Introduction à un domaine de recherche : Caractérisation des structures non dénombrables dans des logiques infinitaires du second-ordre.

Charles VALENTIN

1 Introduction

Un des buts de la logique et de la théorie des modèles est de classifier les structures mathématiques en fonction des énoncés qu'elles vérifient dans une logique donnée. La propriété la plus forte qu'on peut espérer dans cette direction est la caractérisabilité :

Définition 1. *Une structure M est caractérisable dans une logique L s'il existe un énoncé φ de L tel que pour toute structure N, N satisfait φ si et seulement si N est isomorphe à M.*

De ce point de vue, la logique du premier ordre échoue, la logique du premier ordre caractérise uniquement les structures finies, par le théorème de Löwenheim-Skolem. Pour espérer caractériser de plus vastes classes de structures, il faut étendre la logique du premier ordre. Il existe plusieurs manières de faire cela. Il y a les logiques infinitaires, qui permettent des disjonctions, des conjonctions sur une infinité de formules et des quantifications sur des suites infinies de variables, il y a les logiques d'ordre supérieur, qui permettent de quantifier sur l'ensemble des parties des structures, ou l'ensemble des parties de l'ensemble des parties, etc... On peut également introduire de nouveaux quantificateurs, appelés quantificateurs généralisés, pour exprimer de nouvelles propriétés. Ces diverses extensions permettent de caractériser plus de structures que la logique du premier ordre. Typiquement, la plupart des structures mathématiques usuelles sont caractérisables au second ordre, comme les entiers naturels avec leur ordre, ou le corps des réels avec son ordre.

Définition 2. *Etant donné deux cardinaux κ, λ, la logique $L_{\kappa\lambda}$ est la logique où on s'autorise les disjonctions et conjonctions sur un nombre $< \kappa$ de formules, et les quantifications sur les suites de longueurs $< \lambda$ de variables.*

Par exemple, $L_{\omega\omega}$ est la logique du premier ordre. Un exemple important est la logique $L_{\omega_1\omega}$. Son importance se manifeste notamment par le résultat suivant, le théorème d'isomorphisme de Scott (voir [7] p.59) :

Théorème 1. *Pour toute structure dénombrable M dans un langage dénombrable, il existe un énoncé φ de $L_{\omega_1\omega}$ tel que, pour toute structure dénombrable N, N satisfait φ si et seulement si N est isomorphe à M.*

Cependant, les propriétés de ces différentes logiques généralisées peuvent être sensibles au contexte ensembliste dans lequel on se place. Certaines propriétés très simples, qu'on espérerait absolues comme c'est le cas pour la logique du premier ordre, ne le sont pas entre modèles bien fondés de ZFC. Citons quelques exemples fondamentaux, qu'on peut trouver dans [6]. Tout d'abord, dans la logique du second ordre, la relation de satisfaction n'est pas absolu : il existe deux modèles bien fondés $V \subset V'$ de ZFC qui sont en désaccord sur les énoncés du second ordre que vérifient $\mathbb{N}$, muni de sa structure usuelle. Toujours pour le second ordre, la satisfaisabilité d'un énoncé n'est pas absolu : on peut trouver un énoncé du second ordre qui est satisfaisable si et seulement si l'hypothèse du continu est vrai, et un qui est satisfaisable si et seulement si l'hypothèse du continu est fausse. Ensuite, pour les logiques infinitaires, l'ensemble des formules n'est pas absolu, car la notion de cardinal ne l'est pas. On peut également citer l'existence de cardinal de compacité ou de cardinal de Löwemheim-Skolem pour ces différentes logiques, qui est parfois équivalente à l'existence de certains grands cardinaux :

Définition 3. *κ est un cardinal de compacité pour la logique L si pour toute L-théorie T, T a un modèle si et seulement si toute partie de T de taille $< \kappa$ admet un modèle.*

Définition 4. *κ est un cardinal de Löwenheim-Skolem pour la logique L si toute structure admet une L-sous-structure élémentaire de cardinalité $< \kappa$.*

Un certain nombre de résultats de ce type peuvent être trouvés dans [4]. Citons notamment : l'affirmation que deux structures dénombrables vérifiant les mêmes énoncés du second ordre sont nécessairement isomorphes est indépendante de ZFC.

