

Le nec plus ultra du théorème de compacité*

Pierre Cagne

Dernière mise à jour le 26 octobre 2015

Résumé

Le théorème de compacité de la logique du premier ordre énonce qu'une théorie admet un modèle dès que c'est le cas de toutes ses sous-théories finies. Étonnamment, la preuve *classique* de ce résultat passe par le célèbre théorème de complétude de Gödel. On prend ici le contrepied et présentons une preuve restant dans le cadre douillet de la sémantique et s'abstenant ainsi de plonger dans les méandres combinatoires de la théorie de la prouvabilité.

Ce document est soumis à une licence Creative Commons Attribution Share Alike (CC-BY-SA 4.0 International). Texte complet disponible à l'adresse suivante :

<http://creativecommons.org/licenses/by-sa/4.0/>

1 Liminaires

La théorie des modèles ne s'inquiète pas de méta-mathématiques. Elle n'est pas intéressée par ce qu'est ou devrait être une preuve et elle ne raisonne pas sur le raisonnement. Sa raison d'être est de *faire des mathématiques* : elle cherche des méthodes génériques pour raisonner sur les structures que les mathématiciens manipulent au quotidien. Dans cette optique, le slogan de ce document est :

« En théorie des modèles, la syntaxe est au service de la sémantique. »

Le théorème central de la théorie des modèles est le théorème de compacité, qui peut s'énoncer informellement comme suit : un ensemble d'axiomes T est vérifié par au moins une structure mathématique dès que c'est déjà le cas de tous les sous-ensembles finis de T . La preuve la plus classique de ce résultat (et aussi la première historiquement) passe par le théorème de complétude de Gödel. Si l'on en connaît souvent le nom, il reste parfois mal compris. Ce théorème est un pont entre le monde sémantique (la théorie des modèles) et le monde syntaxique (la théorie de la preuve) : il énonce que la tentative de formalisation du concept de preuve par la syntaxe et la déduction naturelle est une réussite. En un sens, il dit que la logique fondamentale peut s'abstraire des modèles (là où les mathématiques ont lieu) sans pour autant obtenir des résultats déraisonnables. Il est ainsi étonnant que la preuve du théorème de compacité, énoncé sémantique par excellence, doive passer par la théorie même qui élimine la sémantique ! On se propose donc ici de donner une preuve du théorème de compacité qui ne quitte pas le monde tellement

*Ce titre est tiré sans vergogne de celui d'une présentation d'Adrien Deloro. Son mérite va bien au-delà de ses traits d'esprit et je le remercie de m'avoir ouvert à ce monde passionnant qu'est la théorie des modèles.

plus simple de la théorie des modèles. Ce sera en particulier l'occasion de découvrir le théorème de Łoś.

On commence par rappeler les notions fondamentales de la logique du premier ordre et de la théorie des modèles afin de donner un sens précis à ce qui a été dit plus haut. Un langage du premier ordre \mathcal{L} est la donnée d'un ensemble $C_{\mathcal{L}}$ et de deux ensembles gradués $F_{\mathcal{L}} = \coprod_{n \geq 0} F_{\mathcal{L}}^n$ et $R_{\mathcal{L}} = \coprod_{n \geq 0} R_{\mathcal{L}}^n$. Les éléments de $C_{\mathcal{L}}$ sont appelés *symboles de constante* de \mathcal{L} , ceux de $F_{\mathcal{L}}^n$ *symboles de fonction n-aire* et ceux de $R_{\mathcal{L}}^n$ *symboles de relation n-aire*. On supposera toujours avoir dans $R_{\mathcal{L}}^2$ un symbole dit d'égalité $\doteq_{\mathcal{L}}$ (souvent abrégé en \doteq) On dispose également d'un ensemble infini fixé V dont les éléments sont appelés *variables*. On peut maintenant définir la syntaxe de la logique du premier ordre.

Définition 1.1 (Termes). Un *terme* sur \mathcal{L} est défini inductivement comme :

- (i) une variable x de V ,
- (ii) une constante c du langage \mathcal{L}
- (iii) $f(t_1, \dots, t_n)$ pour $f \in F_{\mathcal{L}}^n$ et t_i des termes.

