L'architecture de la logique ou la fin du monstre de Gila

Jean-Yves Girard

25 avril 2020

Résumé

La logique, toute la logique, rien que la logique. Sans axiomes ni sémantique et hors de tout système. Comme l'activité universelle, centrée sur la raison pure, qu'elle devrait être.

1 Introduction : le monstre de Gila

1.1 La Théorie des Ensembles

Ce texte est la suite amendée de mon « tract » [2] qui annonçait la fin des systèmes logiques tout en concédant des limites à la « désystémisation », typiquement pour la logique dite du second ordre. Ceci est une erreur : la Théorie des Ensembles permet de sortir, une fois pour toutes de ces bunkers, ces bocaux à poissons que sont les systèmes. Il y a des objections évidentes à cette affirmation, mais nous allons voir qu'elles relèvent d'un scientisme périmé. La première question est donc de nature introspective : pourquoi m'a-t-il fallu autant de temps pour les dépasser ?

« Il ne lâche pas prise, même si on lui coupe la tête, même quand l'homme est mort » : c'est ainsi qu'est décrit, dans *Le trésor de la sierra Madre* (1948), ce gros lézard appelé *monstre de Gila*. Il fournit une métaphore appropriée pour tous les « ismes » qui continuent à sévir alors qu'ils ont fait faillite, en particulier le scientisme 1900 qui survit en logique à travers un anachronique salmigondis, le *réalisme axiomatique* ; pourtant sa tête fut tranchée de façon on ne peut plus nette par Gödel en 1931. Mais le scientisme étant une religion, le besoin de croire se traduit par une fidélité sans faille à cette vision obsolète. On pense à l'abeille qui rebouche *quand même* une alvéole vide ou au soldat japonais perdu sur une île du Pacifique qui continue *quand même* à se battre au nom de l'Empereur.

Le Monstre nous étreint encore de nos jours en insistant sur les vérités apodictiques, indiscutables. Nous devrions pourtant savoir qu'elles sont fondamentalement impossibles ou mensongères et considérer la Théorie des Ensembles non comme un système logique qui aurait besoin de sa propre fondation – ce qu'on pensait il y a un siècle – mais comme les Mathématiques tout court. C'est le système consensuel dans lequel elles s'énoncent et qu'il n'y a aucun sens à mettre en doute. À moins qu'on ne travaille précisement *sur* cette théorie : les questions qui se posent alors ne sont pas fondationnelles au sens ringard du terme, tout projet de ce type ayant fait faillite depuis longtemps, mais touchent à un sujet particulier, l'infini, sans relation directe au raisonnement. Heureux théoriciens des ensembles que le Monstre n'étreint plus de sa mâchoire morte !

Il faut nous libérer de cette manie qui veut ramener toute chose à ses constituants, souvent mineurs. Autant réduire la peinture à la toile, voire au cadre de bois, qui ont servi de support. Et oublier la distinction entre impressionnistes et cubistes pour classer les peintures selon l'essence utilisée pour ce cadre : on distinguerait alors l'école du chêne et celle du pin. Autrement dit, ce n'est parce qu'un ingrédient est indispensable qu'il influe de façon significative sur le produit achevé.

Par exemple, la discussion interne à la Théorie des Ensembles quant aux grands cardinaux n'a pas d'incidence sur la notion logique de *comportement* et ses généralisations, toutes de cardinalités ridiculement faibles. L'invocation des ensembles ne signifie pas non plus que l'on pourrait remplacer la logique par **ZF**. À chaque type d'objet, son type de raisonnement ; les ensembles, entités amorphes que l'on considère à bijection près, se contentent de principes de raisonnement expéditifs, *i.e.,* efficaces et sommaires, ceux de la logique classique qui est à la vraie logique ce que la machette est au scalpel. La machette peut servir à couper le bois qui servira de cadre au tableau, mais mieux vaut l'oublier au moment de peindre. Je rappelle incidemment que, quelque soit son intérêt pratique qu'on ne saurait remettre en cause, la logique classique est fautive, puisque le tiers-exclu et ses variantes comme l'implication $\neg\neg A \Rightarrow A$, sont discutables, *i.e.,* acceptables uniquement dans certains contextes. Par exemple celui de l'étude de l'infini, ou encore chez les mafiosi : Vito Corleone (du *Parrain*) n'est-il pas célèbre pour ses « propositions qu'on ne peut pas refuser » ?

