

# De la vertu négatrice au surmoi logique

Jean-Yves Girard

*Institut de Mathématiques de Luminy, UMR 6206 – CNRS  
163, Avenue de Luminy, Case 907, F-13288 Marseille Cedex 09*

*girard@iml.univ-mrs.fr*

17-18 Octobre 2008

## Introduction : la vertu négatrice

Ce que nous percevons de la négation, c'est — comme dirait Molière — sa « vertu négatrice »; autrement dit, elle nous apparaît à travers un *phénomène*, l'acte de dire « NON ». Aborder la négation sous cet angle (cf. Tarski :  $\neg A$  est vrai quand  $A$  n'est pas vrai), c'est tomber dans une chausse-trappe, se condamner à une réflexion stérile, viciée : c'est ce que j'ai appelé le *point aveugle*<sup>1</sup>.

L'idée fondamentale que nous allons développer, c'est que la négation est un phénomène de la *dualité*; mais aussi que la morphologie, le « surmoi », autrement dit la *normativité logique*, est un phénomène de cette même dualité. La discussion s'organisera donc autour du triptyque

DUALITÉ/NÉGATION/NORME

Avec un résultat paradoxal; cette lecture novatrice, philosophiquement excitante, de la négation est en même temps frustrante. En effet, si c'est un progrès conceptuel que de s'éloigner de la « vertu négatrice », il faut bien constater que les problèmes mathématiques courants réfèrent à la négation « naïve »; ainsi, une question fameuse s'énonce : « montrer que  $P \neq NP$  », et non pas : « établir une dualité entre  $P$  et  $NP$  », ce qui ne veut strictement rien dire. Faute de s'appliquer à la négation de tous les jours — qui reste, en dépit de sa superficialité, celle qui nous intéresse —, la version dualiste peine un peu à convaincre.

## La double négation

La vertu négatrice brille de tous ses feux dans les *algèbres de Boole*. Considérons l'ensemble  $\wp(X)$  des parties d'un ensemble  $X$  : si  $a \subset X$  « représente » la proposition  $A$ , alors le complémentaire  $\complement a := X \setminus a$  représente sa négation  $\neg A$ . C'est ainsi, puisque  $\complement \complement a = a$ , qu'on en arrive à identifier  $\neg\neg A$  avec  $A$ . Cette identification justifie le raisonnement par contraposition (dit encore « par l'absurde »); sa remise en cause est une des pierres de touche de l'*intuitionnisme*.

---

<sup>1</sup>Ce texte se veut le moins technique possible, pour respecter l'audience, i.e., ne pas la terroriser. Mais ce n'est pas non plus un exercice d'esquive : on trouvera des développements détaillés de la plupart des points évoqués dans mon cours en deux volumes [1, 2].

La vie courante remet-elle en cause l'*involutivité*, i.e., le principe  $\neg\neg A = A$ ? La réponse est complexe; prenons l'exemple du discours politique :

- Les États ont de plus en plus tendance à communiquer par double négation : au lieu de « nous avons construit un monde meilleur » on n'entend plus guère que « si vous nous critiquez, vous êtes avec les terroristes ». Et d'ailleurs souvenons-nous du référendum constitutionnel : le slogan « si vous votez non, le ciel vous tombe sur la tête », qui réduit l'argumentaire du « oui » à une double négation, n'a pas su convaincre. L'analyse logique du discours politique se révèle donc d'un certain intérêt, puisqu'elle débusque une figure de rhétorique omniprésente.
- Par contre, le discours politique n'a aucune rigueur et son analyse n'apporte rien à la logique. Ainsi, W. Bush et Al Qaida se renvoient-ils symétriquement des atrocités bien réelles, chacun justifiant son point de vue par double négation : « je suis bien car je me bats contre ce monstre ». Cela ne fonctionne pas logiquement, car, double négation ou pas, pour soutenir W., il faut ne retenir que le 11 septembre en fermant les yeux sur le supplice de la baignoire. En logique, ce type de bidouillage à la tête du client s'appelle « paraconsistance » et c'est un gros mot.

On peut chercher un peu d'air en adoptant une version topologique des algèbres de Boole : on considère les *ouverts* d'un espace  $X$ . La négation est alors interprétée par l'intérieur du complémentaire, ce qui fait que la double négation devient l'intérieur de l'adhérence. Comme, en général,  $(\bar{a})^\circ \neq a$ , on obtient une réfutation convaincante, sinon profonde, de  $\neg\neg A = A$ . Ce faisant, les interprétations topologiques ont ainsi livré leur dernier mot !

Heureusement, il y d'autres outils que cette illusion sémantique — un monde de référence —, en particulier l'étude de la *procéduralité* logique, comme on le fait en *théorie de la démonstration*. Quels sont donc les obstacles procéduraux au raisonnement par l'absurde? Autrement dit, qu'est ce qui empêche de remplacer, étape par étape, une démonstration de  $\neg B \Rightarrow \neg A$  par une démonstration directe de  $A \Rightarrow B$ ? Si on part d'une écriture arborescente dans le style « déduction naturelle » — les hypothèses en haut, la conclusion en bas — on comprend que l'on puisse, récursivement « permuter » la conclusion  $\neg A$  et l'hypothèse  $\neg B$ . Mais ce type d'échange n'est possible que si l'hypothèse  $\neg B$  n'est utilisée qu'une fois. On voit donc que la spécificité du raisonnement par l'absurde réside dans la possibilité de *réutiliser* la conclusion, devenue hypothèse par contraposition. Ceci en accord avec le *théorème de Herbrand* : un énoncé  $\exists x \forall y A[x, y]$  est démontrable si — en introduisant la lettre  $f$  pour dire «  $y = f(x)$  est une variable fraîche » — on peut trouver des termes  $t_1, \dots, t_n$  tels que :

