = F} \\ \mathbf{F\,{}^tF \subset V_2} \\ \mathbf{V_1 \subset {}^tF} \end{array}}$$

($\mathbf{{}^tF}$ représente la transposée de **F** et l'inclusion signifie que chaque "coefficient" de la première matrice est inclus dans le coefficient correspondant de la deuxième).

Finalement, une flèche f de (X_1,V_1) dans (X_2,V_2) est caractérisée par la donnée d'une application F de $X_1 \times X_2$ dans Ouv($\mathbb{R}$), vérifiant les quatre formules ci-dessus (plus généralement, les deux premières formules caractérisent une relation entre X_1 et X_2). Soit $\mathcal{F}(X_1,X_2)$ l'ensemble des matrices **F** associées à ces applications. Il est facile de vérifier que ces axiomes donnent la structure de catégorie à la collection des ensembles empiriques. La composition des flèches correspond au "produit" des "matrices".

(i) $1_{(X,V)}$ correspond à l'application identique sur X. L'application F correspondante est V (car $\forall(x,x') \in X \times X$, $F(x,x') = V(x',f(x)) = V(x',x) = V(x,x')$).
(ii) Si $\mathbf{F} \in \mathcal{F}(X_1, X_2)$ et $\mathbf{G} \in \mathcal{F}(X_2, X_3)$, alors $\mathbf{GF} \in \mathcal{F}(X_1, X_3)$.
(iii) est traduit par $\mathbf{FV_1 = F}$ **et** $\mathbf{V_2F = F}$
(iv) correspond à l'associativité du produit de matrices.

On peut déployer dans cette catégorie toute une théorie similaire à celle de l'algèbre linéaire. Par exemple, l'article de J.Bénabou propose la définition des ensembles empiriques quotients et des parties d'un ensemble empirique. Il propose aussi une logique adaptée à ces ensembles. Soit P une proposition concernant ces ensembles. Posons Val(P) = {$m \in \mathbb{R}$ tels que, pour le réglage m, on puisse "voir" que P est vraie}. Val(P) est un ouvert. On obtient alors une **logique intuitioniste**, où l'on peut définir plusieurs **niveaux de vérité** : P est globalement vraie si Val(P) = $\mathbb{R}$, P est vraie du point de vue de m_0 si $m_0 \in$ Val(P), etc...

Logique classique et logique intuitioniste

Le calcul des propositions considère des propositions p, q, r, etc... connectées par les opérateurs et, ou, non, notés habituellement par $\wedge, \vee, \neg$ On peut aussi utiliser des sous-ensembles P, Q, R, d'un ensemble quelconque U fixé, dont les éléments sont notés x, et remplacer p par $x \in P$. Aux connecteurs correspondent alors respectivement l'intersection, l'union, le complémentaire. Les sous-ensembles, munis des trois connecteurs ci-dessus plus l'implication ($\Rightarrow$) qui doivent vérifier certaines égalités, constituent une algèbre de Boole.

Dans la logique intuitioniste, les sous-ensembles sont remplacés par les ouverts d'un espace topologique. L'interprétation des "et" et "ou" ne change pas, mais le complémentaire de P ouvert n'est pas forcément un ouvert. De même que le complémentaire d'une partie P est la plus grande partie Q disjointe de P, on associe à un ouvert P, pour la négation, le plus grand ouvert disjoint de P : l'intérieur du complémentaire de P. Pour interpréter l'implication $p \Rightarrow q$ on prend le plus grand ouvert W tel que $W \cap P \subset Q$. Moyennant certaines égalités à vérifier, les ouverts constituent une algèbre de Heyting. En particulier, dans une telle algèbre est vérifiée $p \Rightarrow \neg\neg p$ mais pas l'implication réciproque.

C'est Brouwer qui a créé la logique intuitioniste, en 1907. Son élève Heyting a publié dans les années 30.

Si nous reprenons le foncteur de coupure, cette fois entre la catégorie des ensembles empiriques et la catégorie Ens des ensembles habituels, celui à qui tout (X,V) associe X, il est possible de relier de façon cohérente une logique globale (la logique classique associée à la théorie des ensembles) et une ou des logiques locales. Je vais détailler ce point sur un autre exemple de catégorie, plus riche que celle des miroirs, mais plus simple que celle des ensembles empiriques.

