Le Point Aveugle

Cours de logique.

Tome 1 : vers la perfection.

Jean-Yves Girard
Institut de Mathématiques de Luminy, UMR 6206 – **CNRS**
163, Avenue de Luminy, Case 930, F-13288 Marseille Cedex 09

girard@iml.univ-mrs.fr

Table des matières

Plan du tome 2 : vers l'imperfection

Avant-Propos

Ce petit cours de théorie de la démonstration s'adresse *a priori* aussi bien aux mathématiciens qu'aux informaticiens, aux physiciens qu'aux philosophes et aux linguistes ; et, comme nous ne sommes plus au XVI^{e} siècle — ni même au XVIII^{e} —, il est bien évidemment voué à l'échec. À l'inverse d'un cours sur des sous-domaines qui fonctionnent bien (théorie des modèles, théorie des ensembles), très moyennement (logique temporelle, logique modale), voire pas du tout (logique quantique, logique épistémique) et qui s'enracinerait alors dans une excellence technique, ou, plus prosaïquement, dans un cercle d'entraide bien comprise. Cela dit, on peut se donner d'autres objectifs que la réussite pure et simple, par exemple, la révélation d'un désordre dans cet univers apparemment bien rangé, au sein duquel la logique a finalement pris place entre deux pots de confiture et l'almanach des postes et ne dérange plus, ne dérange surtout pas, comme un gros chat qui ronronnerait dans son coin.

Au début du siècle dernier, le chat était plutôt chien-loup et aboyait très fort : le XX^{e} siècle fut celui des totalitarismes en tous genres et en particulier celui du totalitarisme (rebaptisé « tournant ») linguistique. Cette forme extrême de *scientisme* consistait à ramener toute question mathématique (et donc, tout étant supposé mathématisable, toute question) à un problème de protocoles formels, langagiers, bureaucratiques : on attendait Kafka. Moins littérairement, le même scientisme s'attaquait dès 1904 à l'amélioration de la race en Namibie, que de pendus dans ce pays sans arbre ! La logique moderne reste fondamentalement imprégnée de « l'esprit 1900 », cette espèce de prétention à tout simplifier, puisqu'on peut tout résoudre. Quand, après 1930, l'incomplétude viendra secouer cette morgue, on n'observera guère qu'une complexification du discours : au lieu d'expliquer à partir de plus simple, on expliquera à partir du « méta »; commencera alors le temps de la fausse monnaie. C'est aussi à partir de ce moment que la logique, incapable de se réformer, se coupera des mathématiques, de la physique, etc.

Un sophisme typique : a quoi bon chercher de belles structures mathématiques pour la logique ? Il ne peut pas y en avoir, puisque les mathématiques, les bonnes comme les mauvaises, se traduisent en logique : sa structure doit refléter le pire, i.e., ne pas exister, ou, du moins, rester très molle. Donc, si l'on cherche une interprétation topologique, continue, de la logique, on ira directement au pire (e.g., les domaines de Scott) et on en sera même fier ! Parmi les détails révélateurs, cette insistance des logiciens à choisir des symboles contre-intuitifs, pour surtout ne pas suggérer que certaines propriétés, par exemple la distributivité, pourraient être plus importantes que d'autres[1] : « Plus importantes, vraiment ? Ça se définit comment, l'importance ? ». Je pense à cette remarque de ma fille Isabelle — alors très jeune : « Pourquoi ne pas appeler la porte "cuiller" et la cuiller "porte" ? », à laquelle j'avais répondu : « Quand on dit "prends la porte", ce ne doit pas être pris comme une invitation à dîner ». Parmi les erreurs magistrales de la logique, on mentionnera la logique quantique, dont le ridicule ne s'explique que par un sentiment de supériorité du langage — et des idées, même mauvaises, dès qu'elles prennent la forme écrite — sur le monde physique. La logique quantique est, en effet, une espèce de punition infligée à la nature, coupable de ne pas obéir aux préjugés des logiciens... on pense à Xerxes faisant fouetter la mer qui avait emporté son pont de bateaux.

Il y a cent ans, bien rares étaient ceux qui osaient s'opposer aux certitudes scientistes. Après un siècle de massacres, c'est devenu plus facile : bien que les mêmes âneries reviennent sempiternellement — telle que l'idée d'un robot intelligent, fantasme de l'intelligence artificielle, prothèse improbable pour ceux qui en auraient tant besoin —, il est permis de nos jours de se moquer des *Jivaros* de la science. Par exemple de H. Simon, celui qui a fait retrouver la troisième loi de Kepler (les carrés et les cubes) par une machine, oubliant que ce n'est pas la loi qui lie la période au demi-grand axe, mais l'idée-même d'une telle loi, qui est difficile à trouver. Surtout quand on est... astrologue comme Kepler.

Il faudrait aussi remarquer que, malgré son lourd passif scientiste, le bilan de la logique, quoique maigre, n'est pas nul. La théorie des modèles et la théorie des ensembles se portent plutôt bien ; même la théorie de la démonstration a un bilan non négligeable et d'ailleurs de quoi partirais-je sinon, puisque c'est essentiellement de *théorie de la démonstration* qu'il va s'agir ?

Au début du siècle passé, la relativité d'Einstein et, de façon plus radicale, la mécanique quantique, remirent en cause nos « intuitions fondamentales ». La logique, par ses outrances, choisit alors de déboucher sur le vide : la non-structure, le non-signifiant « tout est codable en tout[2] ». Pourtant, dans le « tournant linguistique », l'idée de la prégnance du langage était tout à fait géniale et ne méritait pas de devenir cette *machine à décerveler* que je viens d'évoquer. Si l'on y regarde de plus près, elle contient en germe une autre

[1]Voir par exemple, au début de la logique linéaire, le point d'honneur mis par certains à noter le « par » $+$ et le « avec » $\times$, alors que « par » distribue sur « avec ».

[2]On a envie d'ajouter « et réciproquement ». L'idée de codage mutuel est ancienne et universelle : pensons à des Esseintes et son *orgue à liqueurs* (Huysmans, *À rebours*, 1885) et aussi à l'idée de traduire les images en sons, ou plutôt en borborygmes !

forme de « relativisation », en fait de *déréalisation* de la nature. C'est le point de vue que je vais essayer de développer.

On a parfaitement le droit de trouver ce projet complètement fou et lui préférer des déclarations du genre : « *Un langage est un alphabet fini avec lequel on construit des termes, des énoncés, des démonstrations — la syntaxe — ; le langage est ensuite interprété dans un modèle — la sémantique — ; enfin tout cela est formalisé dans un méta-système.* ». Mais alors on ne fait plus de logique, en tout cas pas de « fondements », on va à la pêche au phoque dans le détroit de Behring : le domaine, tel qu'il s'est ossifié, est tout sauf fou, c'est un cimetière d'idées. La seule excuse, au XXI^e^ siècle, pour faire *quand même* de la *théorie de la démonstration*, des « fondements », c'est le grain de folie.

Un mot sur le titre : c'est en révisant le texte (été 2005) que je me suis aperçu de la récurrence de l'expression « point aveugle ». Le point aveugle c'est ce qu'on (qui) ne voit pas et on ne sait même pas qu'on ne le voit pas[3]. Le point aveugle le plus trivial, c'est cette mauvaise logique modale qu'on justifie par une mauvaise sémantique de Kripke et *vice versa*, on a envie de dire *vide versa* ; mais on trouve des aveuglements similaires dans les interprétations beaucoup plus élaborées. La bonne nouvelle de ce cours, c'est qu'il semble que le point de vue *procédural* soit à même de débusquer le non-dit, le non-vu. Simplement, alors que l'absence de *Hauptsatz* suffit à montrer que la logique **S5** est fausse, il faut travailler nettement plus pour imaginer ce qui pourrait être faux dans les principes justifiant la fonction 2^n.

Quant aux sous-titres des deux tomes, ils sont liés à la distinction fondamentale entre *parfait* (au sens du perfectif des langues slaves) et l'*imparfait*, i.e., l'inachevé, l'infini. C'est la *logique linéaire* qui introduit les opérations parfaites et, partant, cette distinction jusque là inexistante ; c'est l'aboutissement du premier tome. Le point aveugle se réfugie alors dans l'imperfection et le tome 2, en particulier les derniers chapitres, est consacré à une relecture iconoclaste de l'imperfection.

Remerciements

Il s'agit des notes d'un cours donné dans le cadre de l'Université Franco-Italienne, d'octobre à décembre 2004, à l'Università Roma Tre, (Dipartimento di Filosofia). Toute ma reconnaissance va à Michele Abrusci qui est à l'initiative de ce cours. Je remercie particulièrement pour leur aide Lorenzo Tortora de Falco et Marco Pedicini ; et — *last but not least* — celui qui fut la cheville ouvrière de cette entreprise, Roberto Maieli : c'est lui qui s'est occupé de l'organisation des cours, des *streamings* video, de la mise en ligne du texte.

Trois mois de cours avec rédaction simultanée de notes d'à peu près 500 pages, c'est très lourd, cela prend six jours par semaine. Je n'aurai pas tenu

[3]Kreisel en 1984, parlant de certains Américains : « Ils n'ont pas d'âme et ils ne savent pas qu'ils n'en ont pas ».

sans l'auditoire, pas très nombreux, mais fervent. Sans parler de l'attention constante de Francette-Louise.

Deux sections supplémentaires, 15.C et 15.D ont été rédigées pour l'essentiel par Olivier Laurent. La version définitive tient compte de corrections suggérées par Thomas Streicher (chapitre 2), Philip Scott (chapitre 8) et de la longue liste de « typos » trouvées par Akim Demaille.

NON SI NON LA

Première partie

Les bases

Chapitre 1

Existence contre essence

1.1 L'opposition existence/essence

Le célèbre film de Kubrick « 2001 » s'ouvre sur rien moins que la création de l'intelligence : une tribu de singes particulièrement débiles, incapables de faire face aux lions, sangliers, etc., reçoit l'intelligence d'un monolithe tombé du ciel. Remarquons que :

(i) Les « grands galactiques » dépositaires de cette intelligence qui se transmet comme une maladie ont dû eux aussi être singes, ou l'équivalent, dans leur jeunesse et donc...

(ii) L'intelligence est prise comme un attribut absolu, indépendant de toute expérience, de toute interaction, un logiciel qu'on implante sur un *hardware* préexistant. Le même monolithe, tombant au milieu des Galapagos et voilà les reptiles aux commandes du monde... L'Âge de la Tortue ?

En fait, n'était le scientisme affligeant de l'auteur (Arthur C. Clarke), on pourrait y voir une sorte de miracle du type Fatima.

Par rapport à ce type d'« explication », il y a des réactions très marquées, instinctives : soit c'est génial, soit c'est débile, selon qu'on est plutôt *essentialiste* ou *existentialiste*.

Essentialistes

Ce sont ceux qui pensent que tout est déjà là, que l'on ne fait que répéter des archétypes ; ils croient aux soucoupes volantes, surtout sur les pyramides aztèques. En logique, ils sont partisans des « fondements inversés » : un système s'explique par un « méta-système » plus profond ; qui s'explique donc lui-même par un méta-méta-système encore plus profond... et on n'en finit pas[1]. Ce qui pourrait sembler un vice de construction (supposer quelque chose, voire un peu plus) apparaît comme la fascination de l'irréductible. Le père tutélaire de l'essentialisme est le théologien Thomas d'Aquin, et les philosophes thomistes médiévaux ont profondément influencé notre conception de l'immuable, du *nécessaire* (ce qui se retrouve dans les logiques modales) et surtout de l'infini : avec la meilleure volonté du monde, on a du mal à ne pas « voir » les entiers naturels, une liste infinie, alignés sur un mur comme des trophées de chasse. La création de la théorie des ensembles, à partir des travaux de Cantor, Dedekind, etc. fin XIX^{e} siècle, est sûrement un des surgeons les plus inattendus du thomisme.

Existentialistes

Ce terme — d'utilisation un peu délicate, après sa vogue excessive vers le milieu du siècle passé — qualifierait plutôt ceux qui trouvent infantiles les idées de Kubrick (voir aussi la dernière image de « The Shining », qui suggère un éternel retour) et ne croient pas aux soucoupes civilisatrices. Ceux qui ne trouvent pas « profond » de définir la vérité comme l'a soi-disant fait Tarski « $A \wedge B$ est vrai si A est vrai **et** B est vrai ». Si la vérité ne se « définit » qu'à travers un pléonasme ($\& :=$ **et**), c'est qu'elle n'a pas de sens, *même si l'on peut la manipuler mathématiquement*, comme dans le *schéma de réflexion* (section 3.B.4). La meilleure loi n'a de valeur que si on peut la justifier, i.e., montrer l'effet de la non observance.

Il va sans dire que le point de vue de ce cours est *plutôt* celui de l'existence. Ce qui ne va pas sans certaines difficultés :

(i) Le thomisme, qui régente discrètement la culture occidentale, nous ramène toujours dans les mêmes ornières. Par exemple, on a essayé d'opposer à la version officielle de l'infini (l'infini « actuel », tout prêt), un infini dynamique, en construction, « potentiel ». Puis, pour parler de potentialité, on a considéré *tous* les possibles (modèles de Kripke et autres amusettes)... et voilà comment on assassine une idée : si le potentiel est vraiment potentiel, on ne peut pas épingler tous les potentiels possibles comme de vulgaires papillons, cette collection-là n'a pas de sens[2].

(ii) Le monde existentialiste est un monde sans lois, où la remise en cause est totale et constante ; ce n'est pas viable. Ce point de vue ne prend corps que comme « grain de sel » contre la suffisance essentialiste. En

[1] Dans une blague célèbrissime, le monde repose sur une tortue, qui repose elle-même sur une autre tortue... « Turtles all the way down ». On peut aussi évoquer les escroqueries pyramidales, qui consistent à reporter une dette sur un hypothétique suivant.

[2] Dans les logiques modales, la modalité duale de la nécessité, la *possibilité*, s'en tire toujours assez mal. Par exemple, dans la version « mondes possibles », l'interprétation de la phrase « si ma tante en avait » est franchement comique.

fait, plutôt que d'une opposition figée entre deux points de vue (ce qui est encore de l'essentialisme), il faudrait mieux penser à leur interaction réciproque.

L'opposition essence/existence est au cœur du *typage* (voir section 6.F).

1.2 Projets essentialistes et existentialistes

1.2.1 La théorie des ensembles

Le développement de l'analyse mathématique « moderne » dans la seconde partie du XIX^e siècle tendait à clarifier des notions un peu vagues : nombres réels, fonctions continues. Très rapidement on se trouva face à des paradoxes[3] : la courbe sans tangente, ou pire, la courbe de Peano qui « remplit » un carré ; on pouvait craindre le pire, par exemple, que la différence entre courbe et surface ne soit qu'une « illusion d'optique »... sans parler de l'incohérence pure et simple. Pour essayer de résoudre ce problème, la théorie des ensembles (créée pour l'essentiel par Cantor) essaya de tout ramener à un cadre commun. Ce cadre est essentialiste car il suppose une utilisation sans complexe de l'infini. La première théorie des ensembles, dite « naïve », est basée sur le *schéma de compréhension* : toute propriété P définit un ensemble X, noté encore $\{x; P[x]\}$, « l'ensemble des x tels que P » :

$$\exists X\, \forall x\, (x \in X \quad \Leftrightarrow \quad P[x]) \tag{1.1}$$

Dès 1898, on (Burali-Forti) découvre une contradiction, une *antinomie* — pas un simple paradoxe — dans la théorie des ensembles, dont on se rappelle la version simplifiée donnée par Russell en 1902 : soit $a := \{x; x \notin x\}$ (l'ensemble des ensembles qui n'appartiennent pas à eux-mêmes) ; alors $a \in a \Leftrightarrow a \notin a$, contradiction[4]. Ce n'était pas si grave que cela, vu qu'il s'agissait d'un système tout à fait expérimental ; et d'ailleurs, il ne fallut que peu de temps à Zermelo pour formuler des restrictions raisonnables sur le schéma (1.1), restrictions qui sont pour l'essentiel celles du système de théorie des ensembles **ZF**, toujours en vigueur de nos jours, à la satisfaction de tous les utilisateurs.

