Logique, Complexité et Vérification

Richard Lassaigne
Logique mathématique,
CNRS-Université Paris 7

Motivation : rapports entre Informatique et Logique mathématique

Logique, Calculabilité et Décidabilité

A. Church, S. Kleene, A. Turing

Indécidabilité et Incomplétude (K. Gödel)

Logique et Complexité (S. Cook, R. Karp)

P=NP, Algorithmes probabilistes

Logique et Vérification

Model checking et Logiques temporelles

Quelques échantillons du programme de D. Hilbert (1904)

- Problème n°1 : Démonstration de l'hypothèse du continu
  (HC) Il n'existe pas de sous -ensemble infini de $\mathbb{R}$ qui ne soit ni équipotent à $\mathbb{N}$ ni équipotent à $\mathbb{R}$.
  (en termes de cardinaux) $2^{\aleph_0} = \aleph_1$

- Problème n°2 : Démonstration de la cohérence
  (non contradiction) de l'Arithmétique

- Problème n°10 : Décidabilité des problèmes diophantiens
  Existence d'un algorithme permettant de décider
  si un polynôme à cefficients dans $\mathbb{Z}$ possède des racines

Quelques résultats

- Thèse de A. Church : Calculabilité $\equiv$ Récursivité $\equiv$ Calculabilité sur machine de Turing $\equiv$ λ-définissabilité $\equiv \ldots$

- Indécidabilité et Incomplétude de l'Arithmétique
  K. Gödel, 1931

- Indépendance de l'axiome du choix et de l'hypothèse du continu K. Gödel, 1938, P.J. Cohen, 1962

- Indécidabilité des problèmes diophantiens
  Y. Matiyasevich, 1970

Importance des résultats

- Indécidabilité :
  L'Arithmétique de Péano est indécidable

- Incomplétude :
  Toute théorie cohérente et récursivement axiomatisable, contenant l'Arithmétique de Péano, est incomplète

- Indécidabilité :
  La logique du 1er ordre, pour tout langage contenant au moins un symbole de relation binaire, est indécidable

Calculabilité et récursivité

L'ensemble des fonctions récursives (partielles) est le plus petit ensemble

- contenant la fonction constante égale à 0, la fonction successeur et les fonctions projections
- et clos par composition, récurrence et minimisation

Remarque : La fonction (partielle) f est définie par minimisation à partir de g ssi

- s'il existe un y tel que $g(x, y) = 0$ et pour tout $z < y$, $g(x, z)$ est defini et $g(x, z) \neq 0$, alors $f(x) = y$
- sinon, $f(x)$ n'est pas défini

La fonction f est notée : $f(x) = \mu y\ (g(x, y) = 0)$

Calculabilité et récursivité

Théorème :
Une fonction (partielle) est calculable sur machine de Turing
ssi elle est récursive

Le langage de l'Arithmétique

Alphabet Σ constitué de :

- symboles logiques $(\neg, \wedge, \vee, \rightarrow, \leftrightarrow, \exists, \forall)$ et parenthèses,
- ensemble de variables $\mathcal{V} = \{v_i \ / \ i \in \mathbb{N}\}$
- symboles arithmétiques $\mathcal{L}_\mathcal{A} = \{+, ., s, 0\}$
  $+$ et $.$ fonctions binaires, s fonction unaire, 0 constante

Ensemble des termes de $\mathcal{L}_\mathcal{A}$: Le plus petit ensemble tel que

- 0 et les variables sont des termes,
- si t est un terme, alors $s(t)$ est un terme,
- si t_1, t_2 sont des termes, alors $(t_1 + t_2)$ et $(t_1.t_2)$ sont des termes.

Proposition :
Il existe un algorithme permettant de décider si un mot de Σ^* est un terme de l'arithmétique.