Revenons-en à la caractérisabilité. Ici, on s'intéresse aux logiques infinitaires du second-ordre :

Définition 5. *La logique $L^2_{\kappa\lambda}$ est la logique $L_{\kappa\lambda}$, à laquelle on ajoute des quantificateurs du second-ordre.*

Par exemple, la définition d'un bon ordre peut être donnée par l'énoncé du second ordre suivant :

$$\forall X(\exists x(x \in X)) \to \exists x \in X \forall y \in X, x \leq y$$

On peut également définir les ensembles dénombrables par l'énoncé de $L^2_{\omega_1\omega}$ suivant

$$\exists f \exists x \forall y \bigvee_{n \in \mathbb{N}} y = f^n(x)$$

Il est assez facile de démontrer que, dans la logique $L^2_{\kappa^+\omega}$, toute structure de cardinal κ dans un langage de taille $\leq \kappa$ est caractérisable. La question naturelle qui se pose est : peut-on faire mieux ? Peut-on, dans la logique $L^2_{\kappa\omega}$, caractérisé toutes les structures de cardinal κ dans un language de taille $< \kappa$?

Définition 6. *$Car(\kappa)$ est l'assertion que toute structure de taille κ dans un langage de taille $< \kappa$ est caractérisable dans la logique $L^2_{\kappa\omega}$.*

Dans [2], J.Väänänen démontre le résultat suivant :

Théorème 2. *Soit κ un cardinal régulier, tel que $\kappa = \aleph_\alpha$, $\beth_{\omega_1}(|\alpha| + \omega) \leq \kappa$, et $2^\lambda < 2^\kappa$ pour tout $\lambda < \kappa$. Soit T une théorie du premier ordre complète et dénombrable. Alors tout modèle de T de taille κ est caractérisable dans $L^2_{\kappa\omega}$ si et seulement si T est une théorie superstable, shallow sans DOP ou $OTOP$.*

Ce théorème ne sera pas utilisé par la suite, il est juste là pour illustrer le type de théorème que l'on s'attend à avoir : sous certaines hypothèses très fortes sur l'arithmétique cardinale, on obtient une description précise des structures caractérisables dans $L^2_{\kappa\omega}$ en des termes issus de recherches en théorie des modèles. Notre résultat principal est que ces hypothèses sont en quelque sorte inévitables : on démontre que, pour κ successeur d'un cardinal régulier, $Car(\kappa)$ est consistant. Pour ce faire, on mobilise de nombreux concepts et techniques de théorie des ensembles, comme l'univers constructible de Gödel, le forcing, des généralisations de l'axiome de Martin.

2 Indépendance de $Car(\aleph_1)$

Ici, on ne présentera que le cas $\kappa = \aleph_1$, pour simplifier. La démonstration est similaire pour κ successeur d'un cardinal régulier, mais elle nécessite des outils techniques supplémentaires.

Il y a une obstruction simple d'arithmétique cardinale à $Car(\aleph_1)$. Pour un langage contenant un symbole binaire, il y a $2^{\aleph_1}$ structures de cardinalité $\aleph_1$ deux-à-deux non isomorphes. Pour un langage dénombrable, il y a $2^{\aleph_0}$ formules dans la logique $L^2_{\omega_1\omega}$. Ainsi, si $2^{\aleph_0} < 2^{\aleph_1}$, $Car(\aleph_1)$ est faux. L'hypothèse $2^{\aleph_0} < 2^{\aleph_1}$ est consistante, c'est par exemple une conséquence de l'hypothèse du continu, qui est vérifiée dans $\mathbb{L}$, dont on rappelle la définition :