Avant de quitter pour toujours la théorie des ensembles, rappelons que la solution passe par la distinction entre deux sortes d'objets, d'un côté les ensembles, qui sont « petits », de l'autre les classes, qui sont « grandes », typiquement le $\{x; x \notin x\}$ de Russell est une classe, pas un ensemble. Remarquons que la distinction ensemble/classe est nécessaire sans être absolue : comme ce

---

[3] Pas des contradictions, tout simplement des résultats contre le sens commun.

[4] C'est en effet contradictoire dans tous les systèmes « honnêtes », avec une exception notable : les logiques *allégées* **LLL** et **ELL** (voir chapitre 16).

rideau mobile qui sépare la classe affaires de la classe touriste, le départ ensemble/classe s'effectue au niveau d'un cardinal inaccessible assez arbitraire. Question essentialisme, c'est donc un peu raté !

1.2.2 Le projet hilbertien

Le « Programme de Hilbert », qui date pour l'essentiel des années 1920, correspond à la vision *formaliste* des mathématiques. Cette approche est souvent simplifiée en : « Les mathématiques sont un pur jeu de symboles, sans plus de signification que le jeu d'échecs ; tout ce qui compte c'est la cohérence formelle de la règle du jeu », mais il ne faudrait pas prendre ce type de provocation au pied de la lettre : il y a une pensée complexe, quoique réductrice, derrière le formalisme hilbertien. Sans renier un instant l'intérêt de la théorie des ensembles, Hilbert prend en fait à rebrousse-poil l'essentialisme ensembliste, en particulier ses aspects infinis. Ce qui suit est une version non autorisée, mais proche de la pensée de Hilbert.

(i) Nos intuitions sur l'infini nous trompent ; on parle d'objets, et on ne les voit pas. Par contre, le raisonnement, qui s'exprime à travers le formalisme logique, est lui, réel, tangible. Il faut donc comprendre ce qu'est une démonstration, sous son aspect le plus mathématique, i.e., formel ; en fait comprendre *à quoi elle sert*.

(ii) Une démonstration produit un théorème A qui pourra être réutilisé (comme lemme) dans une autre démonstration, à travers des règles genre *Modus Ponens* :

$$\frac{A \quad A \Rightarrow B}{B}$$

B à son tour peut servir de lemme pour produire (moyennant $B \Rightarrow C$) C et ainsi de suite ; ce qui n'apporte rien de neuf. En effet, une variante du *Modus Ponens* nous donne :

$$\frac{A \Rightarrow B \quad B \Rightarrow C}{A \Rightarrow C}$$

autrement dit, on aurait pu faire un *Modus Ponens* entre A et $A \Rightarrow C$, pour obtenir directement C.

(iii) Le seul espoir d'arriver à une explication convaincante par cette méthode consiste à considérer des conséquences B d'une forme particulièrement simple, donc susceptibles d'analyse immédiate.

Le choix le plus évident consiste à prendre $B = \perp$, où $\perp$ est le symbole logique pour l'absurdité. On rappelle que dans ce cas-là, $A \Rightarrow B$ est équivalent à la négation $\neg A$. Que peut-on dire de l'absurdité ? Pas grand-chose, sinon qu'on la rejette et qu'elle ne saurait donc être prouvable. Le *Modus Ponens* prend une forme symétrique

$$\frac{A \quad \neg A}{\perp}$$

une sorte de dualité entre démonstrations de A et démonstrations de $\neg A$, avec ceci de particulier qu'on ne l'énonce que pour mieux la fuir : cette situation ne doit jamais se produire, i.e., le système doit être cohérent, non contradictoire. Il ne doit pas prouver l'absurdité, i.e., un énoncé et sa négation. Voilà l'origine de l'idée de démonstrations de cohérence.

L'inconvénient de cette restriction extrême, c'est qu'elle relie la production de A à sa *non* utilisation (au moyen de $\neg A$). Une version moins frustrante de la même chose consiste à admettre que B puisse être une (in)équation numérique entre entiers, genre $2 + 3 = 5$, $(2 \times 7) + 4 \neq 18$, etc. La contrainte est alors que cette équation doit être vérifiée ; par exemple, on ne peut pas obtenir $(2 \times 7) + 4 \neq 18$. Cette version est équivalente à la cohérence, en effet, elle est plus générale (prendre $0 \neq 0$ comme absurdité) ; mais pas plus, car si $(2 \times 7) + 4 \neq 18$ était démontrable, comme on peut sûrement démontrer $(2 \times 7) + 4 = 18$ (on a évidemment l'appareillage axiomatique pour cela), on obtiendrait une contradiction. Plus généralement, on pourrait admettre que B prend la forme d'une (in)équation universelle, $\forall n (n + 1)^2 = n^2 + 2n + 1$ ou $\forall n \forall p \forall q\ (n^2 + p^2 \neq q^2)$, ce que j'appellerai un énoncé *récessif* (voir section 2.1.4), sans changer grand-chose à la discussion précédente. Pour Hilbert, le récessif est la limite du signifiant *stricto sensu*, le non-récessif est une espèce de douairière qui ne signifie que par délégation ; les énoncés récessifs auraient un sens *finitaire* immédiat et c'est pour cela que le formalisme hilbertien est aussi qualifié de « finitisme ». Observons une contradiction, une tension dans le finitisme : tout est fini, mais ce tout du fini est lui-même infini. On voit bien que cela ne tourne pas tout à fait rond. . .

Un dernier mot pour justifier en quoi la position de Hilbert est loin de l'essentialisme : pour lui, la cohérence est un substitut à l'existence. Au lieu de démontrer l'existence d'un objet a tel que $A[a]$, on démontrera la cohérence de l'énoncé $\exists x A[x]$. Cette très belle idée, qui apparaît dès 1904, se révèle un peu stérile, du fait de l'impossibilité[5] de produire des démonstrations de cohérence. Dans la pratique, on déduit la cohérence de $\exists x A[x]$ à partir de l'objet a et non l'inverse.

1.2.3 Le projet de Brouwer

Dans la lignée des critiques parfois injustes, mais toujours pertinentes, de Poincaré à l'encontre du formalisme, Brouwer propose une relecture totalement anti-logique de l'infini. Si de nos jours, on perçoit ce qui unit Brouwer

[5]Théorème d'incomplétude.

à Hilbert, on doit dire que, dans les années 1920, les relations furent plutôt fraîches, du genre Don Camillo (Brouwer) contre Peppone (Hilbert) — mais sans la moindre sympathie cachée —, le point d'achoppement étant le scientisme. D'ailleurs, l'école fondée par Brouwer, l'*intuitionnisme*, ne proclame-t-elle pas la primauté de l'intuition sur le langage ? Brouwer ne rejette pas l'infini (au contraire du finitisme hilbertien qui n'y voit qu'une façon de parler), par contre, il refuse les aspects les plus thomistes, « actuels » de l'infini : tout particulièrement la théorie des ensembles et l'idée qu'on puisse définir une fonction de la variable réelle point par point. Certains principes, valables dans le domaine fini, cessent de l'être dans le domaine infini, comme le tiers-exclu $A \vee \neg A$. La justification usuelle du *tiers exclu*, le *tertium non datur* est qu'un énoncé A a une valeur de vérité (est vrai ou faux). Cela dit, si l'on peut *calculer* une valeur de vérité dans le cas fini, aucun algorithme ne peut procéder à la vérification d'une infinité d'étapes, et c'est pourquoi Brouwer le met en doute dans le cas infini. On voit que Brouwer est très moderne, dans le sens où la vérité n'existe pas indépendamment des moyens, des protocoles, de vérification ; il faut le rapprocher de la physique quantique et oublier le subjectivisme, voire le solipsisme, dans lequel il s'est parfois égaré.

J'aurai largement l'occasion de reparler de l'intuitionnisme, dont la logique linéaire est avant tout une version « symétrisée ». Concluons cette brève introduction à Brouwer par deux confrontations (imaginaires) avec Hilbert.

Modus Ponens : On a vu que, pour Hilbert, la conséquence logique doit surtout ne pas faire d'erreurs hénaurmes, prouver l'absurdité par exemple ; cela donne un statut de citoyen de seconde zone à la plupart des énoncés logiques : ils n'ont pas de sens en soi, on leur demande de ne pas causer de désordre, c'est tout. Au contraire, pour les intuitionnistes, le *Modus Ponens* n'est pas un administrateur judiciaire, c'est la porte sur un nouveau monde, c'est l'application d'une fonction ($A \Rightarrow B$) à un argument (A) qui donne un résultat (B). Les démonstrations ne sont plus des suites de symboles créées par un bureaucrate fou, ce sont des fonctions, des *morphismes*.

Tertium non datur : Hilbert accepte le tiers-exclu ; non pas qu'il croie à une valeur de vérité préexistante pour A, mais parce que cela lui simplifie la vie : « Voilà un énoncé qui ne veut rien dire, mais qui ne me coûte rien ». Techniquement parlant, l'argument de Hilbert est recevable. On sait en fait depuis 1932 (Gödel, encore lui), que la logique classique se traduit dans la logique intuitionniste de façon fidèle : il suffit de mettre des doubles négations « partout ». En particulier, le tiers-exclu lui-même, une fois traduit en $\neg\neg(A \vee \neg A)$ devient prouvable en logique intuitionniste et il n'est donc pas plus « risqué » de l'ajouter. Cela dit, on peut rétorquer que ce n'est pas parce qu'on est assuré de l'impunité (la cohérence) qu'on a le droit de commettre un crime (énoncer un principe non justifié).

1.3 Gödel et après

1.3.1 Échecs et décadence

Avec le confort et la supériorité facile que donne le temps passé, on voit facilement que et Brouwer et Hilbert disaient des choses très intéressantes, mais étonnamment immatures, le premier indice de cette immaturité étant une invraisemblable animosité, typique du XX^{e} siècle. L'échec de Brouwer était patent dès 1930 ; exclu d'une institution prestigieuse, les *Mathematische Annalen* (par le Kaiser Hilbert), il se contenta de régner sur une chapelle circonscrite à la Hollande. Sa tentative de réécriture de l'analyse a sombré dans l'indifférence générale : trop compliquée, prenant le contre-pied d'une tradition bien établie — et satisfaisante pour l'essentiel, il faut bien l'admettre.

La roche Tarpeïenne du programme de Hilbert, ce fut le théorème d'incomplétude de 1931, qui va faire l'objet du chapitre suivant. Sans déflorer le sujet, rappelons qu'il interdit toute démonstration de cohérence. Pour obtenir la cohérence du système **T**, il faut « plus que **T** ». Comment dit-on « tailler rasibus » en langage fleuri ?

Ç'aurait dû en effet couper court au « formalisme », du moins provoquer une réforme salutaire... Que non pas ! C'était sans compter avec les croyants, avec le « besoin de croire ». On ne peut pas donner de fondation finitiste aux mathématiques, généralisons le finitisme[6] ! Que cette généralisation revienne à supposer le résultat, on trouvera bien une astuce pour faire passer la pilule. Résultat : une industrie, essentiellement allemande, qui va, pendant une cinquantaine d'années, produire des démonstrations de cohérence, de valeur épistémologique à peu près nulle à cause de l'incomplétude. Si les seconds rôles de cette étrange aventure n'ont pas leur place ici, il faut absolument mentionner le protagoniste, Gentzen, à la carrière courte[7] et paradoxale. Gentzen donna, en 1936 et 1938, deux démonstrations de cohérence de l'arithmétique. La première fut très mal reçue, car basée sur une définition de vérité, autrement dit attribuant aux énoncés arithmétiques leur sens naïf, infinitaire, or, on ne doit pas sortir du récessif pour être « finitaire ». La deuxième était beaucoup plus acceptable, car basée sur une induction transfinie jusqu'à un ordinal dénombrable ϵ_0 et c'est cette voie qui fut suivie par les épigones de Gentzen, Schütte en particulier. De la seconde démonstration de cohérence de Gentzen, André Weil a pu dire que « Gentzen a démontré la cohérence de l'arithmétique, i.e., de l'induction jusqu'à l'ordinal ω, au moyen de l'induction jusqu'à ϵ_0 », la méchanceté tenant au fait qu'ϵ_0 est nettement plus grand[8] que ω ; et il y a bien quelque chose de cet ordre, même si c'est moins grave que ça

[6] Le premier à le dire fut... Gödel dans son article de 1931 ; naïveté de jeunesse, ou peur des foudres de Zeus ?

[7] Sa mort prématurée en 1945 est due à son obstination de petit soldat du Reich : on ne quitte pas une ≪ ville allemande ≫ (Prague, où il était *dozent* depuis 1943 !).

[8] ϵ_0 est la plus petite solution de l'équation ordinale $\omega^\alpha = \alpha$.

en a l'air. Si la seconde démonstration de Gentzen, la meilleure, a finalement mal vieilli, la première, celle qui n'avait aucune valeur épistémologique, s'en tire beaucoup mieux. En effet, Gentzen commet la faute irréparable de montrer qu'une démonstration établit une vérité, mais ce qu'il fait n'est pas du tout le truisme que l'on craindrait (les axiomes sont vrais, les règles préservent la vérité, donc...). Il donne, dans une terminologie hésitante, une définition interactive de la vérité, la démonstration apparaissant comme une stratégie gagnante dans une espèce de « jeu de la vérité ». Alors, que cette démonstration prouve encore moins de choses que la seconde, qui déjà ne convainquait que les convaincus, la belle affaire... il s'agit de la première interprétation « interactive » de la logique.

Les créateurs manquent de recul quant à leur propre production ; prenons Kepler pour qui l'œuvre de sa vie était la correspondance qu'il avait cru déceler entre les planètes de l'époque (de Mercure à Saturne) et les polyèdres réguliers, « loi » que la postérité passe gentiment sous silence[9]. De même, pour mener à bien son programme de cohérence, Gentzen dut forger des outils, principalement le « calcul des séquents » (voir chapitre 3). Ce calcul, formulé à l'origine dans un but douteux, reste une des créations majeures de la logique du XX^{e} siècle. Il fera l'objet de notre chapitre 3 et on le retrouvera sous diverses formes — déduction naturelle, réseaux, géométrie de l'interaction — tout au long de ce livre. Le théorème de Gentzen, l'*élimination des coupures*, sera finalement formulé en termes... d'algèbres d'opérateurs (voir chapitre 19).

Parmi les idées un peu décadentes, on mentionnera (pour s'amuser un peu) les itérations de théories. On part d'un système d'arithmétique, disons $\mathcal{T}_0$ et on le « fonde » sur un méta-système $\mathcal{T}_1$ ($\mathcal{T}_0$ auquel on a rajouté la cohérence de $\mathcal{T}_0$). Comme $\mathcal{T}_1$ ne prouve pas sa propre cohérence, on peut la rajouter en formant un « méta-méta-système » $\mathcal{T}_2$, puis $\mathcal{T}_3$... On a reconnu les tortues-gigognes, chacune soutenant la précédente, « Turtles all the way down ». Si vous pensez que personne n'a eu l'idée de se livrer à une telle palinodie, eh bien, détrompez-vous, non seulement on l'a fait, mais en plus on a insisté ! Comme, évidemment, les $\mathcal{T}_n$ ne reposent sur rien, on les assoit sur une $\mathcal{T}_\omega$, puis il faut bien asseoir $\mathcal{T}_\omega$ sur quelque chose, disons $\mathcal{T}_{\omega+1}$. On arrive à des progressions transfinies de méta-théories. La question de savoir « combien » d'étapes transfinies sont permises dans cette fuite en avant relève de la plus haute théologie byzantine. On remarquera au passage qu'une mauvaise idée (les tortues-gigognes) ne s'améliore pas quand on l'itère transfiniment : elle devient simplement une mauvaise idée transfinie.

[9]On se demande pourquoi H. Simon n'a pas lancé son ordinateur là-dessus, le résultat aurait été intéressant, il faut essayer de retrouver aussi les erreurs.

