

9.6 Preuve d'incomplétude des théories contenant l'arithmétique élémentaire

Nous savons déjà que Q (ainsi que l'arithmétique de Peano) sont incohérentes ou incomplètes (corollaire 9.5.1). Mais cette preuve repose sur le lemme 9.3.1, dont la preuve est elle-même effectuée dans une meta-théorie, puisque nous commençons par dire : “ou bien P est récursif, ou bien P n'est pas récursif”. Mais “ P est récursif” n'est pas un énoncé que nous pouvons coder dans l'arithmétique de Peano.

L'objectif est donc ici de donner une preuve d'incomplétude plus élémentaire, qui peut être elle-même codée dans l'arithmétique de Peano. Ceci permettrait de montrer le deuxième théorème d'incomplétude : l'arithmétique de Peano permet d'énoncer sa propre cohérence, mais pas de la prouver (à moins d'être incohérente).

Pour cette partie, on utilise un codage des preuves dans les entiers. Par exemple, on peut utiliser les deux symboles supplémentaires de l'alphabet, que nous avons pris la précaution d'inclure : $\{, \}$, qui permettent de coder les arbres de preuves dans des chaînes de caractères sur un alphabet fixe et donc dans les entiers. Si ϕ est une formule prouvable, en utilisant une preuve $\Pi(\phi)$, on note $\langle \Pi(\phi) \rangle$ l'entier qui code cette preuve.

On supposera dans la suite que l'ensemble A d'axiomes est récursif (ce qui est le cas de l'arithmétique élémentaire et de l'arithmétique de Peano) et que l'ensemble des règles d'inférence est récursif. C'est le cas d'un système de règles d'inférence (sémantiquement) complet pour la logique du premier ordre.

Exercice 195

Expliquer pourquoi, dans ce cas, l'ensemble $\{(n, m) \in \mathbb{N}^2 \mid \exists \phi, n = \langle \phi \rangle \ \& \ m = \langle \Pi(\phi) \rangle\}$ est récursif.

Dans un premier temps, on supposera que T est ω cohérente.

Definition 9.6.1 Une théorie est ω -cohérente si, pour tout $n \in \mathbb{N}$, $T \vdash \phi(\bar{n})$, alors $T \not\vdash \exists x. \neg \phi(x)$.

Exercice 196 (5)

1. Montrer que si T est ω -cohérente, alors T est cohérente.
2. Donner un exemple de théorie qui est cohérente et pas ω -cohérente
3. Existe-t-il une théorie qui satisfait les propriétés de la figure 9.2 et qui est cohérente sans être ω -cohérente ?

Exercice 197 (5)

Montrer que, si T est ω cohérente et vérifie les propriétés (A_+) , (A_{f_+}) , (A_\times) , (A_{f_\times}) , (A_-) , alors elle est incomplète.

Le lemme suivant est l'un des énoncés originaux dûs à Gödel ; la formule Ω est en fait une formule qui exprime sa propre non prouvabilité et donc, intuitivement, ne peut ni être prouvée, ni sa négation être prouvée, sous peine d'incohérence. La preuve, comme celle de Rosser qui suivra, est basée exactement sur la même idée que la preuve du lemme 9.3.1.

Lemme 9.6.1 *Si les fonctions récursives sont représentables dans T et si T est ω -cohérente, alors il existe une formule Ω telle que $T \not\vdash \Omega$ et $T \not\vdash \neg\Omega$.*

Preuve:

Soit $P \stackrel{\text{def}}{=} \{(n, m) \in \mathbb{N}^2 \mid \exists \phi(x). n = \langle \phi(x) \rangle \ \& \ m = \langle \Pi(\phi(\bar{n})) \rangle\}$ P est l'ensemble des paires d'entiers tels que le premier code une formule avec une variable libre et le deuxième est une preuve de la formule appliquée à son propre code.

P est un ensemble récursif : on peut vérifier que n est le code d'une formule $\phi(x)$ et on peut calculer $\phi(\bar{n})$ et vérifier que le second est le code d'une preuve de $\phi(\bar{n})$.

Par hypothèse P est donc représentable par une formule $\phi_P(x, y)$. (**Note** : si nous avons été plus explicites dans le codage des formules et des preuves, nous pourrions donner explicitement ϕ_P). Soit alors $\psi(x) = \forall y. \neg \phi_P(x, y)$ et $\Omega = \psi(\langle \overline{\psi} \rangle)$.