1.2 Modules *vs.* systèmes

J'ai comparé [2] les systèmes à des bunkers, des bocaux à poisson. Pour le monstre de Gila, ils sont au contraire l'essence de la logique, d'où leur prolifération cancéreuse. Voire la formation de métasystèmes, *e.g.,* les *logical* *frameworks* [4], sortes de gigantesques hôpitaux où tout système axiomatique est le bienvenu, pourvu de respecter le confinement qui lui interdit de communiquer avec les autres sous peine d'inconsistance.

Le principal défaut des systèmes est leur fermeture : ce qui est vrai dans $\mathbb{T}$ peut devinir faux dans $\mathbb{U}$ et vice-versa. C'est l'esprit de la secte, incapable de s'accorder avec sa concurrente, surtout si elle dit à peu près la même chose. Je propose donc de remplacer ces systèmes autarciques par des *modules* ouverts. Le maître-mot est celui de *cohérence* dont la consistance n'est qu'une pâle approximation. L'idée fondamentale consiste à définir les signifiants logiques, pour l'essentiel propositions et preuves, hors de tout système, de tout langage. Puis à regrouper, sous forme de modules, certaines configurations logiques remarquables se conformant aux définitions générales. On pourra ainsi fabriquer des modules « Multiplicatifs » ou « Arithmétique ». Et décider, selon les besoins, d'invoquer tel ou tel module dans l'esprit de la commande LaTeX `\usepackage`. Ces modules sont automatiquement compatibles, puisqu'ils se contentent de décliner les principes logiques généraux dans des cas particuliers : $\mathbb{T} + \mathbb{U}$ est en fait une extension conservative des modules $\mathbb{T}$ et $\mathbb{U}$. Ce qui rappelle la propriété de la sous-formule qui exprime la conservativité de tout nouveau connecteur par rapport à ceux déjà connus.

Il n'est pas possible d'envisager un système universel formé de tous les modules. Il en a trop (la puissance du continu) pour qu'ils puissent tous tenir dans le même bocal. Nous donnerons bientôt les grandes lignes de modules correspondant au second ordre ou à l'arithmétique, parties de la logique où les systèmes semblent plus prégnants qu'ailleurs. Contentons-nous de répondre à une objection de base : que se passe-t-il si, par hasard, je déclare deux fois le même module, *e.g.,* les multiplicatifs ? On se trouvera donc avec des doublons, $\otimes, \parr$ et $\otimes', \parr'$. Mais la nature logique des deux modules permet alors d'échanger librement les connecteurs et leurs doublons : $\otimes', \parr'$ apparaissent comme des variantes typographiques de $\otimes, \parr$.

Le Monstre objectera sans doute que je me débarrasse de la multitude des systèmes en les ramenant tous à **ZF**. Ne serait-ce que ça, on aurait obtenu une simplification significative : la possibilité de travailler dans son module de prédilection en pouvant échanger avec les autres. La Théorie des Ensembles qui chapeaute ce petit miracle peut être difficilement qualifiée de système, tant nous en percevons peu les limites, voir section 1.1.

1.3 Programme de travail

Dans cet article, je vais d'abord m'attacher aux grandes lignes de l'architecture : usine et usage, vérité. Puis parer au plus pressé avec le brouillon de modules correspondant au second ordre et à l'arithmétique.

2 L'analytique

2.1 Étoiles et constellations

Ce que j'appelle *analytique* est la base dure, irréfragable de toute connaissance. Comme je l'ai remarqué dans [3], cela suppose une totale évacuation du sens. Car qui dit sens dit engagement et donc possibilité de controverse. Ce qui s'en approche le plus est l'informatique, mais plutôt le bazar de *bits* 0, 1 qui se cache au fond de la machine que ce qui s'affiche à l'écran et qui est pollué par l'interprétation qu'on ne peut éviter d'en faire. Il ne s'agit cependant pas de s'arc-bouter sur des machines existantes ; l'analytique est à chercher parmi les machines abstraites. L'expérience logique (Herbrand) et informatique amène aux definitions suivantes [1].