$$A[t_1, f(t_1)] \vee \dots \vee A[t_n, f(t_n)] \tag{1}$$

Par exemple, si  $A[x, y] := B[x] \Rightarrow B[y]$ ,  $n = 2$ ,  $t_1 := x$ ,  $t_2 = f(x)$  :

$$(B[x] \Rightarrow B[f(x)]) \vee (B[f(x)] \Rightarrow B[f(f(x))]) \tag{2}$$

Ce qui est l'image de l'*intrication* classique : le  $x$  « générique » par rapport à  $A$  est donné de façon indirecte : c'est n'importe quoi ( $t_1 = x$ ) et, au cas où ce n'importe quoi échoue (i.e., si  $B[x]$  mais  $\neg B[y]$  pour un certain  $y$ , ce  $y$ , noté  $f(x)$ , i.e.,  $t_2$ ). Ceci est évidemment une manipulation logique aussi honnête que certaines transactions bancaires... avec un petit bémol : on ne prendra jamais

ce banquier-là la main dans le sac, car il est impossible de prouver qu'il est véreux. La double négation c'est « pas vu, pas pris ».

La *linéarité*, c'est précisément la restriction, tout à fait naturelle du point de vue procédural, à une utilisation unique de chaque hypothèse. On comprend tout de suite que la *logique linéaire* qui est basée sur cette restriction est naturellement dotée d'une négation — la *négation linéaire*  $\sim A$  — involutive :  $\sim\sim A = A$ . La négation linéaire correspond à l'échange entre hypothèse et conclusion. L'implication devient alors *linéaire* : dans  $A \multimap B$ , l'hypothèse n'est utilisée qu'une fois ; on peut l'identifier à  $\sim B \multimap \sim A$ , ce qui justifie la contraposition en mode linéaire. Comme les hypothèses ne sont plus sujettes à duplication, il leur est loisible de migrer toutes en conclusion, produisant ainsi une sorte de démonstration à conclusions multiples, un *réseau de démonstration*. L'étude de ces réseaux est le point de départ de la normativité logique (la négation comme surmoi), voir *infra*.

La linéarité qui est une limitation *a priori* se révèle un espace de liberté ; en effet, il est possible d'instiller de la répétition au moyen d'une primitive de *pérennisation*, l'exponentielle  $!A$ , ce qui permet de replacer la logique classique dans un cadre involutif, cette fois-ci procéduralement correct : il s'agit du système **LC**. C'est en fait plus qu'un simple replâtrage procédural, car la pérennisation, qui faisait partie de ces motifs muraux qu'on ne perçoit pas, devient explicite : s'ouvre alors la question essentielle de la *pérennité de la pérennité*, par exemple du principe  $!A \multimap !!A$ . En modulant les principes de la pérennité, la logique arrive à gérer naturellement la complexité algorithmique (logiques allégées, e.g., **LLL**). La procéduralité hypereffective (complexité polynomiale, etc.) des logiques allégées leur donne une immense prégnance, en complet porte-à-faux avec leur totale absence de dimension sémantique, i.e., d'accord avec nos préjugés sur le fini, l'infini, etc. Ce sont des systèmes paradoxaux, non pas en raison d'un vice de forme, mais en ce qu'ils bousculent l'agencement tarskiste du monde, sans en proposer pour autant un autre. Ce type de hiatus entre deux paradigmes possédant chacun sa légitimité est *a priori* extrêmement fécond.

## Adjonction vs. inversion

Si  $A \Rightarrow B$ , alors les ensembles associés  $a, b$  sont tels que  $a \subset b$ , donc  $\mathbb{C}b \subset \mathbb{C}a$ , ce qui correspond à  $\neg B \Rightarrow \neg A$ , autrement dit la négation est *contravariante*. Mais, une fois de plus, la sémantique n'a rien d'autre dans son sac. Il faut vraiment introduire le point de vue *fonctionnel*<sup>2</sup>.  $A \Rightarrow B$  est l'espace des fonctions de  $A$  dans  $B$ , ce qui fonctionne remarquablement bien pour la logique intuitionniste (remarque de Curry, complétée par Howard). La logique linéaire correspond à une restriction importante : les espaces en cause sont des espaces « linéaires » (espaces cohérents, mais aussi les « vrais » espaces de Banach) et  $A \multimap B$  correspond aux fonctions linéaires de  $A$  dans  $B$ . Réutiliser une hypothèse, c'est vouloir passer d'une fonction bilinéaire  $f(x, y)$  de  $A \times A$  dans  $B$  à une fonction linéaire de  $A$  dans  $B$ , ce qui n'a pas de sens (la solution  $f(x, x)$  étant quadratique, donc tout sauf linéaire). La pérennisation correspond à l'algèbre symétrique.

L'algèbre linéaire nous offre deux paradigmes contravariants, l'*adjonction*

<sup>2</sup>Ou encore catégorique : c'est le niveau **-2**, le niveau **-1** étant celui de la sémantique (vérité, algèbres de Boole) ; quant au niveau **-3**, c'est celui de la procéduralité, à l'œuvre dans le surmoi logique.

et l'*inversion* dont seul le premier est logiquement satisfaisant. La discussion qui suit se place en algèbre linéaire (espaces vectoriels) et s'adapte, *mutatis mutandis*, à tout type d'interprétation linéaire (espaces cohérents, espaces de Banach, etc.).