Formules de l'Arithmétique

Ensemble des formules de $\mathcal{L}_{\mathcal{A}}$: Le plus petit ensemble tel que

- tout mot de la forme $t_1 = t_2$, où t_1, t_2 sont des termes, est une formule (atomique),
- si F est une formule, alors $\neg F$ est une formule,
- si G, H sont des formules, alors $(G \wedge H)$, $(G \vee H)$, $(G \rightarrow H)$ et $(G \leftrightarrow H)$ sont des formules
- si F est une formule et v une variable, alors $\exists v\ F$ et $\forall v\ F$ sont des formules.

Proposition :
Il existe un algorithme permettant de décider si un mot de Σ^* est une formule de l'arithmétique.

Formules de l'Arithmétique

Une variable x est libre dans une formule F si elle a une occurrence qui n'est dans la portée d'aucun quantificateur.
Exemple : x, y sont libres dans la formule $\exists z\ (x + z = y)$

Une formule close est une formule sans variables libres

Proposition :
Il existe un algorithme permettant de décider si un mot de Σ^* est une formule close de l'arithmétique.

Axiomes de Péano

La théorie T_0 est la théorie dont les axiomes sont les suivants :

(A_1) $\forall x\ \neg s(x) = 0$
(A_2) $\forall x \exists y\ (\neg s(x) = 0 \rightarrow x = s(y))$
(A_3) $\forall x \forall y\ (s(x) = s(y) \rightarrow x = y)$
(A_4) $\forall x\ (x + 0 = x)$
(A_5) $\forall x \forall y\ (x + s(y) = s(x + y))$
(A_6) $\forall x\ (x.0 = 0)$
(A_7) $\forall x \forall y\ (x.s(y) = x.y + x)$

L'Arithmétique de Péano est obtenue en rajoutant le Schéma d'induction : pour chaque formule $F(x_0, x_1, \ldots, x_n)$,

$\forall x_1 \ldots \forall x_n((F(0, x_1, \ldots, x_n) \wedge \forall x_0(F(x_0, x_1, \ldots, x_n) \rightarrow$
$F(s(x_0), x_1, \ldots, x_n)) \rightarrow$
$\forall x_0 F(x_0, x_1, \ldots, x_n)))$

Preuves dans l'Arithmétique

Une suite de formules $F_0, F_1, \ldots, F_n$ est une preuve de la formule F dans la théorie T si $F_n = F$ et si pour tout $0 \leq i \leq n$:

- F_i est soit un axiome logique, soit un axiome de la théorie T,
- soit $F_i = \forall v F_j$ est obtenue a partir de F_j (avec$j < i$) par généralisation,
- soit F_i est obtenue à partir de F_j ef $F_k = (F_j \rightarrow F_j)$ (avec $j, k < i$) par modus ponens,

L'existence d'une preuve d'une formule F dans une théorie T est notée : $T \vdash F$

Fonctions représentables

Un ensemble $A \subseteq \mathbb{N}^p$ est représentable s'il existe une formule $F(x_1, \ldots, x_p)$ à p variables libres telle que :
pour tout $a_1, \ldots, a_p \in \mathbb{N}$,

- si $a_1, \ldots, a_p \in A$, alors $T_0 \vdash F(\bar{a}_1, \ldots, \bar{a}_p)$
- si $a_1, \ldots, a_p \notin A$, alors $T_0 \vdash \neg F(\bar{a}_1, \ldots, \bar{a}_p)$

Notation : $\bar{a}$ désigne le terme $s^a(0)$

Une fonction est représentable si son graphe est représentable

Proposition :
Toute fonction calculable sur machine de Turing et totale est représentable

Exemples :

- Les fonctions successeur, addition et multiplication sont

représentables par $v = s(x)$, $v = x_1 + x_2$, $v = x_1.x_2$

- La fonction constante égale a n est représentable par $v = \bar{n}$
- La i-ème projection pr_p^i $(1 \leq i \leq n)$ est représentable par $v = x_i$