Définition 1.2 (Formules). Une *formule* sur \mathcal{L} est définie inductivement comme :

- (i) $R(t_1, \dots, t_n)$ pour $R \in R_{\mathcal{L}}^n$ et t_i des termes,
- (ii) $\neg \varphi$ pour φ une formule,
- (iii) $\varphi \wedge \psi$ pour φ, ψ des formules,
- (iv) $\exists x, \varphi$ pour $x \in V$ et φ une formule.

Les formules de la forme (i) sont dites *atomiques*.

On met en place les raccourcis syntaxiques habituels : $\forall x, \varphi$ pour $\neg(\exists x, \neg \varphi)$, $\varphi \vee \psi$ pour $\neg(\neg \varphi \wedge \neg \psi)$, $\varphi \rightarrow \psi$ pour $\neg(\varphi \wedge \neg \psi)$, et $\varphi \leftrightarrow \psi$ pour $(\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$. C'est tout ce que l'on dira de la syntaxe dans cet exposé : pour être complet, il faudrait parler de substitutions et autres joyeusetés barbantes. Pour notre part, on laissera le lecteur se reposer sur son intuition et sur le sens qu'il met déjà derrière ces formules pour éviter un jeu syntaxique pénible. Par exemple, on définit la notion de *variable liée* d'une formule φ comme étant une variable apparaissant dans φ sous un quantificateur. On définit une *variable libre* comme une variable non liée. Le lecteur comprendra alors aisément que dans

$$(x = y) \rightarrow (\exists x, x = y)$$

la première occurrence de x est libre et la seconde est liée, alors que les deux occurrences de y sont libres. De même, le lecteur identifiera naturellement cette formule avec la suivante

$$(x = y) \rightarrow (\exists z, z = y)$$

car le *sens* que l'on met derrière ces formules est le même.

Sortons précipitamment du monde tortueux où les interrogations précédentes allaient nous plonger. Si ces considérations sont certes nécessaires pour fonder rigoureusement la logique, elle ne le sont point pour comprendre ce qui suit. Le théoricien des modèles lui-même, après s'être assuré que ce qu'il fait est fondé, oublie bien vite le formalisme poussif engendré par les définitions précédentes. En particulier, il note $t(x_1, \dots, x_n)$ pour signifier que les variables apparaissant dans le terme t

sont parmi les x_i ; de même, il note $\varphi(x_1, \dots, x_n)$ si les variables **libres** de la formule φ se trouvent parmi les x_i . Enfin, il s'abstient de dire *symboles de constante* (resp. fonction, resp. relation) puisqu'il pense déjà à de vraies constantes (resp. fonctions, resp. relations).

On en arrive enfin à l'objet central de la théorie des modèles, c'est-à-dire... le modèle ! Pour définir cette notion proprement, il faut déjà dégager celle de structure, dans laquelle on interprètera la logique du premier ordre.

Définition 1.3 (Structures). Une *structure* M sur le langage \mathcal{L} est la donnée d'un ensemble **non vide**, encore noté M , et

- (i) d'un élément $c^M \in M$ pour toute constante c de \mathcal{L} ,
- (ii) d'une fonction $f^M : M^n \rightarrow M$ pour chaque fonction f d'arité n de \mathcal{L} ,
- (iii) d'une relation $R^M \subseteq M^n$ pour toute relation R d'arité n de \mathcal{L} .

On impose de plus que $\overset{M}{=}^{\mathcal{L}}$ soit la diagonale $\{(m, m) : m \in M\}$.

On dit que s^M interprète le symbole s de \mathcal{L} dans M . Cette interprétation des symboles induit par induction une interprétation $t^M(a_1, \dots, a_n)$ dans M d'un terme $t(x_1, \dots, x_n)$ de \mathcal{L} par les éléments $a_1, \dots, a_n \in M$:

- $x^M(a) = a$ pour $x \in \mathcal{V}$ et $a \in M$,
- $(f(t_1, \dots, t_n))^M = f^M(t_1^M, \dots, t_n^M)$ pour f fonction n -aire et t_i des termes.