Une logique des familles.

La logique classique s'appuie sur la théorie des ensembles (dans l'axiomatisation Zermelo-Frankel) dans laquelle les éléments sont distincts et parfaitement discernables. Nous avons vu que la théorie des catégories permet de considérer des objets qui seraient comme des ensembles un peu "mous", où certains éléments sont regroupés dans des n-uples. En vertu de ce regroupement, ils sont indiscernables quand on les regarde d'un certain point de vue. Il faut bien noter que cette ambiguïté et cette indiscernabilité n'ont rien à voir avec le "flou" ni au sens courant, ni au sens probabiliste du terme. En retraduisant les formules de la logique classique dans le langage propre à une catégorie donnée, il est possible de créer des logiques à plusieurs niveaux, où la notion de vérité devient **locale**.

Dans la théorie classique, on dispose d'un ensemble X décrit par les variables, et des formules qui prennent la valeur vrai (v) ou faux (f). Les formules déterminent donc des applications de X^n (si elles mettent en jeu n variables) dans un ensemble $\Omega = \{v,f\}$. Par exemple, soit la formule ".=.". Elle est définie sur X^2, par : ".=."(x,y), ou "$x = y$", qui prend la valeur v si et seulement si $x = y$.

Ensuite, à partir de la formule $\mathcal{F}$, on peut définir (non $\mathcal{F}$) sur X par : non $\mathcal{F}(x)$ prend la valeur v ssi $\mathcal{F}(x)$ prend la valeur f, et ainsi de suite pour ($\mathcal{F}$ ou $\mathcal{G}$), ($\mathcal{F}$ et $\mathcal{G}$), ($\mathcal{F} \Rightarrow \mathcal{G}$), etc. en utilisant des tables de vérité.

Nous allons interpréter ces formules dans une catégorie plus compliquée que celle des ensembles, ce qui donnera des outils plus fins. Prenons comme collection d'objets les flèches de $\mathbb{Ens}$ elles-mêmes. Soit u : F $\rightarrow$ G et u' : F' $\rightarrow$ G'. Pour garder des diagrammes commutatifs, on définit une flèche-flèche φ entre les flèches-objets u et u' par une double application φ : (f : F $\rightarrow$ F', g : G $\rightarrow$ G') telle que le diagramme commute, c'est-à-dire $g \circ u = u' \circ f$.

$$\begin{tikzcd} F \arrow[r, "u"] \arrow[d, Rightarrow, "f"'] \arrow[dr, phantom, "\Downarrow \phi"] & G \arrow[d, Rightarrow, "g"] \\ F' \arrow[r, "u'"'] & G' \end{tikzcd}$$

1_u est évidemment $(1_F\ ,\ 1_G)$. On se convaincra aisément que la composition définie par $\varphi \circ \varphi' = (f \circ f'\ ,\ g \circ g')$ est associative. Cette catégorie va nous servir à définir une logique à deux niveaux.

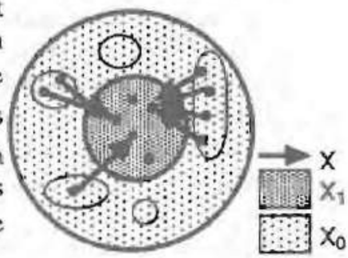
Prenons pour Ω une flèche-objet particulière : son ensemble d'arrivée est $\Omega_1 = \{v, f\}$ et son ensemble de départ est l'ensemble $\Omega_0 = \{V, V', F\}$, où V se lit "globalement vrai", V', "localement vrai" et F, "faux". L'application Ω est définie de Ω_0 dans Ω_1 par : $V \mapsto v$, $V' \mapsto v$ et $F \mapsto f$.
Précisons les tables de la négation et du "et" sur Ω_0 :

A	non A	non(non A)
V	F	V
V'	F	V
F	V	F

et	V	V'	F
V	V	V'	F
V'	V'	V'	F
F	F	F	F

Ces valeurs de vérité entrent dans le cadre de la logique intuitioniste.

