# LE DEUXIÈME THÉORÈME D'INCOMPLÉTUDE DE GÖDEL

*Neil Kennedy, UQÀM, le 31 mars 2006*

## I. TROIS NOTIONS DE COMPLÉTUDE

Le terme « complétude » revêt une forme de polysémie rencontrée souvent en logique et en mathématiques, e.g. dans les termes « algèbre » ou « canonique ». Dans le contexte du travail de Gödel, il y a trois significations de complétude/incomplétude qu'il importe de bien distinguer.

1. <u>Complétude du calcul des prédicats</u>

Le premier sens nous vient d'un théorème que Gödel a démontré en 1929 à l'effet que toute formule logiquement valide d'un calcul des prédicats *K* est démontrable dans *K* (il y a plusieurs calculs des prédicats selon le choix lexical), où « logiquement valide » signifie que la formule est vraie dans tout modèle de *K*. Formellement, une formule $\mathcal{A}$ est logiquement valide si et seulement si pour tout modèle $\mathfrak{M}$ de *K*, $\mathfrak{M} \vDash \mathcal{A}$. Une formule logiquement valide n'est pas nécessairement close (e.g. instanciation d'une tautologie du calcul des propositions). Par ailleurs, il ne s'ensuit pas de la définition de la validité logique qu'une formule de *K* est logiquement valide ou logiquement invalide (fausse dans tous les modèles de *K*). Ces deux catégories – logiquement valide et logiquement invalide – ne sont pas exhaustives des formules de *K*.

2. <u>Complétude d'une théorie du premier ordre</u>

Le second sens de complétude/incomplétude nous vient d'une certaine caractérisation des théories du premier ordre. Une théorie *K* est dite complète si et seulement si, pour toute formule close $\mathcal{A}$ de *K*, $\vdash_K \mathcal{A}$ ou $\vdash_K \neg\mathcal{A}$. Cette deuxième forme de complétude est celle qui est impliquée dans le premier théorème d'incomplétude. En effet, le premier théorème d'incomplétude établit qu'il existe une formule close de *S* qui n'est ni démontrable ni réfutable.

À noter que cette caractérisation porte sur les formules closes de *K*. La complétude de *K* n'entraînerait donc rien sur la prouvabilité des formules non closes. En ce sens, il faut être prudent lorsque nous identifions prouvabilité et vérité dans un système complet et cohérent (ayant un modèle), c'est-à-dire, de conclure que vérité dans le modèle entraîne automatiquement prouvabilité dans la théorie. Ceci n'est vrai que pour les formules closes. Il se pourrait qu'une formule ouverte (non close) soit vraie dans le modèle mais qu'elle soit indémontrable dans la théorie (exemple?).

Par ailleurs, le calcul des prédicats n'est pas complet dans ce second sens. Pour tout calcul des prédicats, il existe une formule close de ce calcul qui est indécidable. En effet, si $A(x_1)$ est un (ou le) symbole de prédicat (unaire) d'un calcul de prédicats *K*, nous pouvons construire deux modèles $\mathfrak{M}$ et $\mathfrak{N}$ (à partir de $\mathbb{N}$) tels que $\mathfrak{M} \vDash \forall x_1 A(x_1)$ et $\mathfrak{N} \vDash \neg\forall x_1 A(x_1)$. $\forall x_1 A(x_1)$ ne pourra donc pas être un théorème.

3. <u>L'incomplétude de la métamathématique</u>

Le troisième sens de complétude/incomplétude requiert que nous rappelions quelques éléments du programme de Hilbert. Pour simplifier, Hilbert avait un projet réductionniste pour la justification des méthodes et des entités mathématiques. Selon lui, au cœur des mathématiques se trouve une arithmétique intuitive (contentuelle) donnée à notre entendement par nos intuitions sur la manipulation des signes concrets. Cette arithmétique ne nécessite pas de justification, elle est saine et exempte de contradictions.

C'est la partie des mathématiques qui n'est pas comprise dans l'arithmétique intuitive qui nécessite des justifications. Pour s'assurer que ces mathématiques ne donnent pas lieu à des contradictions, il faut montrer qu'elles sont cohérentes en utilisant les outils de l'arithmétique intuitive. La réduction que Hilbert opère est un peu particulière : en concevant les théories mathématiques comme des systèmes axiomatiques, c'est-à-dire comme des assemblages de signes concrets, nous pouvons utiliser l'arithmétique intuitive pour chercher à démontrer qu'aucune contradiction ne survient dans la théorie (qu'un signe de formule et un signe de la négation de cette formule ne soient des théorèmes de cette théorie). L'arithmétique intuitive employée de cette manière reçoit le nom de métamathématique et l'objectif que cherche à réaliser la métamathématique a reçu le nom de programme de Hilbert (la démonstration du deuxième problème de Hilbert). L'arithmétisation de Gödel permet une formulation plus rigoureuse de la métamathématique. Au lieu de concevoir les énoncés d'une théorie axiomatisée indirectement comme des nombres (parce qu'ils sont des d'indirectement des nombres comme de symboles, nous parlons directement de nombres de Gödel.