On peut alors interpréter toute la logique sur \mathcal{L} dans une structure M .

Définition 1.4 (Satisfaction). La *satisfaction* d'une formule $\varphi(x_1, \dots, x_n)$ dans M par des éléments (a_1, \dots, a_n) , dénotée $M \models \varphi(a_1, \dots, a_n)$, est définie inductivement par :

- si $\varphi = R(t_1, \dots, t_n)$ est atomique, $M \models \varphi(a_1, \dots, a_n)$ si et seulement si $(t_1^M, \dots, t_n^M) \in R^M$,
- si $\varphi = \neg\psi$, $M \models \varphi(a_1, \dots, a_n)$ si et seulement si $M \not\models \psi(a_1, \dots, a_n)$,
- si $\varphi = \psi \wedge \rho$, $M \models \varphi(a_1, \dots, a_n)$ si et seulement si $M \models \psi(a_1, \dots, a_n)$ et $M \models \rho(a_1, \dots, a_n)$
- si $\varphi = \exists x, \psi(x, x_1, \dots, x_n)$, $M \models \varphi(a_1, \dots, a_n)$ si et seulement s'il existe $a \in M$ tel que $M \models \psi(a, a_1, \dots, a_n)$.

Bref, on rend précis le sens qu'on donnait déjà aux formules. On remarque en particulier que la satisfaction d'une formule φ **sans variables libres** ne dépend d'aucun élément de M mais juste de l'interprétation du langage \mathcal{L} dans M . Ces formules sans variables libres seront appelées *énoncés*.

Une *théorie* T sur \mathcal{L} est simplement un ensemble d'énoncé.

Définition 1.5 (Modèles). Un *modèle* M d'une théorie T sur \mathcal{L} est une structure M telle que $M \models \varphi$ pour tout énoncé $\varphi \in T$.

On notera $T \models \varphi$ pour signifier que tout modèle de T satisfait φ . On dit parfois que φ est *conséquence sémantique* de T . Remarquons cependant qu'un modèle de T satisfait toute conséquence sémantique de T (et réciproquement de manière triviale). Suivant une fois de plus notre slogan initial, les éléments syntaxiques de T ne nous intéressent que par ce qu'ils disent sémantiquement dans les modèles. Ainsi, on identifiera souvent une théorie T et l'ensemble $\{\varphi : T \models \varphi\}$ de ses conséquences sémantiques.

Remarque. Il ne faut pas s'inquiéter outre mesure de l'identification d'une théorie avec ses conséquences sémantiques. En effet, c'est le fameux débat qu'il peut y avoir dans une classe de licence ou de prépa sur la question d'inclure ou non l'axiome « $\forall x, 0 \cdot x = 0$ » parmi les axiomes d'anneaux, puisque cette propriété des anneaux peut-être déduite des autres axiomes. En fait, peu nous chaut, puisque tous les anneaux (définis avec ou sans cet axiome supplémentaire) auront cette propriété.

Une théorie T est *satisfiable* si elle admet au moins un modèle. Elle est *finiment satisfiable* si toute partie finie $T_f \subseteq T$ est satisfiable. Si T et T' sont deux théories telles que $T' \subseteq T$, tout modèle de T est évidemment un modèle de T' . Ainsi, toute théorie satisfiable est finiment satisfiable. Le théorème de compacité nous donne la réciproque.

Théorème 1.6 (Compacité de la logique du premier ordre). *Une théorie est satisfiable si et seulement si elle est finiment satisfiable.*

2 Ultrafiltres et ultraproducts

Sortons quelques instants de la logique du premier ordre et de la théorie des modèles.