Si  $E$  est un espace vectoriel, on sait que tout vecteur  $x \in E$  admet des « coefficients »  $x_i$  par rapport à une base donnée  $(\mathbf{e}_i)$ . L'application  $x \mapsto x_i$  apparaît ainsi comme une *question* que l'on pose aux éléments de  $E$  : « quel est ton coefficient d'index  $i$  ? ». Oubliant la base, l'ensemble des questions possibles, i.e., des application linéaires  $\varphi$  de  $E$  dans  $\mathbb{R}$  — les *formes* linéaires — constitue l'*espace dual*  $E'$ . Si  $u \in \mathcal{L}(E, F)$  est une application linéaire, on peut décrire l'action de  $u$  par un jeu de questions/réponses : si je veux connaître le coefficient  $\varphi(u(x))$ , il suffit de connaître le coefficient  $\psi(x)$  avec  $\psi := \varphi \circ u$ , ce qu'on note  $\psi = u^*(\varphi)$ . L'*adjoint* (ou transposé) de  $u$ ,  $u^* \in \mathcal{L}(F', E')$  correspond donc à un flux dual de questions. L'interprétation de la négation au moyen de l'espace dual rend parfaitement compte de nombreux aspects de la négation, ainsi la contravariance. Savoir si cette version de la négation abolit la version booléenne est une question complexe que je trancherai plus tard ; pour l'instant je me contente de remarquer que :

- Il y a suffisamment de formes, ce qui veut dire que l'on peut *séparer* les vecteurs de  $E$  : si  $x \neq y \in E$ , il existe  $\varphi \in E'$  telle que  $\varphi(x) \neq \varphi(y)$ .
- $E$  se plonge naturellement dans  $E''$  : si  $x \in E$ ,  $\Phi_x(\varphi) := \varphi(x)$  définit une forme sur  $E'$ . Par contre, il est possible que ce plongement ne soit pas surjectif, typiquement les espaces de Banach qui ne sont pas toujours *réflexifs*, i.e., tels que  $E'' = E$  ; les *Espaces de Banach Cohérents* ne gardent en fait qu'une partie de l'espace dual.
- La tradition fonctionnelle identifie un énoncé avec l'ensemble de ses preuves possibles, ou plutôt une version idéalisée d'icelles. La négation se présente comme un ensemble de questions qui vient s'ajouter aux preuves et ne peut s'y réduire (en effet, il y a des points dans  $E$  et dans  $E'$ , mais il n'y a pas de preuve de  $A$  et de  $\sim A$  pour des raisons de cohérence)<sup>3</sup>. La question de la nature des objets supplémentaires, ou plutôt d'une notion d'*épreuve*<sup>4</sup> qui engloberait les preuves est fondamentale.

Les espaces de Banach réflexifs typiques sont les *espaces de Hilbert*. Dans ce cas, l'application qui à  $e \in E$  associe la forme  $\varphi_e(x) := \langle x | e \rangle$  établit une isométrie entre  $E$  et son dual  $E'$ . En particulier, il est possible de transposer un morphisme  $u \in \mathcal{L}(E, F)$  d'espaces de Hilbert en un morphisme  $u^* \in \mathcal{L}(F, E)$  :

$$\langle u(x) | y \rangle = \langle x | u^*(y) \rangle \tag{3}$$

ce qui fait que l'adjonction commence à ressembler à une inversion. . . ce qui est d'ailleurs le cas pour une classe importante de fonctions linéaires, les *unitaires* (applications isométriques) qui vérifient  $u^* = u^{-1}$ . Mais peut-on essayer d'interpréter la négation par l'inversion ? Mathématiquement, l'inversion est bien moins robuste que l'adjonction, ne serait-ce que parce qu'elle n'est pas toujours définie ; il y a cependant une place en logique pour l'inversion, même si

<sup>3</sup>Une pièce de musée, la *réalisabilité* de Kleene, essaye d'interpréter la logique intuitionniste en collant le plus possible aux démonstrations, ce qui fait qu'elle devient à peu près sans intérêt dès que la négation s'en mêle : il y a trop d'ensembles vides. En effet,  $\neg A$  est réalisable ssi  $A$  ne l'est pas, auquel cas n'importe quoi réalise  $\neg A$ .

<sup>4</sup>Suivant l'heureuse expression de Livet.

cette place reste à définir et ne peut être celle de la négation : c'est celle de la *rétrospection*.

Par *rétrospection*, j'entends la réflexion sur l'action d'un système. Ce système peut-être une machine, un être humain (on parlera de diagnostic), la société (on parlera d'enquête) et, bien sûr, un système formel. L'idée est qu'un phénomène a été produit par une cause définie que l'on recherche en explorant des possibles, i.e., en « inversant la vapeur » : à partir de symptômes, d'indices, on essaye de reconstruire une rationalité. Mais cette inversion ne produit au mieux qu'une heuristique sans valeur *déductive* : on sélectionne des suspects sur des présomptions, soit, mais on ne les juge pas là-dessus ! Autrement dit, la rétrospection n'a aucune conséquence automatique, sauf chez Sherlock Holmes (le porte-drapeau de l'« abduction ») et les spécialistes du lynchage (une référence nettement moins consensuelle). L'*abduction* prétend, en effet, expliquer la négation par l'inversion : si  $A \Rightarrow B$ , alors  $\neg A \Rightarrow \neg B$ . Ce qui est lamentable pris à la lettre et qu'on rectifie donc immédiatement avant de bidouiller la rectification, etc. En fait le seul argument en faveur de l'abduction reste une double négation : « il faut bien faire quelque chose » ! Logiquement l'erreur consiste à traiter un problème procédural (les potentialités d'un *système* considéré en tant que système) au moyen d'une banale interprétation booléenne qui en aplatit les aspects cognitifs *qui font ici partie de l'énoncé lui-même* ; on produit ainsi des systèmes contradictoires (voir la « complétion de Clarke » de sinistre mémoire en programmation logique) ou, pire, non déductifs (la « closed world assumption » de Reiter).