Ce qu'est exactement l'arithmétique intuitive n'est pas très clair. Pour les fins de notre discussion, on pourrait dire qu'elle englobe essentiellement les principes que le système arithmétique *S* formalise.

Le concept d'incomplétude tel qu'il figure dans le second théorème d'incomplétude exprime l'incapacité de démontrer la cohérence de l'arithmétique intuitive formalisée (la théorie *S*) avec une métamathématique qui n'utiliserait que l'arithmétique intuitive. L'idée est que si nous trouvons une preuve de la cohérence de *S* dans un métalangage qui n'utilise que les principes de *S*, il devrait être possible de traduire/formaliser cette preuve dans *S*. Puisque la cohérence de *S* est équivalente (dans le métalangage) au fait que $\nvdash_S 0 = \overline{1}$, nous pouvons choisir la formule $\forall x_2 \neg \mathcal{Pf}(x_2, \ulcorner 0 = \overline{1} \urcorner)$ pour « exprimer » la cohérence de *S* dans *S* (nous aurions pu choisir une autre formule, Gödel emploie celle-ci : $\forall x_1 \forall x_2 \forall x_3 \forall x_4 \neg(\mathcal{Pf}(x_1, x_2) \wedge \mathcal{Pf}(x_3, x_4) \wedge \mathcal{Neg}(x_2, x_4)))$. Appelons cette formule $\mathcal{Con}_S$. L'hypothèse est donc que la prouvabilité de la cohérence de *S* dans un métalangage n'utilisant que les méthodes formalisées dans *S* devrait entraîner la prouvabilité de $\mathcal{Con}_S$ dans *S*. Or le deuxième théorème d'incomplétude établit que $\mathcal{Con}_S$ est indémontrable dans *S* si et seulement si *S* est cohérente.

## II. Les conditions de dérivabilité de Hilbert-Bernays

Considérons pour commencer l'implication $\mathcal{Con}_S \Rightarrow \mathcal{G}$. Si nous interprétons cette formule dans le modèle standard et que nous lui redonnons son sens métamathématique, elle affirme essentiellement la chose suivante : si la théorie *S* est cohérente, alors la formule dont le nombre de Gödel est *g*, la formule $\mathcal{G}$ en l'occurrence, est indémontrable. Cette implication est démontrée dans la première partie de la démonstration du premier théorème d'incomplétude. Il s'agit du passage suivant :

> Montrons que $\nvdash_S \mathcal{G}$. Supposons le contraire, $\vdash_S \mathcal{G}$. Soit *g* le nombre de Gödel de $\mathcal{G}$ et *p* le nombre de Gödel d'une preuve de $\mathcal{G}$. Nous avons alors que la relation récursive $\mathrm{Pr}(p, g)$ tient. Puisque $\mathrm{Pr}(x, y)$ est représentable (par $\mathcal{P}\!f(x_1, x_2)$ de surcroît), nous avons que $\vdash_S \mathcal{P}\!f(\overline{p}, \overline{g})$, c'est-à-dire $\vdash_K \mathcal{P}\!f(\overline{p}, \ulcorner\mathcal{G}\urcorner)$. D'un autre côté, par (G), nous pouvons déduire (de la prouvabilité de $\mathcal{G}$) que $\vdash_S \forall x_2 \neg \mathcal{P}\!f(x_2, \ulcorner\mathcal{G}\urcorner)$ ; en particulier, nous pouvons en déduire que $\vdash_K \neg \mathcal{P}\!f(\overline{p}, \ulcorner\mathcal{G}\urcorner)$ par instanciation du quantificateur universel. Mais ceci contredit la cohérence de *S*. D'où $\nvdash_S \mathcal{G}$.

S'il était possible de formaliser cette démonstration dans *S*, nous aurions évidemment que $\vdash_S \mathcal{C}on_S \Rightarrow \mathcal{G}$. Si $\mathcal{C}on_S$ était démontrable dans *S*, il s'ensuivrait par modus ponens que $\mathcal{G}$ serait démontrable dans *S* et, par le premier théorème d'incomplétude, que *S* serait contradictoire.