1.3.2 Relectures et renouveau

Dans l'histoire récente de la logique, la figure de Kreisel reste un peu énigmatique. En effet, il eut, surtout entre 1950 et 1970, une immense influence, bien que ses textes soient assez illisibles et que son apport technique soit honorable, sans plus. Une de ses principales vertus aura été de « casser » les illusions post-formalistes, en particulier en s'attaquant aux sacro-saintes démonstrations de cohérence des épigones de Gentzen. Il avait tendance à déplacer les problèmes du point de vue idéologique au point de vue pragmatique, ainsi « une démonstration de cohérence, cela ne convainc pas, mais cela a peut-être des retombées mathématiques ». En général, il insistait sur les applications de la logique aux mathématiques, ce qui a été réalisé d'ailleurs par la théorie des modèles en géométrie algébrique.

Sa plus belle réalisation est peut-être le *schéma de réflexion* [37] qui est obtenu par formalisation de truismes, du type « la démonstrabilité préserve la vérité ». Ce sophisme, convenablement formalisé, devient un puissant outil métamathématique (voir section 3.B.4).

A contrario, Kreisel n'a jamais vraiment compris l'intuitionnisme, qu'il a voulu réduire à ses aspects de système formel, au moment-même où l'isomorphisme de Curry-Howard suggérait une nouvelle dimension de la logique, son aspect catégorique. Son explication de l'interprétation fonctionnelle des démonstrations [36] est trop formaliste pour être honnête (voir section 5.A). Quant à sa tentative de *revival* de l'analyse de Brouwer, elle tourne au cauchemar bureaucratique. Le vrai renouveau, qui a commencé à la fin des années 1960, est dû à diverses personnes, toutes en relation avec Kreisel, certes, mais s'est fait presque contre lui.

On retrouvera ces noms, Bill Howard, Dag Prawitz, Per Martin-Löf, Dana Scott, Bill Tait et bien d'autres, dont je fais partie. On aura le temps d'examiner ces contributions, voire les critiquer[10].

Il faut à peu près 15 ans à une génération montante pour prendre le pouvoir. Le renouveau de 1970 ne devint manifeste que vers 1985, à la lumière de la « révolution informatique ». Pour moi, c'est une date importante, puisque j'ai alors conçu la *logique linéaire*, qui est un peu au cœur de ce cours.

1.3.3 Et demain ?

Il est hors de question de décrire en peu de lignes une évolution qui s'est étendue sur plus de 40 ans. Je voudrais seulement attirer l'attention sur l'aspect de « pari pascalien » de ce cours. L'hypothèse que je fais est l'inadéquation profonde, absolue, de la logique classique et — du point de vue fondationnel s'entend — des mathématiques classiques. Pour comprendre l'énormité de la

[10] C'est le cas pour les domaines de Scott, qui ont mal vieilli, mais qui restent un moment essentiel du renouveau.

chose, rappelons-nous que Kreisel ne s'est jamais départi d'un essentialisme bien compris et que pour lui, tout avait sa place dans un univers tout à fait tarskien. La logique intuitionniste n'était alors qu'une manière d'obtenir des informations fines quant à la « réalité » classique, typiquement des bornes effectives.

Mon hypothèse est que la logique classique, la vérité classique, ne sont que des illusions essentialistes, qui s'auto-entretiennent. Ainsi, on verra que le théorème d'incomplétude peut être lu comme la non existence de la vérité. De même, une longue familiarité avec la logique classique montre que sa structure interne est loin d'être satisfaisante. La logique linéaire (et rétrospectivement, la logique intuitionniste) peut être vue comme une logique qui abandonnerait la sacro-sainte « réalité », pour se concentrer sur sa propre structure : elle parvient ainsi à toucher ce point aveugle où l'essentialisme nous ment, ou du moins se refuse à toute justification autre que « c'est comme ça ». En 1985, l'outil structurant était les catégories et a permis la mise au jour, au sein de la logique, d'une strate *perfective* (les connecteurs linéaires proprement dits) qui n'est pas obturée par l'essentialisme.

Ce qui reste, la partie *imperfective*, les exponentielles, concentre tout l'aspect essentialiste de la logique et les catégories n'arrivent pas à débrouiller quoi que ce soit là-dedans. En raccourci,

$$\text{essence} = \text{infini} = \text{exponentielles} = \text{modalités}$$

et il est intéressant que la logique linéaire réduise l'essence à un noyau modal opaque, surtout si on se rappelle que les modalités sont une création de la philosophie essentialiste. À ce propos, rien n'est plus arbitraire qu'une logique modale : « *J'ai fini cette logique, je peux en avoir une autre ?* » semble-t-on dire de ce côté-là. Ce qui montre l'insuffisance, le côté menteur de l'essentialisme : à trop invoquer le ciel, on ne convainc plus personne. Il semble que la *géométrie de l'interaction*, et son interprétation dans les algèbres d'opérateurs, soit à même de « briser la coquille ». L'idée serait de penser la logique par rapport à ce phénomène ignoré, méprisé des logiciens (qui l'ont traité par-dessous la jambe avec leur calamiteuse logique quantique), la physique quantique. Imaginer des fondements, sinon « quantiques », du moins dans un esprit quantique : toute proportion gardée, un peu ce que fait Connes avec la géométrie non-commutative. Voilà le projet du moment, de quoi s'occuper. Ce qui prend à rebrousse-poil le préjugé habituel sur la relation logique/quantique : au lieu d'interpréter le quantique dans la logique, on va tenter l'opposé.

Évidemment tout cela est conjectural, c'est l'hypothèse d'une autre régularité, d'une autre logique, vivant sa propre vie, sa propre géométrie, hors de tout cadre *tombé du ciel*, tel le monolithe de « 2001 ». Comme dirait Blaise, si ça marche, c'est tout bon, et si ça ne marche pas, on n'a rien perdu.

1.A Essentialisme et platonisme

(Réponse à la question : « Vous parlez d'essentialisme, ne faut-il pas plutôt parler de platonisme ? »)

1.A.1 Le platonisme

Bonne question, car question embarrassante ; d'autant plus qu'on rapproche traditionnellement Platon de l'existentialisme à cause du côté relativement ouvert des dialogues platoniciens, alors qu'Aristote, plus « carré », est rattaché à l'essentialisme, bien qu'il soit nettement moins « fermé » que Thomas. D'abord, rappelons que je ne suis pas philosophe ; mes considérations « philosophiques » sont plus un pavé dans la mare (on ne voit plus les poissons sous tant de mousses accumulées) qu'une réflexion systématique et encore moins systémique. J'ai toujours suivi avec le plus grand ennui les exposés académiques sur les écoles logiques telles qu'on les voyait du temps de... Bernays et j'ai plutôt cherché mes propres questions. Avec le temps, j'ai eu la curiosité de confronter ce que j'avais compris de la logique à ce qu'on ânonnait à la Sorbonne et j'ai trouvé que cela ne collait pas, mais vraiment pas.

Disons d'abord qu'il est douteux que Platon puisse être qualifié de « platonicien » dans le sens où une certaine épistémologie ossifiée l'entend. On a l'impression que cette expression correspond d'abord à une lecture sommaire du *Mythe de la Caverne*. En tout cas, si l'on suivait cette distinction, tous les bons mathématiciens (et tous les bons scientifiques) seraient platoniciens, car ils croient à ce qu'ils font. Le fait de croire à la « réalité » de ce qu'ils fait... sans donner un sens trop précis à cette expression, c'est avant tout la *responsabilité* du scientifique : il ne raconte pas n'importe quoi, ce qu'il dit correspond à « quelque chose ». L'attitude opposée serait du solipsisme, mais faut-il créer une catégorie spéciale pour le non-solipsisme ? Pour en terminer avec le rapprochement supposé essentialisme/platonisme, observons que l'essentialisme, à force d'invoquer le ciel, peut être un art de l'esquive et du non-sens. Ainsi les logiques modales, triomphe de l'essentialisme et... règne de l'arbitraire le plus absolu. On peut légitimement parier que les fabriquants de logiques modales à la chaîne ne croient pas trop à ce qu'il font, sinon ils ne changeraient pas de système tous les quarts d'heure. À l'opposé de l'attitude « platonicienne » qui exige une certaine honnêteté.

1.A.2 Une question de morphologie

Le vrai débat ne semble pas existentiel, mais plutôt *morphologique*. Ces phénomènes, nous les observons, nous leur donnons forme, mais à quoi correspond cette forme ? J'observe un serpent dans mon jardin ; est-ce Dieu, ou le Diable, venu pour me tenter (version essentialiste), ou y aurait-il une source et dans ce cas l'eau en est-elle buvable, utilisable pour les cultures (version existentialiste) ? Prenons comme exemple une discussion entre deux logiciens connus, Gödel et Bernays, dans les années 1940. Aucun des deux ne peut être qualifié de « platonicien » avec la petite nuance de mépris qui s'attache à ce terme. Ils discutent des notions de proposition, de démonstration ; pour Gödel, il y a une espèce « sauvage », les démonstrations, qu'on domestiquerait (par la logique) de façon à leur faire accomplir des actes compréhensibles : prouver des propositions. Pour Bernays[11], tout est réglé d'avance, comme dans la théologie thomiste, d'abord vient la loi, elle définit les choses licites, les enchaînements autorisés ; le phénomène démonstratif n'est plus primitif, ce n'est que l'activité parfaite d'une armée en marche poursuivant son but. Ce n'est évidemment pas écrit ainsi, la discussion est beaucoup plus ennuyeuse que cela. Au passage, on regrette que Gödel (et ce fut le cas de tous, jusqu'à Kreisel inclus) ne puisse argumenter qu'en utilisant des choses du style « la liste de toutes les démonstrations », « la liste de toutes les propositions » et n'ait pas intégré, au minimum, le point de vue catégorique, qui ne s'impose que trente ans après. D'ailleurs, au milieu des années 1960, Kreisel propose une explication plutôt essentialiste, du genre « on met toutes ces définitions dans un système formel donné à l'avance », donc il est plutôt « côté Bernays ». Cela incite aussi à la prudence devant les arguments philosophiques trop poussés, trop systémiques : ils tournent facilement au sophisme, ils démontrent l'impossibilité, la vanité de n'importe quoi. Un débat philosophique, cela s'alimente comme un feu et le combustible, ce sont les percées techniques, qui permettent de voir la porte cachée dans le labyrinthe. Tout cela pour dire qu'il ne faut peut-être pas pousser trop loin cette discussion sur l'essentialisme. Tout ce qu'on veut au fond, c'est progresser *un peu*.

L'essentialisme apparaît donc avant tout comme un *simplisme morphologique*, par exemple concevoir des ensembles nus qu'on habille ensuite comme des mannequins, avec une structure algébrique, puis topologique[12]. Entre nous, les nombres réels $\mathbb{R}$, auxquels tout le monde (toutes chapelles confondues) croit, est-ce que cela fait sens sans l'addition, sans la multiplication, sans la continuité ? L'essentialisme (la théorie des ensembles) répond « oui », mais y croyons-nous ?

La distinction existence/essence est au cœur de ce cours. Pour ainsi dire, on va reprendre le dialogue Gödel/Bernays, en essayant de justifier le point de vue de Gödel, mais avec des méthodes mathématiques modernes, sérieuses. Dans le rôle du grand méchant loup essentialiste, on trouvera Tarski et sa définition de la vérité et, en général, toutes les définitions logiques qui présupposent la logique.

[11]Logicien d'envergure assez moyenne, surtout si l'on le compare à Gödel.
[12]Ce qui rappelle étonnamment la Genèse et la création en sept jours.

1.B Parfait/imparfait et les catégories

La réponse à cette question anticipe sur la suite du cours, en particulier sur la logique linéaire.

1.B.1 Parfait/imparfait

Un autre aspect de l'opposition existence/essence, c'est l'opposition parfait/imparfait, au sens qu'ont ces termes en linguistique. Le parfait correspond à des actions uniques, bien définies, alors que l'imparfait est le mode de la répétition. C'est seulement avec la logique linéaire (voir chapitre 9) que cette opposition entre dans la logique : le *fragment*[13] multiplicatif/additif, celui qui correspond à la *ludique* (voir chapitres 13-14), illustre le paradigme de « non réutilisation » et on peut dire que, dans la mesure où l'on ne doit pas pousser le point de vue de l'existence jusqu'au système, il correspond parfaitement à nos ambitions anti- (ou plutôt a-) essentialistes.

On sait que la partie purement linéaire doit être complétée par une partie modale, les *exponentielles* $!, ?$, de façon à pouvoir parler de l'infini, ce dont le fragment précédent est rigoureusement incapable. $!A$ énonce la pérennité de A, qui devient donc imparfait et l'infini apparaît comme un attribut de la pérennité : l'infini, c'est ce qui ne s'use pas quand on s'en sert. L'imparfait est le mode des généralités, celui des titres de James Bond : « Les diamants sont éternels », « On ne vit que deux fois », ... et le rapprochement avec l'essentialisme est tout à fait justifié. Ajoutons que le caractère modal des exponentielles ne fait que renforcer cet essentialisme.

1.B.2 Logique linéaire et catégories

La logique linéaire est issue d'une prise en compte systématique de l'interprétation catégorique. En particulier, les *espaces cohérents* (voir chapitre 8), proches des espaces vectoriels ou des groupes commutatifs, font apparaître des structures logiques familières en algèbre linéaire : produits tensoriels (connecteurs multiplicatifs) et sommes ou produits directs (connecteurs additifs). C'est l'origine de la logique linéaire proprement dite (le fragment parfait).

Malheureusement, on ne peut pas se passer de l'imparfait (de l'infini) et donc, techniquement, des exponentielles. Or, il se produit à ce niveau un phénomène déjà évoqué pour les logiques modales et de nature circulaire : l'interprétation catégorique reproduit purement et simplement nos préjugés quant à l'infini. Autrement dit, autant l'interprétation catégorique du monde parfait est originale, satisfaisante, autant, finalement, l'interprétation du monde

[13] Un mot employé à tort et à travers ; l'acception la plus générale est celle d'un sous-langage clos par sous-formules. En aucun cas celle d'un « sous-système » obtenu en affaiblissant les axiomes, notion de peu d'intérêt en logique.

imparfait se révèle-t-elle creuse, décevante. L'infini, même interprété dans les catégories en lieu et place des ensembles, se révèle un noyau dur, primitif qu'on ne sait pas briser.

Cela dit, certains systèmes expérimentaux (logiques allégées) permettent d'envisager un infini qui ne soit pas « infiniment infini » (voir chapitre 16). Il semble que les catégories, pas plus que les ensembles, n'aient la finesse nécessaire pour comprendre ces systèmes complètement atypiques. La solution que je propose se fait en termes de *Géométrie de l'Interaction*, une interprétation au moyen des algèbres d'opérateurs, qui fait l'objet des ultimes chapitres de ce livre.

Bibliographie :[31, 32, 16, 15, 36, 27, 28].

Chapitre 2

Le théorème d'incomplétude

2.1 Énoncé technique

2.1.1 La difficulté du théorème

Il est hors de question de rentrer dans les arcanes techniques du théorème de Gödel[1] et ce, pour plusieurs raisons :

(i) Ce résultat, au fond très facile, ne se perçoit bien, comme les peintures de vieillesse de Claude Monet[2], que d'une certaine distance. De près, ce ne sont que détails fastidieux qu'on n'a pas forcément envie de connaître.

(ii) On n'en a pas besoin non plus, car ce théorème est un cul-de-sac scientifique : il signale une voie sans issue. Puisqu'il n'y a rien à chercher par là, il ne sert à rien d'être expert ès théorème de Gödel.

Mais il est important de connaître le sens général, la structure de la preuve. De plus, comme le théorème est un vrai *paradoxe*[3], on est naturellement tenté de le contourner, de l'éprouver, ne serait-ce que pour mieux le comprendre. L'examen des différentes objections au théorème, toutes fausses, demande plus qu'une simple connaissance détaillée de la démonstration. Plutôt que de m'appesantir sur ces détails fastidieux qui « cachent la forêt », je passerai du temps sur ces objections, des plus ridicules aux moins débiles (aucune ne tient la route).