Définition 7. *$\mathbb{L}$ est l'univers constructible de Gödel. Il est obtenu d'une manière similaire à l'univers de Von Neumann : on commence à $\mathbb{L}_0 = \emptyset$. A l'étape successeur, plutôt que de prendre toutes les parties comme dans la hiérarchie de Von Neumann, on ne prend que celles définissables au premier ordre : $\mathbb{L}_{\alpha+1}$ est l'ensemble des parties définissables au premier ordre de $\mathbb{L}_\alpha$. A l'étape limite, on prend la réunion : $\mathbb{L}_\beta = \cup_{\alpha<\beta}\mathbb{L}_\alpha$.*
$\mathbb{L}$ est alors $\cup_\alpha \mathbb{L}_\alpha$.

On obtient ainsi un modèle de ZFC, qui est très étudié, et possède de nombreuses propriétés remarquables. Notamment, il vérifie l'hypothèse généralisée du continu. On a donc :

Théorème 3. *$\neg Car(\aleph_1)$ est consistant.*

Mais l'énoncé $2^{\aleph_0} < 2^{\aleph_1}$ n'est pas prouvable dans ZFC : il existe des modèles où $2^{\aleph_0} = 2^{\aleph_1}$. Il est donc encore envisageable que $Car(\aleph_1)$ soit vrai dans certains modèles de ZFC. Pour montrer la consistance de $Car(\aleph_1)$, nous avons besoin de coder les sous-ensembles de $\aleph_1$ par des sous-ensembles de $\aleph_0$ d'une manière suffisamment explicite, pour que cela puisse être définissable dans la logique $L^2_{\omega_1\omega}$.

La preuve de $Car(\aleph_1)$ va se faire en trois étapes :
- Trouver une structure suffisamment rigide et caractérisable,
- être capable de décrire chaque sous-ensemble de cette structure,
- généraliser à un langage quelconque

C'est le deuxième point qui n'est pas prouvable dans ZFC. Il faudra des hypothèses supplémentaires, dont la consistance peut être prouvée par la méthode du forcing.

Le premier point se manifeste par ce théorème :

Théorème 4. *La structure* $\langle \mathbb{L}_{\aleph_1}, \in \rangle$ *est caractérisable dans* $L^2_{\omega_1\omega}$.

Définition 8. $(\star)$ *est l'assertion que pour tout sous-ensemble* S *de* $\mathbb{L}_{\aleph_1}$*, il existe un énoncé* φ_S *de* $L^2_{\omega_1\omega}$ *tel que pour tout sous-ensemble* S' *de* $\mathbb{L}_{\aleph_1}$, $\langle \mathbb{L}_{\aleph_1}, \in, S' \rangle \models \varphi_S$ *si et seulement si* $S = S'$.

De ce théorème et de $(\star)$, on peut déduire assez facilement $Car(\aleph_1)$. Il suffit donc de trouver des conditions suffisantes pour que $(\star)$ soit vrai.

Pour ce faire, on a besoin du concept combinatoire de famille presque-disjointe :

Définition 9. *Soit* M *un ensemble. Une famille presque-disjointe est un ensemble* $A \subset P(M)$ *tel que pour tout* $x \in A$, $|x| = |M|$ *et pour tout* $x, y \in A$, $x \neq y$*, on a* $|x \cap y| < |M|$.

Définition 10. *Si* A *est une famille presque-disjointe, on dit qu'une partie* B *de* A *est codée par un sous-ensemble* b *de* M *si pour tout* $x \in A$, $x \in B$ *si et seulement si* $|x \cap b| < |M|$.

Ceci nous permet de coder un sous-ensemble de A, qui peut être de taille $> |M|$, par un sous-ensemble de M d'une manière suffisamment explicite.

Dans notre cas, on a besoin d'une famille presque-disjointe de sous-ensemble de $\aleph_0$, de taille $\aleph_1$ et définissable dans $\mathbb{L}_{\aleph_1}$.