Même si les méthodes employées dans ce passage sont des méthodes arithmétiques, il y a un va et vient entre langage objet et métalangage qui n'est pas facile à capturer dans *S*. Pour abréger l'exposition de la preuve, nous admettrons donc certains résultats sur le prédicat $\exists x_2 \mathcal{P}\!f(x_2, x_1)$ de *S* pour démontrer le second théorème. Si $\mathcal{B}ew(x_1)$ désigne le prédicat $\exists x_2 \mathcal{P}\!f(x_2, x_1)$, nous supposerons que les conditions suivantes sont réalisées :

(HB1) Si $\vdash_S \mathcal{A}$, alors $\vdash_S \mathcal{B}ew(\ulcorner\mathcal{A}\urcorner)$
(HB2) $\vdash_S \mathcal{B}ew(\ulcorner\mathcal{A} \Rightarrow \mathcal{B}\urcorner) \Rightarrow (\mathcal{B}ew(\ulcorner\mathcal{A}\urcorner) \Rightarrow \mathcal{B}ew(\ulcorner\mathcal{B}\urcorner))$
(HB3) $\vdash_S \mathcal{B}ew(\ulcorner\mathcal{A}\urcorner) \Rightarrow \mathcal{B}ew(\mathcal{B}ew(\ulcorner\mathcal{A}\urcorner))$

Les conditions (HB1) et (HB2) se démontrent assez bien dans *S*. La condition la plus difficile est (HB3), elle internalise dans *S* le passage d'un « niveau » linguistique à un autre. Dans sa forme la plus générale, elle s'énonce comme suit :

(HB3a) Si $\mathcal{A}$ est une formule $\Sigma_1$, alors $\vdash_S \mathcal{A} \Rightarrow \mathcal{B}ew(\ulcorner\mathcal{A}\urcorner)$

Puisque $\mathcal{B}ew(x_1)$ est une formule $\Sigma_1$, (HB3) s'ensuit de (HB3a).

### III. Le théorème de Löb

Dans le premier théorème d'incomplétude, nous avons démontré que le point fixe de $\neg\mathcal{B}ew(x_1)$ était indécidable (que la formule $\mathcal{G}$, point fixe de $\forall x_2 \neg \mathcal{P}\!f(x_2, x_1)$, était indécidable). Qu'en est-il d'un point fixe de $\mathcal{B}ew(x_1)$, d'une formule close $\mathcal{H}$ telle que

$\vdash_S \mathcal{H} \Leftrightarrow \mathscr{B}ew(\ulcorner \mathcal{H} \urcorner)$ ? ($\mathcal{H}$ est mis pour Henkin.) Le théorème de Löb montre qu'une formule, comme $\mathcal{H}$, qui affirme d'elle-même qu'elle est démontrable est démontrable.

**Théorème de Löb (3.38).** Soit $\mathcal{A}$ une formule close de *S*. Si $\vdash_S \mathscr{B}ew(\ulcorner \mathcal{A} \urcorner) \Rightarrow \mathcal{A}$ alors $\vdash_S \mathcal{A}$.

PREUVE. Appliquons le théorème du point fixe à la formule $\mathscr{B}ew(x_1) \Rightarrow \mathcal{A}$ pour obtenir une formule close $\mathcal{L}$ telle que

(L) $\vdash_S \mathcal{L} \Leftrightarrow (\mathscr{B}ew(\ulcorner \mathcal{L} \urcorner) \Rightarrow \mathcal{A})$.

Le résultat procède de la déduction suivante :