Oublions donc les grotesques contorsions de l'abduction pour revenir à la question légitime de la rétrospection, i.e., d'une heuristique qui resterait à sa place. Les questions du type « peut-on faire ceci de telle manière ? » sont essentielles en mathématiques depuis l'origine, rappelons-nous  $\sqrt{2}$ ,  $\pi$ , l'équation du cinquième degré : on a, à chaque fois, montré qu'un système (un certain type de méthode) était incapable d'arriver à un certain but. La logique se pose souvent des questions de la même nature sous la forme : « montrer que  $A$  n'est pas démontrable ». Un problème célèbre,  $P \neq NP$ , réfère à toutes nos possibilités de calcul en temps « raisonnable », i.e., polynomial. Nous ne disposons à ce jour d'aucun outil vraiment satisfaisant :

- Les problèmes rétrospectifs les plus simples peuvent être résolus au moyen de *modèles* : on trouve un modèle de  $\mathcal{T}$  dans lequel  $A$  est faux et donc  $\mathcal{T}$  ne démontre pas  $A$ . Cela peut fonctionner pour les problèmes anciens (construction à la règle et au compas, etc.), mais se révèle insuffisant pour des systèmes plus complexes comme les systèmes logiques ou les algorithmes de calcul.
- Les rares résultats rétrospectifs connus en logique sont liés aux travaux de Gentzen et remontent donc aux années 1934 : si  $A \vee B$  est démontrable en logique intuitionniste, alors soit  $A$ , soit  $B$  l'est. C'est un résultat non trivial que l'on n'obtiendra pas avec des joujoux du type « modèles de Kripke ». Malheureusement, trop limité ; en effet la logique linéaire (et sa relecture calculatoire par Andreoli) divise les opérations logiques en deux classes ou *polarités*, les positifs comme la disjonction intuitionniste pour lesquels nous avons des résultats d'introspection, et les négatifs comme l'implication ( $\Rightarrow$  ou  $\multimap$ ) qui sont *inversibles*, autrement dit pour lesquels la déduction est automatique : la rétrospection se fait toujours, mais elle est vide de sens. Or toutes les questions rétrospectives de quelque intérêt

sont implicatives (ou fonctionnelles comme  $P \neq NP$ , mais fonction a valeur d'implication), donc négatives.

- Les logiques procédurales (intuitionniste, linéaire, allégées genre **LLL**) réfutent certains principes (comme le tiers-exclu, la réutilisation, la fonction exponentielle) au nom de la procéduralité. Mais cette réfutation est *externe* : dans **LLL**, on ne peut pas construire la fonction  $n \mapsto 2^n$ . Ne pas pouvoir correspond à un *renforcement* de la rétrospection, mais aussi à un *affaiblissement* des possibilités déductives. Il y a ici un conflit entre les deux aspects de la négation, l'aspect logiquement correct de l'*adjonction*, et l'aspect logiquement incorrect, mais important, de l'*inversion* ; ainsi, la fonction exponentielle est-elle réfutée dans le second sens (on ne peut pas la trouver), mais pas dans le premier (qui demande plus : une réfutation interactive, donc modulaire). Pour mettre les pieds dans le plat, les logiques procédurales sont assises entre deux chaises, car, suivant le point de vue adopté, on veut soit plus, soit moins, que la logique classique ; essayer d'identifier brutalement les deux points de vue (rétrospection et déduction) mène à l'abduction et aux logiques de la Rue du Four.

La très faible efficacité de la rétrospection « à la main » que ne compense pas un statut logique qui permettrait d'approcher celle-ci par des méthodes indirectes explique l'absence de tout progrès sur les questions rétrospectives délicates liées à la logique et au calcul. Il est possible que nos catégories logiques, héritées d'une tradition plus que centenaire et donc non procédurale, soient insuffisantes et qu'il faille en introduire d'autres, mais lesquelles ? Voilà, en tout cas, un domaine de réflexion pour la logique où tout reste à faire ; cela dit, les idées sont rares et mieux vaut se taire que devoir s'abriter derrière un « il faut bien faire quelque chose » !

## Pôle et polaires

Il est temps d'introduire un cadre général dans lequel s'inscrit la dualité logique. Méthodologiquement, il importe de rappeler que de tels cadres ne doivent pas être étudiés pour eux-mêmes (facilité et pêché mignon de la logique, qui induit une dilution de la substance scientifique) : ils ont moins d'intérêt que les cas particuliers qu'ils regroupent.

Pour établir une dualité sur  $A$ , on se donnera une fonction  $x, y \mapsto \langle x | y \rangle$  de  $A \times A$  dans un espace *scalaire*  $S$ , ainsi qu'un *pôle*  $P \subset S$ . On supposera la symétrie  $\langle x | y \rangle = \langle y | x \rangle$ . Étant donné  $a \subset A$ , on définit son *polaire* :

$$\sim a := \{y ; \forall x \in a \quad \langle x | y \rangle \in P\} \quad (4)$$

Un *fait* est un ensemble égal à son bipolaire,  $a = \sim \sim a$  ; comme  $\sim \sim \sim a = \sim a$ , les faits sont exactement les ensembles *polaires*, i.e., ceux de la forme  $\sim b$ .

Ce type de définition par dualité est omniprésent en mathématiques, par exemple, en algèbre, le *commutant* peut s'écrire comme le polaire par rapport à  $\langle x | y \rangle := xy - yx$  et au pôle  $\{0\}$ . En logique, il apparaît en 1986 avec la logique linéaire, dans le cadre très restreint (car de niveau **-1**) des *espaces de phases*. Il se révèle d'un emploi universel ; encore faut-il faire preuve d'imagination et trouver les bons  $A, \langle \cdot | \cdot \rangle, R, P$ . Donnons quelques exemples :

1. Si  $P := \emptyset$ , les seuls ensembles polaires sont le vide et le plein, échangés

par négation,  $\sim \emptyset = A$ ,  $\sim A = \emptyset$ . On est dans la sémantique classique la plus étroite.