2.1.2 L'argument diagonal

Cet argument consiste à partir de fonctions $g(z)$ et $f(x,y)$, à former $h(x) := g(f(x,x))$; si par hasard on peut mettre h sous la forme $h(x) = f(x,a)$,

[1]Cependant on trouvera beaucoup d'informations techniques en annexe de ce chapitre.
[2]Musée Marmottan.
[3]Au sens littéral, d'« extérieur au dogme », à la $\delta \acute{o} \xi \alpha$.

on obtient $h(a) = f(a, a) = g(f(a, a))$, un point fixe de g, ce qui est évidemment inattendu. En fonction du contexte, on en tirera diverses conclusions, le plus souvent paradoxales.

1 : Paradoxe de Cantor : C'est le fait qu'il n'y a pas de bijection entre $\mathbb{N}$ et son ensemble des parties. En effet, soit (X_n) une énumération des parties de $\mathbb{N}$ et soient $f(m, n) := 1$ si $m \in X_n$, 0 sinon et $g(0) = 1, g(1) = 0$. Alors $g(a) = a$, contradiction.

2 : Antinomie de Russell : Même combat, sinon qu'on remplace $\mathbb{N}$ par l'ensemble de tous les ensembles. Les entiers sont remplacés par des ensembles arbitraires et donc $f(x, y) = 1$ ssi $x \in y$ et, avec le même g que précédemment, le point fixe devient $a = \{x; x \notin x\}$.

3 : Point fixe des programmes : Si (f_n) est une énumération de tous les programmes envoyant $\mathbb{N}$ dans $\mathbb{N}$, si g est un des f_n, alors la construction précédente fournit un point fixe pour g. Comme la plupart des fonctions n'ont pas de point fixe, on en conclut que le point fixe correspond souvent à un calcul divergent. Typiquement, partant de $g(n) := n + 1$,
$a = a + 1 = a + 2 = \ldots$, ce qui montre qu'on a en fait affaire à des fonctions *partielles*.

4 : Point fixe du λ-calcul : Si M est un λ-terme et si $\Omega := \lambda x M(x(x))$, alors $\Omega(\Omega)$ est un point fixe de M. Cette version est à la précédente ce que 2 (Russell) est à 1 (Cantor).

5 : Premier théorème d'incomplétude : Le point fixe des programmes (3), mais en utilisant comme langage de programmation une théorie formelle. $f(m, n)$ est le code de $A_n[\bar{m}]$ et g est la non-prouvabilité, ce qui fait que le point fixe est un énoncé disant : « Je ne suis pas prouvable ». Remarquer que le théorème montre aussi que $g(\cdot)$ n'est pas calculable.

À cette série, il convient d'ajouter le paradoxe de Richard, qui sous-tend un peu le théorème de Gödel : « le plus petit entier non définissable en moins de 100 symboles », que je viens de définir en nettement moins de 100 symboles. On sort traditionnellement du paradoxe de Richard en disant que le mot « définir » n'est pas bien défini, qu'il faut préciser le langage. On peut voir le théorème de Gödel comme une version « corrigée » de Richard ; Gödel fait par ailleurs expressément référence à Richard.

2.1.3 Le codage

C'est traditionnellement la partie « difficile » du théorème, celle où certains s'ingénient à égarer le néophyte, peut-être parce qu'ils n'en comprennent pas eux-mêmes la structure. De quoi s'agit-il au juste ? Tout simplement de la « numérisation » du langage, idée tout à fait révolutionnaire en 1931, mais bien éventée depuis, les ordinateurs étant passés par là. Et d'ailleurs, il y a un lien de cause à effet, n'oublions pas la contribution de Turing à l'informatique, contribution qui repose en grande partie sur une relecture du théorème de Gödel ; ainsi, le point fixe des programmes n'est-il rien d'autre que le célèbre théorème d'indécidabilité algorithmique du problème de halte des programmes : il n'y a pas de programme capable de déterminer si un programme va s'arrêter ou non et aucun moyen de contourner cet interdit. Il s'agit là d'une simplification du théorème d'incomplétude qui ne perd pas grand-chose de l'original au contraire de la version de Tarski (voir section 2.D.1).

Toute cette poliorcétique fastidieuse peut se résumer à quelques points, faciles à comprendre au XXI^{e} siècle :

Le formalisme : Toutes les opérations d'un système formel, qui sont la création d'un langage (termes et énoncés), les axiomes, les règles, leur enchaînement pour fournir des démonstrations, le résultat de ces démonstrations (les théorèmes), toutes les opérations bureaucratiques (renommage de variables, substitution de termes pour des variables), peuvent s'écrire informatiquement, en supposant les caractéristiques de l'ordinateur illimitées (pas de limitation de mémoire). Ce n'est fondamentalement pas différent du travail d'un logiciel de traitement de texte, qui va vérifier les parenthèses fermantes, remplacer des mots, etc. Remarquer qu'il s'agit d'une activité vraiment formelle, bureaucratique : la machine ne tolère pas la moindre erreur, par exemple une confusion entre "0" et "O".

La numérisation : On sait tous que le langage se code par des nombres, en binaire ou hexadécimal, qu'importe ; on code même des images et des sons. Ce qu'a fait Gödel en 1931, pour la première fois, a été d'associer à toute expression a du langage un « nombre de Gödel » $\ulcorner a \urcorner$, qui n'est rien d'autre qu'un codage ; ainsi pour Gödel $\ulcorner (\urcorner = 11$, $\ulcorner) \urcorner = 13$, alors que les modernes codes ASCII des parenthèses sont respectivement 40, 41. Le codage utilisé par Gödel est sûrement un peu obsolète, il ne se justifie qu'en l'absence de problème de taille (encombrement mémoire), mais ceci mis à part, il fonctionne de la même façon que les codages modernes. De plus, les opérations (ou les propriétés) d'un système formel seront immédiatement traduites par une fonction (ou une propriété des codes associés) représentable dans l'arithmétique formelle (ou tout système dans lequel on peut traduire l'arithmétique).

En particulier (et c'est l'aspect *réflexif* du théorème), un système d'arithmétique peut se représenter lui-même, « parler » de soi, tout comme un langage de programmation peut se représenter comme un programme particulier de ce même langage. Voyons cela de plus près, toujours en évitant les détails.

2.1.4 L'expansivité et la récessivité

Une propriété des entiers est *expansive* quand elle peut s'écrire $\exists x_1 \ldots \exists x_k A[x_1, \ldots, x_k]$, où A est un énoncé d'arithmétique n'utilisant que des quantificateurs bornés $\forall x < p, \exists x < p$; ces énoncés sont aussi appelés Σ^0_1. La classe duale (propriétés *récessives*, énoncés Π^0_1), est formée des énoncés $\forall x_1 \ldots \forall x_k A[x_1, \ldots, x_k]$, où A est à quantificateurs bornés. Une propriété expansive est « approximable » par des essais systématiques : plus on essaye, plus on a de chances de la vérifier, elle réfère donc à un type particulier de potentialité. *A contrario*, la récessivité, c'est le « jusqu'ici ça va », autrement dit, c'est une propriété qui s'amenuise avec les essais. La démonstrabilité est expansive : plus on fait d'essais, plus on a de théorèmes ; par contre, la cohérence est récessive, plus on essaye, plus on a de chances de trouver une contradiction. L'incomplétude dit fondamentalement que récessif et expansif ne coïncident pas, ne coïncident et ne peuvent coïncider, à aucun prix.

Une propriété qui s'exprime à l'aide de quantificateurs bornés est *algorithmiquement décidable*, dans le sens qu'il y a un algorithme pour savoir si $A[n_1, \ldots, n_k]$ est vraie ou fausse pour chaque choix de valeurs des paramètres $x_1, \ldots, x_k$. Ainsi, pour vérifier une quantification bornée $\forall x < p\ A[x]$, il suffit d'effectuer un nombre fini de vérifications pour les valeurs $x = 0, \ldots, p-1$. Si $\mathcal{T}$ est un système « raisonnable » d'arithmétique, alors le même A devient *décidable dans $\mathcal{T}$*, en d'autres termes, pour chaque valeur $n_1, \ldots, n_k$ des paramètres $x_1, \ldots, x_k$, l'énoncé ou sa négation est prouvable. Pourquoi ? Tout simplement parce qu'un système d'arithmétique raisonnable doit contenir de quoi refléter formellement les calculs. En fait, dans un langage basé sur $0, 1, +, \times, =, <$, quelques axiomes suffisent à cette tâche. Leur liste importe peu, on n'aurait que faire d'un système ne contenant pas ce minimum vital. La décision dans $\mathcal{T}$ se fait dans le même sens que la décision algorithmique, tout simplement parce que la décision dans $\mathcal{T}$ recopie formellement les étapes de l'algorithme. Avec un petite nuance : si $\mathcal{T}$ est contradictoire, alors il « en fait trop », puisqu'il prouve à la fois A et sa négation.

Une propriété expansive $\exists n_1 \ldots \exists n_k A[n_1, \ldots, n_k]$ est (algorithmiquement) semi-décidable. Cela veut dire qu'il existe un semi-algorithme (algorithme qui ne donne pas toujours de réponse) qui rend la réponse « oui » quand l'énoncé est vrai, et ne répond rien sinon. Ce semi-algorithme est facile à trouver : en énumérant les valeurs $n_1, \ldots, n_k$, par exemple pour $k = 2$: $(0,0), (0,1), (1,0), (0,2), (1,1), (2,0), (0,3), (1,2), (2,1), (3,0), (0,4), \ldots,$ en essayant successivement toutes les possibilités, on finira bien par tomber sur le bon choix s'il existe et alors l'algorithme pour A permettra de conclure. Si la propriété est fausse, l'algorithme ne rendra en revanche aucune réponse : c'est pour cela qu'il est « semi », partiel. Du point de vue de la prouvabilité, c'est la même chose, s'il y a un choix de valeurs $n_1, \ldots, n_k$ validant A, alors A est prouvable dans $\mathcal{T}$ pour ce choix et la règle logique du quantificateur existentiel permet sûrement de passer de $A[n_1, \ldots, n_k]$ à $\exists x_1 \ldots \exists x_k A[x_1, \ldots, x_k]$. Comme toujours, avec un système formel, il y a le danger d'en faire trop, par exemple si $\mathcal{T}$ est contradictoire ; il y a même une possibilité plus tordue, à savoir que $\mathcal{T}$ peut être cohérente et prouver des énoncés expansifs faux : ainsi $\mathcal{T} + \neg\mathrm{Coh}(\mathcal{T})$, qui est cohérente si $\mathcal{T}$ l'est (deuxième théorème d'incomplétude) et qui démontre un énoncé expansif faux : la contradiction de $\mathcal{T}$[4].

Dualement, les propriétés récessives ne sont pas visiblement algorithmiques, même « semi ». Il n'y a pas non plus de raison qu'un énoncé récessif vrai soit prouvable. En fait, le théorème d'incomplétude fournit un contre-exemple : l'énoncé de Gödel, récessif et vrai, n'est pas prouvable. Pour l'instant, ce qu'on peut dire, c'est qu'un énoncé récessif prouvable dans $\mathcal{T}$ est forcément vrai, si $\mathcal{T}$ est cohérente. En effet, s'il est faux, sa négation, énoncé expansif et vrai, serait prouvable dans $\mathcal{T}$ et $\mathcal{T}$ serait alors contradictoire.

2.1.5 Le premier théorème

La prouvabilité dans $\mathcal{T}$ est l'énoncé expansif par excellence : « il existe une démonstration de ... dans $\mathcal{T}$ ». Cela demande un travail de fourmi, mais on comprend bien que, s'il existe une démonstration, on l'a bien sous la main et qu'on va pouvoir traduire formellement dans $\mathcal{T}$ le fait que cette suite de symboles suit les règles bureaucratiques. *A contrario*, la non-prouvabilité et en particulier la cohérence de $\mathcal{T}$ sont récessives ; rappelons que la cohérence, c'est le fait que l'absurdité (ou encore $0 \neq 0$) n'est pas démontrable dans $\mathcal{T}$.

L'*énoncé de Gödel G* est obtenu par une technique de diagonalisation dont on a vu le principe et dont les détails précis sont indifférents ; il veut littéralement dire : « Je ne suis pas prouvable dans $\mathcal{T}$ ». Il est donc récessif. Si G était prouvable dans $\mathcal{T}$, il serait donc faux et sa négation $\neg G$, expansive et vraie, serait prouvable dans $\mathcal{T}$. $\mathcal{T}$ prouvant G et $\neg G$ serait alors contradictoire.

Théorème 1 (Premier théorème d'incomplétude)
Si $\mathcal{T}$ est « suffisamment expressive » et cohérente, alors il y a un énoncé G, qui est vrai, mais pas prouvable dans $\mathcal{T}$.

Gödel préfère parler d'un énoncé indécidable (ni démontrable, ni réfutable) dans $\mathcal{T}$, ce qui a pour effet d'éviter toute mention de la notion épistémologiquement suspecte de vérité. Pour la petite histoire, observons qu'il se pourrait très bien que le G construit soit réfutable, tout en étant vrai : on a évoqué plus haut la possibilité que $\mathcal{T}$ puisse démontrer des énoncés expansifs faux, tout en étant cohérente. Ce qui fait que la version originale de Gödel (énoncé indécidable) est formulée avec une hypothèse sur $\mathcal{T}$ plus forte que la simple cohérence[5].

[4] Cet exemple standard d'une théorie cohérente, mais non « 1-cohérente », montre qu'on peut être cohérent et mentir, tout comme dans la vie l'on peut être un bandit et échapper à la justice.

[5] La 1-cohérence.

Quelques années après, Rosser a donné une variante de G, ni démontrable, ni réfutable, sous la simple hypothèse de cohérence (voir section 2.D.3). Mais il s'agit d'un sujet mort et on perd un peu son temps dans tous ces détails.

2.1.6 Le second théorème

Ce résultat est d'une simplicité biblique, mais les détails sont infernaux. On vient de démontrer *rigoureusement* que, si $\mathcal{T}$ est cohérente, alors G n'est pas prouvable dans $\mathcal{T}$: si on l'a démontré, c'est par un raisonnement d'arithmétique et on peut choisir pour $\mathcal{T}$ le système qui formalise ce résultat. $\mathcal{T}$ prouve alors formellement que la cohérence de $\mathcal{T}$, notée $\mathrm{Coh}(\mathcal{T})$ implique la non-prouvabilité de G, i.e., G. En d'autres termes, $\mathcal{T}$ prouve l'implication logique $\mathrm{Coh}(\mathcal{T}) \Rightarrow G$; ne prouvant pas G, il ne prouve pas $\mathrm{Coh}(\mathcal{T})$.

Théorème 2 (Second théorème d'incomplétude)
Si $\mathcal{T}$ est « suffisamment expressive » et cohérente, alors $\mathcal{T}$ ne prouve pas sa propre cohérence.

Quelques petites remarques pour conclure ce bref survol technique :

(i) Le flou artistique du « suffisamment expressive » n'est pas le même dans les deux cas. Pour le premier théorème, on a besoin de très peu de pouvoir expressif dans $\mathcal{T}$, essentiellement de pouvoir décalquer un calcul pas à pas. Pour le second théorème, on a besoin de beaucoup plus, i.e., la possibilité de reproduire les raisonnements (très simples) du théorème 1. Ne perdons pas trop de temps : pour l'essentiel $\mathcal{T}$ doit permettre des raisonnements par induction sur des énoncés de structure assez simple.

(ii) C'est le second théorème qui est sur-médiatisé, alors qu'il est sans doute moins profond que le premier. Évidemment, comme il taille rasibus le programme de Hilbert, on ne voit que lui ; mais le premier est tout aussi radical et plus général.

2.1.7 Version rapide

On appelle (Kreisel) *fonction récursive prouvable* dans $\mathcal{T}$ une fonction récursive (i.e. une fonction donnée par un algorithme) de $\mathbb{N}$ dans $\mathbb{N}$ telle que $\mathcal{T}$ prouve la terminaison de l'algorithme. On peut énumérer les fonctions récursives prouvables : tout simplement si n code une démonstration de terminaison, f_n est la fonction dont on démontre la terminaison, sinon c'est ce qu'on veut, disons la fonction nulle. La diagonalisation construit $f_n(n) + 1$, un exemple de fonction récursive non prouvable dans $\mathcal{T}$[6].

