ENS Rennes & Institut Mathématique de Jussieu - Paris Rive Gauche

Stage de fin de licence de mathématiques

Introduction à la logique mathématique

Roméo Després

encadré par

Boban Velickovic

Juin 2016

Table des matières

1 Introduction

1.1 Histoire

« Tous les hommes sont mortels, Socrate est un homme, donc Socrate est mortel. » Difficile d'imaginer phrase d'amorce plus veillie pour discuter de logique. De fait, il y a plus de vingt siècles que la philosophie se fait de la Logique un de ses objets d'étude favoris. Entre la fin du XIX$^{\text{ème}}$ et le début du XX$^{\text{ème}}$ siècles, les mathématiciens, qui s'étaient pourtant assez éloigné des philosophes, se sentirent soudain concernés par le sujet et décidèrent de s'en emparer. On voulait établir les fameux « fondements des mathématiques ». Mais le paradoxe de Russell, détecté par celui-ci en 1902 [8] dans les premières formalisations de Frege et Cantor fait soudain apparaître la question de la *consistance* des mathématiques : il devient urgent de pouvoir prouver que, si l'on choisit correctement ses axiomes, on ne risque pas de démontrer un jour une contradiction. Hilbert fait alors de la preuve de la consistance de l'arithmétique le second de ses 23 problèmes pour le XX$^{\text{ème}}$ siècle [4], impulsant l'essor de la logique mathématique.

1.2 Philosophie

N.B. Afin de n'offusquer personne, le terme de « philosophie » pourra être entendu ici dans un sens aussi vague que l'on souhaitera.

Pendant toute la durée de mon stage, les questions qui m'ont donné la motivation nécessaire pour m'approprier les notions que j'étudiais furent bien souvent plus philosophiques que mathématiques. À n'en pas douter, la logique d'aujourd'hui est une branche des mathématiques. Cependant, réfléchir au but que l'on se donne et à la fonction que l'on prête à la logique mathématique a fait partie intégrante de mon travail, ainsi que des longues conversations que j'ai pu avoir avec mon maître de stage M. Boban Velickovic. Ces questionnements philosophiques, au fond, se ramènent à celui de savoir ce que peut nous aprendre la logique. Ci-dessous, je retrace le parcours intellectuel qui a été le mien sur cette question pendant le stage. Il ne s'agit donc, au fond, que d'opinions personnelles d'une valeur discutable. Ainsi, par exemple, mon travail n'ayant concerné que les liens entre la logique et les mathématiques, je ne parlerai pas de ses applications à l'informatique. Pourtant, elles sont paraît-il nombreuses et essentielles pour ce qui est de motiver la recherche en logique ; mais je n'y connais absolument rien.

1.2.1 Fondements

Que peut donc nous apprendre la logique mathématique ? *A priori*, on pourrait imaginer comme Frege dans son livre *Begriffsschrift* (1879) qu'elle permet de fonder les mathématiques en en donnant une formalisation. C'est en tout cas ainsi que je l'envisageais en débutant mon stage, influencé principalement par la lecture de Bourbaki [1]. Mais intervient alors naturellement la question de savoir comment on fonde la logique ; autrement dit, on a seulement déplacé le problème. Le philosophe Anton Dumitriu [2] résume ainsi cette objection :

> Nous nous trouvons en face du dilemme suivant : ou bien tout système formel, mathématiquement constitué, a besoin d'une justification logique de son schéma formel et alors la logique comme système formel reste sans nulle justification de ce genre ; ou bien il faut accepter le système formel de la logique comme se satisfaisant en soi, et alors on ne voit pas pourquoi les autres systèmes n'ont pas le même droit. Quoi qu'il en soit, si l'on accepte la construction de la logique de la manière dont elle a été faite par les mathématiciens modernes, elle perd la raison même en vue de laquelle elle a été constituée.

Or, si l'on considère la formalisation contemporaine de la logique du premier ordre, il est évident que les mathématiciens n'y cherchent plus à établir les fondements de leur discipline. En effet, la plupart de ses définitions suppose le monde mathématique déjà construit, et admet la vérité de ses théorèmes. À titre d'exemple, S. M. Srivastava [9] propose une démonstration du théorème de compacité (3.2.1) en logique propositionnelle s'appuyant sur le lemme de Zorn, et donc sur une théorie des ensembles bien établie. On peut ainsi supposer que les logiciens ont d'autant plus facilement accepté l'objection ci-dessus qu'ils se sont aperçu qu'ils trouvaient de bien meilleurs théorèmes logiques lorsqu'ils faisaient des mathématiques leur outil plutôt que leur produit.

Mais alors, si cette nouvelle logique qui est *incluse* dans les mathématiques ne peut en assurer les fondements, à quoi sert-elle ? Il s'agit d'une nouvelle théorie mathématique, qui étudie des objets appelés *théorèmes* et qui émet des méta-théorèmes à leur propos. Dans quelle mesure cette théorie qui utilise les mathématiques peut-elle nous donner des propriétés *sur* les mathématiques ?

1.2.2 Modélisation

Bien sûr, on ne saurait exiger des théories mathématiques qu'elles justifient systématiquement leur utilité en proposant des applications ; la logique peut très bien être une belle théorie, point final. Cependant, on ne peut nier qu'elle tient une place bien particulière au sein de l'édifice mathématique. Un résultat comme le théorème de complétude (4.2.1) qui, grossièrement, stipule qu'une proposition est vraie si et seulement si elle est démontrable, ne peut laisser indifférent un mathématicien. En effet celui-ci sent bien que ce qu'elle dit dépasse le cadre des théorèmes formels vus comme des objets mathématiques, et peut s'appliquer à n'importe quel théorème mathématique. Autrement dit, la logique fonctionne comme une *modélisation* des mathématiques.

On peut alors s'interroger sur la valeur de la connaissance qu'elle nous apporte. Il ne s'agit plus d'une connaissance déductive établie comme celles des mathématiques, mais d'une connaissance inductive, plus proche de celle des sciences expérimentales. Cette modélisation est-elle vraiment juste ? En particulier, la notion « extérieure » de vérité (celle des mathématiques préexistantes, dont elle admet comme *vrais* les théorèmes) à laquelle elle fait appel n'est-elle pas suspecte d'une forme de réalisme mathématique ? Et faut-il s'attendre à ce que cette modélisation soit réfutée un jour, comme l'auraient peut-être prévu Karl Popper [7] ou Thomas Kuhn [5] ? Quoi qu'il en soit, ces questions méritent d'être posées.

1.2.3 Applications

Libre à chacun de déterminer s'il préfère interpréter la logique comme une modélisation des mathématiques, ou simplement comme une théorie parmi d'autres, avec ses propres objets indépendants du reste des mathématiques, et dont la justification n'est pas à faire. Toutefois, sa structure lui permet de redéfinir de nouvelles mathématiques, à l'intérieur de la théorie logique, et donc de parler d'ensembles, d'arithmétique, de géométrie, etc., sans l'aide d'éléments extérieurs. Cette particularité autorise des théorèmes qui semblent tenir lieu d'exemples d'« application » de la logique aux mathématiques. Le théorème de Paris-Harrington (5.2.2) [6], qui était l'objet d'étude initial de mon stage, donne ainsi un exemple d'un théorème d'arithmétique vrai dans $\mathbb{N}$ mais non démontrable dans l'arithmétique de Peano. L'idée sous-jacente est donc de convaincre les mathématiciens qu'il est inutile de chercher à construire une preuve de ce théorème qui n'utiliserait que des notions arithmétiques, cela ayant été prouvé impossible. Il n'est pas certain, au vu des sections précédentes, que ce raisonnement soit exact. Cependant, ce théorème incarne bien une forme d'application de la logique à un domaine des mathématiques classiques, l'arithmétique.

1.3 Conclusion et remerciements

Alors que nous apprend la logique ? Au fond, elle semble se comporter comme n'importe quelle théorie mathématique. L'esprit mathématique, tout emprunt de rigueur, sait qu'elle ne nous apprend rien, et n'est que spéculation irréaliste ; mais le cœur du mathématicien, convaincu de la vaste portée de ses théorèmes, se surprend à s'en émerveiller. Toutefois, sa particularité d'être capable de produire un modèle des mathématiques rend son interprétation particulièrement complexe, et c'est pourquoi ces questionnements introductifs qui peuvent sembler abstraits dans un premier temps sont nécessairement présents à l'esprit quand on étudie la logique mathématique.