2. Si  $R := \{0, 1\}$ ,  $P := \{0\}$  et  $\langle x | y \rangle := 1$  si  $x = y$ ,  $\langle x | y \rangle := 0$  sinon, tous les ensembles sont polaires et  $\sim a = \complement a$ .
3. Si  $A$  est un monoïde commutatif, si  $R := A$  et si  $\langle x | y \rangle := xy$ , l'interprétation qui en résulte est la *sémantique des phases*, sémantique complète pour la logique linéaire. Ceci dit, objet de second choix, la logique linéaire trouvant son origine dans l'exemple suivant, les *espaces cohérents*. En même temps, la première interprétation logique du type « polarité ».
4. Si  $A := \wp(X)$ ,  $R := \sharp(A)$ ,  $P := \{0, 1\}$  et  $\langle x | y \rangle := \sharp(x \cap y)$ , un ensemble polaire est appelé *espace cohérent* de trame  $X$ . Si  $\mathfrak{X}$  est un espace cohérent de trame  $X$  considérons la relation binaire (réflexive et symétrique) sur  $X$  définie par  $x \subset y \iff \{x, y\} \in \mathfrak{X}$ ; alors les ensembles de  $\mathfrak{X}$  sont exactement les  $a \subset X$  tels que

$$\forall x, y \in a \quad x \subset y \tag{5}$$

(les *cliques*). La négation  $\sim \mathfrak{X}$  correspond à  $\succ$ , avec :

$$x \succ y \iff x \not\subset y \text{ ou } x = y \tag{6}$$

C'est sous la forme « clique » que sont apparus les espaces cohérents (et cela reste l'approche la plus commode) dont est issue la logique linéaire. La définition par polarité est tardive : elle s'inspire de la sémantique des phases qui n'est pourtant que de niveau **-1**, alors qu'ici on est au niveau **-2**.

5. Si  $A := \mathcal{H}(E)$  est l'espace des opérateurs hermitiens (= auto-adjoints) sur un espace de Hilbert  $E$  de dimension finie, si  $R := \mathbb{R}$ ,  $P := [0, 1]$  et  $\langle x | y \rangle := \text{tr}(xy)$ , les ensembles polaires sont appelés *espaces cohérents quantiques*. On est toujours au niveau **-2**; les ECQ sont en fait un exemple d'espace de Banach cohérent.
6. La ludique, la géométrie de l'interaction sont des paradigmes de niveau **-3** qui rentrent toujours dans ce cadre. Ainsi, la GdI est-elle basée sur la quantité  $\langle x | y \rangle := \det(1 - xy)$ . Mais cette énumération est lassante : soit on connaît déjà et le truc est émoussé, soit on s'est perdu dans cette énumération.

## Le surmoi logique

La logique repose sur un certain nombre d'*interdits* : par exemple, tout ce qui a trait au langage, depuis le parenthésage jusqu'au fait qu'une règle a plusieurs prémisses mais une seule conclusion. Toutes ces restrictions, quelqu'en soit le bien-fondé, font partie de la *morphologie*. Le tarskisme, vision essentialiste de la logique, traite la morphologie à l'instar des Tables de la Loi dans le célèbre film de Cecil B. de Mille où elles sont gravées par un éclair divin.

Ce *Deus ex machina* ne convainc que si l'on ne veut pas parler de logique, ou, ce qui revient au même, si l'on adopte une attitude normative et sectaire, pour réduire la logique à un des nombreux « ismes » dont elle a été proluxe. En fait la morphologie se présente souvent sous deux aspects, cf. la discussion

d'il y a une centaine d'années (Russell vs. Zermelo) sur le *typage*. Il s'agissait de théorie des ensembles et de ses paradoxes (Burali-Forti 1898 simplifié par Russell 1902) que l'on essaye de corriger en traitant l'auto-appartenance  $x \in x$  :

**Russell** : Le typage *explicite* : les objets viennent avec un *type*, une espèce de pedigree qui gouverne les relations sociales comme sous l'*apartheid* : on n'a pas droit d'écrire  $x \in y$  que si le type de  $x$  est strictement inférieur au type de  $y$ . Exit  $x \in x$ .

**Zermelo** : Le typage est *implicite* : on se restreint aux ensembles *x bien fondés*, i.e., tels qu'il n'y ait pas de suite  $\dots x_3 \in x_2 \in x_1 \in x_0 = x$ .

La version de Russell (qui n'a pas survécu) est *essentialiste* : la morphologie (le typage) y est primitif. À quoi s'oppose l'*existentialisme* de la version de Zermelo (pour l'essentiel, la théorie des ensembles que nous connaissons) : on soumet les ensembles à des tests (la recherche de suites décroissantes

( $x = x_0, x_1, x_2, x_3, \dots$ )), et, s'ils les passent, on les accepte. On peut alors reconstituer le typage comme la hauteur ordinaire de l'arbre des suites *finies* décroissantes partant de  $x$  (la hiérarchie  $V_\alpha$ ). Pour Russell, la justification de la loi (le typage) est que l'on doit la respecter, pour Zermelo, c'est que s'en éloigner produit des effets indésirables (des suites infinies décroissantes). Remarquons au passage que la condition de bonne fondation est plus générale que  $x \notin x$ , autrement dit que l'on ne fait pas de la morphologie interactive si facilement que cela ; ce qui explique le penchant pour les Tables de la Loi : dans ce cas, il suffit d'évoquer Dieu qui n'a jamais contredit personne, sauf chez Cecil B. de Mille !

