

Diagonalisation

Nous rassemblons tous les outils construits à présent afin de démontrer l'étape cruciale de l'incomplétude : le lemme de diagonalisation.

1. Résumé des épisodes précédents

L'objectif final n'est maintenant plus très loin : démontrer le théorème d'incomplétude. Nous sommes ici dans un chapitre encore technique avec un lemme crucial qui nécessite tous les chapitres précédents. Avant de nous plonger dans les détails, revenons brièvement sur toutes les notions utiles.

1.1. Arithmétique de Peano

- L'arithmétique de Peano est basée sur le langage $\mathcal{L}_0 = (0, S, +, \times)$, dans lequel on construit une théorie $\mathcal{PA}$ (les axiomes de Peano).
- Les entiers naturels $\mathbb{N}$ sont un modèle de cette théorie : les axiomes de Peano sont satisfaits sur $\mathbb{N}$.
- Pour $n \in \mathbb{N}$, son numéral $\bar{n} = SS \dots S0$ est son écriture dans $\mathcal{L}_0$ (il y a exactement n symboles S).

1.2. Codage de Gödel

- À toute formule de $\mathcal{L}_0$, on associe un entier dans $\mathbb{N}$: son nombre de Gödel. On le note aussi $n = \ulcorner \mathcal{P} \urcorner$.
- Dans $\mathcal{L}_0$ son numéral est donc $\bar{n} = \overline{\ulcorner \mathcal{P} \urcorner}$, que l'on notera souvent aussi par $\ulcorner \mathcal{P} \urcorner$ (c'est un peu abusif, mais il ne peut pas y avoir de confusion dans les formules).
- On a déjà donné un aperçu de ce qu'est la diagonalisation d'une formule lors de la preuve informelle de l'incomplétude, l'idée est de considérer $\mathcal{G} = \mathcal{P}(\ulcorner \mathcal{G} \urcorner)$.
- Le but du chapitre est de formaliser de façon rigoureuse cette formule $\mathcal{G}$.

1.3. Fonctions récursives primitives

- On rappelle que les fonctions récursives primitives sont des fonctions $f : \mathbb{N}^p \rightarrow \mathbb{N}$ facilement calculables.
- Le point important pour ce chapitre est que la fonction substitution $(n, v, k) : \mathbb{N}^3 \rightarrow \mathbb{N}$ est récursive primitive, n désigne le nombre de Gödel d'une formule $\mathcal{P}$, v le code d'une variable x , et k est la valeur que l'on souhaite appliquer à x . La fonction renvoie donc $\ulcorner \mathcal{P}(k) \urcorner \in \mathbb{N}$.

1.4. Fonctions représentables

- Toute fonction récursive primitive est représentable dans $\mathcal{PA}$. On peut alors traduire tous nos concepts mathématiques dans $\mathcal{L}_0$.
- Pour $f : \mathbb{N} \rightarrow \mathbb{N}$ récursive primitive, cela signifie qu'il existe une formule $\mathcal{P}(x, y)$ de $\mathcal{L}_0$ telle que pour tout $n \in \mathbb{N}$, on ait une preuve dans $\mathcal{PA}$ de l'équivalence :

$$\forall y \quad \mathcal{P}(\bar{n}, y) \leftrightarrow y = \overline{f(n)}$$

- Autrement dit, on caractérise la valeur $f(n)$ de façon unique dans $\mathcal{L}_0$ grâce à $\mathcal{P}$.

2. Lemme de diagonalisation

Attaquons-nous à la clé de voûte de la preuve du théorème d'incomplétude.

2.1. Énoncé

Lemme 1 (de diagonalisation).

Pour toute formule $\mathcal{P}(x)$ (où x est la seule variable libre), il existe une formule fermée $\mathcal{G}$ telle que :

$$\boxed{\mathcal{PA} \vdash (\mathcal{G} \leftrightarrow \mathcal{P}(\ulcorner \mathcal{G} \urcorner))}$$

- Ainsi $\mathcal{G}$ apparaît deux fois dans l'énoncé, une fois en tant que formule logique, l'autre via son nombre de Gödel. La formule $\mathcal{G}$ vérifie donc une propriété sur elle-même.
- On peut considérer le lemme de diagonalisation comme un théorème de point fixe : au lieu d'obtenir $f(x) = x$, on a $\mathcal{P}(\ulcorner \mathcal{G} \urcorner) \leftrightarrow \mathcal{G}$.
- L'équivalence est la façon syntaxiquement correcte de dire que $\mathcal{G}$ et $\mathcal{P}(\ulcorner \mathcal{G} \urcorner)$ sont des énoncés équivalents (mais ce ne sont pas des formules égales dans $\mathcal{L}_0$). Par contre on verra dans la preuve ci-dessous que $\mathcal{G}$ est exactement défini par $\mathcal{G} = \mathcal{H}(\ulcorner \mathcal{H} \urcorner)$.