Lemme 1 :

L'ensemble des fonctions représentables est clos par composition

On suppose que $f(x) = h(g(x))$ et que g, h sont représentables par des formules $G(x, z)$ et $H(z, y)$:

$y = f(x)$ peut s'exprimer par $\exists z\ (G(x, z) \wedge H(z, y))$

Lemme 2 :
L'ensemble des fonctions représentables est clos par minimisation

Soit g une fonction a 2 variables telle que pour tout entiers a, il existe un entier b satisfaisant $g(a,b) = 0$. On suppose que g est représentable par une formule $G(x,y,v)$.
On veut montrer que la fonction f définie par :
$f(a) = \mu b\ g(a,b) = 0$ est représentable.
On considère la formule $F(x,y)$: $G(x,y,0) \wedge \forall z\ (z < y \rightarrow \neg G(x,z,0))$,
où $z < y$ est une abréviation pour la formule $\exists v\ (v \neq 0 \wedge y = z + v)$

Lemme 3 (Théorème du reste chinois) :
Si $b_0, b_1, \ldots, b_k$ sont des entiers 2 à 2 premiers entre eux et $p = \prod_{i=0}^{k} b_i$, alors $\mathbb{Z}/p\mathbb{Z} \cong \prod_{i=0}^{k} \mathbb{Z}/b_i\mathbb{Z}$

Corollaire (Fonction β de Gödel) :
Il existe une fonction β à 3 variables, calculable sur MT, représentable et satisfaisant :
pour toute suite d'entiers $n_1, n_2, \ldots, n_k$, il existe deux entiers a, b tels que pour tout $0 \leq i \leq k$, $\beta(i, a, b) = n_i$

On choisit $m \geq k+1$ et $a = m! \cdot k$ tels que $a \geq n_i$ $(0 \leq i \leq k)$
Les entiers $b_i = a \cdot ((i+1)+1)(0 \leq i \leq k)$ sont 2 à 2 premiers
D'après le théorème du reste chinois, il existe un entier b tel que :
pour tout $0 \leq i \leq k$, $b \equiv n_i$ modulo b_i. La valeur de la fonction $\beta(i, a, b)$ peut être définie comme le reste de la division de b par b_i
La fonction β estreprésentable par la formule $B(x_0, x_1, x_2, v)$:

$$\exists y \ (x_2 = y(x_1(x_0+1)+1) + v \wedge v < x_1(x_0+1)+1)$$

Lemme 4 :
L'ensemble des fonctions représentables est clos par récurrence

Soit g, h des fonctions représentables par des formules G, H et f une fonction définie par récurrence à partir de g, h :

$$f(x, 0) = g(x)$$
$$f(x, z+1) = h(x, z, f(x, z))$$

Pour exprimer que $y = f(x, z)$, on écrit qu'il existe une suite d'entiers $n_0, n_1, \ldots, n_z$ tels que :
$n_0 = g(x)$, $n_z = y$ et pour tout $0 \leq i \leq z-1$, $n_{i+1} = h(x, i, n_i)$
Les valeurs de cette suite d'entiers peuvent être representées comme les valeurs de la fonction $\beta(i, a, b)$ $(0 \leq i \leq z-1)$ pour 2 entiers bien choisis a, b

La fonction f est représentable par la formule $F(x, z, y)$:

$\exists a \exists b\ (\exists x_0 (B(0, a, b, x_0) \wedge G(x, x_0) \wedge B(z, a, b, y) \wedge$
$(\forall i < z) \exists x_1 \exists x_2\ (B(i, a, b, x_1) \wedge B(s(i), a, b, x_2) \wedge H(x, z, x_1, x_2)))$

On commence par coder les suites d'entiers

Proposition :
La fonction α_2 définie sur $\mathbb{N}^2$ par : $\alpha_2(x,y) = \frac{(x+y)(x+y+1)}{2} + y$ est calculable et bijective. La fonction inverse peut être définie en utilisant 2 fonctions calculables π_2^1 et π_2^2

