Logical Analysis of (some) Proofs in Algebra

Thierry Coquand

Aug. 30, 2006

Content of the course

Lecture I: Hilbert's program, intuitionistic/classical logic, negative translation

Lecture II: coherent logic, dynamical proof

The lecture II presents a *new approach* to constructive mathematics

Content of the course

Coste, Lombardi, Roy

"*Dynamical Method in Algebra*",

Ann. Pure Appl. Logic 111 (2001), no. 3, 203–256.

Prawitz

"*Ideas and results in proof theory*"

in the Proceedings of the Second Scandinavian Logic Symposium, 1971

Hilbert's program

In mathematics, success of non effective methods to prove concrete statements

concrete: existence of a "finitary" object satisfying a decidable property

Hilbert's program

Theorem (Krivine): *If $P \in \mathbb{Q}[x_1, \ldots, x_k]$ is > 0 on $[0,1]^n$ then it can be written as a polynomial in $x_i, 1 - x_i$ with rational* positive *coefficients*

This is also proved with the Axiom of Choice

It is not true if P is only ≥ 0: take $(2x-1)^2$

(but it works for $(2x-1)^2 + \epsilon$ if $\epsilon > 0$)

Hilbert's program

Theorem (Kronecker): *Given* $n+2$ *polynomials in* $\mathbb{Q}[X_1,\dots,X_n]$ $P_1,\dots,P_{n+2}$ *there exist* $n+1$ *polynomials* $Q_1,\dots,Q_{n+1}$ *such that*

$$Z(P_1,\dots,P_{n+2}) = Z(Q_1,\dots,Q_{n+1}) \subseteq \mathbb{C}^n$$

One proof (van der Waerden) uses the notion of Noetherian rings and Krull dimension

Other examples

(Counter) Example: any polynomial P of degree ≥ 1 in $K[X]$ has an irreducible factor

If P is not irreducible, $P = QR$ with $1 \leq d(Q) < d(R)$ and we can find an irreducible factor

This looks like an algorithm but even if K is concretely given it can be shown that there is *no* such algorithm in general

The property: "to be irreducible" is not decidable in general

Other examples

How is it that we can use such properties without problems?

This notion of irreducible polynomials is used without comment in the work of Abel and Galois (early use of classical reasoning?)

Hilbert's program

Whenever we use "ideal methods" to prove a concrete statement we should be able to explain the use of these ideal methods and replace this argument by a proof which has a direct algorithmic content

In particular, if we prove the existence of an object, this proof should give us a way to find this object

"ideal methods": Axiom of Choice, prime ideals

Role of classical arguments?

Brouwer-Heyting-Kolmogorov: constructive proofs can be seen directly as algorithms

Question: has the use of ideal/non effective arguments in mathematics some computational relevance??

We shall see that it has some connection with the idea of *lazy* computation

(We cannot compute completely an infinite object but we can use partial finite amount of information about this object during a computation.)

Fragments of First-Order Logic

Equational logic (much larger fragment than it seems: one important theorem of Serre 1958 can be formulated in this fragment)

Coherent logic (for which intuitionistic and classical logic coincide)

Intuitionistic logic (for which we have the BHK interpretation)

Classical first-order logic

Geometric logic

Some properties cannot be stated in first-order logic but belongs to a logic for which intuitionistic and classical logic coincide (extension of coherent logic with countable disjunction)

nilpotent $\bigvee_{n\in\mathbb{N}} x^n = 0$

flat module $\Sigma r_i u_i = 0 \rightarrow \exists B.\exists \vec{v}.\vec{u} = B\vec{v} \wedge (r_1, \ldots, r_k)B = 0$

to be integral $\bigvee_{n\in\mathbb{N}} \exists u_1, \ldots, u_n.x^n + u_1 x^{n-1} + \cdots + u_n$

Completeness theorems

Theorem: (Birkhoff's theorem) *If an equation is a semantical consequence of an equational theory then it can be deduced purely by equational reasoning*

Theorem: (Skolem, Gödel) *If a first-order statement is a semantical consequence of a first-order theory then it can be deduced purely by first-order reasoning*

Theorem: (Deligne?) *If a coherent first-order statement is a semantical consequence of a coherent first-order theory then it can be deduced purely by a dynamical proof*

Completeness theorems

This is an indication that Hilbert's program should hold in algebra since most statements there can be formulated in an equational or coherent way