Un objet quelconque de la catégorie est une application X d'un ensemble X_0 vers un ensemble X_1. Par exemple, supposons que les éléments de X_0 puissent être regroupés en familles, et que X associe un nom à chaque élément de X_0, l'ensemble des noms étant X_1. Le triplet $[X, X_0, X_1]$ caractérise une association découpant X_0 en familles.



Pour interpréter une formule logique $\mathcal{F}_1$ de la logique classique, on choisit un triplet $[X, X_0, X_1]$ et l'on complète par une application $\mathcal{F}_0$ telle que le diagramme ci-dessous soit commutatif.

$$\begin{tikzcd} X_0 \arrow[r, "\mathcal{F}_0"] \arrow[d, "X"'] & \Omega_0 = \{V, V', F\} \arrow[d, "\Omega"] \\ X_1 \arrow[r, "\mathcal{F}_1"'] & \Omega_1 = \{\ v,\ f\} \end{tikzcd}$$

Pour tout x_0 de X_0, on a donc $\mathcal{F}_1[X(x_0)] = \Omega[\mathcal{F}_0(x_0)]$.
Les variables de la formule $\mathcal{F}$ peuvent être localisées au niveau 0 (elles sont alors désignées par une minuscule) ou au niveau 1. Si elles sont choisies au niveau 0, la valeur de vérité de $\mathcal{F}$ peut être examinée au niveau 0 ou au niveau 1. Si elles sont prises au niveau 1, la valeur de vérité de $\mathcal{F}$ ne peut être

examinée qu'à ce niveau, et nous retrouvons la logique classique.

Prenons par exemple ".=." pour la formule $\mathcal{F}_1$. Soient x_0 et y_0 des éléments de X_0.

Au niveau 1, niveau global :

"$(X(x_0) = X(y_0)$"$_1$ = v si $X(x_0) = X(y_0)$ (ce qui correspond pour les variables du niveau 0 à $[x_0 = y_0$ ou $(x_0 \neq y_0$ et $X(x_0) = X(y_0))])$

= f si $X(x_0) \neq X(y_0)$ (ce qui correspond à $[x_0 \neq y_0$ et $X(x_0) \neq X(y_0)\,])$

Au niveau 0, niveau local :

"$x_0 = y_0$ "$_0$ = V si $x_0 = y_0$,

= V' si $x_0 \neq y_0$ et $X(x_0) = X(y_0)$

= F si $x_0 \neq y_0$ et $X(x_0) \neq X(y_0)$.

Au niveau 1, niveau général, on peut considérer que deux individus distincts mais appartenant à une même famille sont indiscernables, alors qu'au niveau local, ils sont indiscernables quelles que soient leurs familles respectives.

Un regard sur un auteur classique

Pour échapper au vertige des empilements symboliques, mais pas au trouble de l'ambiguïté, faisons une excursion syntaxique :

«[esprit de géométrie :] il faudrait avoir tout à fait l'esprit ***faux*** *pour* ***mal*** *raisonner sur des principes si gros qu'il est presque* ***impossible*** *qu'ils* ***échappent.*** *[...] [esprit de finesse:] les principes sont si déliés et en si grand nombre qu'il est presque* ***impossible*** *qu'il* ***n****'en* ***échappe****.[...] Il y a un certain modèle d'agrément [...]. Tout ce qui* ***n'est point*** *fait sur ce modèle* ***déplaît*** *à ceux qui ont le goût bon.[...] On* ***n'****apprend* ***pas*** *aux hommes à être honnêtes hommes et on leur apprend tout le* ***reste ;*** *et ils* ***ne*** *se piquent* ***jamais*** *tant de savoir rien du* ***reste*** *comme d'être honnêtes hommes. Ils* ***ne*** *se piquent de savoir* ***que*** *la seule chose qu'ils* ***n'****apprennent* ***point****.[...] [l'imagination], cette maîtresse d'****erreur*** *et de* ***fausseté****, et d'autant plus* ***fourbe*** *qu'elle* ***ne*** *l'est* ***pas*** *toujours ; car elle serait règle* ***infaillible*** *de vérité, si elle l'était* ***infaillible*** *du* ***mensonge****.»* [7]