Pour les suites finies de longueur $p \geq 2$, on peut définir la numérotation α_p par récurrence sur p :

$$\alpha_{p+1}(x_1, x_2, \ldots, x_{p+1}) = \alpha_2(\alpha_p(x_1, x_2, \ldots, x_p), x_{p+1})$$

Le codage des termes du langage de l'Arithmétique peut être défini par induction sur le terme t :

- si $t = 0$, alors $\sharp(t) = 0$
- si $t = v_n$, alors $\sharp(t) = \alpha_3(n+1, 0, 0)$
- si $t = s(t_1)$, alors $\sharp(t) = \alpha_3(\sharp(t_1), 0, 1)$
- si $t = t_1 + t_2$, alors $\sharp(t) = \alpha_3(\sharp(t_1), \sharp(t_2), 2)$
- si $t = t_1.t_2$, alors $\sharp(t) = \alpha_3(\sharp(t_1), \sharp(t_2), 3)$

On peut définir de manière analogue le codage des formules de l'Arithmétique

A toute suite finie $d = (F_0, F_1, \ldots, F_n)$ de formules de l'Arithmétique, on peut associer l'entier $\sharp(d)$ défini par :

$$\sharp(d) = \prod_{i=0}^{n} p(i)^{\sharp(F_i)}$$

où $p(i)$ est le $(i+1)$-ème nombre premier et $\sharp(F_i)$ le numéro de Gödel de la formule F_i

L'entier $\sharp(d)$ est appelé le numéro de Gödel de la suite d

Exercice :

A partir d'un entier z, on peut facilement retrouver la suite $(\sharp(F_i))_{0 \leq i \leq n}$ dont z est le numéro de Gödel à l'aide des 2 fonctions calculables suivantes

- $lg(z)$ est la longueur de la suite dont z est le numéro de Gödel
- $\delta(i, z)$ est l'exposant de $p(i)$ dans la décomposition en facteurs premiers de z

Théories décidables

• Une théorie T est récursivement axiomatisable s'il existe un algorithme permettant de déterminer si une formule est un axiome de cette theorie

• Une théorie T est décidable s'il existe un algorithme permettant de déterminer si une formule est un théorème de la théorie T

Exemples :

• Toute théorie finiment axiomatisable (T_0) est récursivement axiomatisable

• L'Arithmétique de Péano est récursivement axiomatisable

Proposition :
Si T est une théorie récursivement axiomatisable, alors il existe un algorithme permettant de décider l'ensemble

$$Dem(T) = \{(n,m) \in \mathbb{N}^2 \ / \ n = \sharp(F),\ m = \sharp(d),$$
F est une formule et d est une preuve de F a partir de $T\}$

La définition d'une preuve fournit un algorithme pour décider si une suite finie de formules est une preuve

$(n,m) \in Dem(T)$ ssi les conditions suivantes sont satisfaites :

1. pour tout $i < lg(m)$, $\delta(i,m)$ est le numéro de Gödel d'une formule,
2. $\delta(lg(m),m) = n$,
3. pour tout $i < lg(m)$, $\delta(i,m)$ est le numéro de Gödel d'un axiome logique, d'un axiome de T ou d'une formule obtenue à partir de formules la précédant dans la suite à l'aide de la règle de généralisation ou de celle de modus ponens.