Avant de débuter le corps de ce rapport, je tiens à remercier mon maître de stage M. Boban Velickovic, pour sa générosité, sa gentillesse, et ses bienveillantes explications qui ont su me guider dans la compréhension du sujet tout en m'en laissant explorer moi-même le chemin.

Je précise, enfin, que la majorité des connaissances présentées dans ce rapport provient du livre de S. M. Srivastava *A Course on Mathematical Logic* [9] ainsi que des notes de cours de Jaap von Oosten *Introduction to Peano Arithmetic* [10].

2 Syntaxe et sémantique de la logique du premier ordre

Comme précisé dans l'introduction, on suppose connus et utilisables la formalisation et les premiers résultats d'une théorie des ensembles classique. Ainsi, contrairement à ce qu'on pourrait croire, ce qui suit ne consiste aucunement en une tentative de fondement des mathématiques.

2.1 Syntaxe

2.1.1 Langage

Difficile de parler de théories, propositions, théorèmes, si on ne sait pas les exprimer. Un langage, ça sert à s'exprimer. Le premier des objets logiques sera donc le langage.

Définition (Langage du premier ordre).
*Un **langage du premier ordre**, ou **langage**, est un ensemble L contenant :*
- *l'ensemble des symboles logiques $\{\exists, \neg, \vee, =\} \cup \{x_0, x_1, x_2, ...\}$, commun à tous les langages, où $x_0, x_1, ...$ sont appelés « variables » ;*
- *un ensemble de symboles de constantes $\{c_i, i \in I\}$;*
- *pour tout $n \in \mathbb{N}$, un ensemble $\{f_j, j \in J_n\}$ de symboles de fonctions de n variables et un ensemble $\{p_k, k \in K_n\}$ de symboles de relations n-aires.*

Par exemple, le langage de la théorie des groupes comprend un symbole de constante e et un symbole de fonction binaire $\cdot$.

2.1.2 Termes

Dorénavant, on se donne un langage L. Les termes de L correspondent aux « objets » de la future « théorie ».

Définition (Termes d'un langage).
*L'ensemble T des **termes** de L est définit par induction comme le plus petit ensemble contenant toutes les variables et toutes les constantes de L et tel que si $t_1, ..., t_n \in T$ et f est une fonction à n variables de L, alors $f(t_1, ..., t_n) \in T$. Une constante ou une variable sera dite de rang 0 ; et si le maximum des rangs de $t_1, ..., t_n$ est k, alors $f(t_1, ..., t_n)$ sera dit de rang $k+1$.*

Si s est un terme contenant des variables parmi $x_1, ..., x_n$, on notera $s[x_1, ...x_n,]$; alors, si $t_1, ..., t_n$ sont des termes, on notera $s[t_1, ..., t_n]$ l'expression obtenue en remplaçant respectivement les occurences de $x_1, ..., x_n$ dans s par $t_1, ..., t_n$.

Proposition 2.1.1.
Si $s[x_1, ..., x_n]$ et $t_1, ..., t_n$ sont des termes de L, alors $s[t_1, ..., t_n]$ est un terme de L.

Preuve.
On le prouve par récurrence sur le rang de s. Si s est une constante, alors $s = s[t_1, ..., t_n]$ donc s est un terme ; de même si s est une variable ne figurant pas parmi $x_1, ..., x_n$. Si s est l'un des x_i, alors $s[t_1, ..., t_n] = t_i$ donc s est un terme.

Supposons maintenant que le rang k de s soit supérieur ou égal à 1. Alors $s[x_1, ..., x_n]$ est de la forme $f(s_1[x_1, ..., x_n], ..., s_m[x_1, ...x_n])$. Donc $s[t_1, ..., t_n] = f(s_1[t_1, ..., t_n], ..., s_m[t_1, ..., t_n])$. Or par hypothèse de récurrence, les $s_i[t_1, ..., t_n]$ sont des termes ; donc $s[t_1, ..., t_n]$ est un terme.

2.1.3 Formules

Définition (Formules d'un langage).
On définit tout d'abord une ***formule atomique*** *(ou* ***atome****) :*

— *si t et s sont des termes, alors $t = s$ est un atome ;*
— *si p est une relation n-aire, et si $t_1, ...t_n$ sont des termes, alors $p(t_1, ..., t_n)$ est un atome.*

On définit ensuite par récurrence une ***formule*** *(et son rang) :*

— *tout atome est une formule de rang 0 ;*
— *si A est une formule de rang k, et v une variable, alors $\exists v A$ et $\neg A$ sont des formules de rang $k+1$. Si B est une formule de rang inférieur ou égal à k, $A \vee B$ et $B \vee A$ sont des formules de rang $k+1$.*

On peut alors définir nos premières abréviations. Si A est une formule et v une variable, on notera par exemple $\forall v A$ pour abréger $\neg \exists v \neg A$. Si B est une formule, on notera $A \to B$ pour abréger $B \vee \neg A$, et $A \wedge B$ pour $\neg(\neg A \vee \neg B)$. On aurait pu prendre ces symboles comme données de base plutôt que de les définir ainsi, et ce choix n'est pas sans conséquence ; toutefois cela aurait demandé de nombreuses distinctions de cas dans les preuves, pour un résultat identique puisque nous nous intéresserons finalement à une logique classique, admettant implicitement les principes de non-contradiction et du tiers-exclu.

Une occurence d'une variable v dans une formule A est dite *liée* si elle est contenue dans une sous-formule de A de la forme $\exists v B$; dans le cas contraire, elle est dite *libre*.

Soient t un terme, v une variable, et A une formule. On dit que t est **substituable** à v dans A si et seulement si pour toute variable w apparaissant dans t, aucune sous-formule de A de la forme $\exists w B$ ne contient d'occurence libre de v dans A. Dans ce cas, on notera $A_v[t]$ l'expression obtenue en remplaçant chaque occurence libre de v dans A par t. Si $t_1, ...t_n$ sont substituables respectivement à $v_1, ..., v_n$ dans A, on notera $A[t_1, ..., t_n]$ l'expression obtenue en appliquant le procédé à chaque variable.

Proposition 2.1.2.
$A[t_1, ..., t_n]$ est alors une formule de L.

Preuve.
On le prouve, comme pour la proposition 2.1.1, par récurrence sur le rang de A.

2.1.4 Théories

Les notions précédentes suffisent pour définir une théorie, censée modéliser la notion de théorie mathématique.

Définition (Théorie).
Une ***théorie du premier ordre****, ou simplement* ***théorie****, est composée d'un langage L, et d'un ensemble de formules de L appelées axiomes non-logiques.*

Par exemple, la théorie des groupes est la théorie dont les symboles non-logiques sont un symbole de constante e et un symbole de fonction binaire $\cdot$, et dont les axiomes non-logiques sont (x,y et z sont ici des variables) :

— $\forall x(x \cdot e = x) \wedge (e \cdot x = x)$
— $\forall x \exists y(x \cdot y = e) \wedge (y \cdot x = e)$
— $\forall x \forall y \forall z x \cdot (y \cdot z) = (x \cdot y) \cdot z$

Un second exemple important est celui de l'arithmétique de Peano (cf. [10]). Cette théorie a pour symboles non-logiques un symbole de constante 0, un symbole de fonction unaire S, et deux symboles de fonctions binaires $+$ et $\cdot$. Ses axiomes sont :

— $\forall x \neg(S(x) = 0)$
— $\forall x \forall y(S(x) = S(y) \rightarrow x = y)$
— $\forall x(x + 0 = x)$
— $\forall x \forall y(x + S(y) = S(x + y))$
— $\forall x(x \cdot 0 = 0)$
— $\forall x \forall y(x \cdot S(y) = (x \cdot y) + x)$
— pour toute formule $A[v]$, $(A_v[0] \wedge (\forall v(A \rightarrow A_v[S(v)]))) \rightarrow A$

On note PA cette théorie (qui comprend donc un nombre dénombrable d'axiomes).

On peut, de même, définir la « théorie des ensembles » en tant que théorie du premier ordre. Ainsi, l'édifice mathématique complet peut être représenté *à l'intérieur* de la logique et devenir l'un de ses objets d'étude.

2.2 Sémantique

Dans cette section, on définit la notion de « vérité » dans une théorie. Il s'agit d'offrir aux arides objets formels (symboles de constantes, de fonctions, de relations, et formules) un sens dans le « monde réel ». Le monde réel, c'est les vraies mathématiques ; autrement dit, les ensembles.