Les excès de la morphologie essentialiste sont bien connus. Le plus simple est de retourner Rue du Four : considérons, par exemple, les « logiques » paraconsistantes. On a des règles logiques, mais il est interdit de conclure à l'absurdité. C'est comme un chèque sans provision que l'on n'aurait pas le droit de mettre à l'encaissement. . . cela dit, la solution de la crise financière est un peu du type « paraconsistant », reste à savoir où sont les gogos : dans le cas de la logique paraconsistante, ce sont ceux qui avalent de telles sornettes.

Si l'on part du principe qu'une loi ne se justifie que par les conséquences négatives de sa non-observance, on aboutit à l'image suivante : à l'origine, tout est amorphe, ce qui fait que n'importe quoi peut se produire. La morphologie, la loi, est ce qui canalise ce « n'importe quoi ». En termes logiques, un énoncé **A** (en anglais, *sentence*, ce qui nous ramène à la loi), identifié à l'ensemble de ses preuves (il y a donc normativité, puisque les preuves sont soumises à de fortes contraintes), est ce qui passe un certain nombre de tests ; soit **A'** l'ensemble de ces tests, soit  $R$  l'espace des résultats et  $P \subset R$  le sous-espace formé des résultats « positifs », alors :

$$\pi \in \mathbf{A} \quad :\Leftrightarrow \quad \forall \tau \in \mathbf{A}' \quad \tau(\pi) \in P \quad (7)$$

Ce qui est, en fait insuffisant. Par exemple, les *labelled deductive systems* de Gabbay<sup>5</sup> sont définis comme des systèmes syntaxiques arbitraires sur lesquels on met des restrictions à la tête du client, ce qui permet d'accepter les logiques paraconsistantes, mais qui rentre dans le cadre ci-dessus. Il y a d'ailleurs un malentendu persistant sur la procéduralité : l'intuitionnisme analyse la disjonction non pas au moyen d'une quelconque sémantique de Kripke qui ne donne

<sup>5</sup>Magnifique exemple de *novlangue* : « labelled » veut dire « non déductif » !



rien d'intéressant, mais au moyen de sa « procéduralité » : démontrer une disjonction revient *in fine* à démontrer l'un des deux. En partant de ce principe, on fabrique une logique procédurale, mais elle est tenue à vérifier le principe dont elle part, ici la propriété de la disjonction : les règles de la logique et la logique des règles doivent coïncider. C'est donc tout, sauf un paradigme laxiste à la Gabbay ! Et qui échoue donc facilement ; ainsi, l'interprétation procédurale de la négation : «  $\neg A$  est prouvable quand  $A$  n'est pas prouvable » échoue : faire coïncider les règles de la logique avec le logique des règles amène à l'identification « vrai=cohérent », ce qui est réfuté par le théorème de Gödel. Nous avons vu que la bonne procéduralité de la négation est à chercher dans la dualité, i.e., dans l'adjonction.

En fait, la marque distinctive d'une approche existentialiste, c'est de mettre la loi sur un pied d'égalité, autrement dit, quand on forme  $\tau(\pi)$ ,  $\tau$  juge  $\pi$ , mais  $\pi$  juge  $\tau$  en même temps : quand tu regardes l'abîme, celui-ci regarde en toi... Ou encore, les juges seront jugés. Il faut alors remplacer la notation  $\tau(\pi)$  par l'expression symétrique  $\langle \pi | \tau \rangle$ , ce qui nous ramène à la polarité :

$$\mathbf{A} = \sim \mathbf{A}' \quad (8)$$

$\mathbf{A}$  étant polaire, on peut supposer que  $\mathbf{A}' = \sim \mathbf{A}$ , ce qui fait que la négation apparaît comme la forme logique du *surmoi* !

Un exemple typique a déjà été donné par les *espaces cohérents*. Ils sont apparus au départ comme le résultat d'une contrainte morphologique, être une *clique*, voir (5) *supra*, par rapport à une *relation de cohérence*  $\supset$  tombée du ciel. Beaucoup plus tard est apparue une version essentialiste de la même chose, basée sur la polarité :

$$\sharp(a \cap b) \leq 1 \quad (9)$$

Cette définition est devenue nécessaire à un certain moment. En effet, la morphologie est une prison et la plus sévère des prisons est celle dont on ne voit pas les murs. Dans les espaces cohérents, même « existentialisés », reste un non-dit subliminal : l'utilisation des ensembles. Dans une certaine vision « non commutative » du monde, les atomes ensemblistes deviennent des espaces de dimension 1 et l'écriture d'un ensemble fini sous la forme  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  devient le choix d'une *base*. On s'aperçoit que le sujet s'est fait couleur de muraille, qu'il se dissimule derrière des choix morphologiques qui *vont de soi* : « je vais vous faire une proposition que vous ne pourrez pas refuser ». Si l'on essaye d'adapter la version essentialiste (la relation de cohérence) à un espace sans base distinguée, on arrive strictement à rien ; par contre, si l'on remplace les sous-ensembles de la base  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  par les sous-espaces engendrés, et donc par les projecteurs orthogonaux associés, (9) devient :

$$0 \leq \text{tr}(ab) \leq 1 \quad (10)$$

Ce qui est la forme sous laquelle se présentent les espaces cohérents quantiques<sup>6</sup>.

Les paradigmes de niveau **-3** qui sont les seuls à rendre pleinement compte de l'idée de morphologie interactive, sont en même temps les plus difficiles à expliquer : *ludique, géométrie de l'interaction*. L'idée de surmoi logique m'est apparue au moment de la rédaction de mon article sur la logique linéaire, en

<sup>6</sup>Si  $a, b$  sont positifs, e.g., des projecteurs,  $\text{tr}(ab) \geq 0$  ; si  $a, b$  sont hermitiens, on a seulement  $\text{tr}(ab) \in \mathbb{R}$ , d'où la forme de (10), puisque  $a, b$  sont des hermitiens arbitraires.