Si $T \vdash \Omega$, alors par définition de P , il existe un entier m tel que $(\langle \overline{\psi} \rangle, m) \in P$. Donc $T \vdash \phi_P(\langle \overline{\psi} \rangle, \bar{m})$ et donc $T \vdash \exists y. \phi_P(\bar{\psi}, y)$, soit $T \vdash \neg \psi(\langle \overline{\psi} \rangle)$, c'est-à-dire $T \vdash \neg \Omega$, ce qui contredit la cohérence de la théorie.

Donc $T \not\vdash \Omega$.

On a besoin de la ω -cohérence pour la réciproque : si $T \vdash \neg \psi(\langle \overline{\psi} \rangle)$, alors $T \vdash \exists y. \phi_P(\langle \overline{\psi} \rangle, y)$ par définition de ψ .

Par ω -cohérence, il existe un entier $n \in \mathbb{N}$ tel que $T \not\vdash \neg \phi_P(\langle \overline{\psi} \rangle, \bar{n})$

Alors, ou bien $T \vdash \phi_P(\langle \overline{\psi} \rangle, \bar{n})$, ou bien la théorie T est incomplète : il suffit de choisir $\Omega' = \phi_P(\langle \overline{\psi} \rangle, \bar{n})$.

Dans le premier cas, on a par ailleurs $T \vdash \psi(\langle \overline{\psi} \rangle)$ par définition de P . Et, en conclusion $T \vdash \Omega$ ssi $T \vdash \neg \Omega$. Si la théorie est cohérente, $T \not\vdash \Omega$ et $T \not\vdash \neg \Omega$.

Cette preuve n'est pas satisfaisante pour deux raisons. D'abord, elle fait l'hypothèse de ω -cohérence, qui est (a priori) plus forte que la cohérence. Ensuite, elle n'est pas totalement constructive puisqu'elle laisse le choix entre plusieurs énoncés Ω qui ne sont pas prouvables et dont la négation n'est pas prouvable.

L'étape suivante est une amélioration de ce résultat (Rosser 1936) dans laquelle on n'utilise plus l' ω -cohérence, mais seulement la cohérence.

Théorème 9.6.1 *Si T est cohérente et satisfait les propriétés de la figure 9.2, alors on peut construire une formule Ω telle que $T \not\vdash \Omega$ et $T \not\vdash \neg \Omega$.*

Preuve:

On définit les deux prédicats suivant :

$$P \stackrel{\text{def}}{=} \{(n, m) \in \mathbb{N}^2 \mid \exists \phi(x). n = \langle \phi(x) \rangle \ \& \ m = \langle \Pi(\neg \phi(\bar{n})) \rangle\}$$

$$Q \stackrel{\text{def}}{=} \{(n, m) \in \mathbb{N}^2 \mid \exists \phi(x). n = \langle \phi(x) \rangle \ \& \ m = \langle \Pi(\phi(\bar{n})) \rangle\}$$

P et Q sont deux prédicats récursifs, d'après l'exercice 195 et les propriétés de clôture de l'ensemble des fonctions récursives. D'après le lemme 9.5.1, nos hypothèses entraînent la représentabilité dans T des fonctions récursives. Donc

9.6. PREUVE D'INCOMPLÉTUDE DES THÉORIES CONTENANT L'ARITHMÉTIQUE ÉLÉMENTAIRE

en particulier de P, Q : il existe des formules ϕ_P et ϕ_Q qui représentent respectivement P et Q . (**Note** : Si nous avons été plus explicites sur le codage des preuves et des formules, nous pourrions donner explicitement ces formules au lieu de donner une définition ensembliste de P, Q). On définit alors :

$$\theta(x) \stackrel{\text{def}}{=} \forall y. (\phi_Q(x, y) \rightarrow \exists z. z < y \wedge \phi_P(x, z))$$

Et $\Omega \stackrel{\text{def}}{=} \theta(\overline{\langle \theta \rangle})$. Intuitivement, Ω exprime que, si elle est prouvable, il existe une preuve plus petite de sa négation.