On se donne toujours un langage du premier ordre L.

2.2.1 Structures

Définition (Structures d'un langage).
*Une **structure**, ou **interprétation** de L est la donnée d'un ensemble M (appelé **univers** de la structure) et d'une application qui associe à tout symbole de constante c un élément $c_M \in M$; à tout symbole de fonction de n variables f une fonction $f_M : M^n \rightarrow M$; et à tout symbole de relation n-aire p, une relation $p_M \subset M^n$. Notons que l'interprétation du symbole « $=$ » sera toujours la relation d'égalité dans M.*

Ainsi, tout groupe est une structure naturelle pour le langage de la théorie des groupes. Toutefois il n'importe pas encore de savoir si les axiomes sont correctement vérifiés ; autrement dit, un monoïde fait également une très bonne structure de ce même langage.

Notons que si t est un terme sans variable, on peut aisément définir son interprétation t_M dans M par récurrence sur son rang (il suffit d'itérer un certain nombre de fonctions sur une constante).

2.2.2 Vérité dans une structure

Le procédé d'interprétation nous permet presque de donner une valeur de vérité à une formule dans une structure M. Il reste tout de même à donner un sens à $\exists xA$. Comment « injecter » un élément quelconque de M dans la formule A pour tester sa véracité ? À chaque fois qu'on parlera de vérité dans M, on admettra donc implicitement qu'on aura fait la construction suivante.

On note L_M le langage du premier ordre obtenu en ajoutant pour tout $a \in M$ un symbole constante i_a (distinct de ceux pré-existants) dans L. i_a sera appelé le nom de a. Puis on étend naturellement la structure sur M à L_M de sorte à ce que l'interprétation de chaque i_a soit a. On peut alors définir convenablement la vérité d'une formule.

On dira qu'une formule est close si elle ne comporte aucune occurence libre de variable. Si A est une formule, on définit sa fermeture comme la formule $\forall x_n ... \forall x_1 A$, où les variables avec des occurences libres dans A sont parmi $x_1, ..., x_n$, classées dans l'ordre alphabétique, avec x_n ayant effectivement une occurence libre dans A.

On définit dans un premier temps la vérité (ou validité) d'une formule close, par récurrence sur son rang. Soient $t_1, ..., t_n$ des termes sans variables, et p un symbole de relation n-aire. Alors on dira que $p(t_1, ..., t_n)$ est vraie dans M si et seulement si $((t_1)_M, ..., (t_n)_M) \in p_M$; sinon, cette formule sera dite fausse. Une formule close de la forme $\neg A$ sera dite vraie si et seulement si A est fausse. Une formule close de la forme $A \vee B$ sera dite vraie si et seulement la formule A est vraie ou la formule B est vraie. Enfin, une formule close de la forme $\exists xA$ sera dite vraie si et seulement si la formule $A_x[i_a]$ est vraie pour au moins un $a \in M$.

On dira finalement qu'une formule de L est vraie (resp. fausse) dans M si et seulement si sa fermeture (resp. sa négation) est vraie dans M. Si A est vraie dans M, on note

$$M \models A.$$

Remarquons qu'une formule non close peut n'être ni vraie ni fausse. Par exemple, si on considère la structure $\mathbb{N}$ pour le langage de l'arithmétique de Peano, la formule $x = 0$ n'est pas vraie (car $\forall x(x = 0)$ est fausse dans $\mathbb{N}$), mais n'est pas fausse non plus (car $\forall x(\neg x = 0)$ est fausse dans $\mathbb{N}$). De même, la formule $x = 0 \vee x \neq 0$ est vraie dans $\mathbb{N}$ mais aucune des deux formules $x = 0$ et $x \neq 0$ n'est vraie dans $\mathbb{N}$.

2.2.3 Modèles

Définition (Modèle d'une théorie).
Soient T une théorie du premier ordre sur le langage L et M une structure sur L. On dit que M est un modèle de T si et seulement si tout axiome non-logique de T est vrai dans M.

Définition (Validité dans une théorie).
Une formule A est dite valide une théorie T si et seulement si elle est vraie dans tout modèle de T. On note alors $T \models A$.

Par exemple, tout groupe est un modèle de la théorie des groupes. De même, l'ensemble $\mathbb{N}$ muni de la fonction successeur $S : x \mapsto x + 1$ ainsi que de l'addition et la multiplication usuelles est un modèle de PA. On l'appelle le modèle standard de l'arithmétique de Peano.

3 Preuves et tautologies en logique du premier ordre

Dans ce chapitre, on introduit la notion de preuve dans une théorie du premier ordre. On fixe un langage du premier ordre L.

3.1 Preuves

On définit tout d'abord les **axiomes logiques** de L :

— les axiomes propositionnels, formules de la forme $A \vee \neg A$;

— les axiomes d'identité, formules de la forme $x = x$ où x est un terme ;

— les axiomes d'égalité, formules de la forme

$$(x_1 = y_1 \wedge ... \wedge x_n = y_n) \rightarrow f(x_1, ..., x_n) = f(y_1, ..., y_n)$$

ou

$$(x_1 = y_1 \wedge ... \wedge x_n = y_n) \rightarrow p(x_1, ..., x_n) \rightarrow p(y_1, ..., y_n)$$

où f est un symbole de fonction et p un symbole de relation ;

— les axiomes de substitution, formules de la forme $A_x[t] \rightarrow \exists x A$, où t est un terme substituable à x dans A.

Puis on définit les règles d'inférence :

— Expansion : on infère $B \vee A$ de A ;

— Contraction : on infère A de $A \vee A$;

— Associativité : on infère $(A \vee B) \vee C$ de $A \vee (B \vee C)$;

— Coupure : on infère $B \vee C$ de $A \vee B$ et $\neg A \vee C$;

— Existence : si x n'est pas libre dans B, on infère $\exists x A \leftarrow B$.

Par construction, les axiomes logiques sont vrais dans toute structure de L ; et toute formule obtenue par une règle d'inférence à partir de prémisses vraies dans une structure de L est vraie dans cette structure.

Définition (Preuves et théorèmes).
Soit T une théorie du premier ordre. On appelle ***preuve*** *dans T toute séquence finie $A_1, ..., A_n$ de formules telles que pour tout $1 \leq i \leq n$, A_i est un axiome de T (logique ou non-logique) ou peut être inférée de $A_1, ..., A_{i-1}$ par une règle d'inférence. On dit alors que $A_1, ..., A_n$ est une preuve de A_n dans T. Toute formule ayant une preuve dans T est appelée* ***théorème*** *de T. Si A est un théorème de T, on note $T \vdash A$.*

De nombreuses autres règles d'inférence peuvent être déduites de celles présentées ici, illustrant par exemple l'associativité et la commutativité du $\vee$; l'inférence de $\neg\neg A$ à partir de A ; ou bien celle de $s = t$ à partir de $t = s$. Leurs démonstrations ne seront pas détaillées ici en raison de leur intérêt sommaire, ainsi qu'afin de privilégier la clarté de ce rapport. La totalité de ces démonstrations peut être trouvée par exemple dans [9].

Théorème 3.1.1 (Théorème de validité).
Tout théorème de T est valide dans T.

Preuve.
On le montre par récurrence sur la longueur de la preuve. Soit A un théorème de T. Soit $A_1, ..., A_n$ une preuve de A ($A_n = A$). Si $n = 1$, alors A est un axiome de T, donc A est vrai dans tout modèle de T. Sinon, A est obtenu par inférence à partir des A_i pour $1 \leq i \leq n - 1$. Or pour tout $1 \leq i \leq n - 1$, $A_1, ..., A_i$ est une preuve de A_i. Donc A_i est valide dans T. Donc A est valide dans T.

La réciproque de ce théorème est vraie également ; c'est le « théorème de complétude » de Gödel, que nous aborderons plus tard.

3.2 Valuations

3.2.1 Définitions

On dispose actuellement de deux notions différentes caractérisant la « vérité » d'une formule. Une formule peut être prouvée (théorème) *ou* être valide (vraie dans tout modèle). On introduit maintenant une troisième forme de « vérité » : la tautologie. L'aboutissement de ce rapport sera de prouver qu'elles sont toutes trois équivalentes.