| | | |
|---|---|---|
| 1. | $\mathcal{L} \Rightarrow (\mathscr{B}ew(\ulcorner \mathcal{L} \urcorner) \Rightarrow \mathcal{A})$ | (L) |
| 2. | $\mathscr{B}ew(\ulcorner \mathcal{L} \Rightarrow (\mathscr{B}ew(\ulcorner \mathcal{L} \urcorner) \Rightarrow \mathcal{A}) \urcorner)$ | (HB1) sur 1 |
| 3. | $\mathscr{B}ew(\ulcorner \mathcal{L} \urcorner) \Rightarrow \mathscr{B}ew(\ulcorner \mathscr{B}ew(\ulcorner \mathcal{L} \urcorner) \Rightarrow \mathcal{A} \urcorner)$ | (HB2) et MP |
| 4. | $\mathscr{B}ew(\ulcorner \mathscr{B}ew(\ulcorner \mathcal{L} \urcorner) \Rightarrow \mathcal{A} \urcorner) \Rightarrow (\mathscr{B}ew(\ulcorner \mathscr{B}ew(\ulcorner \mathcal{L} \urcorner) \urcorner) \Rightarrow \mathscr{B}ew(\ulcorner \mathcal{A} \urcorner))$ | (HB2) |
| 5. | $\mathscr{B}ew(\ulcorner \mathcal{L} \urcorner) \Rightarrow (\mathscr{B}ew(\ulcorner \mathscr{B}ew(\ulcorner \mathcal{L} \urcorner) \urcorner) \Rightarrow \mathscr{B}ew(\ulcorner \mathcal{A} \urcorner))$ | 4, 5 et tautologie |
| 6. | $\mathscr{B}ew(\ulcorner \mathcal{L} \urcorner) \Rightarrow \mathscr{B}ew(\mathscr{B}ew(\ulcorner \mathcal{L} \urcorner))$ | (HB3) |
| 7. | $\mathscr{B}ew(\ulcorner \mathcal{L} \urcorner) \Rightarrow \mathscr{B}ew(\ulcorner \mathcal{A} \urcorner)$ | 5, 6 et tautologie |
| 8. | $\mathscr{B}ew(\ulcorner \mathcal{A} \urcorner) \Rightarrow \mathcal{A}$ | Hypothèse |
| 9. | $\mathscr{B}ew(\ulcorner \mathcal{L} \urcorner) \Rightarrow \mathcal{A}$ | 7, 8 et tautologie |
| 10. | $\mathcal{L}$ | (L) |
| 11. | $\mathscr{B}ew(\ulcorner \mathcal{L} \urcorner)$ | (HB1) sur 10 |
| 12. | $\mathcal{A}$ | MP sur 9, 11 |

Nous avons donc que $\vdash_S \mathcal{A}$. ∎

**Corollaire (3.39).** Soit $\mathcal{H}$ une formule close telle que $\vdash_S \mathcal{H} \Leftrightarrow \mathscr{B}ew(\ulcorner \mathcal{H} \urcorner)$. Alors $\vdash_S \mathcal{H}$ et $\mathcal{H}$ est vraie dans le modèle standard (en fait, $\mathcal{H}$ est vraie dans tout modèle de *S*).

PREUVE. Immédiate. Ça ne vaut même pas la peine d'en parler et encore moins de faire une remarque à l'effet qu'elle est immédiate. ∎

## IV. LE SECOND THÉORÈME DE GÖDEL

Nous arrivons enfin au but :

**Second théorème d'incomplétude de Gödel (3.40).** Si *S* est cohérente alors $\nvdash_S \mathscr{C}on_S$ (la réciproque est vraie aussi).

PREUVE. Supposons que *S* est non contradictoire. Puisque $\vdash_S 0 \neq \overline{1}$, par la cohérence de *S*, il s'ensuit que $\nvdash_S 0 = \overline{1}$. Nous avons donc que $\nvdash_S \mathscr{B}ew(\ulcorner 0 = \overline{1} \urcorner) \Rightarrow 0 = \overline{1}$ car sinon, par le théorème de Löb, nous aurions que $\vdash_S 0 = \overline{1}$. Par conséquent, nous ne pouvons pas avoir que $\vdash_S \neg \mathscr{B}ew(\ulcorner 0 = \overline{1} \urcorner)$ car sinon, par la tautologie $\neg \mathcal{A} \Rightarrow (\mathcal{A} \Rightarrow \mathcal{B})$, nous aurions que $\vdash_S \mathcal{Bew}(\ulcorner 0 = \overline{1} \urcorner) \Rightarrow 0 = \overline{1}$. Mais si $\mathcal{Con}_S$ est $\forall x_2 \neg \mathcal{Pf}(x_2, \ulcorner 0 = \overline{1} \urcorner)$, ceci entraîne que $\nvdash_S \mathcal{Con}_S$. Si $\mathcal{Con}_S$ est $\forall x_1 \forall x_2 \forall x_3 \forall x_4 \neg(\mathcal{Pf}(x_1, x_2) \wedge \mathcal{Pf}(x_3, x_4) \wedge \mathcal{Neg}(x_2, x_4))$, il suffit d'observer, en employant (HB1), que $\vdash_S \mathcal{Con}_S \Rightarrow \neg \mathcal{Bew}(\ulcorner 0 = \overline{1} \urcorner)$. La prouvabilité de $\mathcal{Con}_S$ entraînerait celle de $\neg \mathcal{Bew}(\ulcorner 0 = \overline{1} \urcorner)$. □