1986 ; jusque là, j'étais sur des positions trinitaires sémantique/syntaxe/méta (relativement) standard. Voilà qu'apparaissent les réseaux de démonstration et subitement un problème de syntaxe : comme on a plusieurs conclusions, on n'a pas indique de dernière règle et il se peut qu'il n'y en ait pas du tout, i.e., que la « démonstration » écrite soit logiquement fausse. Trouver le critère permettant de discriminer les *bonnes* démonstrations des autres revient à soumettre la prétendue démonstration à un certain nombre de tests interactifs, e.g., certains modes de parcours de graphe. Ces tests se révèlent, à l'analyse, complètement homogènes à des démonstrations de la négation, en particulier, ils peuvent être manipulés de façon déductive (grande différence avec la thématique de l'inversion, *supra*, où rien de ce que l'on peut faire n'est déductif). Il s'agit bien de *surmoi* logique, puisque c'est la négation qui décide de la possibilité d'écrire ou de ne pas écrire une règle ! La géométrie de l'interaction et la ludique sont les développements de cette remarque. Par exemple, la GdI *ante literam* met en dualité les permutations d'un même ensemble fini  $X$ , avec pour polarité :

$$\sigma\tau \text{ cyclique} \quad (11)$$

La *ludique* repose sur les *desseins* qui sont des espèces de stratégies dans un jeu très laxiste : quand deux desseins  $\mathcal{D}$ ,  $\mathcal{E}$  sont en situation d'interagir, leur interaction se traduit par une *dispute*  $[\mathcal{D} \rightleftharpoons \mathcal{E}]$  qui produit, soit un résultat, le *daimon*  $\mathfrak{X}$ , soit qui ne produit rien : le « résultat »  $\langle \mathcal{D} | \mathcal{E} \rangle$  est alors noté  $\Omega$ , c'est la *foi* (en ce qui n'arrivera pas). On prend pour espace scalaire  $R = \{\mathfrak{X}, \Omega\}$  et pour pôle  $P = \{\mathfrak{X}\}$ . La dualité s'articule autour de :

$$\langle \mathcal{D} | \mathcal{E} \rangle = \mathfrak{X} \quad (12)$$

Les ensembles polaires de desseins sont appelés *comportements*. La ludique n'est pas un « vrai » jeu, avec règle du jeu, arbitre : on a trop de liberté pour « jouer ». Cela dit, la possibilité que la dispute ne mène à rien (soit parce que l'un des deux traîne à répondre, soit à cause d'une scène de ménage, i.e., d'une dispute infinie) va produire un surmoi logique. Par exemple, si  $\mathbf{A}$  est positif (i.e., « joue » en premier) et que l'on veuille lui interdire un certain coup (ou *action*)  $\kappa$ , il suffit que  $\sim \mathbf{A}$  abrite un dessin  $\mathcal{E}$  qui ne répond rien à  $\kappa$ , en effet, si  $\mathcal{D}$  commence par  $\kappa$ , alors  $\langle \mathcal{D} | \mathcal{E} \rangle = \Omega \notin P$ . Par contre, si  $\mathcal{D}$  ne commence pas par  $\kappa$ , alors  $\mathcal{E}$  « abandonne », i.e.,  $\langle \mathcal{D} | \mathcal{E} \rangle = \mathfrak{X}$ . Le (contre-)dessin  $\mathcal{E}$  est un pur « chien de garde », dont la seule fonction est négative, normative. Il ne faut pas croire que  $\sim \mathbf{A}$  soit pour autant un *comportement* débile. En effet, il peut tout à fait abriter des desseins qui ne sont pas des chiens de garde. Le chien de garde est purement dissuasif : si tu ne suis pas ma règle, je t'embrouille dans une scène de ménage. Et d'ailleurs, par symétrie,  $\mathbf{A}$  possède aussi des chiens de garde dont le rôle est de brider  $\sim \mathbf{A}$ .

Dans un comportement, un dessin perd un peu de ses moyens, tout comme Sarah Palin qui a peut-être une cervelle après tout, mais qui, en tant que candidate républicaine, n'utilise guère que ses canines. C'est la théorie de l'*incarnation* : deux desseins d'un même comportement seront équivalents s'ils induisent les mêmes disputes,  $[\mathcal{D} \rightleftharpoons \mathcal{E}] = [\mathcal{D}' \rightleftharpoons \mathcal{E}]$  contre tout élément de  $\mathcal{E} \in \sim \mathbf{A}$ . Dans les desseins, cette équivalence se traduit par l'ablation des organes (ici, des actions) qui ne seront jamais utilisés, et donc, plus un comportement est grand, plus son surmoi est petit, et donc plus il devient difficile de distinguer entre desseins : l'incarnation est contravariante.

L'idée que le comportement de l'adversaire (i.e., l'ensemble de toutes les stratégies qu'il peut adopter) définit en fait la « vraie » règle du jeu, se retrouve dans des situations familières : si vous jouez aux échecs avec un débutant, vous ne lui sortez pas la variante Capablanca de l'ouverture Nimzo-indienne, vous lui faites un bon coup du berger, ça marche et c'est plus drôle ! Autrement dit, comme l'opposant a moins de possibilités du fait de son ignorance, c'est comme si on avait changé de règle du jeu.

Il est temps de revenir sur la négation proprement dite : j'ai été amené à la négation comme adjoint (ou dual, polaire, etc.) par diverses raisons que je pense avoir amplement discutées et justifiées. Ce faisant, j'ai dû plonger les preuves dans un espace plus large d'*épreuves* qui inclut, par exemple, les « chiens de garde » de la ludique. La question qui se pose alors est la séparation du bon grain de l'ivraie : comment, parmi les épreuves, caractériser les preuves. Autrement dit : le chien du jardinier n'a pas droit à la salade.