Skolem and Gödel proved completeness w.r.t. "cut-free" provability (this is not the case for Henkin's proof)

Example: Jacobson radical

Classically one defines $J \subseteq R$ as the intersection of all maximal ideals of R

One can prove $x \in J \leftrightarrow \forall z.inv(1-xz)$ where $inv(u) \equiv \exists y.uy = 1$

It follows that we have

$$\forall z.inv(1-uz) \wedge \forall z.inv(1-vz) \quad \rightarrow \quad \forall z.inv(1-(u+v)z)$$

This is a *first-order tautology* and hence it can be proved in first-order logic

Furthermore the proof cannot be "too complicated"

Kronecker's theorem

We consider a two sorted theory: theory of commutative rings and theory of distributive lattice with $D : R \to L$ satisfying

$$D(0) = 0 \qquad D(1) = 1 \qquad D(uv) = D(u) \wedge D(v) \qquad D(u+v) \leq D(u,v)$$

where $D(u_1, \ldots, u_n)$ denotes $D(u_1) \vee \cdots \vee D(u_n)$

Key example: L is the lattice of finitely generated radical ideals of R

Kronecker's theorem

In this theory, the following property holds

$$D(uv) = 0 \rightarrow D(u+v) = D(u,v)$$

Since this is first-order the proof cannot be too complex: it follows from

$$D(u+v, uv) = D(u,v)$$

Kronecker's theorem

We say that $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ are complementary iff

$$D(a_1, b_1) = 1,\ D(a_1 b_1) \leq D(a_2, b_2),\ \ \ldots, D(a_n b_n) = 0$$

For $n = 1$ this means that $D(a_1)$ is the complement of $D(b_1)$

R, D is of dimension $< n$ iff any n-ary sequence has a complementary sequence

Kronecker's theorem

Theorem: *if R, D is of dimension $< n$ then for any $u_0, u_1, \ldots, u_n$ there exists $v_1, \ldots, v_n$ such that $D(u_0, \ldots, u_n) = D(v_1, \ldots, v_n)$*

This is a generalisation of Kronecker's Theorem

Since it is formulated as a *first-order schema* the proof cannot be complicated *a priori*

Forster's theorem

Let M be a rectangular matrix and $\Delta_n(M)$ be $\vee_\nu D(\nu)$ where ν ranges over the $n \times n$ minors of M

Theorem: *If $\Delta_n(M) = 1$ and R, D is of dimension $< n$ then there exists an unimodular combination of the column vectors of M*

This is a non Noetherian version of Forster's 1964 Theorem

Since it is formulated as a *first-order schema* the proof cannot be complicated *a priori*

For a given n and given size of the matrix, one expects to have an algorithm which produces the unimodular combination

Negative translation

We shall explain an important example of a conservativity result which has a completely elementary proof (this is not the case for the cut-elimination results)

Conservativity of classical logic over intuitionistic logic, for a large class of statements

Kolmogorov (1925), Gödel (1932), Bernays, Gentzen, Friedman-Dragalin

Negative translation

The fact that there is such a simple translation between *classical* and *intuitionistic* logic gives an *intuitionistic* proof of consistency of classical arithmetic

This result probably surprised Gödel, Bernays (1932) who introduced since a distinction between "finitary" and "intuitionistic"

There is another distinction: difference between "feasible" and "finitary". The negative translation is feasible, normalization in natural deduction is *not*.

A-translation for first-order logic

We fix a first-order language

$$A, B \ ::= \ R(t_1, \ldots, t_k) \mid A \wedge A \mid A \rightarrow A \mid \ \perp \ \mid A \vee A \mid \forall x.A \mid \exists x.A$$

As "usual" we define $\neg A$ to be $A \rightarrow \perp$

Natural deduction

Prawitz/Gentzen $\Gamma \vdash A$ if $A \in \Gamma$

$$\frac{\Gamma \vdash A \quad \Gamma \vdash B}{\Gamma \vdash A \wedge B} \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A} \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash B}$$

$$\frac{\Gamma, A \vdash C \quad \Gamma, B \vdash C \quad \Gamma \vdash A \vee B}{\Gamma \vdash C} \qquad \frac{\Gamma \vdash A}{\Gamma \vdash A \vee B} \qquad \frac{\Gamma \vdash B}{\Gamma \vdash A \vee B}$$

$$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \rightarrow B} \qquad \frac{\Gamma \vdash A \rightarrow B \quad \Gamma \vdash A}{\Gamma \vdash B}$$