L'accumulation des négatifs est essentielle à la transmission de la pensée de Pascal. L'usage des doubles négations, joint à des "affaiblisseurs" (conditionnels, presque) montre bien que l'auteur hésite à énoncer brutalement une propriété (P), mais préfère la remplacer par la proposition plus faible $\neg\neg P$. A la place de la règle de la logique classique : $\neg\neg P \Leftrightarrow P$, l'auteur utilise plutôt

7) Découpé et souligné par moi.

ici la règle : $P \Rightarrow \neg\neg P$. Il se place dans une logique de **persuasion** ou d'**inscription**. Or, nous savons bien que dans d'autres situations, comme le fameux texte du pari, ou ses travaux mathématiques, Pascal utilise rigoureusement la logique classique. Il se situe alors dans une logique de **décision.**
On passe d'une logique à l'autre, comme de la catégorie **Fam** à **Ens**, par le foncteur de coupure.

Un autre regard sur une théorie classique : les groupes de Galois.

Que l'on rencontre de l'indiscernable et de l'ambiguïté dans le discours, le droit ou la psychologie ne surprendra personne, mais en mathématiques ? Prenons un problème classique : peut-on exprimer dans un langage algébrique donné, celui des radicaux par exemple, les racines d'une équation ? L'idée géniale de Lagrange, mise au point par Galois, est qu'il faut s'intéresser aux permutations qui laissent invariantes certaines relations entre les racines [8]. Ce qui signifie que, vis-à-vis de ces relations, les racines sont indiscernables. Le rôle que joue la notion d'ambiguïté n'est pas une reconstruction historique de ma part. A la fin de sa fameuse lettre-testament du 29 Mai 1832, Evariste Galois l'écrit explicitement : *«...Mes principales méditations, depuis quelques temps, étaient dirigées sur l'application de l'analyse transcendante de la théorie de l'ambiguïté. Il s'agissait de voir a priori, dans une relation entre des quantités ou fonctions transcendantes, quels échanges on pouvait faire, quelles quantités ou pouvait substituer aux quantités données, sans que la relation pût cesser d'avoir lieu...»* Le texte intégral se trouve dans une brochure de l'APMEP [9], suivi d'un commentaire mathématique d'Amy Dahan-Dalmedico, dont j'extrais cette phrase : *«Ainsi Galois se sert du groupe d'une équation comme d'un miroir dans lequel se reflètent les difficultés de résolution de celles-ci. Le groupe permet de mesurer ce que Verriest [en 1897] a par la suite appelé "l'indiscernable" des racines sur le corps»* [10]

Plus précisément, le groupe de Galois rend compte des limites du langage algébrique réduit à l'écriture des nombres entiers, de lettres pour désigner des quantités inconnues et des quatre opérations élémentaires entre ces objets. Ce langage ne permet d'écrire que des polynômes. Un polynôme permet de désigner l'ensemble de ses racines. Pour discerner ***une*** racine, il faut

---

8) Le groupe de Galois d'une équation est le groupe formé par ces permutations;

9) *Présence d'Evariste Galois,* Brochure APMEP n°48, 1982, p.35.

10) Op.cit. p.47

une famille de polynômes dont elle soit la seule racine commune. La mesure de cette impossibilité, c'est-à-dire de l'indiscernabilité algébrique qui règne au sein des racines, est donnée par le groupe de Galois de l'équation.

Citons également Dieudonné : *«En donnant l'exemple d'une manière canonique d'associer, à des objets mathématiques d'une certaine nature, comme les corps, des objets d'une autre nature, comme les groupes, [la théorie de Galois] a servi de modèle dans des théories diverses (revêtements, équations différentielles algébriques, Théorie des nombres algébriques) et il n'est peut-être pas exagéré d'y voir le premier exemple de la notion de* foncteur.» [11]

La prise en compte de ce fait négatif qu'est l'indiscernabilité a pour effet de créer un langage plus vaste, qui va permettre, non seulement de conduire la résolution complète par radicaux quand elle est possible, ou de déterminer quand cette résolution est impossible, mais de construire des concepts très féconds dans d'autres domaines.