Théorème 5. *Si* $(\aleph_1)^{\mathbb{L}} = \aleph_1$*, alors il y a une famille* $(s_i)_{i<\aleph_1}$ *presque-disjointe de sous-ensembles de* $\aleph_0$ *définissable dans* $\mathbb{L}_{\aleph_1}$ *dans la logique* $L^2_{\omega_1\omega}$.

Ici, on ne peut pas se passer de l'hypothèse $(\aleph_1)^{\mathbb{L}} = \aleph_1$, car dans $\mathbb{L}$, il y a $(2^{\aleph_0})^{\mathbb{L}} = (\aleph_1)^{\mathbb{L}}$ sous-ensembles de $\aleph_0$. Ainsi, si l'on souhaite avoir une famille presque-disjointe de taille $\aleph_1$, il faut $(\aleph_1)^{\mathbb{L}} = \aleph_1$.

Théorème 6. *Si tout sous-ensemble de* $(s_i)_{i<\aleph_1}$ *admet un code, alors* $(\star)$ *est vérifiée.*

Démonstration. Soit S une partie de $\mathbb{L}_{\aleph_1}$. Il y a un bon ordre définissable de type $\aleph_1$ sur $\mathbb{L}_{\aleph_1}$, on peut donc supposer que S est une partie de $(s_i)_{i<\aleph_1}$. Ainsi, S admet un code d_S. Soit S' une partie $(s_i)_{i<\aleph_1}$. Alors $S = S'$ si et seulement si

$$\forall i (S'(i) \leftrightarrow |d_S \cap s_i| = \aleph_0)$$

On peut vérifier que cette dernière propriété peut être exprimée dans la logique $L^2_{\omega_1\omega}$, on en déduit le résultat souhaité. □

Tout réside donc dans l'existence de ce codage. C'est là qu'intervient le forcing, et l'axiome de Martin. Tout d'abord, quelques définitions :

Définition 11. *Soit* $\mathbb{P}$ *un ensemble partiellement ordonné. Une partie* D *de* $\mathbb{P}$ *est dite dense si pour tout* $p \in \mathbb{P}$*, il existe* $q \in D$ *tel que* $q \leq p$.
Une partie G *de* $\mathbb{P}$ *est un filtre si pour tout* $p \in G$, $q \in \mathbb{P}$*, si* $p \leq q$ *alors* $q \in G$ *et si pour tout* $p, p' \in G$*, il existe* $p'' \in G$ *tel que* $p'' \leq p$ *et* $p'' \leq p'$.
Si $\mathfrak{D}$ *est un ensemble de parties denses, et* G *un filtre, on dit que* G *est* $\mathfrak{D}$*-générique si pour tout* $D \in \mathfrak{D}$, $D \cap G$ *est non vide.*

Le forcing permet, étant donné un modèle V de ZFC et un ensemble partiellement ordonné $\mathbb{P}$, de construire une extension de V, appelé extension générique, qui contient un filtre G qui intersecte toutes les parties denses de $\mathbb{P}$ qui sont dans V. En général, le complémentaire d'un filtre est une partie dense, donc ce filtre est nécessairement en dehors de V : on a ainsi construit un modèle, noté $V[G]$, qui contient strictement V. Suivant le choix de l'ensemble ordonné $\mathbb{P}$, ce modèle peut avoir des propriétés très différentes des celles de V. C'est l'un des outils les plus prolifiques pour établir des résultats d'indépendance. On peut se demander ce qu'il se passe si on ne demande pas à G d'intersecter tous les ensembles denses, mais seulement certains d'entre eux. Une première réponse est donnée par le lemme de Rasiowa-Sikorski :

Théorème 7. *Si $\mathfrak{D}$ est un ensemble dénombrable de partie denses, alors il existe un filtre $\mathfrak{D}$-générique.*

Ce lemme peut se voir comme une généralisation du lemme de Baire :

Théorème 8. *Soit $(X, \mathfrak{T})$ un espace topologique ayant de bonnes propriétés (typiquement : métrique complet ou localement compact). Alors toute intersection dénombrable d'ouverts dense est non vide.*