On se fixe un langage fonctionnel bâti à partir de variables $x, y, z, \ldots$ en quantité dénombrable à l'aide d'un nombre fini de symboles fonctionnels $a, b, f, g, \ldots$ d'arités arbitraires. Le choix de ces symboles ne revêt pas la moinde importance dès qu'il y en a un d'arité nulle et un d'arité au moins 2. Les termes fonctionnels, appelés *rayons*, sont gérés par unification : étant donnés t, u, on cherche à résoudre l'équation $t = u$ en donnant des « valeurs » aux variables, ce qui peut s'écrire $t\theta = u\theta$, θ étant une substitution. Le résultat fondamental d'Herbrand nous dit que, si t et u sont unifiables, alors ils ont un pgcu (plus général commun unificateur) θ_0, sorte de mère de toutes les unifications : tout unificateur θ s'écrit $\theta = \theta_0\theta'$ pour un θ' unique. L'opération de base de notre analytique est en fait une variante de l'unification, le *filtrage*, solution la plus générale de $t\theta = uv$; ce qui correspond à modifier les variables de t et u de façon à les rendre distinctes, puis à unifier t' et u' ainsi obtenues. Typiquement, filtrer x et $f(x)$ revient à unifier x et $f(y)$.

À chaque rayon t on peut associer l'ensemble $|t|$ des $t\theta$ clos (sans variables). Si t et u sont filtrables, alors $v = t\theta_0 = uv_0$ obtenu au moyen du pgcf (θ, v) est tel que $|v| = |t| \cap |u|$. C'est pourquoi deux rayons non filtrables sont réputés *disjoints*.

Une *étoile* est un ensemble $[\![\, t_1, \ldots, t_n \,]\!]$ de rayons (c'est moi qui les numérote, ils sont donnés en vrac) soumis aux restrictions suivantes :

— $n \neq 0$: $[\![\ ]\!]$ est interdit.

— Les t_i utilisent *exactement* les mêmes variables : $[\![\, x, y \,]\!]$ est interdit.

— Les t_i sont deux à deux *disjoints* : $[\![\, x, f(x) \,]\!]$ est interdit.

Une *constellation* (ou *dessein*) est un ensemble fini d'étoiles dont les rayons sont deux à deux disjoints.

On comprendra, sans qu'il soit nécessaire de le formaliser plus que ça, que les étoiles (et les constellations) sont considérées aux variables près : il n'y a pas lieu de distinguer entre $[\![\, f(x), g(x) \,]\!]$ et $[\![\, f(y), g(y) \,]\!]$. Autrement dit, les variables de étoiles sont *muettes*, tout comme x dans $\forall x A[x]$ ou $\sum f(x)dx$. Ce mutisme des variables explique l'emploi du filtrage : on ne peut pas unifier des termes définis au nom de leurs variables près.

2.2 Constat et performance

La distinction fondamentale entre statique (donnée, résultat) et dynamique (programme) est obtenue en introduisant les deux « couleurs » complémentaires `vert`/`magenta` qui sont en fait des symboles fonctionnels unaires distingués. Les rayons de la forme $\mathtt{vert}(t)$ ou $\mathtt{magenta}(u)$ sont dits *colorés* et notés [1] t, u.

Si les rayons d'une constellation sont tous incolores, c'est un *constat* ; c'est une *performance* sinon. Une performance peut être transformée, sous réserve de certaines conditions, en constat, l'idée fondamentale consistant à raccorder des étoiles à travers des rayons de couleurs complémentaires : $[\![t_1, \ldots, t_m, t]\!]$ et $[\![u, u_1, \ldots, u_n]\!]$ produisent ainsi $[\![t_1\theta, \ldots, t_m\theta, , u_1\theta, \ldots, u_n\theta]\!]$, où θ est le pgcu de t, u. Cela suppose que les variables aient renommées correctement pour éviter les conflits.

La définition est un peu compliquée par le fait qu'on ne peut pas remplacer $[\![t_1, \ldots, t_m, t]\!] + [\![u, u_1, \ldots, u_n]\!]$ par $[\![t_1\theta, \ldots, t_m\theta, , u_1\theta, \ldots, u_n\theta]\!]$ car $[\![t_1, \ldots, t_m, t]\!]$ est peut être raccordable à d'autres étoiles. D'où la notion de diagramme qui correspond à un raccordement itéré.

Soit $\mathcal{C}$ un constellation. On appelle *diagramme* de $\mathcal{C}$ un graphe connexe et acyclique (un arbre topologique) Δ dont les sommets sont des étoiles de $\mathcal{C}$ sans variable en commun [2] et les arêtes entre deux sommets $\mathcal{S}$ et $\mathcal{T}$ un couple de rayons de couleurs opposées, disons t de $\mathcal{S}$ et u de $\mathcal{T}$; une telle arête est notée $t \rightleftharpoons u$. Les rayons des divers sommets peuvent être utilisés au plus une fois pour former des arêtes ; ceux qui ne cotribuent pas à des artêes sont dits *libres*. La *taille* d'un diagramme est son nombre N d'arêtes ; il a alors $N+1$ sommets. Comme un diagramme peut réutiliser librement les étoiles, une constellation en engendre en général une infinité.