On peut observer une situation du même genre avec les théories de la mécanique : des systèmes en translation uniforme sont indiscernables du point de vue des lois de la physique classique. L'ambiguïté qui en résulte peut être cernée par le groupe de Galilée. Dans le cadre de la relativité restreinte, ce rôle est joué par le groupe de Lorentz. Pour être acceptable, une loi de la physique doit avoir une expression invariante par ce groupe.

Plus généralement, dès que l'on se place à l'intérieur d'un système formel clos, et à cause de cette limitation, il y a des entités distinctes indiscernables, et qui de ce fait conduisent souvent à des paradoxes. Tout l'art du mathématicien consiste à créer un système élargi dans lequel ces quantités ne sont plus impossibles à penser, c'est-à-dire deviennent objet d'un calcul. La théorie des catégories offre aussi des modèles pour décrire la structuration de tels systèmes, le jeu de leurs logiques internes.

Les propriétés universelles

Elles permettent, quand elles sont vérifiées, de définir (construire) un objet d'une catégorie **par ses relations aux autres objets**. Ceci est un élément important des schémas de pensée catégoriques. Cet objet, quand il existe dans la catégorie, est unique à un unique isomorphisme près. Pour chaque propriété, nous déterminerons, à titre d'exemple, l'objet qu'elle permet de construire dans la catégorie $\mathbb{E}\mathbf{ns}$ des ensembles, les morphismes étant alors simplement les applications. Cette catégorie est un topos, ainsi que celle des miroirs.

11) op.cit. p.42.

Pourquoi s'intéresser aux relations entre objets ? Dans un ensemble, un objet est caractérisé par son nom, si l'ensemble est défini en extension, ou par ses propriétés, si l'ensemble est défini en compréhension. Mais supposons que vous cherchiez la rue Descartes, à Paris. Ce n'est ni son nom ni ses propriétés intéressantes (elle est en pente, elle n'est pas rectiligne, elle mesure tant en largeur, tant en longueur, etc...), qui vous permettront de la trouver, mais plutôt ses relations avec d'autres rues.

Objet initial I :

Il est caractérisé par la propriété P : Pour tout objet C, il existe un unique isomorphisme de source I et but C.

Exemple :montrons que, pour $\mathbb{C}$ = **Ens**, cette propriété définit comme objet initial l'ensemble vide $\varnothing$.

Tout d'abord $\varnothing$ vérifie P1 :

Soit E un ensemble. Toute application d'un ensemble A dans un ensemble B est définie par son graphe, qui est un sous-ensemble du produit cartésien $A \times B$. Le graphe d'un morphisme f de domaine $\varnothing$ est un sous-ensemble de $\varnothing \times E$, qui est lui-même $\varnothing$, donc le graphe de f est $\varnothing$. Ce qui donne l'existence et l'unicité de f.

Vérifions par l'absurde que $\varnothing$ est le seul ensemble à posséder P1 :

Soit A non vide. Il possède au moins un élément *a*. Prenons pour E un ensemble qui possède deux éléments E = {*b*,*c*}. On peut alors construire deux applications distinctes qui envoient tout élément de A respectivement sur *b* et *c* (ce qui ne serait pas possible si *a* n'existait pas).

Dans une catégorie générale, montrons que la Propriété 1 définit un élément unique à un isomorphisme près :

Soient I et I' deux objets initiaux, f l'unique morphisme de I dans I' (puisque I est initial) et f' l'unique morphisme de I' dans I (en considérant I' comme initial). Alors f et f' sont inverses l'un de l'autre. En effet, $f' \circ f$ est un morphisme de I dans I, unique puisque I est initial, or 1_I est un morphisme de I dans I donc $f' \circ f = 1_I$. On a de même $f \circ f' = 1_{I'}$. Ce qui prouve bien que I et I' sont isomorphes par f.

Remarque : ceci exclut la possibilité d'être un topos pour toutes les collections d'ensembles qui n'admettent pas l'ensemble vide parmi eux, la catégorie de groupes, par exemple.

Objet final F :

Il est caractérisé par la propriété P2 : Pour tout objet C de $\mathbb{C}$, il existe une flèche unique de source C et but F.