Définition 2.1 (Filtres). Un *filtre* sur X est un ensemble $\mathcal{F} \subseteq \mathcal{P}(X)$ de parties de X respectant les trois propriétés suivantes :

- (i) $\emptyset \notin \mathcal{F}$,
- (ii) si $A \subseteq B$ et $A \in \mathcal{F}$, alors $B \in \mathcal{F}$ également,
- (iii) si $A, B \in \mathcal{F}$, alors $A \cap B \in \mathcal{F}$.

Un *ultrafiltre* est un filtre maximal pour l'inclusion.

Les exemples de filtres pullulent. Par exemple, l'ensemble des voisinages d'un point x dans un espace topologique X est un filtre sur X . On peut également exhiber facilement des ultrafiltres : pour chaque élément x d'un ensemble X , on a un ultrafiltre, dit *principal en x* ,

$$\delta_x = \{Y \subseteq X : x \in Y\}.$$

En revanche, ce sont les seuls ultrafiltres explicites ! Pour être plus précis, l'existence d'autres ultrafiltres repose sur l'axiome du choix. Qu'à cela ne tienne, l'axiome du choix nous va très bien¹. Le lemme suivant nous assure qu'on en a alors pléthore.

Lemme 2.2. *Tout filtre est inclus dans un ultrafiltre.*

Preuve. Il suffit d'appliquer le lemme de Zorn à l'ensemble des filtres sur X . □

On donne maintenant deux lemmes, utiles par la suite. Le premier explique comment construire un filtre à partir d'une famille d'ensemble ayant de bonnes propriétés. Le second donne un critère pour reconnaître les filtres parmi les ultrafiltres.

1. On pourrait croire qu'ici la démonstration sémantique demande une logique plus forte que la démonstration classique. Ce n'est pas le cas, car la démonstration du théorème de complétude de Gödel repose également sur une forme d'axiome du choix.

Lemme 2.3. Soit $\mathcal{B} \subseteq \mathcal{P}(X)$ une famille de X ayant la propriété des intersections finies, i.e. $\bigcap_{i=1}^n B_i \neq \emptyset$ pour tous $B_1, \dots, B_n \in \mathcal{B}$. Alors, il existe un filtre contenant \mathcal{B} .

Preuve. On vérifie facilement que

$$\mathcal{F} = \{Y \subseteq X : \exists B_1, \dots, B_n \in \mathcal{B}, \bigcap_{i=1}^n B_i \subseteq Y\}$$

est le filtre cherché. □

En particulier, comme tout filtre est contenu dans un ultrafiltre, toute famille ayant la propriété des intersections finies aussi.

Lemme 2.4. Un filtre \mathcal{U} sur X est un ultrafiltre si et seulement si pour tout $A \subseteq X$, $A \in \mathcal{U}$ ou $X \setminus A \in \mathcal{U}$.

Preuve. Soit \mathcal{U} un ultrafiltre sur X . Si $B \subseteq X$ vérifie que $B \cap Y \neq \emptyset$ pour tout $Y \in \mathcal{U}$, alors $U \cup \{B\}$ admet la propriété des intersections finies et est ainsi incluse dans un filtre : la maximalité de \mathcal{U} assure alors que $B \in \mathcal{U}$. Supposons maintenant qu'il existe $A \subseteq X$ tel que ni A ni $X \setminus A$ ne soit dans \mathcal{U} : alors par la propriété précédente, il existe $Y_1, Y_2 \in \mathcal{U}$ tel que $A \cap Y_1 = \emptyset = Y_2 \setminus A$. En posant $Y = Y_1 \cap Y_2 \in \mathcal{U}$, on a encore $A \cap Y = \emptyset = Y \setminus A$. On a alors la contradiction suivante :

$$\mathcal{U} \ni Y = (A \cap Y) \cup (Y \setminus A) = \emptyset.$$

Réciproquement, on suppose que le filtre \mathcal{U} contient toute partie ou son complémentaire. Soient \mathcal{V} un filtre contenant \mathcal{U} et $A \in \mathcal{V}$: alors $X \setminus A \notin \mathcal{U}$ (sinon $\emptyset = A \cap (X \setminus A) \in \mathcal{V}$), donc $A \in \mathcal{U}$. Autrement dit, \mathcal{U} est maximal. □