Pour voir cette analogie, on associe à X l'ensemble $\mathfrak{T}$ de ses ouverts, ordonné par $\subseteq$. Etant donné un ouvert U dense dans X, on peut considérer $\tilde{U} = \{V \in \mathfrak{T} | V \subseteq U\}$. Alors $\tilde{U}$ est dense dans $(\mathfrak{T}, \subseteq)$, au sens de la définition précédente. Ainsi, si $(U_n)_{n\in\mathbb{N}}$ est une suite d'ouverts denses, le lemme de Rasiowa-Sikorski nous donne un filtre qui intersecte tous les $(\tilde{U}_n)_{n\in\mathbb{N}}$. Etant donné un point x de X, on peut définir le filtre $F_x = \{U \in \mathfrak{T} | x \in U\}$. On aimerait une sorte de réciproque, c'est-à-dire être capable d'obtenir un point de l'espace à partir d'un filtre. Ce sont les conditions imposées à l'espace X qui permettent cela. Ainsi, sous ces conditions, le filtre donné par le lemme de Rasiowa-Sikorski nous donne un point de l'espace qui appartient à tous les $(U_n)_{n\in\mathbb{N}}$, ce qui donne le lemme de Baire.

On peut se demander si on peut améliorer le lemme de Rasiowa-Sikorski.

Définition 12. *Une antichaine dans $\mathbb{P}$ est une partie A de $\mathbb{P}$ telle que, pour tout $x, y \in A$, $x \neq y$, il n'existe pas de $p \in \mathbb{P}$ tel que $p \leq x$ et $p \leq y$. $\mathbb{P}$ a la propriété d'antichaines dénombrables (c.a.c.) si toute antichaine est dénombrable.*

Définition 13. *L'axiome de Martin pour $\aleph_1$, noté $MA(\aleph_1)$, est l'énoncé suivant : Pour tout ensemble partiellement ordonné $\mathbb{P}$ ayant la c.a.c., et pour toute famille $\mathfrak{D}$ d'au plus $\aleph_1$ ensembles denses, il existe un filtre $\mathfrak{D}$-générique.*

On ne peut pas se passer d'une hypothèse restrictive sur les $\mathbb{P}$ qu'on considère. En effet, on veut éviter certains $\mathbb{P}$ pathologiques, dont on peut prouver qu'il n'existe pas nécessairement de filtre suffisamment générique. La consistance de l'axiome de Martin a été prouvé dans [11]. C'est une des premières utilisations du forcing itéré. L'axiome de Martin a de très nombreuses conséquences, pas seulement en théorie des ensembles. On peut trouver dans [1] un nombre important de ces conséquences. L'une d'entre elles va particulièrement nous intéresser : $MA(\aleph_1)$ entraîne l'existence des codes que nous cherchions. Pour prouver cela, il faut trouver un $\mathbb{P}$ convenable :

Définition 14. *Soit A une famille presque-disjointe de parties de $\aleph_0$, et soit B une partie de A. On définit $\mathbb{P}_{A,B} := \{(s,F) | s \subset \aleph_0, F \subset A - B, s \text{ et } F \text{ finis}\}$. On ordonne $\mathbb{P}_{A,B}$ par $(s',F') \leq (s,F)$ si $s \subset s'$, $F \subset F'$ et pour tout $x \in F'$, $x \cap s' \subset s$.*

Définition 15. *Si G est un filtre sur $\mathbb{P}_{A,B}$, on pose $d_G := \bigcup\{s | \exists F, (s,F) \in G\}$*

Théorème 9. *Pour $x \in A - B$, l'ensemble $D_x := \{(s,F) | x \in F\}$ est dense. Pour $y \in B$ et $n \in \omega$, l'ensemble $E_{y,n} := \{(s,F) | s \cap y \not\subseteq n\}$ est dense. Si G intersecte tous les D_x et $E_{y,n}$, alors d_G est un code pour B.*