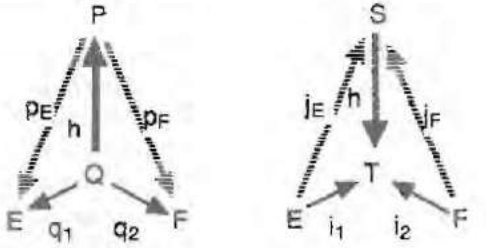
Dans l'exemple **Ens**, les objets finaux sont les singletons, tous isomorphes entre eux.

Vérifions que $\{a\}$ convient : il est nécessaire et suffisant de prendre pour morphisme f la fonction constante de valeur a sur l'ensemble C. L'unicité à un isomorphisme près de l'objet terminal se démontre dans le cas général comme celle de l'objet initial, mutatis mutandis.

La notion d'objet final est la notion duale de celle d'objet initial au sens où l'objet final d'une catégorie est l'objet initial de la catégorie duale.

Produit cartésien de deux objets E et F :
Ce n'est pas un objet de la collection, mais un objet-muni-de-deux-flèches-dont-il-est-source, de buts respectifs E et F, noté par le triplet $[P, p_E, p_F]$, et il est caractérisé par la propriété universelle P3 : Pour tout triplet $[Q, q_1, q_2]$ de même genre, il existe une flèche h unique de source Q et but P telle que $p_E \circ h = q_1$ et $p_F \circ h = q_2$.

Pour y voir plus clair, on dessine un schéma. Les deux dernières conditions s'expriment par le fait que le schéma est commutatif.



produit coproduit

Exemple : Dans **Ens**, on vérifie que si l'on prend pour P le produit cartésien de E par F, et pour p_E et p_F les projections respectives de P sur E et F, la fonction h de Q dans P définie par $h(x) = (q_1(x), q_2(x))$ convient et est unique.
L'objet dual, le **coproduit** est plus difficile à interpréter dans **Ens**. C'est la réunion disjointe, c'est-à-dire que l'on compte deux fois les éléments de l'intersection.

Noyau d'un couple de flèches parallèles (u,v) de source E et but F.

C'est un objet-muni-d'une-flèche dont il est source et dont le but est E. Il est défini par P5 : le noyau [K,i] est tel que $u \circ i = v \circ i$ et pour tout autre couple [K',i'] vérifiant $u \circ i' = v \circ i'$, il existe une flèche unique h de K' dans K telle que $i' = i \circ h$.

$$\begin{tikzcd} K \arrow[dr, "i"] & & \\ & E \arrow[r, shift left, "u"] \arrow[r, shift right, "v"'] & F \\ K' \arrow[uu, "h"] \arrow[ur, dashed, "i'"'] & & \end{tikzcd}$$

noyau

$$\begin{tikzcd} & & Q \arrow[dd, "h"] \\ E \arrow[r, shift left, "u"] \arrow[r, shift right, "v"'] & F \arrow[ur, "q"] \arrow[dr, dashed, "q'"'] & \\ & & Q' \end{tikzcd}$$

conoyau

Dans la catégorie des espaces vectoriels munis des applications linéaires, le noyau du couple (u,v) d'applications linéaires de E dans F est Ker(u–v) muni de l'injection canonique.

Le conoyau est défini en inversant les flèches et correspond à la structure quotient.

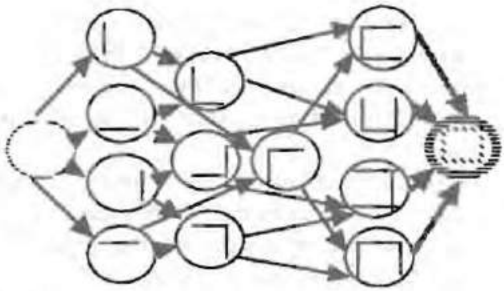
Les autres propriétés universelles font intervenir des objets encore plus complexes, mais ceux que nous avons rencontrés nous donnent une idée des types de raisonnement utilisés.

Limite projective et limite inductive.