L'opération fondamentale d'*actualisation* d'un diagramme Δ consiste à résoudre les équations suggérées par les arêtes : les arâtes $t_1 \rightleftharpoons u_1, \ldots, t_N \rightleftharpoons u_N$ donnent lieu au système d'équations $t_1 = u_1, \ldots, t_N = u_N$, ce qui est un banal problème d'unification. Il est donc résolu au moyen d'une substitution θ telle que $t_1\theta = u_1\theta, \ldots, t_N\theta = u_N\theta$, parmi lesquelles un pgcu θ_0, auquel cas

1. J'ai un peu trafiqué les couleurs, mon magenta est en fait un rouge plus visible à l'écran.

2. Ce qui permet à la même étoile d'y apparaître plusieurs fois avec un choix différent de variables.

on dit que le diagramme est *correct* ; si le système n'a pas de solution, on dit que le diagramme *échoue*.

Les conditions de *normalisation* régente la production d'un constat à partir d'une performance.

1. $\mathcal{C}$ ne produit qu'un nombre fini de diagrammes corrects. En d'autres termes, tous les diagrammes d'une certaine taille N échouent ; et donc tous ceux de taille supérieure, puisqu'un diagramme de taille $N+1$ contient un sous-diagramme de taille N.

2. Il n'y a pas de diagramme correct fermé, *i.e.,* sans rayon libre.

L'étoile résiduelle d'un diagramme correct dont les rayons libres sont $t_1, \ldots, t_n$ est par définition $[\![t_1\theta, \ldots, t_n\theta]\!]$. Si $\mathcal{C}$ vérifie les consitions de normalisation, on appelle *forme normale* de $\mathcal{C}$ l'ensemble (fini) de ses étoiles résiduelles non colorées.

La forme normale de $\mathcal{C}$ est une constellation de type constat. La seconde condition de normalisation exclut le cas $n = 0$ d'une étoile résiduelle sans rayon. Mais il faut surtout vérifier que les rayons de la forme normale sont dux à deus disjoints.

Take two diagrams whose free rays are uncoloured, with (before actualisation) a common free ray s ; starting from s, the two diagrams must first disagree on some edge, *e.g.,* $t = u$ *vs.* $t = v$, thus inducing actualisations $t\theta, t\theta'$. Since u, v are disjoint, so are the eventual actualisations of s in both diagrams : here we use the fact that the variables are exactly the same. Take now the case where the two s (say, s_1, s_2) are part of the same diagram : there is a path leading from s_1 to s_2 through edges $T_1/U_1, \ldots, T_n/U_n$, with $T_i, U_i = t_i, u_i$ or t_i, u_i. We can handle this case as if s_1, s_2 were in distinct diagrams, unless there is an automorphism σ of the diagram such that $\sigma(s_1) = s_2$; if $s_3 := \sigma(s_2) \neq s_1$, then $s_1, s_2, s_3, \ldots, s_k = s_1$ would form a cycle in the diagram. If $\sigma(s_2) = s_1$, then $T_1 = U_n, T_2 = U_{n-1}, \ldots, T_n = U_1$. If $n = 2m$, then $T_m = U_m$, which is impossible, since T_m, U_m are of distinct colours ; $n = 2m+1$ is impossible too, since there would be a star with equal rays U_m, T_{m+1}.

Références

[1] J.-Y. Girard. **Transcendental syntax 1 : deterministic case**. *Mathematical Structures in Computer Science*, pages 1–23, 2015. *Computing with lambda-terms. A special issue dedicated to Corrado Böhm for his 90th birthday.*

[2] J.-Y. Girard. **Un tract anti-système**. Technical report, http ://girard.perso.math.cnrs.fr/systeme.pdf, 2019.

[3] J.-Y. Girard. **Le Fantôme de la transparence**. Allia, Paris, Septembre 2016. 248 pp.

[4] R. Harper, F. Honsell, and G. Plotkin. **A framework for defining logics**. *LFCS report series, Edinburgh*, 162, 1991.