Si $T \vdash \Omega$ alors il existe $m \in \mathbb{N}$ tel que $(\langle \theta \rangle, m) \in Q$, par définition de Q (il existe une preuve de θ appliquée à son propre code). Donc

$$T \vdash \phi_Q(\overline{\langle \theta \rangle}, \overline{m}) \quad \text{Par définissabilité de } Q$$

$$T \vdash \exists z. z < \overline{m} \wedge \phi_P(\overline{\langle \theta \rangle}, z) \quad \text{Par définition de } \theta$$

$$T \vdash \phi_P(\overline{\langle \theta \rangle}, \overline{0}) \vee \dots \vee \phi_P(\overline{\langle \theta \rangle}, \overline{m-1}) \quad \text{En utilisant } (A_{\leq})$$

Par représentabilité de P , il existe donc $e \in \mathbb{N}$, $e < m$ tel que $(\langle \theta \rangle, e) \in P$, et donc

$$T \vdash \neg\theta(\overline{\langle \theta \rangle}) \quad \text{Par définition de } P$$

Soit $T \vdash \neg\Omega$

Si $T \vdash \neg\Omega$ soit $e = \langle \Pi(\neg\Omega) \rangle$. Par définition de P , $(\langle \theta \rangle, e) \in P$. Par représentabilité de P ,

$$T \vdash \phi_P(\overline{\langle \theta \rangle}, \overline{e})$$

$$T \vdash \forall x. (\overline{e} < x \rightarrow \exists z. z < x \wedge \phi_P(\overline{\langle \theta \rangle}, z)) \quad \text{Par complétude sémantique}$$

$$(4) \quad T \vdash \forall x. (x \leq \overline{e} \vee \exists z. z < x \wedge \phi_P(\overline{\langle \theta \rangle}, z)) \quad \text{Par } (A_{<>})$$

Par ailleurs, par cohérence, $T \not\vdash \theta(\overline{\langle \theta \rangle})$, donc, pour tout $m \in \mathbb{N}$, $(\langle \theta \rangle, m) \notin Q$:

$$T \vdash \neg\phi_Q(\overline{\langle \theta \rangle}, \overline{m}) \quad \text{Pour tout } m \in \mathbb{N}, \text{ par représentabilité de } Q$$

En particulier pour $m \leq e$:

$$T \vdash \forall x. (x \leq \overline{e} \rightarrow \neg\phi_Q(\overline{\langle \theta \rangle}, x)) \quad \text{En utilisant } (A_{\leq})$$

$$T \vdash \forall x. (\neg\phi_Q(\overline{\langle \theta \rangle}, x) \vee (\exists z. z < x \wedge \phi_P(\overline{\langle \theta \rangle}, z))) \quad \text{En utilisant } (4)$$

Mais cette dernière formule est exactement θ :

$$T \vdash \theta(\overline{\langle \theta \rangle})$$

d'où la contradiction.

Dans la preuve, nous avons utilisé, pour la bonne compréhension, des raisonnements sur les entiers et les ensembles d'entiers, mais il faut noter que tous ces raisonnements peuvent être formalisés dans PA .

Exercice 198 (6)

Φ est l'ensemble des formules du premier ordre sans variable libre sur les prédicats $\{=\}$ et les symboles de fonction $\mathcal{F} = \{0(0), S(1), +(2), \times(2)\}$. Si Π est une preuve de $Q \vdash \phi$, on note $\langle \Pi(\phi) \rangle \in \mathbb{N}$ son code. \mathbb{N} est le modèle de Q dans lequel les fonctions ont leur interprétation habituelle.

Soit $D = \{(n, m) \in \mathbb{N}^2 \mid \exists \phi \in \Phi, n = \langle \phi \rangle \ \& \ m = \langle \Pi(\phi) \rangle\}$ et $\phi_D(x, y)$ une formule qui représente D dans Q . (En supposant Q cohérente), parmi les affirmations suivantes, lesquelles sont vraies pour tous les énoncés $\phi \in \Phi$? Justifier.

1. $\mathbb{N} \models \phi \rightarrow (\exists x. \phi_D(\overline{\langle \phi \rangle}, x))$
2. $Q \vdash \phi \rightarrow (\exists x. \phi_D(\overline{\langle \phi \rangle}, x))$
3. $\mathbb{N} \models (\exists x. \phi_D(\overline{\langle \phi \rangle}, x)) \rightarrow \phi$
4. $Q \vdash (\exists x. \phi_D(\overline{\langle \phi \rangle}, x)) \rightarrow \phi$.

9.7 Deuxième théorème d'incomplétude

On montre ici que, si T est cohérente et T est contenue dans PA , alors la cohérence de PA ne peut être prouvée dans T . Ce résultat est connu sous le nom de "deuxième théorème d'incomplétude de Gödel". Il implique en particulier l'échec du programme de Hilbert : on ne peut pas prouver la cohérence de PA dans Q .