Si nous reprenons les exemples précédents, nous voyons qu'une propriété universelle permet de construire un objet de la catégorie à partir d'un diagramme, c'est-à-dire une sous-catégorie. Par exemple, pour le noyau et le conoyau, ce diagramme est une double flèche. L'objet ainsi créé a des liens avec chaque objet du diagramme. De façon générale, dans une catégorie $\mathbb{C}$, une propriété universelle définit, à partir d'un diagramme $\mathbb{C}'$, un objet P de $\mathbb{C}$ qui est source (resp. but) d'une flèche avec chacun des objets de $\mathbb{C}'$, de façon cohérente et commutative avec les liens du diagramme. P muni de ses liens avec les objets du diagramme s'appelle un cône projectif (resp. inductif). On impose de plus à ce cône d'être optimal dans un certain sens : il existe un morphisme h unique tel que pour toute flèche f' d'un autre cône P', il existe une flèche f de P vérifiant $f' = f \circ h$ (resp. $f' = h \circ f$). L'objet P ainsi défini s'appelle **limite projective** (resp.**inductive**) du diagramme.

Un carré à trois côtés ?

Je vais essayer d'expliquer comment ce procédé permet d'«écrire» des objets impossibles à écrire a priori dans un langage donné. Prenons un exemple élémentaire. Je dispose de trois allumettes et je veux former un carré (sans astuce, sans les couper), le carré étant un objet que je connais dans un langage plus puissant (quatre allumettes !). C'est impossible. Comment le démontrer ? en montrant que toutes les tentatives possibles échouent. Pour être sûr de n'en oublier aucune, on peut les disposer dans un diagramme.



Le carré "idéal", qui ne peut pas s'écrire dans le langage des trois allumettes, apparaît ainsi comme la limite inductive du diagramme. Celui-ci se présente comme la trace de toutes les tentatives pour écrire l'objet. Il indique aussi le mot qu'il faut ajouter au langage de départ pour écrire l'objet désiré. Ainsi se construit un nouvel univers. Dans ce cas d'ailleurs, le carré, en tant qu'objet géométrique combinatoire, est identique à l'organisation déployée pour tenter de le construire.

Si nous reprenons le problème des racines d'une équation, l'objet manquant au langage algébrique des quatre opérations n'est autre que le groupe de Galois. Celui qui manque au langage de la mécanique classique pour écrire une théorie de la relativité restreinte est le groupe de Lorentz.

En guise de conclusion, ou plutôt d'ouverture...

Depuis que j'ai découvert ces théories, je suis fascinée par cette construction à l'intérieur des mathématiques que l'on peut utiliser, même sans trop y croire, pour examiner, de l'extérieur, la pensée mathématique et même la pensée tout court. C'est pour tous les frustrés de la logique classique que j'ai écrit ces lignes...

J'espère au moins qu'elles éclaireront ces quelques phrases relevées lors d'une conférence d'Alain Badiou [12].

«La théorie des topos est la seule qui se présente, à l'intérieur du mouvement des mathématiques, comme alternative déployée aux propos ensemblistes. Elle a une ambition de même nature et vise aussi à une reformulation langagière du corps des mathématiques, et à en dévoiler la logique essentielle. Non pas comme une formulation syntaxique du langage, mais comme une proposition d'univers dotés de leur langage immanent...

12) Conférence du 17 Juin 1993, au Collège Internationalde Philosophie. *«Ensembles ou catégories ? platonisme et/ou aristotélisme en mathématiques».*

Le concept central de topos ne propose pas *un* univers, mais un réseau de définitions aboutissant à la description d'univers possibles, à partir d'une axiomatique minimale : la définition d'une catégorie "presque rien", alors que la théorie des ensembles choisit le critère d'extension maximale du pensable au consistant…

La logique du topos est définie de façon interne, la détermination d'un objet est extrinsèque, par le réseau des flèches qui le visent. Elle procède plus par schémas de construction que par axiomes. Elle accorde le primat au virtuel sur l'actuel, à la construction sur la décision…

La théorie des catégories permet la mise en visibilité des contraintes de la décision logique, le penser de son propre univers de possibilités. C'est une théorie de la corrélation entre la construction et la décision, alors qu'une théorie de l'événement, de la décision, relève de la théorie des ensembles.»