[6] Hypothèse : la 1-cohérence de $\mathcal{T}$, qui exclut le gag d'une fonction partielle que $\mathcal{T}$ prétendrait totale.

Cette version un peu simplifiée montre que $\mathcal{T}$ n'est pas capable de reconnaître toutes les fonctions totales. Celles dont la complexité algorithmique[7] est trop grande ne sont pas reconnues par $\mathcal{T}$ comme des fonctions totales. En particulier, si la solution d'un problème algorithmique admet une borne inférieure non prouvable dans $\mathcal{T}$, on obtient une forme concrète d'incomplétude.

2.2 Hilbert face à l'incomplétude

2.2.1 Le programme

Même s'il s'en défend dans son article en évoquant de prétendues méthodes finitistes généralisées, c'est bien à une exécution capitale que s'est livré Gödel.

Le programme de Hilbert se proposait de justifier mathématiquement les mathématiques. Mauvais départ, dira-t-on, puisque cela ressemble beaucoup à de l'auto-amnistie : si les mathématiques sont contradictoires et si elles prouvent leur propre cohérence, cela ne nous avance guère. Objection à laquelle Hilbert avait répondu de façon apparemment convaincante : on ne demande pas l'avis du Parlement sur le Parlement, mais celui du Conseil Constitutionnel. En l'occurrence, l'avis d'un petit morceau de mathématiques à part, en dehors du débat, les *métamathématiques*, supposées s'intéresser aux propriétés des langages formels, avec une restriction drastique des méthodes, essentiellement la combinatoire finie. Un argument métamathématique typique aurait pu être le suivant : montrer par une induction simple que tous les théorèmes ont un nombre pair de symboles (on l'aurait vérifié pour les axiomes, puis montré que les règles de déduction préservent cette propriété) ; on aurait conclu en remarquant que l'absurdité $\perp$ a un symbole et n'est donc pas prouvable.

Parallèlement à cette méthode de justification, Hilbert développe une ontologie dans laquelle la plupart des êtres mathématiques n'ont pas de vraie place, il n'y a que les entiers, conformément à la célèbre phrase de Kronecker : « Dieu a créé les entiers, l'homme a fait le reste ». De plus, la plupart des propriétés des entiers n'ont aucun sens : il ne reconnaît de signification qu'aux identités, aux énoncés quantifiés universellement, qu'on a déjà rencontrés, sous le nom de *récessifs*. Tout le reste n'est que façon de parler. Le programme de Hilbert consiste à établir métamathématiquement que les mathématiques générales (disons la théorie des ensembles **ZF**, mais on commencera avec l'arithmétique de Peano **AP**, voire les *Principia* de Whitehead & Russell) sont *conservatives* sur les énoncés récessifs. Un résultat prouvé par des méthodes « infinies », peut, *modulo* des transformations pas forcément très sympathiques — mais possibles « en principe » — s'établir dans un cadre strictement finitaire. C'est donc un principe de *pureté des méthodes*, tiré de la pratique de la théorie des nombres (démonstrations élémentaires), que Hilbert cherche à mécaniser.

[7] Attention : il s'agit de très grosses complexités, pas du tout P, NP etc.

On remarquera que, si ce projet échoue du fait de l'incomplétude, l'élimination des coupures réalise un idéal assez proche, la *propriété de la sous-formule* (voir 3.3.1).

Par quelques considérations générales, il est facile de montrer que cette conservation est équivalente à la cohérence formelle. Elle implique la cohérence, puisque l'absurdité, qui peut s'écrire $0 \neq 0$, est récessive ; si elle est démontrable, elle l'est aussi de façon « finitaire » et on n'a pas de doute quant au « finitaire ». Réciproquement, la remarque (section 2.1.4) qu'une théorie cohérente ne prouve pas d'énoncé récessif faux, justifie tout théorème récessif à partir d'une démonstration de cohérence de la théorie $\mathcal{T}$ qui le prouve ; une démonstration « finitaire » de cohérence permet donc de transformer toute démonstration en une démonstration finitaire, pourvu que le résultat soit récessif.

D'un certain point de vue, l'immense travail accompli par Gödel n'est pas surprenant. Car ces métamathématiques, ces méta-méthodes finies, font partie des mathématiques. Soyons clair : si l'on n'avait pas pu traduire les métamathématiques dans les mathématiques, ç'aurait été une incomplétude autrement fâcheuse que celle de Gödel. On doit voir le travail de Gödel comme la fin du processus enclenché par Cantor, Dedekind, ... de codage des mathématiques : les ancêtres avaient codé les réels par des ensembles de rationnels, eux-mêmes codables par des entiers, Gödel ferme la boucle en s'attaquant au langage.

Dernière remarque : on admirera, même si cela ne fonctionne pas, la haute tenue du point de vue de Hilbert. Il ramène tout à la cohérence, qui apparaît, *modulo* le codage de Gödel, comme un énoncé récessif, du type donc, très restreint, auquel il daigne accorder un sens. Donc ce programme est faux, réducteur, tout ce qu'on veut, mais pas stupide. On ne peut pas en dire autant des *remakes* affligeants proposés par l'intelligence artificielle.

2.2.2 Le coup de grâce

Par le second théorème, la cohérence de $\mathcal{T}$ n'est pas prouvable dans $\mathcal{T}$. Et donc, même la version la plus brutale, l'« auto-amnistie » ne fonctionne pas. On ne pourra pas démontrer la cohérence par des méthodes finitaires, vu que les méthodes (*a priori* bien plus puissantes) disponibles dans $\mathcal{T}$ ne suffisent pas. Remarquons que le premier théorème d'incomplétude suffit à la réfutation de la version « conservation » : en effet, si $\mathcal{T}$ dénote les mathématiques « finitistes », G fournit l'exemple d'un énoncé récessif prouvable par des méthodes « infinies » mais non prouvable dans $\mathcal{T}$.

En écrivant la dernière ligne j'ai un doute sur la rigueur de mon enchaînement, je passe un peu trop vite de « vrai » à « prouvable par des méthodes infinies ». Plutôt que de rationaliser tout cela, remarquons plutôt que la différence entre vrai et prouvable (dans les formalisations « courantes » des mathématiques) ne faisait pas partie du paysage de l'époque (1930) ; alors, à quoi bon chercher la petite bête quand on discute l'échec du programme de Hilbert. La plupart des nuances qu'on peut faire sont *postérieures* à l'incomplétude et, quelque bonne volonté qu'on ait, on ne remettra jamais debout ce programme.

Une dernière chose : on a remarqué (voir section 1.2.2) que la règle de *Modus Ponens* établit une dualité entre preuves de A et preuves de $\neg A$. Une telle dualité suggère une forme de complétude « interne », de la forme « A est prouvable ssi $\neg A$ n'est pas prouvable », sans référence à la vérité. Par contre, en présence d'une telle complétude, on définirait facilement la vérité comme étant la prouvabilité. Or, que dit le théorème d'incomplétude, le premier, à ce sujet ? G n'est pas démontrable, bien que sa négation $\neg G$ ne le soit pas[8]. Autrement dit, il y a quelque chose de pourri au Royaume du Formalisme.

2.3 L'incomplétude n'est pas un manque

Quand quelque chose ne tourne pas rond, on a tendance à vouloir le réparer ; c'est ce qu'on faisait jadis pour les fous : on les lobotomisait, chimisait, ... alors que, dans la plupart des cas, il n'y avait rien à faire. L'incomplétude est, de même, une maladie qui ne se soigne pas et il est ridicule de chercher la « pièce manquante du puzzle ». On va maintenant parcourir plusieurs propositions, en commençant par les plus indigentes. Il ne s'agit pas tant de tératologie que d'une exploration amusante des facettes de l'incomplétude.

2.3.1 Manque de rigueur ?

Le manque le plus radical est évidemment le manque de rigueur. C'est pourquoi, périodiquement, fleurissent des réfutations, cela vient par période (ainsi, l'an 2000 fut très giboyeux). Il n'y a pas grand-chose à en dire, elles sont toutes sur le même moule : on affirme que, puisque G est vrai, il est forcément prouvable. Ce qui est grave là-dedans ce n'est pas que quelques attardés ressassent les mêmes erreurs, c'est la complaisance coupable dont jouissent ces bricoleurs. Le milieu de l'intelligence artificielle leur fait une publicité insidieuse, courriers électroniques du style « on n'osait pas le dire », « tiens, tiens... », invitations à parler dans des séminaires au nom du « on ne peut rien exclure »... Il est hors de question de réprimer la production de telles âneries, surtout quand leurs auteurs sont des crétins. Par contre, les *révisionnistes* qui leur servent de caisse de résonance sont tout sauf des idiots ; ce sont même des surdoués dans leur genre, la nage entre deux eaux. À ces gens-là il faudrait dire que :

(i) Que cela plaise ou non, il n'y a pas d'erreur dans la démonstration de l'incomplétude, qui a même fait l'objet d'une vérification par ordinateur.

[8]Rigoureusement, si $\mathcal{T}$ est 1-cohérente, ou alors en remplaçant G par la variante de Rosser, qui n'est ni prouvable ni réfutable.

(ii) Une réfutation du théorème (ce qu'on ne peut évidemment pas exclure *a priori*, même si c'est très improbable) produirait donc une contradiction dans les mathématiques.

(iii) Mais l'hypothèse du théorème, c'est que les mathématiques sont cohérentes. Une réfutation donnerait, en fait, un argument encore plus fort.

Autrement dit, *ce qui me tue me renforce*. Étrange résultat, n'est-ce pas, que cette incomplétude ? Un des rares à être absolument irréfragable.

2.3.2 Manque d'imagination ?

Ici, l'on s'attaque au second théorème en remettant en cause la définition de la cohérence... Soyons créatifs ! La cohérence, c'est l'impossibilité de prouver l'absurdité ; en réalité, on utilise le fait que l'absurdité implique n'importe quoi : « *ex falsum quod libet* ». Un système est cohérent quand il y a un énoncé qu'il ne démontre pas. On part donc de cette définition, on rajoute un énoncé, par exemple, une « super-absurdité » et on s'arrange pour qu'il n'y ait *aucun moyen* de la démontrer : par exemple, en interdisant purement et simplement de l'écrire. Évidemment, un système basé sur une telle « logique paraconsistante » est trivialement non contradictoire. On obtient ainsi au plus bas prix une réfutation du deuxième théorème d'incomplétude. C'est un peu comme cela qu'on a fait baisser la criminalité : en ne prenant plus les déclarations de vol de mobiles. Il faut pourtant que les choses aient un sens honnête et la cohérence, c'est la forme la plus primaire d'honnêteté ; si l'on commence à donner des définitions malhonnêtes, incohérentes, de l'honnêteté, de la cohérence, tout est possible, mais cela rime à quoi au juste[9] ?

En résumé : le théorème s'applique à un système *déductif*, pas à un machin dans lequel le raisonnement se fait à la tête du client.

2.3.3 Manque cognitif ?

Les logiques « épistémiques » sont supposées illustrer des principes de raisonnement « abductifs » : du fait qu'on ne sait pas, on en déduit quelque chose. Ces logiques sont basées sur des métaphores un peu infantiles (et en fait limitées à l'axiomatisation de leurs métaphores), comme celle des cocus de Bagdad, qui tuent leur femme parce que... Plutôt que de reproduire cette ânerie, donnons-en une variante inédite : *les cocus de Houston*. Ils sont deux, V. et W., ils savent tout l'un sur l'autre et aussi qu'un des deux l'est ; en fait il n'y en a qu'un, W.[10], qui sait qu'il y a un cocu, qui sait que ce n'est pas V., mais qui n'en tire aucune conséquence, car il est un peu... lent. En revanche,

[9] Les « paraconsistants » sont des formalistes primaires, mais intéressés, qui dissimulent leurs grossiers stratagèmes derrière des définitions plus indéchiffrables que le plan de Shinjuku pour un *gaijin* : on arrive rarement au bas de la première page.

[10] Dans la version « officielle » les deux le sont.

V., qui est très futé et a un PhD ès logique épistémique, se dit : « Tiens, tiens, si je n'étais pas cocu, W. aurait conclu qu'il l'est et tué sa femme ». Donc V. tue sa douce moitié ; moralité, la logique épistémique est dangereuse pour la santé.

Qu'illustre donc au juste cette blague retournée comme un gant ? Elle illustre la différence entre *constatation* et *déduction*. W. doit faire une déduction simple, or c'est un crétin et il ne la fait pas. Du point de vue algorithmique, c'est la différence entre le calcul à l'ancienne (lire un résultat dans une table de fonctions) et le calcul moderne, exécuter un programme, ce qui peut prendre un temps fou, en tout cas imprévisible.

Les logiques épistémiques sont basées sur une identification entre « ne pas savoir » et « savoir que pas ». Si une telle identification était déductivement viable, il suffirait d'ajouter les axiomes idoines. On voit bien que c'est impossible, vu qu'alors, G étant non prouvable, ce fait serait prouvable (ce qui s'exprime par G), contradiction. On objectera que, dans certaines provinces, on continue à soutenir des thèses en logique épistémique et que les résultats sans être vraiment renversants, ne sont pas faux. Or, ces systèmes devraient être contradictoires au nom du théorème de Gödel... Petite énigme, dont voici la solution : l'incomplétude ne s'applique qu'à des systèmes avec un *minimum* de pouvoir expressif (vraiment pas grand-chose). Ce minimum de pouvoir expressif interdit de prédire à l'avance qu'un énoncé n'est pas démontrable. Les logiques épistémiques n'ont pas droit à ce minimum d'expressivité et c'est pour cela qu'elles se réduisent à la logique de leur propre (et pesante) métaphore.

2.3.4 Manque d'axiomes ?

Le pas que ne franchit pas la logique épistémique, la logique non monotone l'a franchi sans complexes. On ajoute un principe du style « si $\neg A$ n'est pas prouvable, alors A l'est ». Si l'on prend les précautions idoines, ces système sont non contradictoires et complets, alors de quoi se plaint-on ? Ils sont tout simplement non déductifs, car il n'y a aucun moyen d'activer le principe supplémentaire. Il ne s'agit plus de système formel, puisqu'il est impossible de savoir qu'un énoncé n'est pas prouvable (c'est déjà impossible dans un système déductif, alors dans un machin comme ça, bonjour !). À ce propos, réfutons directement l'analogue algorithmique de la « non-monotonie ». On pourrait *compléter* tout algorithme de la façon suivante : si l'algorithme donne une réponse, disons « oui », on répond « oui », s'il dit « non », on répond « non », s'il ne répond pas, on répond ce qu'on veut, « oui » ou « non », voire « je ne sais pas ». C'est impossible, car l'indécidabilité du problème d'arrêt (Turing) dit qu'il n'y a pas d'algorithme pour savoir qu'on ne sait pas.

Exit les logiques non monotones. Il peut être amusant de regarder comment ces non-systèmes se comportent par rapport au deuxième théorème. Encore une fois, ce n'est qu'un exercice stérile. Fondamentalement, il ne s'agit que d'un ensemble cohérent et complet d'énoncés arithmétiques, i.e., essentiellement d'un modèle. Tout dépend de la complexité logique d'un tel ensemble :

(i) Il existe des ensembles cohérents et complets qui sont exprimables par un énoncé de l'arithmétique[11]. Dans ce cas-là, la cohérence s'exprime en arithmétique. Malgré la non-calculabilité de la déduction, on peut reproduire l'argument de point fixe et obtenir un « énoncé de Gödel » G. G est vrai ssi G n'est pas prouvable, i.e., ssi $\neg G$ est prouvable : on voit que la théorie fait des erreurs et qu'on ne peut pas lui faire confiance.

(ii) La plupart des solutions ne sont pas définissables ; ainsi, si l'on prend les énoncés vrais, il est immédiat[12] que cet ensemble ne s'exprime pas par un énoncé. Dans ce cas, on ne peut même pas exprimer la cohérence.