Passons maintenant aux ultraproducts et revenons dans le cadre de la théorie des modèles. Si l'on dispose d'une famille d'ensembles $(X_i)_{i \in I}$ indexée par I et d'un ultrafiltre \mathcal{U} sur I , alors on a une relation d'équivalence naturelle sur $\prod_{i \in I} X_i$:

$$(m_i) \sim (n_i) \quad \text{si et seulement si} \quad \{i \in I : m_i = n_i\} \in \mathcal{U}.$$

On notera $[(m_i)]$ la classe modulo \sim de (m_i) . Un parallèle intéressant à tacer ici est avec la théorie de la mesure : il faut penser les éléments $A \subseteq X$ qui ne sont pas dans le filtre \mathcal{U} comme des ensembles de mesure nulle. La relation d'équivalence ci-dessus se lit alors comme suit : on identifie deux vecteurs quand ils sont égaux presque partout.

Définition 2.5 (Ultraproduits). L'ultraproduit des X_i par l'ultrafiltre \mathcal{U} , noté $\prod_{i \in I} X_i / \mathcal{U}$, est le quotient de $\prod_{i \in I} X_i$ par la relation d'équivalence \sim .

La pertinence des ultraproducts en théorie des modèles vient de la construction suivante. Supposons qu'on ait une famille $(M_i)_{i \in I}$ de structures, toutes sur un même langage \mathcal{L} . Alors l'ultraproduit $M^* = \prod_{i \in I} M_i / \mathcal{U}$ est naturellement une structure sur \mathcal{L} . Il suffit pour cela d'interpréter :

- une constante c par $c^{M^*} = [(c^{M_i})]$,
- une fonction n -aire f par

$$f^{M^*} : ([(m_i^1)], \dots, [(m_i^n)]) \mapsto [(f^{M_i}(m_i^1, \dots, m_i^n))]$$

— une relation n -aire R par

$$([(m_i^1)], \dots, [(m_i^n)]) \in R^{M^*} \iff \{i \in I : (m_i^1, \dots, m_i^n) \in R^{M_i}\} \in \mathcal{U}$$

Il faut bien entendu vérifier que ces interprétations sont bien définies, i.e. que nos formules ne dépendent pas des représentants choisis. On laisse cet exercice fastidieux au lecteur motivé. Les autres types de lecteur se contenteront de se convaincre que quelqu'un a déjà vérifié tout cela. Le plus simple est de se rappeler que c'est l'interprétation « presque partout » : un élément est l'interprétation d'une constante si c'est en presque toutes coordonnées l'interprétation d'une constante ; un élément est l'image par l'interprétation d'une fonction d'un n -uplet si c'est le cas en presque toutes coordonnées ; enfin des éléments sont en relation dans l'ultraproduit s'ils le sont en presque toute coordonnées.

3 Preuve (ultrasémantique) du théorème de compacité

On est maintenant en mesure de démontrer le résultat central de ce document : le théorème de Łoś. Ce théorème dit ni plus ni moins qu'une formule est vraie dans un ultraproduit si et seulement si elle est vraie en presque toutes coordonnées. On en déduira certes le théorème de compacité mais c'est un théorème autrement puissant, qui permet de démontrer des théorèmes non triviaux en dehors du domaine de la théorie des modèles (par exemple, le théorème d'Ax). Il permet également de distinguer les parvenus, qui écorcheront le nom du mathématicien polonais en le prononçant « l'os », des aficionados de la théorie des modèles, qui sauront lui donner sa sonorité approximativement exacte « ouoch ».

On dispose toujours d'une famille (M_i) de structures sur le langage \mathcal{L} indexée par I et d'un ultrafiltre \mathcal{U} sur I . On continue de noter M^* la structure sur \mathcal{L} sus-décrite.