On appelle formule élémentaire tout formule atomique ou de la forme $\exists x B$. Une **valuation** de L est une fonction v de l'ensemble des formules élémentaires de L vers {vrai, faux}. On peut alors étendre v à l'ensemble des formules de L comme suit. Soit A une formule. Si $A = B \vee C$, $v(A) =$ vrai si et seulement si $v(B) =$ vrai ou $v(C) =$ vrai. Si $A = \neg B$, alors $v(A) =$ vrai si et seulement si $v(B) =$ faux.

Soit $\mathcal{A}$ un ensemble de formules. On dit qu'une valuation v satisfait $\mathcal{A}$, ou que v est un modèle de $\mathcal{A}$, si et seulement si pour tout $A \in \mathcal{A}$, $v(A) =$ vrai ; on le note $v \models \mathcal{A}$. Si A est une formule, on dit que A est une conséquence tautologique de $\mathcal{A}$ si et seulement si A est vraie dans tout modèle de $\mathcal{A}$.

Proposition 3.2.1.
Soit A une formule.

Si A peut être inférée à partir d'un ensemble de prémisses $\mathcal{A}$, alors A est une conséquence tautologique de $\mathcal{A}$.

Par conséquent, si A est un théorème d'une théorie T, alors A est une conséquence tautologique de l'ensemble des axiomes non-logiques de T.

Preuve.
Il suffit de le vérifier sur chacune des règles d'inférence et des formes de formule possibles.

3.2.2 Théorème de compacité

Un ensemble de formules $\mathcal{A}$ est dit **satisfiable** s'il admet un modèle. Il est dit **finiment satisfiable** si chacun de ses sous-ensembles finis est satisfiable. Un ensemble satisfiable est donc clairement finiment satisfiable. Mais la réciproque est également vraie : c'est ce qu'affirme le théorème de compacité.

Lemme 3.2.1 (Lemme de Zorn).
Si $(\mathbb{P}, \leq)$ est un ensemble non vide partiellement ordonné tel que tout sous-ensemble totalement ordonné de $\mathbb{P}$ admet un majorant, alors $\mathbb{P}$ admet un élément maximal.

Lemme 3.2.2.
Si $\mathcal{A}$ est un ensemble de formules finiment satisfiable et A une formule, alors $\mathcal{A} \cup \{A\}$ ou $\mathcal{A} \cup \{\neg A\}$ est finiment satisfiable.

Preuve.
Supposons que $\mathcal{A} \cup \{\neg A\}$ ne soit pas finiment satisfiable. Il existe un certain $\mathcal{B} \subset \mathcal{A}$ fini tel que $\mathcal{B} \cup \{\neg A\}$ ne soit pas satisfiable. Alors A est une conséquence tautologique de $\mathcal{B}$. Soit $\mathcal{C} \subset \mathcal{A}$ fini. $\mathcal{B} \cup \mathcal{C}$ est finiment satisfiable. Cela nous fournit une valuation v qui satisfait $\mathcal{B}$, donc qui satisfait $\{A\}$, donc qui satisfait $\mathcal{C} \cup \{A\}$. Ainsi pour tout $\mathcal{C} \subset \mathcal{A}$ fini, $\mathcal{C} \cup \{A\}$ est satisfiable. Donc $\mathcal{A} \cup \{A\}$ est finiment satisfiable.

Théorème 3.2.1 (Théorème de compacité).
Un ensemble de formules $\mathcal{A}$ est satisfiable si et seulement si il est finiment satisfiable.

Preuve.
Il suffit de prouver que si $\mathcal{A}$ est finiment satisfiable, alors il est satisfiable. Supposons donc que $\mathcal{A}$ soit finiment satisfiable. Soit $\mathbb{P}$ l'ensemble de tous les ensembles de formules finiment satisfiables contenant $\mathcal{A}$, et munissons-le de l'ordre partiel $\subset$. $(\mathbb{P}, \subset)$ forme donc bien un ensemble partiellement ordonné et non vide (car contenant $\mathcal{A}$). Soit $\mathbb{Q}$ un sous-ensemble totalement ordonné de $\mathbb{P}$. $\bigcup_{q \in \mathbb{Q}} q$ est alors finiment satisfiable, donc un majorant de $\mathbb{Q}$. On peut ainsi appliquer le lemme de Zorn à $\mathbb{P}$ pour obtenir un élément maximal $\mathcal{M}$. Cet élément a en particulier, d'après le lemme précédent, la propriété que pour toute formule A, $A \in \mathcal{M}$ ou $\neg A \in \mathcal{M}$.

Notons alors v la valuation définie par

$$v(A) = \text{vrai} \Leftrightarrow A \in \mathcal{M} \tag{3.1}$$

lorsque A est une formule élémentaire (et étendue à l'ensemble des formules de L). Montrons par récurrence que l'équation 3.1 vaut pour toute formule A.

- Si A est une formule élémentaire, le résultat est vrai par construction.
- Si A est de la forme $\neg B$ où B vérifie l'équation 3.1, alors $v(A) = \text{vrai}$ ssi $v(B) = \text{faux}$ ssi $B \notin \mathcal{M}$ ssi $A \in \mathcal{M}$.
- Si A est de la forme $B \vee C$, avec B et C vérifiant l'équation 3.1, alors supposons que $v(A) = \text{faux}$. Alors $v(\neg B) = v(\neg C) = \text{vrai}$; donc d'après le point précédent, $\neg B$ et $\neg C$ appartiennent à $\mathcal{M}$. Ce dernier étant finiment satisfiable, il est impossible que $A \in \mathcal{M}$. Donc $A \notin \mathcal{M}$.

On a ainsi mis en évidence une valuation satisfaisant $\mathcal{M}$. Elle satisfait donc en particulier $\mathcal{A}$.

3.2.3 Théorème de complétude sur les valuations

On prouve ici, à l'aide du théorème de compacité, un premier théorème de complétude, qui permet de faire concorder deux des trois « formes de vérité » introduites : celle des preuves, et celle des tautologies.

Lemme 3.2.3 (Règle de détachement).
Si $T \vdash A_1, ..., T \vdash A_n$ et $T \vdash (A_1 \wedge ... \wedge A_n) \to A$, alors $T \vdash A$.

Preuve.
Comme précédemment, on admettra cette règle d'inférence.

Lemme 3.2.4.
Soit $n \geq 2$, et $A_1, ..., A_n$ des formules de L. Si $X = A_1 \vee ... \vee A_n$ est une tautologie (i.e une conséquence tautologique de $\emptyset$), alors c'est un théorème (de $\emptyset$).

Preuve.
On le démontre par récurrence sur la somme S des longeurs de $A_1, ..., A_n$. La longueur d'une formule A vaut 1 si c'est une formule élémentaire, longueur de B+longueur de C+1 si $A = B \vee C$, et longueur de $B + 1$ si $A = \neg B$.

Si $A_1, ..., A_n$ sont toutes des formules élémentaires, on peut fixer comme on le souhaite leur valeur de vérité dans une valuation, et par conséquent X n'est pas une tautologie.

Supposons que tous les A_i soient de longueur ≤ 2. Alors si X est une tautologie, il existe une formule élémentaire B telle que B et $\neg B$ apparaissent toutes les deux parmi les A_i. Or $B \vee \neg B$ est un théorème ; donc par expansion, X est un théorème.

Supposons maintenant qu'il existe un A_i (et on supposera même, sans perte de généralité, que $i = 1$) de longueur > 2.

— Si $A_1 = B \vee C$, alors $(B \vee C) \vee A_2 \vee ... \vee A_n$ est une tautologie. Par hypothèse de récurrence, c'est donc un théorème. Donc X est un théorème par associativité du $\vee$.
— Si $A_1 = \neg\neg B$, alors $B \vee A_2 \vee ... \vee A_n$ est une tautologie. Par hypothèse de récurrence, c'est donc un théorème. On peut alors en inférer X.
— Si $A_1 = \neg(B \vee C)$, alors $\neg B \vee A_2 5 \vee ... \vee A_n$ et $\neg C \vee A_2 5 \vee ... \vee A_n$ sont des tautologies, donc des théorèmes. On en déduit encore une fois que X est un théorème.

On remarque qu'avec $A_1 = ... = A_n = A$, on obtient par contraction que A est une tautologie si et seulement si c'est un théorème.

Théorème 3.2.2 (Théorème de complétude sur les valuations).
Soit $\mathcal{A}$ un ensemble de formules de L, et A une formule de L. Alors A est un théorème de $\mathcal{A}$ si et seulement si c'en est une conséquence tautologique.