2.3.5 Manque de vérité ?

Dans l'échelle ouverte de la médiocrité, ce qui suit ne commet aucune erreur grossière de logique ; mais il s'agit d'une interprétation académique qui ne rend pas justice à la radicalité de l'incomplétude. L'incomplétude, ce serait du « vrai pas prouvable ». On va dire que je charrie, vu que j'ai énoncé le théorème sous cette forme !

Voilà le point crucial, il faut faire le départ entre un énoncé mathématique et son interprétation. Ainsi, il n'y a aucun problème à formuler la notion de vérité arithmétique (même si elle ne s'exprime pas en arithmétique, elle s'exprime en théorie des ensembles) et on peut l'utiliser dans des démonstrations mathématiques, ou même pour formuler des résultats style incomplétude. Cela dit, ce n'est pas parce qu'on peut définir une notion en théorie des ensembles qu'elle a un sens. C'est particulièrement flagrant pour la vérité, puisque Tarski l'a définie par un pléonasme, ainsi :

$$\forall x A[x] \text{ est vrai si } A[n] \text{ est vrai pour tout entier } n$$

la vérité de A, c'est A, c'est ce que j'ai qualifié d'essentialisme. On doit légitimement douter d'une notion qui se révèle à un tel point opaque. Au lieu de l'interprétation académique du « déficit de vérité », je proposerai la version plus stimulante « la vérité n'a pas de sens » (je n'ai pas dit « n'est pas définissable en arithmétique », mais bien « pas de sens »). Ce qui ne signifie pas que j'ai eu tort de dire que G est vraie (on l'a établi !) ; simplement que, de même qu'il n'y a pas de définition *générale* du beau, du bien, etc., il n'y a pas de notion *générale* de vérité. Ce qu'on manipule en logique sous le vocable « vrai » est une coquille vide. Un dernier mot : il ne faudrait pas non plus oublier que l'énoncé de Gödel, ce machin tarabiscoté, avant de vouloir dire « je ne suis pas prouvable », dit « je ne veux rien dire ». On s'approche avec l'incomplétude

[11] Plus complexes que récessif ou expansif, au minimum Δ^0_2.

[12] Théorème de Tarski, variante un peu faiblarde de l'incomplétude (voir section 2.D.1).

des limbes du signifiant. *Dixit* René Thom : « la limite du vrai ce n'est pas le faux, c'est l'insignifiant ».

Revenons à la lecture académique, « tarskienne », de l'incomplétude et encore une fois, rappelons qu'il n'y a là aucune erreur *technique*, ce qui la place quand même nettement au-dessus des « lectures » précédentes. Au lieu de compléter brutalement, à la « non-monotone », un système en un système complet, mais non déductif, on peut envisager des complétions *partielles*. Par exemple, ajouter la cohérence, puis recommencer. C'est ainsi qu'on arrive au mythe des métas emboîtés, des progressions transfinies de théories, approche stérile s'il en fut. Cela correspond au fantasme essentialiste de l'escroquerie pyramidale, et nous ramène à Kubrick et aux soucoupes volantes aztèques.

Du point de vue théologique, la version académique dominante est assez étonnante. En effet, c'est du vrai non prouvable qu'on peut « touiller » avec du méta. On pourrait voir la vérité (sémantique) comme le Père, la prouvabilité (syntaxe) comme le Fils (ou encore Verbe) et le méta comme l'entremetteur, le Saint-Esprit. Mais alors l'incomplétude apparaît comme la non-consubstantialité du Fils, une sorte de Nestorianisme logique. Ça sent le fagot et nous sommes à Rome, arrêtons-nous avant qu'il ne soit trop tard !

2.4 Lectures métaphoriques

2.4.1 Le théorème de Blair

L'actualité récente invite à considérer un « théorème de Blair », pas de Baire, lequel Blair a dit récemment (2004) : « les armes de destruction massive (ADM) existent, mais on ne pourra jamais les trouver ». L'analogie avec l'incomplétude est trompeuse. En effet, sur le terrain, l'armée d'occupation a employé *tous* les moyens pour « faire parler ». Trouver des caches est de nature expansive, plus on torture, plus on en connaît et si l'on n'a rien trouvé c'est qu'il n'y avait rien. On est donc plus près du paradoxe du menteur : « je mens » que de l'incomplétude.

On s'approche davantage de la bonne métaphore en tentant un « théorème de Saddam ». En effet, ce dernier clamait l'absence d'ADM, mais mettait toute une série d'obstacles aux inspections. Ce qui fait que l'absence d'ADM, énoncé récessif vrai, était invérifiable.

2.4.2 L'anti-mécanisme

Si l'on veut élargir le débat, on voit bien que l'incomplétude s'oppose à la vision mécaniste du monde, celle de Hilbert, ou celle de l'intelligence artificielle. C'est pourquoi ce résultat est tellement détesté (sous les fleurs, que d'épines !).

Franchissons carrément une étape et admettons avec Régis Debray que l'incomplétude est un résultat anti-totalitaire. À ceci près que l'argumentaire de Debray (système politique = système formel) est franchement douteux. Malgré tout, cet argument frappe juste. N'oublions pas que les totalitarismes du XX^{e} siècle se sont tous basés sur des illusions scientistes, théories raciales, matérialisme historique et que leurs crimes monstrueux se réclamaient d'une science infaillible[13]. Très loin des remugles totalitaires, le débat interne à la logique touchait à la possibilité d'atteindre des « solutions finales[14] » aux débats scientifiques les plus abstraits. Ces débats n'avaient aucune incidence sur la possible extermination de tel ou tel groupe. Et pourtant, si la réponse au programme de Hilbert avait été « oui », signifiant que toute question formelle est solvable — en principe, mais pas forcément en pratique (on aurait pu, par exemple, imaginer une mécanisabilité de type « chaotique », impraticable car trop coûteuse) —, quel argument pour les Hitler, les Staline ! Grâce à l'incomplétude, il n'est plus possible de prétendre que toute question a sa réponse.

Il convient cependant d'observer que le théorème de Gödel n'est pas la panacée anti-scientiste. Par exemple, face aux élucubrations de H. Simon (son programme « Keplerophagique »), un peu de bon sens épistémologique suffit : rappelons-nous que la science recherche avant tout des questions !

2.4.3 Digression : l'intelligence artificielle

(Réponse à la question : « Clarifiez votre position sur l'intelligence artificielle. »)

C'est avant tout une question de nuances. Si l'on entend par là qu'on peut mécaniser l'activité de certaines zones du cerveau, par exemple automatiser la reconnaissance de l'espace et fabriquer des robots qui retombent sur leurs pattes, tout comme des chats, ou qui corrigent les déviations de trajectoire d'un véhicule au conducteur somnolent… bien sûr. Mais doit-on parler d'intelligence, ou plutôt d'*instinct* ? L'« instinct artificiel », cela marche sans problème.

Si l'on veut vraiment de l'intelligence, de la créativité, c'est plus complexe et franchement voué à l'échec. Le théorème de Gödel s'y oppose, mais, plus en réfutant la métaphore totalitaire d'une science mécanisable et finale, que dans les détails. Pour les détails, utilisons plutôt le bon sens. Ce qui fait l'intelligence au sens créatif, c'est qu'on ne l'attend pas, qu'elle se place en position d'inconfort, c'est donc un certain « grain de folie », déjà mentionné *supra*, dont la condition *sine qua non* est la possibilité d'erreurs, le *droit à l'erreur*. L'in-

[13] On remarquera au passage que cette identification entre *constatation* et *réflexion* qui sous-tend la logique épistémique et autres *paralogiques* est une définition plausible du totalitarisme : l'idéal de ces « logiques explicites » qui refusent tout raisonnement hypothétique, c'est le PV de gendarmerie.

[14] C'est le joli terme employé par Hilbert et consorts. Ce terme est choquant, mais ce n'était qu'un poncif scientiste, recyclé par le III^{e} Reich.

telligence suit des chemins déviants, inattendus, y compris ceux du préjugé, de l'ambition, de la colère, les sept péchés capitaux et même pire. Cela demeure tolérable dans la société, car le pouvoir individuel restant limité, cette « cuisine interne » n'a pas d'effet trop dramatique. Quand un individu très intelligent atteint au pouvoir suprême, c'est immanquablement pour faire « bénéficier » le monde des errements de sa pensée. Imaginons qu'on puisse confier à un robot tout-puissant la possibilité de décréter sur un coup de tête l'éradication de l'humanité, pour d'ailleurs trouver une meilleure idée le lendemain ! Tant qu'à faire, on préfère être gouverné par un W. Bush, dont le manque d'envergure limite la nocivité.

Donc, en résumé : il faut avant tout éviter le simplisme, les exagérations. Au lieu de chercher un logiciel qui traduirait automatiquement Proust dans le style de Proust[15], cherchons plutôt, comme Gérard Huet, à faire un dictionnaire de sanscrit, qui donne le plus économiquement, le plus efficacement possible, la morphologie des expressions, etc. (voir `http ://sanskrit.inria.fr`). Par rapport au fantasme d'un robot-Einstein, cela peut sembler limité, mais la *véritable* ambition se pare des atours les plus modestes.

2.4.4 L'extinction du popperisme

Le philosophe Popper a proposé une épistémologie des propriétés « falsifiables », qu'on pourrait traduire par « réfutables », ou encore par « récessives ». Il s'agit de lois générales, sujettes à des vérifications partielles, des tests, par exemple expériences particulières, vérification jusqu'à une certaine décimale... On reconnaît sans peine l'influence de Hilbert, puisque cela ressemble à une mise en perspective de l'ontologie hilbertienne par rapport au champ complet du discours scientifique. Le côté sympathique de l'affaire, c'est le positionnement anti-essentialiste : les lois ne tombent pas du ciel. Par contre, puisque Gödel réfute Hilbert, il réfute *a fortiori* Popper qui le démarque.

On remarquera que Popper ne donne aucun statut au théorème de Gödel ; on a vu, en effet, qu'il est irréfragable et donc en aucune façon « falsifiable ». C'est bien commode, puisque, cet énoncé n'ayant pas de statut, on peut l'ignorer comme « métaphysique ». Ce qui cloche, dans cette ontologie du test, c'est tout simplement son asymétrie : d'un côté le test, absolu, en face la loi, relative et testable. Ce qui amène immanquablement à une ontologie du récessif. Mais on pourrait imaginer des situations plus intriquées, où la qualité du test puisse être mise en cause : si la loi teste le test pendant que le test teste la loi : « Quand tu regardes l'abîme longuement, l'abîme regarde aussi en toi », on entre dans un monde beaucoup plus ouvert. C'est ce qui se passera par exemple en *ludique*, où il n'y a pas d'arbitres assermentés et où tout peut être remis en cause.

[15] On en est tellement loin : il n'y a même pas de logiciel de rédaction automatique de thèses en logique épistémique.

2.4.5 Divers

En guise de détente, quelques interprétations particulièrement salées :

(i) L'impossibilité de se penser soi-même : le système $\mathcal{T}$ ne pourrait pas parler de lui-même. Réduire la pensée à une activité formelle, ce n'est pas très gentil pour les poètes, philosophes et même les mathématiciens. Il y a de plus un contresens absolu, puisque l'incomplétude est basée sur le fait que $\mathcal{T}$ peut coder des faits le concernant. Et la cohérence, tel le talon d'Achille, est à peu près la seule chose concernant $\mathcal{T}$ qui « échappe » à $\mathcal{T}$. Entre nous, le fait qu'on puisse tout faire avec ses lunettes sur le nez, sauf les revisser, mérite-t-il des tartines sur... les méta-lunettes ?

(ii) L'auto-référence : la pièce dans la pièce, la peinture exposée dans la méta-galerie (les toilettes), sans oublier la pièce montée de service, le cultissime Gödel-Escher-Bach ! Que fait donc l'entarteur ?

2.A Davantage sur la classification des prédicats

2.A.1 Au premier ordre

On compte le nombre d'alternances de quantificateurs d'un énoncé d'arithmétique, supposé en forme prénexe (les quantificateurs non bornés d'abord) ; ainsi, $\forall\forall\exists\exists\exists$ correspond-il à une alternance 2. Pour $n \geq 1$, les énoncés Σ^0_n (Π^0_n) sont ceux avec une alternance n et commençant par un groupe existentiel (universel). Comme on peut toujours rajouter des quantificateurs « bidons », un énoncé d'alternance n pourra, au choix, être vu comme Σ^0_m ou Π^0_m, dès que $m > n$. En particulier, un énoncé à quantificateurs bornés sera à la fois Σ^0_1 et Π^0_1, i.e., expansif et récessif.

Cette classification s'étend aux *ensembles* définis par de tels énoncés. De ce point de vue apparaît une troisième hiérarchie, celle des Δ^0_n, i.e., des ensembles qui admettent les deux formes, Σ^0_n et Π^0_n. Ainsi, les ensembles Δ^0_1 sont-ils les ensembles récursifs (décidables par un algorithme), alors que les Σ^0_1 sont les ensembles semi-récursifs (reconnaissables par un semi-algorithme). On dit encore *récursivement énumérable* (r.e.) dans ce cas, car un ensemble semi-récursif non vide est l'image d'une fonction récursive (calculable). La hiérarchie des Δ^0_n n'est pas une hiérarchie formelle, vu qu'elle suppose deux écritures équivalentes : cette équivalence peut être immédiate (cas d'une implication entre deux Σ^0_1, qui admet deux formes prénexes, $\forall\exists$ et $\exists\forall$, ce qui la rend immédiatement Δ^0_2) ou franchement indémontrable (cas d'un ensemble *a priori* semi-récursif, dont le caractère récursif, Δ^0_1, dépend d'une démonstration de décidabilité).

Le résultat fondamental, obtenu par codage, est que chaque classe Σ^0_n (Π^0_n) possède un élément *universel* (ou complet). Par exemple, pour $n = 1$, la prouvabilité formelle dans un système cohérent d'arithmétique est complète, dans le sens que tout ensemble Σ^0_1 peut se factoriser à travers la prouvabilité. Il suffit de prendre la définition de l'ensemble : en effet, $A[n]$ est vrai ssi $A[\bar{n}]$ est prouvable, du moins quand le système est 1-cohérent ; s'il est seulement cohérent, il faut faire une modification à la Rosser.

Par diagonalisation, chaque classe est distincte de sa duale. Ainsi, si $f(m,n)$ $(m \in F_n)$ est une énumération Σ^0_1 des ensembles Σ^0_1, alors l'ensemble Π^0_1 $\{n; n \notin F_n\}$ n'est pas Σ^0_1. C'est le théorème de Turing, ou encore, si l'on utilise comme élément universel la prouvabilité, le premier théorème d'incomplétude.

2.A.2 Au second ordre

La logique peut se formuler au second ordre en admettant des quantifications sur les prédicats. Il convient de ne pas confondre :

La hiérarchie projective : On utilise des quantifications au premier ordre sur les entiers, au second ordre sur les ensembles d'entiers. Le résultat est classifié en Σ^1_n et Π^1_n, $(n \geq 1)$, où n compte le nombre d'alternances de quantificateurs du second ordre. La classe $\Delta^1_1 := \Sigma^1_1 \cap \Pi^1_1$ (ensembles *hyperarithmétiques*) est déjà strictement plus grande que tous les Σ^0_n et Π^0_n. Un ensemble hyperarithmétique non arithmétique typique est l'ensemble des (nombres de Gödel d') énoncés arithmétiques vrais.

La hiérarchie logique : La même chose qu'avant, mais les quantificateurs du premier ordre ne sont pas numériques. La notation est (faute de mieux) Σ^n, Π^n, toujours en ne comptant que les alternances au second ordre. Cette hiérarchie, plus fine que la précédente, est essentiellement décalée d'une case. Elle est bien adaptée à la théorie de la démonstration, à l'opposé de la précédente, plutôt ensembliste.