2.2. Preuve

Diagonalisation sur les entiers

On va construire une fonction $\text{diag} : \mathbb{N} \rightarrow \mathbb{N}$ récursive primitive.

- Pour $\mathcal{Q}(x)$, on sait lui associer son nombre de Gödel $n = \ulcorner \mathcal{Q} \urcorner \in \mathbb{N}$. On va s'intéresser à l'opération inverse.
- Pour $n \in \mathbb{N}$, on sait décider si cet entier correspond à une formule correctement formée $\mathcal{Q}(x)$ de $\mathcal{L}_0$, ayant x comme seule variable libre. Si ce n'est pas le cas, on pose $\text{diag}(n) = 0$.
- Si n représente la formule $\mathcal{Q}(x)$, alors on peut évaluer $\mathcal{Q}$ en n , c'est-à-dire écrire la substitution $\mathcal{Q}(\bar{n})$. C'est une formule de $\mathcal{L}_0$, qui a pour nombre de Gödel $\ulcorner \mathcal{Q}(\bar{n}) \urcorner$. On définit alors l'entier :

$$\text{diag}(n) = \ulcorner \mathcal{Q}(\bar{n}) \urcorner$$

C'est-à-dire $\text{diag}(n) = \ulcorner \mathcal{Q}(\ulcorner \mathcal{Q} \urcorner) \urcorner$.

- Résumons la construction par un schéma :

$$n \in \mathbb{N} \xrightarrow{\text{inversion du nb de Gödel}} \mathcal{Q} \in \mathcal{L}_0 \xrightarrow{\text{substitution}} \mathcal{Q}(\bar{n}) \in \mathcal{L}_0 \xrightarrow{\text{nb de Gödel}} \text{diag}(n) = \ulcorner \mathcal{Q}(\bar{n}) \urcorner \in \mathbb{N}$$

- Pour une formule $\mathcal{Q}(x)$ de nombre de Gödel $n = \ulcorner \mathcal{Q} \urcorner$, le calcul de $\mathcal{Q}(k)$ se fait par la fonction $\text{substitution}(n, \ulcorner x \urcorner, k)$ qui, pour la formule $\mathcal{Q}$ de nombre de Gödel n , remplace les occurrences de la variable x par $\bar{k}$, puis renvoie le nombre de Gödel de la formule obtenue, c'est-à-dire $\ulcorner \mathcal{Q}(\bar{k}) \urcorner$.
Remarque : dans toute la suite la variable à substituer sera toujours la variable x .

- Décider si un entier n désigne une formule $\mathcal{Q}(x)$ correctement formée ayant x pour seule variable libre est une opération récursive primitive. La substitution l'est aussi (voir les chapitres « Codage de Gödel » et « Fonctions récursives primitives »). Ainsi la fonction $\text{diag} : \mathbb{N} \rightarrow \mathbb{N}$ est aussi une fonction récursive primitive.

Diagonalisation dans le langage

- Toute fonction récursive primitive est représentable (voir le chapitre « Fonctions représentables »), donc pour notre fonction $\text{diag} : \mathbb{N} \rightarrow \mathbb{N}$, il existe une formule $\text{Diag}(x, y)$ de $\mathcal{L}_0$ telle que, quel que soit $n \in \mathbb{N}$ on ait :

$$\mathcal{P.A} \vdash \left(\forall y \text{ Diag}(\bar{n}, y) \leftrightarrow y = \overline{\text{diag}(n)} \right)$$

Formule $\mathcal{G}$

- On fixe la formule $\mathcal{P}(x)$ de l'énoncé et on définit une nouvelle formule $\mathcal{H}$ par :

$$\mathcal{H}(x) : \exists y \text{ Diag}(x, y) \wedge \mathcal{P}(y)$$

Notons p le nombre de Gödel de cette dernière formule : $p = \ulcorner \mathcal{H} \urcorner$.

- On définit la formule de Gödel par :

$$\mathcal{G} = \mathcal{H}(\bar{p})$$

Ainsi par définition, $\mathcal{G}$ désigne « $\mathcal{H}(\bar{p})$ » et donc aussi « $\exists y \text{ Diag}(\bar{p}, y) \wedge \mathcal{P}(y)$ ».