Preuve.
On a déjà vu (Proposition 3.2.1) que si A est un théorème de $\mathcal{A}$ alors c'en est une conséquence tautologique. Réciproquement, supposons que A soit une conséquence tautologique de $\mathcal{A}$. Alors $\mathcal{A} \cup \{\neg A\}$ n'est pas satisfiable. Par le théorème de compacité, on déduit qu'il n'est pas finiment satisfiable. Donc il existe $\mathcal{B} = \{B_1, ..., B_n\} \subset \mathcal{A}$ fini tel que $\mathcal{B} \cup \{\neg A\}$ ne soit pas satisfiable. On en déduit que A est une conséquence tautologique de $\mathcal{B}$. Il est aisé de voir alors que $(B_1 \wedge ... \wedge B_n) \to A))$ est une tautologie. D'après la remarque ci-dessus, c'est donc un théorème. Soit P une preuve de ce théorème. D'après la règle de détachement (Lemme 3.2.3), la séquence $B_1, ..., B_n, P, A$ forme alors une preuve de A dans $\mathcal{B}$, donc dans $\mathcal{A}$.

3.3 Exemple de méta-théorème en logique du premier ordre : théorème de déduction

Avant de clore ce chapitre, on présente ci-dessous un exemple de méta-théorème plus concret, portant sur le mécanisme effectif de la démonstration mathématique. Bien qu'étant probablement d'une signification moindre que, par exemple, le théorème de complétude, il a ceci d'important que sa démonstration permet de bien intégrer le fonctionnement des objets manipulés. En effet leur apparente simplicité intuitive s'emploie souvent à nous faire perdre de vue leur grande rigidité formelle...

Si T est une théorie et A une formule, on désigne par $T[A]$ la théorie obtenue en ajoutant A aux axiomes non-logiques de T.

Proposition 3.3.1 (Règle de déduction).
Soit A une formule close. Alors

$$T \vdash A \to B$$

si et seulement si

$$T[A] \vdash B$$

.

Preuve.
Supposons que

$$T \vdash A \to B.$$

Alors,

$$T[A] \vdash A \to B.$$

De plus,

$$T[A] \vdash A$$

. Donc

$$T[A] \vdash B$$

par la règle de détachement.

Supposons maintenant que

$$T[A] \vdash B$$

. Soit $A_1, ..., A_n$ une preuve de B dans $T[A]$. Montrons par récurrence sur i que

$$T \vdash A \to A_i$$

pour $1 \leq i \leq n$, ce qui prouvera le résultat.

A_1 est un axiome de $T[A]$. Si c'est un axiome de T, alors

$$T \vdash A_1$$

et donc

$$T \vdash A \to A_1$$

d'après la règle d'expansion. Si $A_1 = A$, alors

$$T \vdash A \to A_1$$

(c'est l'axiome propositionnel).

Admettons que l'hypothèse soit vraie pour tout $j < i$. Si A_i est un axiome de $T[A]$, on a bien $T \vdash A \to A_i$ par un raisonnement identique au précédent. Si A_i est obtenue à partir des A_j pour $j < i$ par une autre règle d'inférence que celle d'introduction, alors A_i est une conséquence tautologique des A_j. Dans ce cas, $A \to A_i$ est une conséquence tautologique de $\{A \to A_j, j < i\}$. Par hypothèse de récurrence, ces formules sont toutes des théorèmes de T. Alors d'après le théorème de complétude sur les valuations (3.2.2),

$$T \vdash A \to A_i.$$

On suppose enfin que A_i est inférée d'une certaine formule A_j par la règle d'introduction. Par conséquent il existe des formules B et C et une variable v sans occurence libre dans C telles que $A_j = B \to C$ et $A_i = \exists v(B \to C)$. Par hypothèse de récurrence,

$$T \vdash A \to B \to C$$

. Alors par commutativité du $\vee$,
$$T \vdash B \to A \to C.$$
Comme A est close, v n'a pas d'occurence libre dans $A \to C$. Donc par règle d'introduction,
$$T \vdash \exists v B \to A \to C.$$
À nouveau par commutativité du $\vee$, on a alors
$$T \vdash A \to \exists v B \to C$$
c'est-à-dire
$$T \vdash A \to A_i.$$
Cela conclut la preuve.

4 Consistance et complétude

Dans quelle mesure risque-t-on un jour de découvrir que les mathématiques telles qu'on les a construites sont contradictoires ? Peut-on tout démontrer ? Y a-t-il une réponse à chaque problème mathématique ? Ces questions motivent les notions de consistance et de complétude que nous allons aborder dans ce chapitre, et sont au cœur du programme de Hilbert [4] et des travaux de Gödel [3].

4.1 Consistance

Définition (Consistance d'une théorie).
Une théorie T est dite ***inconsistante****, ou* ***incohérente****, si toute formule y est un théorème. Dans le cas contraire, on dit que T est* ***consistante****, ou* ***cohérente****.*

Proposition 4.1.1.
Une théorie T est inconsistante si et seulement si il existe une formule A telle que A et $\neg A$ soient des théorèmes de T.

Preuve.
Le sens direct est évident. Prouvons la réciproque. Supposons donc qu'il existe A telle que $T \vdash A$ et $T \vdash \neg A$. Soit B une formule quelconque. Par la règle d'expansion, on a $T \vdash A \vee B$ et $T \vdash \neg A \vee B$. Donc par la règle de coupure, $T \vdash B \vee B$. Enfin, par la règle de contraction, $T \vdash B$.

Cette preuve est très courte, mais le résultat reste étonnant. Prenons l'exemple de la théorie naïve des ensembles telle que formalisée par Frege. Elle comporte le schéma d'axiomes suivant : pour toute formule A du langage de la théorie des ensembles, et toute variable v de A, « $\exists X \forall x (x \in X \leftrightarrow A_v[x])$ » est un axiome. Si l'on considère alors la formule $v \notin v$ et qu'on y applique cet axiome, on peut montrer $\exists X (X \in X \wedge X \notin X)$. La négation de cette formule étant une tautologie de la théorie, on a montré que cette dernière était inconsistante. Par conséquent toute formule peut être prouvée dans la théorie naïve des ensembles. Comment se fait-il, dans ce cas, qu'il ait fallu une trentaine d'années pour que Russell [8] finisse par trouver le raisonnement ci-dessus ? Ainsi, contrairement à ce que suggère la définition de l'inconsistance, celle-ci n'est pas nécessairement facile à détecter.

On appelle **partie finiment axiomatisée** d'une théorie T toute théorie sur le même langage que T, ne comportant qu'un nombre fini d'axiomes qui sont tous des théorèmes de T. La preuve d'un théorème ne nécessitant jamais qu'un nombre fini d'axiomes, on remarque alors qu'une théorie est consistante si et seulement si toute ses parties finiment axiomatisées sont consistantes. En effet, si T est inconsistante et si A est une formule de T, il existe une preuve de A dans T et une preuve de $\neg A$ dans T. Si l'on prend la partie finiment axiomatisée de T qui a pour axiomes non-logiques toutes les formules intervenant dans ces preuves, celle-ci démontre A et $\neg A$ donc est inconsistante. La réciproque, quant à elle, est immédiate.

Proposition 4.1.2.
Si une théorie admet un modèle, alors elle est consistante.

Preuve.
Soit M un modèle d'une théorie T. Supposons que T soit inconsistante. Soit A une formule close. A et $\neg A$ sont alors des théorèmes de T ; donc par le théorème de validité 3.1.1, A et $\neg A$ sont valides dans M. C'est une contradiction.

Théorème 4.1.1 (Théorème de fermeture).
Soient A une formule, et B sa fermeture. Alors $T \vdash A$ si et seulement si $T \vdash B$.

Preuve.
On le montre en utilisant la règle d'existence.

Proposition 4.1.3.
Soient A une formule, et B sa fermeture. Alors $T \vdash A$ si et seulement si $T[\neg B]$ est inconsistante.

Preuve.
Supposons que $T \vdash A$. Alors $T \vdash B$; et donc $T[\neg B] \vdash B$. Ainsi $T[\neg B]$ est inconsistante.

Réciproquement, supposons $T[\neg B]$ inconsistante. Donc d'après la règle de déduction (Proposition 3.3.1), $T \vdash \neg B \to B$, *i.e.* $T \vdash B$. Donc $T \vdash A$.

4.2 Enoncés du théorème de complétude

Ci-dessous sont présentés deux énoncés du théorème de complétude de Gödel, dont on montre qu'ils sont équivalents. On rappelle qu'une formule est dite valide dans une théorie si et seulement si elle est vraie dans tout modèle de cette théorie.