La relation entre les deux hiérarchies s'établit au moyen de la définition de Dedekind des entiers naturels : « Le plus petit ensemble contenant 0 et clos par rapport au successeur S ». Ce qui s'exprime par une quantification sur un prédicat unaire X :

$$x \in \mathbb{N} : \quad \Leftrightarrow \quad \forall X((X(0) \wedge \forall y(X(y) \Rightarrow X(Sy))) \Rightarrow X(x))$$

Un énoncé Σ^0_1, disons $\exists n A[n]$, qui utilise une quantification numérique $\exists n$ peut se traduire au second ordre en $\exists x(x \in \mathbb{N} \wedge A[x])$, dont on voit qu'il est Π^1. De même, les énoncés récessifs sont Σ^1. Plus généralement, un énoncé Π^1_n peut s'écrire sous forme Π^{n+1}.

2.A.3 Complétude vs. incomplétude

Ces notions se discutent à deux niveaux :

Techniquement : La complétude du calcul des prédicats (Gödel, 1930), s'énonce ainsi : « un énoncé B clos et vrai dans tous les modèles est prouvable ».

Disons que $B = B[P]$ a un seul symbole de prédicat P ; alors l'énoncé du second ordre $A := \forall X B[X]$ est vraiment clos (plus aucune constante), il est Π^1. De plus $\forall X B[X]$ est vrai ssi B est vrai dans tout modèle, i.e., pour tout choix d'une interprétation pour P. La complétude se récrit donc en « un énoncé A clos, Π^1 et vrai est prouvable ». Cette relation « vrai $\Rightarrow$ prouvable » ne subsiste pas au-delà de la classe Π^1. Ainsi, elle est fausse pour les énoncés Σ^1, puisque l'énoncé Π^0_1 $(= \Sigma^1)$ de Gödel est vrai sans être prouvable, résultat justement nommé incomplétude.

Généralement : Il importe de ne pas réduire les notions à leur expression technique. La complétude est une plénitude, une sorte d'équilibre interne. On la formule par rapport à la vérité, parce que c'est commode, mais je préfère, pour ma part, la forme de complétude venant de l'élimination des coupures (voir chapitre 3). Pour la même classe d'énoncés (Π^1), la propriété de la sous-formule (voir section 3.3.1), dit que « tout est là ». Tant qu'on respecte l'élimination des coupures, on ne peut rien prouver de plus. Au mieux, peut-on donner des versions « plus rapides » des « mêmes[16] » démonstrations. Pour la classe duale $(= \Sigma^1)$, l'absence de propriété de sous-formule fait qu'on ne contrôle plus rien.
La *ludique* généralise cette approche en définissant une notion de complétude interne (voir section 13.8.4).

2.A.4 Expansivité et récessivité

Restreinte[17] au raisonnement, l'expansivité correspond bien au fonctionnement déductif, alors que la récessivité correspond au mode inductif, généralisation à partir de cas particuliers, éventuellement infirmable par la prise en compte d'un nouvel exemple, le « jusqu'ici ça va ». La terminologie (toute personnelle) employée s'attache plus à la dynamique en ignorant le contenu (raisonnement, nombre, image).

Pour ce qui est de la relation à la *polarité*, on distinguera entre :
positif/réponse/explicite/expansif et : négatif/question/implicite/récessif.
Pour éviter le systémisme, il importe de remarquer que la polarité vient d'une idée d'« oignon logique », dont on enlève les peaux successives. Ainsi, l'opposition réponse/question semble-t-elle bien recouvrir l'opposition expansif/récessif. Par contre, positif/négatif ne concerne que la peau la plus externe et, de même, explicite/implicite semble ne faire vraiment sens qu'avec un seul niveau.

[16]Ce ne sont les mêmes que *modulo* la lecture catégorique des démonstrations.
[17]En anticipant sur la polarité (voir chapitre 12).

2.B L'arithmétique formelle

2.B.1 Le système RR

Le premier théorème d'incomplétude est habituellement formulé pour des systèmes basés sur un langage simpliste, dû à Peano. Il y a quatre symboles fonctionnels, la constante 0 (ou $\bar{0}$, comme on veut) le successeur S (unaire) et $+, \times$ (binaires) ; ainsi l'entier 5 sera-t-il codé par le terme $\bar{5} := SSSSS0$ et, en général, n par $\bar{n}$, n symboles S suivi d'un zéro. À noter, la régression qui accompagne souvent le progrès, ici, le formalisme qui nous ramène à la numération pré-babylonienne. Il y a deux symboles binaires de prédicats, $=, <$. Les axiomes du système **RR**[18] se répartissent en trois groupes :

Égalité : Ces axiomes expriment que l'égalité est une congruence : $x = x$; $x = y \Rightarrow y = x; x = y \wedge y = z \Rightarrow x = z; x = y \wedge z = t \wedge x < z \Rightarrow y < t$; $x = y \Rightarrow Sx = Sy; x = y \wedge z = t \Rightarrow x + z = y + t$; $x = y \wedge z = t \Rightarrow x \times z = y \times t$. Ces axiomes sont suffisants pour démontrer $x = y \wedge A[x] \Rightarrow A[y]$.

Définitions : Ces axiomes permettent de démontrer les équations et inéquations de base. Le groupe $x + 0 = x; x + Sy = S(x + y); x \times 0 = 0$; $x \times Sy = (x \times y) + x$ permet de démontrer, pour tout terme clos t dont la valeur est n, l'égalité $t = \bar{n}$ et donc toutes les égalités $t = u$ entre termes clos qui sont vraies. Le groupe $Sx \neq 0; Sx = Sy \Rightarrow x = y$ (troisième et quatrième axiomes de Peano) permet de démontrer toutes les inéquations $t \neq u$ entre termes clos qui sont vraies. Noter que ces deux axiomes supposent un domaine infini (sans eux on pourrait avoir $\overline{10} = \bar{0}$). Enfin le groupe $\neg(x < 0); x < Sy \Leftrightarrow x < y \vee x = y$ permet de démontrer toutes les inégalités ou non inégalités $t < u, \neg(t < u)$ entre termes clos qui sont vraies. Il permet de démontrer l'équivalence $x < n \Leftrightarrow x = \overline{0} \vee x = \overline{1} \vee \ldots \vee x = \overline{n-1}$ et ainsi de passer l'étape cruciale des quantificateurs bornés.

Un dernier axiome : L'axiome $x < y \vee x = y \vee y < x$ est de nature légèrement différente des autres, car il n'est pas nécessaire à l'incomplétude, puisque la représentation des propriétés expansives est assurée par les axiomes de définition. Il sert à la représentation des fonctions récursives et donc à l'*indécidabilité* algorithmique de **RR** et de toutes ses extensions cohérentes. On l'utilise aussi dans la variante de Rosser.

Ces axiomes doivent être manipulés au moyen de règles de logique. Je me permets de remettre ce point au chapitre suivant. En effet, il y a autant d'axiomatisations de la logique que l'on veut, toutes sans aucun intérêt pratique : on n'en a pas besoin pour raisonner. Donc, tant qu'à apprendre des règles,

[18]La principale vertu de cette axiomatique due à R. Robinson est d'être sans quantificateurs.

apprenons-les sous une forme qui présente un intérêt théorique, par exemple le *calcul des séquents* (voir chapitre 3). L'incomplétude utilise ordinairement la logique classique, mais il n'y a pas d'inconvénient à utiliser la logique intuitionniste, voire la logique linéaire.

Donc, tout système cohérent contenant ces quelques axiomes va être incomplet et, partant, indécidable. Par indécidable, on entend que le système ne décide pas toutes les propriétés, mais surtout qu'il n'y a pas d'algorithme pour savoir si un énoncé est prouvable ou non. Pour être décidable, il faut donc frapper très fort, par exemple :

(i) Remplacer le troisième axiome de Peano par, disons $\overline{10} = \bar{0}$. On est alors dans le fini (les entiers *modulo* 10) et le système devient décidable ; si on ne le remplace pas, on reste dans l'indécidable[19].

(ii) Se priver d'un symbole, typiquement le produit (arithmétique de Pressburger, décidable). C'est raisonnable, vu la part prépondérante prise par la multiplication dans la représentation des suites finies.

Ces systèmes sont décidables parce qu'ils ne sont bons à rien ou presque.

2.B.2 L'arithmétique de Peano

Originellement, Peano avait formulé son arithmétique au moyen des axiomes précédents (ou une variante) et du schéma d'*induction* :

$$A[0] \wedge \forall y(A[y] \Rightarrow A[Sy]) \Rightarrow \forall x A[x]$$

formulé pour une *propriété* arbitraire des entiers, avec tout le vague qui s'attache à ce mot. Le paradoxe de Richard (1905) montre que ce n'est pas tenable : il faut que A fasse partie d'un langage formel bien défini[20]. Ce qu'on appelle de nos jours *arithmétique de Peano*, c'est le système formel **AP** obtenu en rajoutant aux axiomes de **RR** (on peut se passer du dernier), le schéma d'induction restreint aux propriétés arithmétiques, i.e., aux énoncés du langage. Pendant qu'on y est, il y a aussi une arithmétique de Heyting **AH**, la même chose, mais sur un substrat intuitionniste.

Le second théorème d'incomplétude est pour l'essentiel une formalisation du premier théorème. La démonstration informelle du premier théorème se fait par des inductions et donc le second théorème s'établit dans l'arithmétique de Peano (ou sa variante intuitionniste). En fait, c'est beaucoup trop, car les inductions se font sur des énoncés de prouvabilité, donc expansifs, Σ^0_1. Un bon substrat pour le second théorème est donc l'induction restreinte aux énoncés Σ^0_1. C'est loin d'être optimal, mais cela donne une bonne idée. En réalité, le

[19] A est démontrable dans le système « de base » ssi $T \Rightarrow A$ est démontrable dans le système sans le troisième axiome, avec $T := \forall x(Sx \neq 0)$.

[20] C'est la technique du « rideau mobile », qui sépare assez arbitrairement la classe touriste — les choses bien définies — de la classe affaires qui contient le reste.

« bon système » est obtenu en rajoutant des primitives fonctionnelles (e.g., la fonction exponentielle), de façon à pouvoir dans certains cas remplacer des quantifications non bornées par des quantifications bornées. On restreint alors l'induction aux énoncés sans quantificateurs non bornés.

Finalement, revenons à la phrase dévastatrice d'André Weil au sujet de Gentzen. L'induction peut aussi se formuler avec deux variables emboîtées, trois, etc. et ces inductions généralisées apparaissent comme des inductions transfinies, par exemple, à deux variables jusqu'à ω^2, à trois jusqu'à ω^3 etc. Il y a donc *a priori* deux paramètres pour une induction, d'une part la longueur de l'induction, d'autre part la complexité (alternance de quantificateurs) de l'énoncé « induit ». Dans le cas de la seconde démonstration de Gentzen, l'induction de départ est de longueur courte (ω), mais de complexité arbitraire (n'importe quel Σ^0_n), alors que sa démonstration de cohérence utilise une induction longue (ϵ_0), mais de basse complexité (sans quantificateurs). C'était pour nuancer le désastre, qui n'est pas si total que cela !

2.B.3 Systèmes plus généraux

L'arithmétique de Peano fournit un exemple de système non finiment axiomatisé et d'ailleurs non finiment axiomatisable. Il faut évidemment mettre des limites à l'utilisation de listes infinies d'axiomes, sinon on risque de tomber dans le travers des logiques non-monotones, i.e., ne plus avoir de notion de démonstration. Une définition naturelle est celle d'une famille décidable d'axiomes (on peut savoir algorithmiquement si un axiome est accepté ou non). Une fausse généralisation : une axiomatisation expansive, i.e., récursivement énumérable, ne produit rien de plus. En effet, il suffit de remplacer l'axiome A_n fourni à l'étape n par une variante équivalente de taille au moins n (e.g. la conjonction ou la disjonction de n copies de A_n) pour remplacer une axiomatique r.e. par une axiomatique récursive.

Si l'on abandonne les restrictions récursives sur la démonstrabilité, on perd l'expansivité et le lien avec le calcul[21]. Par contre, toute la machinerie est intacte, il faut seulement considérer la complexité logique de la « déduction », ne pas se priver de n-cohérences... et l'on atteint ainsi facilement une généralisation incolore, insipide et sans saveur.

[21] Détail piquant pour les logiques non-monotones, censées résoudre des problèmes informatiques : l'exigence de calculabilité est sacrifiée sur l'autel de l'efficacité. C'est l'aveuglement idéologique : les nationalistes français collaboraient avec les Allemands, les communistes ont instauré le système le plus inégalitaire de l'histoire, etc.

2.C Techniques de l'incomplétude

2.C.1 Le point fixe : Russell

L'antinomie de Russell, qui n'est que la version du théorème de Cantor appliquée à « l'ensemble de tous les ensembles », produit immédiatement une antinomie : si $a := \{x; x \notin x\}$, alors $y \in a \Leftrightarrow y \notin y$, donc $a \in a \Leftrightarrow a \notin a$. Dans les systèmes « normaux », à déduction non artificiellement bridée (comme les logiques paraconsistantes), ceci mène à contradiction. Il est important de remarquer que le point fixe logique (l'incohérence) est le même que le point fixe du λ-calcul, celui-là évitant la contradiction par une boucle infinie des calculs.

C'est plus qu'une remarque pédagogique : dans son livre « Natural Deduction » [47], Prawitz introduisait une déduction naturelle pour le schéma de compréhension « naïf » (voir section 4.C.2), donc pour un système logique incohérent. Cette déduction naturelle admet une notion de normalisation, mais *qui ne converge pas.* En fait, le processus de normalisation de l'antinomie de Russell n'est autre que le point fixe du λ-calcul. Ce qui établit une relation importante :

$$\text{incohérence} \sim \text{non terminaison}$$

Les logiques « normales » ont une vision trop brutale (essentialiste) de l'infini, en d'autres termes, elles définissent des fonctions qui « croissent trop vite », exponentiellement ou pire ; ces fonctions sont pour l'essentiel le temps de normalisation (voir chapitre 4). La question est d'isoler le principe qui produit l'« explosion » combinatoire dans la normalisation. Or, du point de vue d'un système formel, il n'y a pas de différence entre un algorithme dont on ne peut pas prouver la terminaison et une boucle du genre de celle induite par l'antinomie de Russell. Moralité, pour étudier la grosse complexité, étudions la divergence, l'antinomie de Russell ; c'est ce qui mènera aux logiques *allégées* du chapitre 16.

L'antinomie de Russell se décompose en deux parties, $a \in a \vee a \in a$ et $a \notin a \vee a \notin a$. Elle ne devient vraiment contradictoire que *modulo* la règle de *contraction*, i.e., la possibilité de remplacer une disjonction $A \vee A$ par A. Ceci indique la direction à suivre, la règle de contraction. Cela dit, son élimination pure et simple serait trop radicale et il faut plutôt la brider.

Enfin, une remarque sur le méta. Mis à toutes les sauces, ce machin ne veut plus rien dire : *grosso modo*, rien ne distingue un système d'un méta-système. Pourtant, il y a peut-être des choses à dire, pourvu qu'on s'accorde un peu de finesse géométrique. Ainsi, dans la diagonalisation, la fonction de deux arguments $f(x, y)$ est en fait de la forme $f_y(x)$, l'indice y étant en quelque sorte en position « méta » par rapport à x. Avec une vision absolue, à la Kronecker, des entiers, faire $x = y$ est très naturel, mais quand on voit la boîte de Pandore ouverte par cet acte anodin, ne faut-il pas y réfléchir à deux fois ? Du point de vue de la logique linéaire, cette opération ne respecte pas la profondeur des boîtes (exponentielles, pas de Pandore !). Donc hypothèse, chantier, etc. : boîtes et « méta ».

2.C.2 Le point fixe : Gödel

Le codage de la syntaxe fait apparaître de nombreuses fonctions récursives et propriétés expansives, par exemple $\mathrm{Sub}(m, n)$ tel que $\mathrm{Sub}(\ulcorner A \urcorner, \ulcorner \bar{q} \urcorner) = \ulcorner A[\bar{q}] \urcorner$ et qui représente donc, au niveau des codes le résultat de la substitution du terme codant l'entier q dans l'énoncé A et pour être rigoureux, il faudrait encore dire pour quelle variable s'est effectuée la substitution (disons y_0) (sans oublier une tartine sur les variables libres ou muettes). De même, on introduira $\mathrm{Dem}(m, n)$, tel que $\mathrm{Dem}(\ulcorner A \urcorner, \ulcorner D \urcorner)$ quand D est une démonstration de A, etc. Finalement $\mathrm{Coh}(\mathcal{T})$, la cohérence de $\mathcal{T}$ (propriété récessive) s'exprime comme $\forall x \neg \mathrm{Dem}(\ulcorner 0 \neq 0 \urcorner, x)$.