Natural deduction

$$\frac{\Gamma \vdash A(x)}{\Gamma \vdash \forall x. A(x)}\ (*) \qquad \frac{\Gamma \vdash \forall x. A(x)}{\Gamma \vdash A(t)}$$

$$\frac{\Gamma \vdash A(t)}{\Gamma \vdash \exists x. A(x)} \qquad \frac{\Gamma, A(x) \vdash B \qquad \Gamma \vdash \exists x. A(x)}{\Gamma \vdash B}\ (*)$$

Natural deduction

$$\frac{}{\Gamma, \bot \vdash A}$$

Without this rule we have *minimal* logic

With this rule we have *intuitionistic* logic

To get *classical* logic, we add the rule

$$\frac{\Gamma, \neg A \vdash \bot}{\Gamma \vdash A}$$

Notice that classical logic contains intuitionistic logic

A-translation for first-order logic

$\neg_A B = B \to A$

$B^* = \neg_A \neg_A B$, B atomic

$(B_1 \wedge B_2)^* = B_1^* \wedge B_2^*$

$(B_1 \to B_2)^* = B_1^* \to B_2^*$

$(\perp)^* = A$

$(B_1 \vee B_2)^* = \neg_A \neg_A (B_1^* \vee B_2^*)$

A-translation for first-order logic

Since $\vdash_i \neg_A(B \vee C) \leftrightarrow \neg_A B \wedge \neg_A C$ and $\vdash_i \neg_A(\exists x.B) \leftrightarrow \forall x.\neg_A B$ this amounts to take away $\vee$ and $\exists$ and to *define* then

$$B \vee C \ = \ \neg(\neg B \wedge \neg C)$$

$$\exists x.B \ = \ \neg(\forall x.\neg B)$$

This is what Prawitz does (and he restricts the inference of $\vdash B$ from $\neg B \vdash \perp$ to *atomic* formulae B)

A-translation for first-order logic

Already in Kolmogorov (1925) on the form

$B^* = \neg_A \neg_A B$, B atomic

$(B_1 \wedge B_2)^* = \neg_A \neg_A (B_1^* \wedge B_2^*)$

$(B_1 \rightarrow B_2)^* = \neg_A \neg_A (B_1^* \rightarrow B_2^*)$

$(B_1 \vee B_2)^* = \neg_A \neg_A (B_1^* \vee B_2^*)$ (by translation)

$(\perp)^* = A$

A-translation for first-order logic

$(\forall x.B)^* = \forall x.B^*$

$(\exists x.B)^* = \neg_A \neg_A (\exists x.B^*)$

Lemma: $\neg_A \neg_A B^* \vdash_i B^*$ *for any formula* B

Theorem: *If* $B_1, \ldots, B_k \vdash_c B$ *then* $B_1^*, \ldots, B_k^* \vdash_i B^*$

Assume that Σ is a set of formulae such that $\Sigma \vdash_i B^*$ if $B \in \Sigma$ and $\vdash_i C^* \rightarrow C$

Corollary: *If* $\Sigma \vdash_c C$ *then* $\Sigma \vdash_i C$

Application: arithmetic

The first application is when Σ is the theory of natural numbers. The only relation symbol is equality.

$$0 \neq x+1 \qquad x+1 = y+1 \rightarrow x = y$$

For each of these formulae we have $\vdash_i A \rightarrow A^*$

The induction schema $I(A)$ is the formula

$$A(0) \wedge \forall x.(A(x) \rightarrow A(x+1)) \rightarrow \forall x.A(x)$$

We have $I(A)^* = I(A^*)$

Application: arithmetic

Peano Arithmetic is Σ with classical logic, *Heyting Arithmetic* is Σ with intuitionistic logic

Theorem: *If* PA $\vdash C$ *then* HA $\vdash C^*$

If C is a formula $\exists y.B(x,y)$ with B atomic and we take $A = \exists x.B(x,y)$ we get

Theorem: *If* PA $\vdash \exists y.B(x,y)$ *then* HA $\vdash \exists y.B(x,y)$

Application: arithmetic

This can be considered as a solution of Hilbert's program for arithmetic: if we prove *classically* the existence of a number satisfying some equations then we have also an *intuitionistic* proof