Théorème 4.2.1 (Théorème de complétude - première forme).
Une formule d'une théorie T y est un théorème si et seulement elle y est valide.

Théorème 4.2.2 (Théorème de complétude - seconde forme).
Une théorie est consistante si et seulement si elle admet un modèle.

Proposition 4.2.1.
Les deux formes du théorème de complétude sont équivalentes.

Preuve.
Admettons la première forme du théorème de complétude. On sait, de plus, qu'une théorie admettant un modèle est nécessairement consistante. Fixons donc T une théorie consistante. Il existe alors une formule A qui n'est pas un théorème de T ; donc, par la première forme du théorème de complétude, A n'est pas valide dans T. Par définition, cela nous fournit un modèle de T.

Admettons maintenant la seconde forme du théorème de complétude. Soient T une théorie et A une formule de T. D'après le théorème de fermeture (4.1.1), on peut supposer sans perte de généralité que A est close. On sait d'après la proposition 4.1.3 que A est un théorème si et seulement si $T[\neg A]$ est inconsistante. D'après la seconde forme du théorème d'incomplétude, A est donc un théorème si et seulement si $T[\neg A]$ n'admet aucun modèle. Or, comme A est close, les modèles de $T[\neg A]$ sont exactement les modèles de T dans lesquels A est fausse. Autrement dit, A est un théorème si et seulement si elle est vraie dans tout modèle de T.

4.3 Preuve du théorème de complétude

Prouvons maintenant le théorème de complétude. On en prouvera la seconde forme (Théorème 4.2.2). La première preuve en a été donnée par Gögel en 1930 ; mais celle qui est présentée ici est due à Leo Henkin.

On se donne donc une théorie T que l'on suppose consistante. Les seuls objets dont nous disposions sont syntaxiques ; c'est donc à partir d'eux qu'il va falloir construire un modèle de T.

4.3.1 Structure canonique

Tout d'abord, on suppose que T possède au moins une constante. Si ce n'est pas le cas, il suffit de lui en ajouter une sans modifier les axiomes, ce qui ne modifie pas sa consistance, et nous donnera bien un modèle de la théorie d'origine. Notons N l'ensemble des termes sans variables de T (qui est donc non-vide). Si a et b sont deux termes sans variables tels que $T \vdash a = b$, alors les interprétations de a et b doivent être les mêmes dans tout modèle de T. On définit par conséquent une relation binaire $\sim$ sur N de la manière suivante :

$$\text{pour tous } a, b \in N, a \sim b \text{ ssi } T \vdash a = b.$$

Lemme 4.3.1.
La relation $\sim$ est une relation d'équivalence sur N.

Preuve.
— **Réflexivité** : pour tout $a \in N$, on a $T \vdash a = a$ par l'axiome logique d'identité ;
— **Symétrie** : pour tous $a, b \in N$, si $T \vdash a = b$ alors $T \vdash b = a$;
— **Transitivité** : pour tous $a, b, c \in N$, on a $T \vdash a = b \to b = c \to a = c$; donc par règle de détachement, si $T \vdash a = b$ et $T \vdash b = c$, alors $T \vdash a = c$.

On note alors M l'ensemble-quotient de N par la relation d'équivalence $\sim$. Pour $a \in N$, on notera $[a]$ la classe de a. M sera l'univers de notre modèle.

On définit naturellement l'interprétation des symboles non-logiques de T dans M de la façon suivante :

$$c_M = [c]$$

$$f_M([a_1], ..., [a_n]) = [f(a_1, ..., a_n)]$$

et

$$p_M([a_1], ..., [a_n]) \text{ si et seulement si } T \vdash p(a_1, ..., a_n)$$

où c est un symbole de constante ; $a_1, ..., a_n$ sont des termes sans variable ; f est un symbole de fonction à n variables ; et p est un symbole de relation n-aire.

La structure M que nous venons de définir sur le langage L de T est appelée **structure canonique** sur le langage de T.

Par récurrence sur la longueur des expressions considérées, on prouve facilement le lemme suivant :

Lemme 4.3.2.
Soient a un terme sans variable et A une formule sans variable de T. Alors

$$a_M = [a]$$

et

$$T \vdash A \Leftrightarrow M \models A.$$

4.3.2 Théorie de Henkin

A priori, il n'y a pas tellement de raison pour que la structure canonique sur T en soit un modèle. En particulier, cela est d'autant moins probable que T a peu de constantes. En effet, pour tout théorème de la forme $\exists vA$, il faut trouver un terme sans variable t tel que $A_v[t_M]$ est vraie. On introduit donc un type particulier de théorie, pour lequel la structure canonique sera effectivement un modèle.

Définition (Théorie de Henkin).
*On appelle **théorie de Henkin** toute théorie T telle que pour toute formule close de la forme $\exists vA$, il existe un symbole de constante c tel que*

$$T \vdash \exists vA \to A_x[c].$$

Définition.
*Une formule A d'une théorie T est dite **décidable** dans T si et seulement si $T \vdash A$ ou $T \vdash \neg A$. Une théorie consistante dans laquelle toute formule close est décidable est dite **complète**.*

Théorème 4.3.1.
Si T est une théorie de Henkin complète, alors la structure canonique sur T est un modèle de T.

Preuve.
Montrons que, pour toute formule close A,

$$T \vdash A \Leftrightarrow M \models A.$$

Montrons cette propriété par récurrence sur le nombre de fois que les symboles logiques $\vee, \neg$ et $\exists$ apparaissent dans A. Ce nombre est appelé la **hauteur** de A.

Soit A de hauteur 0, c'est-à-dire que A est une formule atomique, alors soit A est de la forme $t = s$ où t et s sont des termes sans variable (A étant close), et dans ce cas $T \vdash A \Leftrightarrow M \models A$ par construction de M ; soit A est de la forme $p(t_1, ..., t_n)$ où $t_1, ..., t_n$ sont des termes sans variable, et dans ce cas également, $T \vdash A \Leftrightarrow M \models A$ par construction de l'interprétation p_M du symbole de relation p.

Soit A une formule de hauteur > 0 et supposons que le résultat soit vrai pour toute formule close de hauteur strictement inférieure à celle de A. Si A est de la forme $\neg B$ où B est une formule close et si $T \vdash A$, alors comme T est consistante, $T \nvdash B$. Par hypothèse de récurrence, $M \not\models B$. Donc $M \models A$. Réciproquement, si $M \models A$, alors $M \not\models B$, et donc par hypothèse de récurrence, $T \nvdash B$. Comme T est complète, on a alors nécessairement $T \vdash A$.

Supposons A de la forme $B \vee C$. Si $T \vdash A$, comme B et C sont closes et comme T est complète, on a nécessairement $T \vdash B$ ou $T \vdash C$. Donc par hypothèse de récurrence, $M \models B$ ou $M \models C$, donc $M \models A$. Réciproquement, si $M \models A$, alors par définition on a $M \models B$ ou $M \models C$. Donc, par hypothèse de récurrence, $T \vdash B$ ou $T \vdash C$. Alors d'après la règle d'expansion, on a bien que $T \vdash A$.

Supposons enfin que A soit de la forme $\exists xB$, et que $T \vdash A$. Comme T est une théorie de Henkin, il existe un symbole de constante c tel que

$$T \vdash \exists xB \to B_x[c].$$

D'après la règle de détachement, on a donc

$$T \vdash B_x[c].$$

Comme la hauteur de $B_x[c]$ est strictement inférieure à celle de A, et comme $B_x[c]$ est close car A l'est (donc il n'existe pas d'autre variable ayant une occurence libre dans B que x), on peut appliquer l'hypothèse de récurrence et obtenir que

$$M \models B_x[c].$$

Alors en notant $m = c_M$, on a, comme $c = i_m$, $M \models B_x[i_m]$. Par définition, on a donc

$$M \models A.$$

Réciproquement, supposons que $M \models A$. Alors il existe $m \in M$ tel que

$$M \models B_x[i_m].$$

Par construction de M, il existe un terme sans variable a tel que $m = [a]$. D'après le lemme 4.3.2, $a_M = (i_m)_M = m$. Autrement dit,

$$M \models a = i_M.$$

Donc,

$$M \models B_x[a].$$

Par hypothèse de récurrence on a alors

$$T \vdash B_x[a].$$

Or, $B_x[a] \to \exists x B$ est un axiome de T (axiome de substitution). Donc par la règle de détachement, on obtient que

$$T \vdash A.$$

On a donc bien prouvé que M était un modèle de T.