- De plus, $\text{diag}(p) = \ulcorner \mathcal{G} \urcorner$. En effet, reprenons le schéma ci-dessus, cette fois pour l'entier p :

$$p \in \mathbb{N} \xrightarrow{\text{inversion du nb de Gödel}} \mathcal{H}(x) \in \mathcal{L}_0 \xrightarrow{\text{substitution}} \mathcal{H}(\bar{p}) \in \mathcal{L}_0 \xrightarrow{\text{nb de Gödel}} \text{diag}(p) = \ulcorner \mathcal{H}(\bar{p}) \urcorner = \ulcorner \mathcal{G} \urcorner \in \mathbb{N}$$

Fin de la preuve

Rappelons que le but est de montrer :

$$\mathcal{P.A} \vdash \left(\mathcal{G} \leftrightarrow \mathcal{P}(\overline{\ulcorner \mathcal{G} \urcorner}) \right)$$

On se place donc dans la théorie $\mathcal{P.A}$. On rappelle que $p = \ulcorner \mathcal{H} \urcorner$, $\mathcal{G} = \mathcal{H}(\bar{p})$ et $\text{diag}(p) = \ulcorner \mathcal{G} \urcorner$.

- **Sens $\rightarrow$**

Supposons $\mathcal{G}$

donc $\mathcal{H}(\bar{p})$ par définition de $\mathcal{G}$

donc $\exists y \text{ Diag}(\bar{p}, y) \wedge \mathcal{P}(y)$ par définition de $\mathcal{H}$

donc $\exists y (y = \overline{\text{diag}(p)}) \wedge (\mathcal{P}(y))$ par la représentabilité de diag par Diag dans $\mathcal{P.A}$

donc $\mathcal{P}(\overline{\text{diag}(p)})$ car l'égalité force la valeur de y

donc $\mathcal{P}(\overline{\ulcorner \mathcal{G} \urcorner})$

Ainsi $\mathcal{G} \rightarrow \mathcal{P}(\overline{\ulcorner \mathcal{G} \urcorner})$.

- **Sens $\leftarrow$**

Supposons $\mathcal{P}(\overline{\ulcorner \mathcal{G} \urcorner})$. Il s'agit d'en déduire $\mathcal{G} = \mathcal{H}(\bar{p})$, c'est-à-dire :

$$\exists y \text{ Diag}(\bar{p}, y) \wedge \mathcal{P}(y)$$

On va montrer que $y = \overline{\ulcorner \mathcal{G} \urcorner}$ convient.

— D'une part, $y = \overline{\ulcorner \mathcal{G} \urcorner} = \overline{\text{diag}(p)}$, donc $\text{Diag}(\bar{p}, y)$ est vérifié par la représentabilité.

— De plus, $\mathcal{P}(y) = \mathcal{P}(\overline{\ulcorner \mathcal{G} \urcorner})$ est notre hypothèse.

— Donc, pour ce y , on a bien $\text{Diag}(\bar{p}, y) \wedge \mathcal{P}(y)$.

Par introduction de $\exists$, on a :

$$\mathcal{P}(\overline{\ulcorner \mathcal{G} \urcorner}) \rightarrow [\exists y \text{ Diag}(\overline{p}, y) \wedge \mathcal{P}(y)]$$

ce qui est exactement $\mathcal{P}(\overline{\ulcorner \mathcal{G} \urcorner}) \rightarrow \mathcal{G}$.

On a terminé la preuve du lemme de diagonalisation, on a montré :

$$\mathcal{PA} \vdash (\mathcal{G} \leftrightarrow \mathcal{P}(\overline{\ulcorner \mathcal{G} \urcorner}))$$

3. Théorème de Tarski

3.1. Motivation

Le reste de ce chapitre n'est pas utile pour la preuve du théorème d'incomplétude. Cependant, c'est une excellente application du lemme de diagonalisation et c'est aussi l'occasion de souligner une nouvelle fois la différence entre le vrai et le démontrable. Le théorème de Tarski est une version puissante du paradoxe du menteur. Je dis la phrase suivante :

« Je suis un menteur. »

- Si je suis effectivement un menteur, alors la phrase est vraie, donc je dis la vérité, donc je ne suis pas un menteur. Contradiction.
- Si je ne suis pas un menteur, alors je dis la vérité, donc d'après ma phrase je suis un menteur. Contradiction à nouveau.