Pour obtenir l'énoncé de Gödel, on forme donc $\forall x \neg \mathrm{Dem}(\mathrm{Sub}(y_0, y_0), x)$, qui est donc un énoncé $A[y_0]$, et on pose $G := A[\ulcorner A \urcorner]$. G exprime alors sa non-prouvabilité.

2.C.3 Le codage des suites

La représentation des énoncés récessifs repose pour l'essentiel sur le codage des suites d'entiers par des entiers. Le codage des suites de longueur donnée se fait au moyen de polynômes, ainsi, en longueur 2, le couple (m, n) peut-il être codé par $(m + n)(m + n + 1)/2 + n$, ce qui correspond à l'énumération $(0, 0), (0, 1), (1, 0), (0, 2), (1, 1), (2, 0), (0, 3), (1, 2), (2, 1), (3, 0), (0, 4), \ldots$; au passage, remarquons que ce codage permet de réduire à 1 le nombre de quantificateurs non bornés dans un énoncé expansif (récessif).

Mais il faut pouvoir coder des suites de longueur arbitraire, typiquement pour représenter les démonstrations. Pour cela, Gödel utilise une technique qui n'est pas vraiment originale, puisqu'on sait maintenant qu'il l'a reprise d'un cours sur le corps de classes qu'il venait de suivre... Quoi qu'il en soit, il fallait y penser. Épistémologiquement, on est quand même obligé de remarquer que, si ce codage n'avait pas été possible, on aurait alors introduit une primitive spéciale pour le codage des suites. Avec une petite nuance : la possibilité du codage dans un langage avec aussi peu de primitives rend le théorème beaucoup plus spectaculaire, un dilemme entre incomplétude et inexpressivité... ce qui serait moins flagrant si le codage des suites était plus coûteux, puisqu'on aurait alors une zone intermédiaire de systèmes décidables et moyennement expressifs.

Le point de départ est le théorème du Chinois qui s'énonce $\mathbb{Z}/pq\mathbb{Z} \sim (\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/q\mathbb{Z})$ quand p, q sont premiers entre eux. Concrètement, étant donnés $a < p, b < q$, on peut trouver un (unique) $c < pq$ tel que les restes respectifs de la division de c par p et q soient a et b. Ce résultat s'étend évidemment à un nombre quelconque d'entiers deux à deux premiers. Pour représenter la suite $(a_1, \ldots, a_n)$, on choisit un entier N strictement plus grand que les a_i et la longueur n de la suite. On vérifie facilement que les nombres $N! + 1, N!.2 + 1, \ldots, N!.n + 1$ sont deux à deux premiers. Il existe alors un entier s dont les restes respectifs de la division par $N!+1, N!.2+1, \ldots, N!.n+1$ sont $a_1, a_2, \ldots, a_n$. On peut représenter la suite au moyen des trois nombres N, n, s. La propriété d'être premiers entre eux, le reste de la division, tout cela s'exprime en termes de $0, S, +, \times, <, =$.

La possibilité de coder des suites de longueur arbitraire explique pourquoi on ne peut pas brider à l'avance la dynamique et forcer, comme dans les logiques épistémiques, les démonstrations à être de longueur, disons 25 (cas des 25 cocus de Bagdad). La décidabilité de l'arithmétique de Pressburger est évidemment liée au rôle crucial du produit dans le codage des suites de longueur inconnue.

2.D Incomplétude et vérité

2.D.1 Le théorème de Tarski

Le théorème de Tarski dit que la vérité arithmétique n'est pas définissable en arithmétique. De façon plus précise, il n'y a pas d'énoncé arithmétique $V[n]$ tel que l'équivalence $V[\ulcorner A \urcorner] \Leftrightarrow A$ soit vraie pour tout énoncé clos (sans variables libres) A. En d'autres termes, la vérité est hyperarithmétique sans être arithmétique. Voir aussi la section 3.B.3.

Cela résulte d'une diagonalisation triviale : dans le théorème d'incomplétude, remplacer "prouvable" par "vrai". Il y a fort à penser que Gödel (comme tout le monde, incapable de concevoir l'incomplétude *ante literam*) a essayé de *définir* la vérité par la prouvabilité et a cru obtenir ainsi une contradiction, pas si éloignée que cela de l'antinomie de Russell. Une seconde lecture lui a fait mesurer que sa définition de vérité pouvait être incorrecte.

Pourquoi ce résultat est-il attribué à Tarski ? Il illustre surtout sa propension à *faire les poubelles* (voir aussi ce théorème de point fixe particulièrement tarte auquel il a tenu à attacher son nom).

Il ne faut pas croire que la complétion d'une théorie cohérente soit forcément hyperarithmétique, ou pire. Avec un peu de savoir-faire, on fabrique des modèles Δ^0_2 de l'arithmétique. Ils sont bien évidemment « non-standards », en cela qu'ils vérifient des tas de choses fausses.

2.D.2 La 1-cohérence

Dans quelle mesure la cohérence signifie-t-elle l'« honnêteté » d'un système ? Cela signifie sûrement quelque chose, sinon il n'y aurait pas le puissant barrage du théorème d'incomplétude. Ce sentiment d'honnêteté est renforcé par le côté fabriqué des logiques paraconsistantes : comme je l'ai dit, il faut définir l'honnêteté honnêtement. Mais cela ne va que jusqu'à un certain point : prenons l'idée (sûrement honnête) que cohérence vaut existence. Elle n'est en fait valable que dans le monde algébrique, un monde d'équations : le monde récessif cher à Hilbert. Ce qui pose problème, c'est la cohérence de systèmes fautifs, par exemple $\mathcal{T} + \neg\mathrm{Coh}(\mathcal{T})$, quand $\mathcal{T}$ est une théorie cohérente. On voit ici la limite de l'exercice : la cohérence ne concerne que l'observance de la loi, pas de son esprit. Donc elle n'est qu'une forme *primaire* d'honnêteté.

On introduit la notion de Σ^0_n- (ou Π^0_n-) *fidélité*, pour dire que tout théorème clos de la classe en question est vrai. Il est immédiat que :

(i) La Π^0_1-fidélité n'est autre que la cohérence.

(ii) La Σ^0_n-fidélité est équivalente à la Π^0_{n+1}-fidélité, *a priori* plus générale.

On définit donc la n-cohérence comme la Π^0_{n+1}-fidélité, en particulier la 0-cohérence n'est autre que la cohérence tout court. Il est facile de voir que la $n+1$-cohérence est strictement plus forte que la n-cohérence, on l'a vu pour $n = 0$ et cela se généralise sans peine.

2.D.3 La variante de Rosser

La variante de Rosser du premier théorème d'incomplétude résout un petit problème laissé ouvert par Gödel, i.e., l'hypothèse de 1-cohérence[22] nécessaire pour démontrer que $\neg G$ n'est pas prouvable. Le système de départ $\mathcal{T}$ est artificiellement bridé par la contrainte : on n'a le droit de démontrer A que si $\neg A$ n'a pas été démontré (dans $\mathcal{T}$) avec une démonstration de code (nombre de Gödel) plus petit. Autrement dit $\mathrm{Dem}'(m, n) := \mathrm{Dem}(m, n) \wedge \forall p < n \neg\mathrm{Dem}(\mathrm{Neg}(m), p)$, avec $\mathrm{Neg}(\ulcorner A \urcorner) = \ulcorner \neg A \urcorner$. De notre point de vue, $\mathcal{T}$ étant supposée cohérente, il n'y a aucune différence entre $\mathcal{T}$ et sa version bridée ; de plus, la version bridée reste expansive, puisqu'elle diffère de la version « normale » par une quantification bornée (il faut vérifier que les entiers plus petits qu'un certain n ne sont pas des codes de démonstration de $\neg A$).

Formellement, $\mathcal{T}$ démontre, pour chaque énoncé A non prouvable dans $\mathcal{T}$ et pour chaque entier n l'énoncé $\mathrm{Dem}(\overline{\ulcorner \neg A \urcorner}, \bar{n}) \Rightarrow \neg\mathrm{Dem}'(\overline{\ulcorner A \urcorner}, x)$. En effet, le « dernier axiome » donne $x < \bar{n} \vee x = \bar{n} \vee \bar{n} < x$, soit encore $x = \bar{0} \vee x = \bar{1} \vee \ldots \vee x = \bar{n} \vee \bar{n} < x$ et l'on procède ensuite par cas.

Soit R l'énoncé de Rosser, i.e., l'énoncé de Gödel du système bridé. Aucune différence notable en cas de prouvabilité : R n'est pas prouvable, i.e., est vraie. Si $\neg R$ est prouvable avec une démonstration de code n, ce fait est prouvable dans $\mathcal{T}$; par ce qui précède, on conclut (dans $\mathcal{T}$) $\forall x \neg\mathrm{Dem}'(\ulcorner R \urcorner, x)$, la non-prouvabilité (bridée) de R, soit R, qui n'est pas prouvable. $\neg R$ n'est donc pas prouvable.

[22] À l'époque appelée ω-cohérence, terminologie obsolète.

La variante de Rosser est, on l'a peut-être reconnu, un système paraconsistant. Cette manière de choisir, entre A et $\neg A$, celui qui a la plus petite démonstration n'a évidemment aucun sens déductif, puisque, par rapport à la moindre conséquence logique, cette plus petite démonstration risque de changer de bord. Il y a donc tromperie sur la marchandise si l'on prend cette astuce technique au pied de la lettre. Ce qui fait la valeur de ce « trucage », c'est que c'est le système formel $\mathcal{T}$ que l'on trompe et non pas son prochain comme le font les « paraconsistants ».

L'astuce de Rosser se décline en techniques variées pour représenter les fonctions etc. dans le cas de théories infidèles. On peut se passer d'y aller voir, le paysage est sans surprise ; comme tout ce qui touche à l'incomplétude, l'exploration systématique débouche sur l'ennui, la non-substance. Encore une fois, on vient de parcourir un bras mort de la logique et il est temps d'aller dans des zones plus animées.

2.E L'indécidabilité

2.E.1 L'indécidable

L'incomplétude est une indécidabilité déductive (on ne sait pas décider toute question A en la prouvant ou en la réfutant). Mais elle implique aussi une indécidabilité *algorithmique*, dans le sens où il n'y a pas d'algorithme pour décider si oui ou non un énoncé donné est démontrable.

La manière standard de montrer l'indécidabilité est de représenter les fonctions calculables (récursives) dans **RR** : un argument de diagonalisation permet (comme toujours) de conclure. Je me contente de donner la définition de la notion de représentation : une fonction f est représentable par un énoncé $A[x, y]$ quand on peut démontrer, pour chaque entier m :

$$A[\bar{m}, y] \quad \Leftrightarrow \quad y = \overline{f(m)}$$

Une forme plus faible de représentabilité serait de demander que $A[\bar{m}, \bar{n}]$ soit démontrable quand $f(m) = n$, $\neg A[\bar{m}, \bar{n}]$ soit démontrable quand $f(m) \neq n$. Cette forme faible se convertit en forme forte (celle qui est vraiment utile) en remplaçant A par $B[x, y] := A[x, y] \wedge \forall z < y \neg A[x, z]$. On utilise le « dernier axiome » pour effectuer le passage de la forme faible à la forme forte.

Observons finalement que, puisque **RR** est finiment axiomatisable, le calcul des prédicats est lui-même indécidable.

2.E.2 L'inséparabilité

Dans le système **RR**, il n'y a même pas d'algorithme de *séparation*, i.e., qui répond « oui » quand A est démontrable, « non » quand $\neg A$ est démontrable, n'importe quoi autrement, *mais quelque chose* ; autrement dit, les théorèmes et les anti-théorèmes ne peuvent pas être *récursivement séparés.* Ceci reste trivialement vrai de toute extension cohérente $\mathcal{T}$ de **RR**, puisqu'un algorithme de séparation pour $\mathcal{T}$ en serait aussi un pour **RR**. En d'autres termes, toute théorie « suffisamment forte » est récursivement inséparable.

Ce n'est pas le cas du calcul des prédicats, indécidable sans être inséparable. Par exemple, on peut séparer les théorèmes au moyen d'un modèle fini. Ma méthode de séparation préférée est la suivante : oublier le premier ordre ; étant donné un énoncé A du calcul des prédicats, définir A^- en :

- Remplaçant les symboles de prédicats $P(\ldots), Q(\ldots)$ par des symboles propositionnels P, Q.
- Ignorant les quantifications.

On obtient ainsi une interprétation du calcul des prédicats dans le calcul propositionnel et il est immédiat que cette interprétation préserve les démonstrations, ce que l'on peut vérifier par exemple, avec le calcul des séquents.

2.E.3 Fonctions ambiguës

On prend un système auquel s'applique l'incomplétude, disons **ZF** ; on se donne deux fonctions récursives f, g et un énoncé récessif $B = \forall n A[n]$. On définit h comme étant f tant qu'on n'a pas réfuté B et g « après » :

$$\begin{aligned} h(n) &:= f(n) \text{ si } \forall m \leq n A[m] \\ h(n) &=: g(n) \text{ si } \exists m \leq n \neg A[m] \end{aligned} \tag{2.1}$$

De façon évidente, h est soit égale à f (si B est vraie), soit égale à g (si B est fausse), sauf pour un nombre fini de valeurs. Pour aller vite, disons que h est *à peu près égale à* f ou g. Si, maintenant, il se trouve que B est indécidable dans **ZF**, on ne pourra pas décider laquelle de f, g est à peu près égale à h.

Ainsi, si f, g sont des fonctions constantes égales à 0 et 1, h est une suite de Cauchy dont on ne peut pas décider dans **ZF** si elle tend vers 0 ou 1.

Cette technique éculée a été utilisée encore et encore pour produire des contre-exemples superficiels à telle ou telle question de décidabilité. Ainsi :

Il existe des réels récursifs a, b, c tels que la solvabilité dans $\mathbb{R}$ de l'équation $ax^2 + bx + c = 0$ soit indécidable dans **ZF**.

La démonstration ne demande aucune connaissance approfondie des équations algébriques ; on prend tout simplement $a = 1, b = 0$, c étant le réel défini par la suite de Cauchy ambiguë mentionnée plus haut. L'équation est soit $x^2 = 0$ (une solution) soit $x^2 + 1 = 0$ (pas de solution).

Ces exemples modestes ayant été épuisés, la technique a été appliquée à la production d'énoncés plus prétentieux, typiquement :

Il existe une fonction récursive h telle que la question $P \stackrel{?}{=} NP$ *relativisée à h soit indécidable dans la théorie des ensembles* **ZF**.

Ce qui est un corollaire trivial d'un résultat célèbre de Solovay : on peut trouver des fonctions récursives f et g telles que le problème $\mathrm{P} \stackrel{?}{=} \mathrm{NP}$ relativisé à f et g soit vrai ou faux suivant le cas. En prenant un h ambigu, on obtient tout de suite le résultat. Pour conclure, point n'est besoin de connaître la démonstration du théorème de Solovay, ni même quoi que ce soit en théorie de la complexité ; tout ce qu'on a besoin de savoir, c'est que la relativisation est la même pour des fonctions presque égales. En particulier, prétendre que ce « théorème » montre la difficulté de la question $\mathrm{P} \stackrel{?}{=} \mathrm{NP}$ est une ânerie : il est obtenu par une imitation mécanique de l'argument concernant l'équation du second degré et personne ne croit que la solution de cette équation soit un problème « difficile ».

Cette sorte de sinistre pantalonnade est une illustration parfaite de l'état actuel du domaine « incomplétude » : un cimetière.

Bibliographie : [24], [22]. xs