4.3.3 Extension de Henkin

On dit qu'une théorie T' est une **extension** de T si et seulement si le langage de T' est une extension de celui de T (*i.e.* chaque symbole respectivement de constante, fonction ou relation du langage de T est un symbole respectivement de constante, fonction ou relation du langage de T') et tout axiome non-logique de T est un théorème de T'. De plus, on appellera T' une **extension conservative** de T si et seulement si c'est une extension de T telle que toute formule de T qui est un théorème de T' soit un théorème de T.

Proposition 4.3.1.
Si T est consistante et T' est une extension conservative de T, alors T' est consistante.

Preuve.
Si T' est inconsistante, toute formule de T' est un théorème de T'. En particulier, toute formule de T est un théorème de T', donc un théorème de T puisque T' en est une extension conservative. Donc T est inconsistante.

Théorème 4.3.2.
Toute théorie admet une extension de Henkin conservative.

Preuve.
Soit T une théorie. Soit $T_0 = T$. Pour $n \in \mathbb{N}$, on définit une extension T_{n+1} de T_n comme suit : pour toute formule close de T_n de la forme $\exists x A$, on ajoute un nouveau symbole de constante $c_{\exists x A}$ ainsi qu'un nouvel axiome $\exists x A \to A_x[c_{\exists x A}]$. Les nouvelles constantes ainsi ajoutées seront appelées **constantes spéciales de niveau** $n+1$. Si c est une constante spéciale, l'axiome $\exists x A \to A_x[c]$ sera appelé **axiome spécial de** c.

Soit alors T' la théorie dont le langage est « l'union » de tous les langages des T_n, c'est-à-dire que les symboles de constante (resp. de fonction, resp. de relation) de T' sont exactement les symboles α tels qu'il existe $n \in \mathbb{N}$ pour lequel α est un symbole de constante (resp. de fonction, resp. de relation) de T_n. L'ensemble des axiomes non-logiques de T' est, quant à lui, l'union (au sens strict) des ensembles d'axiomes non-logiques de toutes les théories T_n. Il est facile de vérifier que T' est alors une théorie de Henkin. Il reste à montrer qu'elle est conservative en tant qu'extension de T.

Montrons que, pour tout $n \in \mathbb{N}$, T_{n+1} est une extension conservative de T_n. Plus précisément, si on note T_n^c la théorie obtenue à partir de T_n en y ajoutant les constantes spéciales de niveau $n+1$, mais aucun axiome non-logique, T_n^c est une extension conservative de T_n. Il suffit donc de montrer que T_{n+1} est une extension conservative de T_n^c.

Soit A une formule de T_n^c telle que $T_{n+1} \vdash A$. Soient $B_1, ...B_k$ les axiomes spéciaux distincts associés à des constantes spéciales de niveau $n+1$ apparaissant dans la preuve de A dans T_{n+1}. Alors on a

$$T_n^c \vdash (B_1 \wedge ... \wedge B_k) \to A.$$

Notons $B_1 = \exists x C \to C_x[c]$ (c est donc la constante spéciale associée à $\exists x C$). Puisque les B_i sont distincts, c n'apparaît ni dans B_2, ..., ni dans B_n.

Soit y une variable n'apparaissant pas dans $(B_1 \wedge ... \wedge B_k) \to A$. Alors on a

$$T_n^c \vdash (\exists x C \to C_x[y]) \to (B_2 \wedge ... \wedge B_n) \to A.$$

Par la règle d'introduction on a alors

$$T_n^c \vdash \exists y(\exists x C \to C_x[y]) \to (B_2 \wedge ... \wedge B_n) \to A.$$

Ainsi,

$$T_n^c \vdash \exists x C \to \exists y C_x[y];$$

donc comme y n'apparaît pas dans $\exists x C$, on a

$$T_n^c \vdash \exists y(\exists x C \to C_x[y]).$$

Enfin, d'après la règle de détachement, on a

$$T_n^c \vdash (B_2 \wedge ... \wedge B_n) \to A.$$

En répétant ce procédé k fois, on obtient que

$$T_n^c \vdash A.$$

Ainsi pour tout $n \in \mathbb{N}$, T_{n+1} est une extension conservative de T_n. Si maintenant A est une formule de T telle que $T' \vdash A$, comme la preuve de A dans T' ne fait appel qu'à un nombre fini d'axiomes et de constantes, il existe $n \in \mathbb{N}$ tel que $T_n \vdash A$. Or, par une récurrence immédiate, on montre que T_n est une extension conservative de $T = T_0$. Donc $T \vdash A$.

Théorème 4.3.3.
Toute théorie consistante admet une extension complète.

Preuve.
Soit T une théorie consistante. Notons $\mathbb{P}$ l'ensemble des extensions consistantes de T utilisant le même langage que T. On munit $\mathbb{P}$ de la relation d'ordre « être une extension de » (cette relation est en particulier bien antisymétrique car toutes les théories de $\mathbb{P}$ ont le même langage). Soit $\mathbb{Q}$ un sous-ensemble totalement ordonné de $\mathbb{P}$. Soit $T_\mathbb{Q}$ la théorie constituée du même langage que T et dont l'ensemble des axiomes est l'union des ensembles d'axiomes de tous les éléments de $\mathbb{Q}$. $T_\mathbb{Q}$ est bien une extension de T utilisant le même langage. Supposons, par l'absurde, qu'elle soit inconsistante. Soit A une formule telle que $T_\mathbb{Q} \vdash A$ et $T_\mathbb{Q} \vdash \neg A$. Soient $A_1, ..., A_n$ les axiomes apparaissant dans les preuves de A et $\neg A$ dans $T_\mathbb{Q}$. Par définition, pour tout $1 \leq i \leq n$, il existe une théorie $T_i \in \mathbb{Q}$ telle que A_i soit un axiome de T_i. $\mathbb{Q}$ étant totalement ordonné, il

existe $1 \leq i_0 \leq n$ tel que $T_{i_0} = \max\{T_i, 1 \leq i_0 \leq n\}$. Alors les preuves de A et $\neg A$ dans $T_\mathbb{Q}$ sont aussi des preuves de A et $\neg A$ dans T_{i_0}. Donc T_{i_0} est inconsistante : absurde. Ainsi $T_\mathbb{Q}$ est consistante ; elle appartient donc à $\mathbb{P}$. C'est, de plus, clairement un majorant de $\mathbb{Q}$.

Ainsi tout sous-ensemble totalement ordonné de $\mathbb{P}$ admet un majorant. On peut donc appliquer le lemme de Zorn à $\mathbb{P}$ et en obtenir un élément maximal T'. Montrons alors que T' est complète. Soit A une formule close telle que $\neg A$ ne soit pas un théorème de T'. D'après la propriété 4.1.3, $T'[A]$ est alors consistante. Ainsi $T'[A] \in \mathbb{P}$. De plus $T'[A]$ est évidemment une extension de T'. Par maximalité de T', on a donc $T'[A] = T'$; donc A est un théorème de T'.

On a donc bien obtenu une extension T' de T complète.

La proposition suivante est une évidence au vu de la définition d'une théorie de Henkin.

Proposition 4.3.2.
Toute extension d'une théorie de Henkin conservant le même langage est une théorie de Henkin.

Il ne reste maintenant qu'à conclure. Soit T une théorie consistante. D'après le théorème 4.3.2, T admet une extension de Henkin conservative T'. L'extension étant conservative, d'après la proposition 4.3.1, T' est consistante. Le théorème 4.3.3 nous en fournit donc une extension complète T''. De plus, au vu de la preuve de ce théorème, cette extension possède le même langage, et donc par la propriété 4.3.2, T'' est une théorie de Henkin. Enfin, d'après le théorème 4.3.1, la structure canonique M de T'' est un modèle de T''. Comme T'' est une extension de T, M est un modèle de T. Le théorème de complétude est donc prouvé.

Ainsi, les trois définitions différentes que l'on peut imaginer de la « véracité » d'une formule concordent. En particulier, ce théorème nous assure que toute formule empiriquement vraie (c'est-à-dire vérifiée dans tous les modèles de la théorie) est prouvable dans la théorie. En somme, tout ce qu'il y a à trouver est trouvable... Voilà qui a de quoi rassurer le mathéticien ; il n'y a qu'à se mettre au travail.