Furthermore we have an explicit way to get this proof from the classical argument

Theorem: *If* PA $\vdash\perp$ *then* HA $\vdash\perp$

This is a simple *constructive* proof of the *consistency* of Peano Arithmetic

Why it is not considered as a solution to Hilbert's program: distinction between *finitary* and *intuitionistic* (Gödel, Bernays 1932)

Application I: coherent theories

We assume that all formulae in Σ are the form

$$H ::= R(t_1, \ldots, t_k) \mid H \wedge H \mid H \vee H \mid \perp \mid \top \mid \exists x.H \qquad I ::= H \rightarrow H \mid \forall x.I$$

The formulae H are called *positive* formulae

Lemma: *For any positive formula H we have $\vdash_m H^* \leftrightarrow \neg_A \neg_A H$*

Lemma: *For any coherent formula I we have $\vdash_m I \rightarrow I^*$*

Theorem: *If all formulae of Σ are coherent, I is coherent and $\Sigma \vdash_c I$ then $\Sigma \vdash_m I$*

Application I: coherent theories

An example of a coherent theory is the theory of local rings

Axiom $(\exists y.xy = 1) \vee (\exists y.(1 - x)y = 1)$

We define $inv(x) \equiv \exists y.xy = 1$, the axiom can be written

$inv(x) \vee inv(1 - x)$

We have $inv(xy) \leftrightarrow inv(x) \wedge inv(y)$ hence the axiom implies

$inv(x) \vee inv(1 - xy)$

Define $J(x) \equiv \forall y.inv(1 - xy)$ we have *classically*

$inv(x) \vee J(x)$

Application I: coherent theories

J defines an ideal of R

Classically we can show $inv(x) \vee J(x)$ and $k = R/J$ is a field

$inv(x) \vee J(x)$ expresses that we have a local ring with a *detachable* maximal ideal

One can *prove* that this is not provable intuitionistically (using the technique presented in the second lecture)

Application I: coherent theories

Using that R/J is a field one can prove

Lemma: If F is an idempotent square matrix over a local ring R then F is similar to a matrix of the form

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$$

By our metatheorem this can also be proved intuitionistically

In *practice*, in most examples, the direct intuitionistic proof is not more complicated than the classical proof

Application II: example

(This example is due to U. Berger and H. Schwichtenberg.)

Suppose that we have proved $5 < f(3)$ and $f(0) \leq 5$

We show $\exists n. f(n) \leq 5 < f(n+1)$ classically

Application II: example

If this is not the case, we have $\forall n. f(n) \leq 5 \to f(n+1) \leq 5$

Since $f(0) \leq 5$ by induction we have $f(n) \leq 5$ for all n

This contradicts $5 < f(3)$

A constructive proof will be to prove directly, from $f(0) \leq 5 < f(3)$

$(f(0) \leq 5 < f(1)) \quad \vee \quad (f(1) \leq 5 < f(2)) \quad \vee \quad (f(2) \leq 5 < f(3))$

What do we get by negative translation?

Application II: counter-example

We have a proof in PA of a statement $\exists n. \forall m. f(n+m) \neq 0$

It states that the equation $f(x) = 0$ has only a finite number of solutions

In general, from a proof of this statement, it will *not* be possible to compute a bound for the solutions

Concrete instance: Mordell's conjecture, which states that a large class of polynomial equations has only a finite number of rational solutions

Problems in constructive algebra

Proposition: *There is no irreducibility test for $k[X]$ even if k is discrete*

We reduce the problem to a decision $\forall n.\alpha_n = 0 \vee \exists n.\alpha_n = 1$

Is $X^2 + 1$ irreducible over $k[X]$ where k is the field generated by the elements $\alpha_n i,\ n \in \mathbb{N}$??

This field k is well-defined: its elements are polynomials $f(\alpha_0 i, \ldots, \alpha_n i)$ and it is *discrete*

$X^2 + 1$ is irreducible over $k[X]$ iff $\forall n.\alpha_n = 0$

References

R. Zach (presentation of Hilbert's program)

Berger, Schwichtenberg (examples of A-translation)

Gödel collected work III, IV (difference between finitary and intuitionistic)

D. Prawitz "Ideas and results in proof theory" Proceedings of the Second Scandinavian Logic Symposium (Univ. Oslo, Oslo, 1970), pp. 235–307.

U. Kohlenbach monotone Dialectica/realisability interpretation