Commençons par définir la cohérence de PA dans PA . Soit

$$N \stackrel{\text{def}}{=} \{(n, m) \in \mathbb{N}^2 \mid \exists \phi. n = \langle \phi \rangle \ \& \ m = \langle \neg \phi \rangle\}$$

Ce prédicat est récursif. De même que précédemment, on définit aussi le prédicat de prouvabilité :

$$PR \stackrel{\text{def}}{=} \{(n, m) \in \mathbb{N}^2 \mid \exists \phi. n = \langle \phi \rangle \ \& \ m = \langle \Pi(\phi) \rangle\}$$

PR est aussi récursif. Par représentabilité des fonctions récursives dans PA , il existe donc des formules ϕ_N et ϕ_{PR} qui représentent N et PR respectivement dans PA . A nouveau, nous pourrions donner explicitement ces formules, si nous avons donné explicitement le codage des formules et des preuves.

PA est cohérente ssi, pour toute formule ϕ , $PA \not\vdash \phi$ ou $PA \not\vdash \neg \phi$, et on définit donc

$$C \stackrel{\text{def}}{=} \forall x, y, z, w. (\neg \phi_N(x, y) \vee \neg \phi_{PR}(x, z) \vee \neg \phi_{PR}(y, w))$$

C représente (intuitivement) le fait que PA est cohérente.

Théorème 9.7.1 *Si PA est cohérente, alors $PA \not\vdash C$.*

Preuve:

Nous ne pouvons pas faire la preuve complète sans rentrer dans le détail des codages des formules et des preuves. Notons néanmoins que la preuve du fait que l'arithmétique élémentaire satisfait les propriétés de la figure 9.2 (théorème

9.5.2 peut être formalisée dans PA (mais pas dans Q !) En particulier, on peut montrer dans PA que PA satisfait les propriétés de la figure 9.2.

Considérons maintenant la preuve du théorème 9.6.1. Elle peut-être formalisée dans PA (ce que nous ne faisons pas!) :

$$PA \vdash C \rightarrow \forall x. \neg \phi_{PR}(\overline{\langle \Omega \rangle}, x)$$

Mais $\Omega \stackrel{\text{def}}{=} \theta(\overline{\langle \theta \rangle})$ et $\theta(x) \stackrel{\text{def}}{=} \forall y. (\phi_Q(x, y) \rightarrow \exists z. z < y \wedge \phi_P(x, z))$ et on peut choisir pour ϕ_Q la formule : $\phi_Q(x, y) \stackrel{\text{def}}{=} \exists z. \phi_{PR}(z, y) \wedge \phi_F(x, z)$ où ϕ_F est une formule qui représente l'ensemble $\{(n, m) \in \mathbb{N}^2 \mid \exists \phi. n = \langle \phi(x) \rangle, m = \langle \phi(\overline{\langle \phi(x) \rangle}) \rangle\}$. Comme, par définition, $PA \vdash \phi_F(\overline{\langle \theta \rangle}, \overline{\langle \Omega \rangle})$ (et $PA \vdash \forall z. \phi_F(\overline{\langle \theta \rangle}, z) \rightarrow z = \overline{\langle \Omega \rangle}$),

$$PA \vdash \forall y. (\neg \phi_{PR}(\overline{\langle \Omega \rangle}, y) \rightarrow (\phi_Q(\overline{\langle \theta \rangle}, y) \rightarrow (\exists z. z < y \wedge \phi_P(\overline{\langle \theta \rangle}, z))))$$

Comme $PA \vdash C \rightarrow \forall x. \neg \phi_{PR}(\overline{\langle \Omega \rangle}, x)$ alors

$$PA \vdash C \rightarrow \forall y. \phi_Q(\overline{\langle \theta \rangle}, y) \rightarrow (\exists z. z < y \wedge \phi_P(\overline{\langle \theta \rangle}, z))$$

C'est à dire, par définition, $PA \vdash C \rightarrow \theta(\overline{\langle \theta \rangle})$ ou encore :

$$PA \vdash C \rightarrow \Omega$$

On ne peut donc pas avoir $PA \vdash C$, sans quoi $PA \vdash \Omega$.

Ce résultat entraîne que si T est contenue dans PA , T ne permet pas de démontrer la cohérence de PA , à moins que T soit incohérente.

Exercice 199

Montrer que PA est cohérente ssi elle est ω -cohérente. Cette preuve peut-elle être formalisée dans PA ?

Exercice 200

Montrer que, si T est une théorie contenant PA et T permet de prouver sa propre cohérence, alors T est incohérente.