5 Suites possibles : conséquences du théorème de complétude et théorèmes d'incomplétude

On s'intéresse dans ce chapitre à un certain nombre de résultats que nous ne démontrerons pas toujours, avec pour objectif de donner une idée de ce que peut être la « suite » de la connaissance en logique ; alors que le théorème de complétude concluait un débat sur le bien-fondé des définitions d'une preuve, d'une théorie, d'un modèle, etc., ces résultats introduisent de nouveaux questionnements et nous engagent sur de nouvelles voies.

5.1 Conséquences du théorème de complétude

On présente ici deux théorèmes importants qui découlent aisément du théorème de complétude, afin de donner une idée de sa portée.

5.1.1 Théorème de Löwenheim-Skolem descendant

Théorème 5.1.1 (Théorème de Löwenheim-Skolem descendant).
Soit κ un cardinal infini et T une théorie consistante dont le langage a au plus κ symboles non-logiques. Alors il existe un modèle M de T tel que $|M| \leq \kappa$.

Preuve.
Il faut reprendre la preuve du théorème de complétude et vérifier que le cardinal du modèle construit est inférieur ou égal à κ. On a construit une extension de Henkin de T en définissant une suite de théories $(T_n)_{n\in\mathbb{N}}$. Supposons que le langage de T_n ait au plus κ symboles non-logiques. On ajoute alors à T_n au plus autant de symboles de constante qu'il y a de formules de T_n ; donc au plus κ. Par conséquent T_{n+1} possède également au plus κ symboles non-logiques. On a ensuite pris comme langage pour l'extension de Henkin T' de T « l'union » de tous les langages des T_n. Donc le langage de T' a au plus κ symboles non-logiques. On en a ensuite fourni une extension complète ayant le même langage, puis on a utilisé comme modèle pour T la structure canonique sur ce langage. Il est facile, d'après sa construction, de constater que la structure canonique sur un langage a elle-même un cardinal inférieur ou égal à celui de l'ensemble des symboles non-logiques de ce langage. On a donc bien un modèle de T de cardinal au plus κ.

Ce théorème a ceci de remarquable qu'il implique ce qu'on appelle le « paradoxe de Skolem ». Au fond, tout mathématicien est persuadé que la théorie dans laquelle il travaille est consistante. Si elle ne l'était pas, absolument tout serait vrai ; et on peut trouver de mutliples raisons pour montrer que c'est insensé (à commencer par les innombrables applications des mathématiques aux autres sciences). La théorie des ensembles, qu'on utilise souvent pour fonder les mathématiques, ne possède pourtant qu'un nombre fini de symboles non-logiques. Donc, si elle est consistante, elle admet un modèle dénombrable, c'est-à-dire un ensemble dénombrable contenant tous les ensembles, donc tous les nombres réels, qui sont en nombre indénombrable ! Le paradoxe réside dans le fait que l'on utilise deux notions distinctes de « dénombrabilité » : une *interne* à la théorie des ensembles formelle, et une *externe*, celle des « méta-mathématiques » qui permettent de formaliser la théorie des ensembles.

5.1.2 Théorème de compacité pour les théories du premier ordre

La démonstration de ce théorème est absolument triviale maintenant que l'on dispose du théorème de complétude ; le réultat, néanmoins, vaut la peine d'être mentionné.

Théorème 5.1.2 (Théorème de compacité pour les théories du premier ordre). *Une théorie admet un modèle si et seulement si chacune de ses parties finiement axiomatisées admet un modèle.*

Preuve.
T admet un modèle $\Leftrightarrow$ T est consistante $\Leftrightarrow$ chacune de ses parties finiement axiomatisées est consistante $\Leftrightarrow$ chacune de ses parties finiement axiomatisées admet un modèle.

À l'aide de ce théorème, on peut par exemple montrer qu'il n'existe aucune théorie dont les modèles seraient précisément les ensembles finis ; ou bien construire des modèles non-standards arbitrairement grands de l'arithmétique de Peano. Ceux-ci sont en particulier utiles pour les théorèmes d'incomplétude présentés dans la section suivante.

5.2 Théorèmes d'incomplétude

La question de savoir si l'on peut prouver que les mathématiques (et en particulier l'arithmétique de Peano) sont consistantes, posée par Hilbert en 1902 [4], a trouvé une forme de réponse avec les théorèmes d'incomplétude de Gödel en 1931 [3]. On n'en trouvera pas ici les énoncés exacts, car leur compréhension nécessite d'avancer au-delà de ce que j'ai appris durant mon stage dans la théorie de la logique mathématique. Cependant, leur contenu peut être exposé de manière informelle comme suit :

— Premier théorème d'incomplétude : Toute théorie consistante (et dont les axiomes peuvent être reconnus par un procédé algorithmique) capable de formaliser l'arithmétique de Peano est incomplète.
— Second théorème d'incomplétude : Dans toute théorie satisfaisant les hypothèses précédentes, sa consistance peut s'exprimer par une formule, et celle-ci est indécidable.

Une théorie suffisamment élaborée pour formaliser les mathématiques est donc nécessairement incomplète, et en particulier incapable de prouver sa propre consistance. On peut alors se demander « dans quelle mesure » elle l'est ; y a-t-il beaucoup d'énoncés indécidables, et qui sont-ils ? Par exemple, les diverses conjectures que l'on ne parvient pas pour l'instant à prouver malgré la grasse rémunération promise en font-elles partie ? Et pourquoi rencontre-t-on si peu souvent des énoncés indécidables ? Ces questions motivent une part importante de la recherche actuelle en logique. Pour conclure ce rapport, j'énonce un dernier théorème, dont l'étude était l'objectif de mon stage à son début mais que je n'ai pas eu le temps d'aboutir. Je le cite tout de même, sans démonstration, en tant qu'il illustre un type de résultats que peuvent produire ces diverses questions. Il s'agit de donner un énoncé purement combinatoire (comprendre : dénué de toute notion logique), qui est vrai dans $\mathbb{N}$ mais pas dans un certain modèle non-standard de PA, et qui, par conséquent, est indécidable dans PA. Comme quoi, des énoncés indécidables existent ; ils ne sont pas seulement des lubies de logiciens ; on peut tomber dessus, un jour, sans s'en douter. Mieux vaut alors être au courant avant de s'escrimer en vain.

Théorème 5.2.1.
Soient e, r, k trois entiers. Il existe $M \in \mathbb{N}$ tel que pour toute fonction $P : \mathcal{P}_e(M) \to r$ de $\mathcal{P}_e(M)$ (ensemble des parties à e éléments de M), il existe un ensemble $H \subset M$ tel que les trois propositions suivantes soient vérifiées :

— $|H| \geq \min(H)$;
— $|H| \geq k$;
— $\forall A, B \in \mathcal{P}_e(H), P(A) = P(B)$.

Théorème 5.2.2 (Théorème de Paris-Harrington).
Le théorème 5.2.1 ci-dessus est formalisable dans PA, mais y est indécidable.

Bibliographie

[1] Nicolas Bourbaki. *Eléments de mathematique*. Springer, 1970.

[2] Anton Dumitriu. La science de la logique. *Notre Dame Journal of Formal Logic*, page 386, oct 1971.

[3] Kurt Gödel. *On formally indecidable propositions of Principia Mathemica and related systems*. Dover, 1962. Traduction anglaise par Bernard Meltzer de l'article original publié en 1931.

[4] David Hilbert. Mathematical problems : lecture delivered before the international congress of mathematicians at paris in 1900. *Bulletin of the American Mathematical Society*, pages 437–479, 1903. traduction en anglais par Dr. Maby Winton Newson, consultation en ligne en 2016 : http ://www2.clarku.edu/ djoyce/hilbert/.

[5] Thomas Kuhn. *La structure des révolutions scientifiques*. Flammarion, 1962.

[6] Jeff Paris and Leo Harrington. A mathematical incompleteness in peano arithmetic. In Jon Barwise, editor, *Handbook of Mathematical Logic*, pages 1133–1142. North-Holland Publishing Company, 1977.

[7] Karl Popper. *Conjectures and Refutations : The Growth of Scientific Knowledge*. Routledge, 1963.

[8] Bertrand Russell. *The Principles of Mathematics*. Cambridge university Press, 1903.

[9] Shashi Mohan Srivastava. *A Course on Mathematical Logic*. Springer, 2008.

[10] Jaap van Oosten. Introduction to peano arithmetic. june 1999.