En ludique, la solution est extrêmement simple : on demande que le dessein  $\mathcal{D} \in \mathbf{A}$  soit *gagnant*, autrement dit que, pour tout contre-dessein  $\mathcal{E} \in \sim\mathbf{A}$ , l'abandon par lequel se conclut la dispute  $[\mathcal{D} \rightleftharpoons \mathcal{E}]$  soit dû à  $\mathcal{E}$ . Autrement dit, que  $\mathcal{D}$  soit une stratégie *gagnante* ; en se restreignant à l'incarnation, cela veut dire que  $\mathcal{D}$  n'utilise pas le *daimon*  $\mathfrak{X}$ . Les preuves étant identifiées aux desseins gagnants, la vérité d'un comportement devient l'existence d'une preuve. Ce qui se décline en des résultats de complétude ou d'incomplétude, selon que ces preuves correspondent ou non aux habituels artefacts syntaxiques, les *démonstrations*. Cette définition a pour vertu immédiate d'assurer la cohérence logique : en effet,  $\mathcal{D} \in \mathbf{A}$  et  $\mathcal{E} \in \sim\mathbf{A}$  ne peuvent pas être tous deux gagnants !

La géométrie de l'interaction est, *mutatis mutandis*, à la ludique, ce que les espaces cohérents quantiques sont aux espaces cohérents. La différence est que, alors que les ECQ fonctionnent dans un cadre plus restreint que les espaces cohérents, la GdI est plus universelle que la ludique. C'est un sujet trop technique ; je me contenterai d'un cas particulier. Si l'on était en dimension finie  $n$ , la polarité entre deux matrices hermitiennes de norme au plus 1 s'énoncerait au moyen des deux conditions :

$$\varrho(ab) < 1 \tag{13}$$

$$\det(1 - ab) \neq 1 \tag{14}$$

(13) porte sur le rayon spectral, (14) est proche, par son esprit, de la polarité (10) des espaces cohérents quantiques. La question qui se pose est celle du *gain* : peut-on, dans certains cas, déterminer un gagnant entre  $a$  et  $b$  ? La seule réponse raisonnable consiste à demander que  $a$  soit une symétrie partielle ( $u^3 = u$ ) et que, par rapport à une base *fixée une fois pour toutes*,  $a$  apparaisse comme une matrice à coefficients 0 ou 1. Si  $a$  satisfait cette condition,  $b$  ne le peut pas, car, en effet, les itérées  $(ab)^k$  seraient à coefficients 0 ou 1, et soit un coefficient 1 finit par apparaître en diagonale, auquel cas  $\varrho(ab) = 1$ , soit  $(ab)^n = 0$ , auquel cas  $\det(1 - ab) = 1$ .

L'intérêt de cette courte digression « non commutative » est qu'elle définit la vérité par rapport à une base, un *point de vue*. La vérité devient ainsi *subjective*. Cela dit, cela n'ouvre guère de perspective à une quelconque forme de révisionnisme ; en effet, la relativité de la vérité s'applique à un énoncé isolé, qui peut être vrai ou faux selon le point de vue, en aucune façon à deux énoncés (car il faut pouvoir harmoniser les points de vue). De même la relativité de la

vitesse n'empêche pas que les collisions de véhicules à vitesse constante (et donc relativistiquement parlant, individuellement nullifiables) soient mortelles. La relativité de la vérité ne met pas non plus en question les théorèmes familiers des mathématiques, car leur signification réelle suppose un découpage privilégié de l'espace de représentation de la pensée et donc, sinon un point de vue unique, du moins une très forte restriction sur les *points de vue* signifiants.

## En guise de conclusion

J'ai essayé de faire le tour des principaux aspects de la négation. J'ai cependant négligé la *polarisation* que nous avons déjà rencontrée ; il s'agit de la reconnaissance de deux polarités (positif/droit/actif vs. négatif/gauche/passif) échangées par négation. Typiquement, dans un jeu « je commence » (positif) vs. « tu commences » (négatif), la négation linéaire échangeant les joueurs. La restriction courante aux seuls jeux négatifs « tu commences » produit donc des négations non involutives, puisqu'il faut ajouter un coup « bidon » pour échanger les joueurs. La polarité apparaît avec la distinction due à Andreoli entre énoncés négatifs *asynchrones* (inversibles) et positifs *synchrones* (sujets à une rétrospection non triviale). De même, cette distinction est essentielle dans le système classique **LC**, les positifs étant duplicables, alors que les négatifs sont des dupicateurs ; le sous-système de **LC** connu sous le nom de  $\lambda\mu$ -calcul qui ne garde que la partie négative de **LC** n'est pas vraiment classique, puisqu'il n'a pas de négation involutive. Classiquement, il est impossible de fondre les deux polarités : en effet, si chacun a le droit de dupliquer l'autre, cela se traduira par une dichotomie infinie des tâches et l'impossibilité de passer à l'acte.

Les travaux les plus récents de GdI reposent eux aussi sur une polarisation ; celle-ci ne recouvre pas les notions déjà existantes, ce qui fait que la polarisation est un chantier dont il serait prématuré d'entreprendre la visite guidée.

NON SI NON LA

## Références

- [1] J.-Y. Girard. **Le point aveugle, tome 1 : vers la perfection**. Visions des Sciences. Hermann, Paris, 2006. 296 pp.
- [2] J.-Y. Girard. **Le point aveugle, tome 2 : vers l'imperfection**. Visions des Sciences. Hermann, Paris, 2007. 299 pp.