Théorème 3.1 (Łoś). Soient $\varphi(x_1, \dots, x_n)$ une formule sur \mathcal{L} et $m^* =([(m_i^1)], \dots, [(m_i^n)])$ un élément de $(M^*)^n$. Alors

$$M^* \models \varphi(m^*) \iff \{i \in I : M_i \models \varphi(m_i^1, \dots, m_i^n)\} \in \mathcal{U}$$

Preuve. La preuve se fait bien entendu par induction sur φ . Pour plus de simplicité, on note $I_\chi = \{i \in I : M_i \models \chi\}$ pour une formule χ quelconque.

- Si φ est atomique, c'est simplement la définition de l'interprétation des relations de \mathcal{L} dans M^* .
- Si $\varphi = \neg\psi$, alors $I_\varphi = I \setminus I_\psi$ et, par propriété des ultrafiltres, $I_\varphi \in \mathcal{U}$ si et seulement si $I_\psi \notin \mathcal{U}$. Par hypothèse d'induction, $I_\psi \in \mathcal{U}$ si et seulement si $M^* \models \psi(m^*)$, c'est-à-dire si et seulement si $M^* \not\models \varphi(m^*)$.
- Si $\varphi = \psi \wedge \rho$, alors $I_\varphi = I_\psi \cap I_\rho$, et par propriété des filtres, $I_\varphi \in \mathcal{U}$ si et seulement si $I_\psi, I_\rho \in \mathcal{U}$. Par hypothèse d'induction, $I_\psi \in \mathcal{U}$ si et seulement si $M^* \models \psi(m^*)$ et $M^* \models \rho(m^*)$, c'est-à-dire si et seulement si $M^* \models \varphi(m^*)$.
- Enfin supposons $\varphi = \exists x, \psi(x, x_1, \dots, x_n)$. Si $M^* \models \varphi(m^*)$, c'est qu'il existe $n^* = [(n_i)]$ tel que $M^* \models \psi(n^*, m^*)$. Par hypothèse d'induction, on a alors $J = \{i \in I : M_i \models \psi(n_i, m_i^1, \dots, m_i^n)\} \in \mathcal{U}$. Le sur-ensemble $I_\varphi \supseteq J$ est alors également dans \mathcal{U} .

Réciproquement, profitons de l'existence de $n_i \in M_i$ tel que $M_i \models \psi(n_i, m_i^1, \dots, m_i^n)$ pour chacun des $i \in I_\varphi$ pour construire $n^* = [(n_i)]$ en complétant arbitrairement pour les $i \notin I_\varphi$. Alors comme $I_\varphi \in \mathcal{U}$, cela décide dans M^* du fait que $\psi(n^*, m^*)$ est satisfaite. Autrement dit, on a créé un témoin de l'affirmation $M^* \models \varphi(m^*)$.

□

Le théorème de compacité arrive alors comme un corollaire fort simple du théorème de Łoś.

Théorème 3.2. *Toute théorie finiment satisfiable est satisfiable.*

Preuve. Soit T une théorie et posons I l'ensemble des ses parties finies. L'hypothèse du théorème est qu'il existe un modèle $M_i \models i$ pour tout $i \in I$. Pour tout énoncé φ sur \mathcal{L} , on peut alors considérer l'ensemble $I_\varphi = \{i \in I : M_i \models \varphi\}$. C'est là que tout se joue : $\{I_\varphi : \varphi \in T\}$ admet la propriété des intersections finies. En effet, pour $\varphi_1, \dots, \varphi_n \in T$, $\bigcap_{j=1}^n I_{\varphi_j} = I_{\varphi_1 \wedge \dots \wedge \varphi_n}$ qui est non vide puisqu'il contient au moins le singleton $\{\varphi_1 \wedge \dots \wedge \varphi_n\}$. Il existe donc un ultrafiltre \mathcal{U} sur I contenant la famille $\{I_\varphi : \varphi \in T\}$. L'ultraproduit M^* créé par l'ultrafiltre \mathcal{U} est un modèle de T : si $\varphi \in T$, alors $\{i \in I : M_i \models \varphi\} = I_\varphi \in \mathcal{U}$, et le théorème de Łoś conclut que $M^* \models \varphi$. □