Une variante est :

« Cette phrase est fausse. »

3.2. La vérité n'est pas définissable

De façon informelle, le théorème de Tarski (1933) dit qu'il n'existe pas de formule $\mathcal{V}(x)$ qui détecte les formules vraies, c'est-à-dire que $\mathcal{V}(x)$ serait une formule telle que $\mathcal{V}(\overline{n})$ est vraie exactement lorsque la formule fermée $\mathcal{P}$ de nombre de Gödel n est vraie.

Dans notre contexte, une formule fermée $\mathcal{P}$ est vraie si $\models_{\mathbb{N}} \mathcal{P}$, c'est-à-dire si elle est satisfaite dans le modèle standard des entiers naturels $\mathbb{N}$.

Voici l'énoncé précis.

Théorème 1 (Tarski).

Il n'existe pas de formule $\mathcal{V}(x)$ de $\mathcal{L}_0$ telle que pour toute formule fermée $\mathcal{P}$ de $\mathcal{L}_0$ on ait :

$$\models_{\mathbb{N}} \mathcal{P} \text{ si et seulement si } \models_{\mathbb{N}} \mathcal{V}(\overline{\ulcorner \mathcal{P} \urcorner})$$

Une reformulation est « la vérité n'est pas définissable ». En effet, si on note A l'ensemble des entiers $n = \ulcorner \mathcal{P} \urcorner$ tels que $\models_{\mathbb{N}} \mathcal{P}$ (c'est-à-dire les nombres de Gödel des formules vraies), alors il n'existe pas de formule $\mathcal{V}(x)$ telle que $\models_{\mathbb{N}} \mathcal{V}(\overline{n})$ pour tous les $n \in A$, et uniquement ceux-là. On renvoie au paragraphe « Ensembles définissables » du chapitre « Arithmétique de Peano » pour cette notion de « définissable ». Par contre, on verra dans le chapitre suivant que « être démontrable » est définissable.

3.3. Preuve

Par l'absurde, supposons qu'une telle formule $\mathcal{V}(x)$ existe. On applique le lemme de diagonalisation à $\neg \mathcal{V}(x)$ (à la place de $\mathcal{P}(x)$ dans l'énoncé du lemme). Ainsi il existe une formule fermée $\mathcal{G}$ telle que :

$$\mathcal{PA} \vdash (\mathcal{G} \leftrightarrow \neg \mathcal{V}(\overline{\ulcorner \mathcal{G} \urcorner}))$$

Cette phrase $\mathcal{G}$ est notre version du menteur qui parle de lui-même. On se place sur la structure $\mathbb{N}$ pour la théorie $\mathcal{PA}$ sur $\mathcal{L}_0$. Par le théorème de validité, si $\mathcal{PA} \vdash \mathcal{Q}$ alors $\models_{\mathbb{N}} \mathcal{Q}$: tout ce qui est prouvable est

vrai. Ainsi l'équivalence ci-dessus est vraie :

$$\mathcal{G} \leftrightarrow \neg \mathcal{V}(\overline{\neg \mathcal{G}})$$

Vérifions que la phrase $\mathcal{G}$ contredit l'équivalence du Théorème 1 voulue par Tarski (où $\mathcal{G}$ joue le rôle de la phrase $\mathcal{P}$).

- Si $\mathcal{G}$ est vraie, c'est-à-dire $\models_{\mathbb{N}} \mathcal{G}$, alors par l'équivalence du lemme de diagonalisation on a aussi $\models_{\mathbb{N}} \neg \mathcal{V}(\overline{\neg \mathcal{G}})$, c'est-à-dire que $\mathcal{V}(\overline{\neg \mathcal{G}})$ est faux, donc on n'a pas $\models_{\mathbb{N}} \mathcal{V}(\overline{\neg \mathcal{G}})$. Ainsi dans ce cas, $\mathcal{V}$ ne détecte pas que $\mathcal{G}$ est vraie.
- Si $\mathcal{G}$ est fausse (c'est-à-dire on n'a pas $\models_{\mathbb{N}} \mathcal{G}$), alors par l'équivalence du lemme de diagonalisation $\neg \mathcal{V}(\overline{\neg \mathcal{G}})$ est faux, donc $\mathcal{V}(\overline{\neg \mathcal{G}})$ est vrai. Ainsi on a $\models_{\mathbb{N}} \mathcal{V}(\overline{\neg \mathcal{G}})$. Par conséquent dans ce cas, $\mathcal{V}$ ne détecte pas que $\mathcal{G}$ est fausse.
- Ainsi, sans savoir si $\mathcal{G}$ est vraie ou fausse, $\mathcal{G}$ est quand même un contre-exemple à l'équivalence recherchée.