Corollaire 1 :
Si T est une théorie récursivement axiomatisable, alors l'ensemble $Thm(T)$ des numéros de Gödel des théoremes de T est récursivement énumerable, ou semi-décidable, c'est-à-dire il existe un semi-algorithme permettant de les énumérer

Un entier n est le numéro de Gödel d'un théorème de T ssi :

- n est le numéro de Gödel d'une formule close
- et il existe un entier m tel que $(n, m) \in Dem(T)$

Corollaire 1(suite) :
Si T est une théorie récursivement axiomatisable, alors l'ensemble $Thm(T)$ des numéros de Gödel des théorèmes de T est récursivement énumerable, ou semi-décidable, c'est-à-dire il existe un semi-algorithme permettant de les énumérer

L'ensemble des numéros de Gödel des formules closes est décidable.
Comme T est une théorie récursivement axiomatisable, l'ensemble $Dem(T)$ est décidable
L'ensemble $\{n \ / \ $ il existe m tel que $(n, m) \in Dem(T)\}$ est semi-décidable, comme projection d'un ensemble décidable
L'ensemble $Thm(T)$, intersection d'un ensemble décidable et d'un ensemble semi-décidable, est semi-décidable.

Une théorie T est complète ssi pour toute formule close F :

soit $T \vdash F$, soit $T \vdash \neg F$

Corollaire 2 :

Si T est une théorie récursivement axiomatisable et complète, alors T est décidable

Il suffit de montrer qu'il existe un semi-algorithme permettant d'énumérer le complémentaire de $Thm(T)$

Corollaire 2(suite) :
Si T est une théorie récursivement axiomatisable et complète,
alors T est décidable

Un entier n n'est pas le numéro de Gödel d'un théorème de T ssi :

- n n'est pas le numéro de Gödel d'une formule close
- ou si n est le numéro de Gödel d'une formule close qui n'est pas un théorème de T

Comme T est complète, alors pour toute formule close F :
F n'est pas un théorème de T ssi $\neg F$ est un théorème de T

Le complémentaire de $Thm(T)$, réunion d'un ensemble décidable
et d'un ensemble semi-décidable, est semi-décidable

Theorème :
Si T est une théorie cohérente et contenant T_0, alors T est indécidable

Soit T une théorie cohérente et décidable, contenant T_0
$A = \{(m,n) \in \mathbb{N}^2 \ / \ n = \sharp(F(v_0)), F$ à 1 variable libre et $T \vdash F(\bar{n})\}$
est décidable
$B = \{n \ / \ (n,n) \notin \mathbb{N}\}$ est aussi décidable, donc représentable par une formule $G(v)$ à 1 variable libre, c'est-à-dire :

- si $n \in B$, alors $T_0 \vdash G(\bar{n})$ et $T \vdash G(\bar{n})$,
- si $n \notin B$, alors $T_0 \vdash \neg G(\bar{n})$ et $T \vdash \neg G(\bar{n})$

Soit a le numéro de Gödel de la formule $G(v)$. Si l'on applique la définition de B à l'entier a, on obtient une contradiction avec la propriété précédente :

- si $a \in B$, alors $(a,a) \notin A$ et $T \nvdash G(\bar{a})$,
- si $a \notin B$, alors $(a,a) \in A$ et $T \vdash G(\bar{a})$
  et la théorie T est incohérente

Théorème (1er théorème d'incomplétude de Gödel) :
Si T est une théorie cohérente, contenant l'arithmétique de Péano, et récursivement axiomatisable, alors T est incomplète

Si T était une theorie complète, alors elle serait décidable

Corollaire :
Il n'existe pas d'algorithme permettant de décider si une formule close d'un langage du 1er ordre, contenant celui de l'arithmétique, est une formule valide

Soit $\mathcal{V}al$ l'ensemble des formules closes valides et F_0 la conjonction des axiomes de la theorie T_0. Pour toute formule close G du langage de l'arithmétique, la condition suivante est satisfaite :

$$T_0 \vdash G \text{ ssi } (F_A \rightarrow G) \in \mathcal{V}al$$

S'il existait un tel algorithme, alors la théorie T_0 serait décidable

Références

- R. Cori et D. Lascar : *Logique mathématique, tome II*
  Masson, 1993

- R. Lassaigne et M. de Rougemont : *Logic and Complexity*
  Springer-Verlag, 2004