Dans le cas qui nous intéresse, il y a $\aleph_1$ ensembles denses à considérer. Il est facile de voir que $\mathbb{P}_{A,B}$ vérifie la propriété d'antichaines dénombrables, on peut donc appliquer $MA(\aleph_1)$. Ainsi, en combinant tous les résultats précédents, on obtient :

Théorème 10. *Si $(\aleph_1)^{\mathbb{L}} = \aleph_1$ et si $MA(\aleph_1)$, alors $Car(\aleph_1)$.*

La méthode pour prouver la consistance de $MA(\aleph_1)$ préserve les cardinaux, et s'applique à $\mathbb{L}$, ainsi $(\aleph_1)^{\mathbb{L}} = \aleph_1 + MA(\aleph_1)$ est consistant, d'où

Théorème 11. *$Car(\aleph_1)$ est consistant.*

Pour le cas général de $Car(\lambda^+)$ avec λ régulier, la démonstration est en grande partie identique. Cependant, les ensembles partiellement ordonnés que l'on considère pour obtenir le codage ne vérifient plus la condition d'antichaines dénombrables. La généralisation la plus naturelle de l'axiome de Martin, obtenue en remplaçant "antichaines dénombrables" par "antichaines de taille $< \lambda$ et λ-clôture" est malheureusement inconsistante, résultat prouvé dans [9]. Il est donc nécessaire de trouver des généralisations adéquates de l'axiome de Martin, dont on puisse prouver la consistance et qui permettent d'obtenir les codages que l'on souhaite.

3 Conclusion

L'étude des différentes extensions de la logique du premier ordre amène assez rapidement à des considérations fortement liées à la théorie des ensembles. En effet, certaines propriétés de ces logiques dépendent du contexte ensembliste. Ici, on s'est particulièrement intéressé à la capacité, pour la logique $L^2_{\kappa\omega}$ de caractériser toutes les structures de cardinal κ. Si κ est le successeur d'un cardinal régulier, alors cette propriété est indépendante de ZFC. Pour démontrer ceci, on utilise différentes techniques et outils de théorie des ensembles, notamment les propriétés très fortes de l'univers des constructibles de Gödel, ainsi que l'axiome de Martin et des généralisations adéquates, dont la consistance nécessite la technique du forcing itéré.

Cependant, les techniques utilisées ici échouent complètement pour traiter les cas où κ n'est pas successeur d'un cardinal régulier.

Références

[1] D. H. Fremlin. *Consequences of Martin's Axiom.* Cambridge University Press, 1984. Cambridge Books Online.

[2] T. Hyttinen, K. Kangas, and J. Väänänen. On second-order characterizability. *Logic Journal of the IGPL*, 21(5) :767–787, 2013.

[3] T. Jech. *Set Theory.* Springer, 2003.

[4] L. Keskinen. Characterizing all models in infinite cardinalities. *ILLC Dissertation Series*, 2011-05.

[5] K. Kunen. *Set Theory.* Studies in Logic. College Publications, 2011.

[6] M. Magidor. On the set theory of generalized logics. *Freiburg Conference on Set Theory*, 2014.

[7] D. Marker. *Model Theory : An Introduction.* Graduate Texts in Mathematics. Springer New York, 2006.

[8] S. Shelah. A weak generalization of MA to higher cardinals. *Israel Journal of Mathematics*, 30(4) :297–306, 1978.

[9] S. Shelah and L. Stanley. Generalized Martin's axiom and Souslin's hypothesis for higher cardinals. *Israel Journal of Mathematics*, 43(3) :225–236, 1982.

[10] S. Shelah and J. Väänänen. Stationary Sets and Infinitary Logic. *Journal of Symbolic Logic*, 65(3) :1311–1320, 2000.

[11] R. M. Solovay and S. Tennenbaum. Iterated Cohen extensions and Souslin's problem. *Annals of Mathematics*, 94(2) :pp. 201–245, 1971.