The Second Incompleteness Theorem In General

Albert Visser

Philosophy, Faculty of Humanities, Utrecht University

Tribute to Kurt Gödel 2020

Brno, January 13, 2020

Universiteit Utrecht

Overview

Preludium

The Interpretation Version of the Theorem

A Fefermanian Variation

Coordinate Free G2?

Overview

Preludium

The Interpretation Version of the Theorem

A Fefermanian Variation

Coordinate Free G2?

Overview

Universiteit Utrecht

Overview

Preludium

The Interpretation Version of the Theorem

A Fefermanian Variation

Coordinate Free G2?

Overview

Preludium

The Interpretation Version of the Theorem

A Fefermanian Variation

Coordinate Free G2?

Preludium 1

G2: A sufficiently strong theory does not prove its own consistency.

But what does that mean?

We will zoom in on the following questions:

A. Can we give a decent statement of the theorem that does justice to our understanding?

B. How general is the theorem?

Since the defects of contemporary statements of the theorem can be viewed as symptoms of insufficient generality, our main target is (A).

Preludium 2

There are (at least) two approaches to generality. I will discuss *the interpretation version* and say nothing about the approach involving *the Löb Conditions*.

We will discuss three stages of generality:

i. interpretations,
ii. axiomatisations,
iii. variations in arithmetisation.

We will touch on philosophical questions too:

- What do we mean when we say that a certain arithmetical statement expresses the consistency of a theory?
- Can we get some grip on the totality of the sentences expressing consistency?

Overview

Preludium

The Interpretation Version of the Theorem

A Fefermanian Variation

Coordinate Free G2?

Interpretations

An *interpretation* K of a theory V in a theory U is based on a *translation* τ_K of the predicate symbols of V to formulas of U.

τ_K can be lifted to the full language of V by making it commute with the logical connectives. In case of the quantifiers we allow *domain relativisation*.

Translations can come with all kinds of special features: they can be multi-dimensional; they can have parameters; they can build up the domain from pieces . . .

τ_K supports the interpretation $K : V \lhd U$ iff, for all V-sentences, if $V \vdash A$, then $U \vdash A^{\tau_K}$.

Interpretability

We write:

- $V \lhd U$ or $U \rhd V$ for: there is a K with $K : V \lhd U$.

Weak Theories

Our favourite weak theory for the verification of G2 is Buss' theory S^1_2. This is a very weak sub-theory of Peano Arithmetic PA. S^1_2 is a theory of p-time computability.

S^1_2 does not prove that exponentiation is total.

S^1_2 has the advantage that the usual proof of G2 goes through without any extra trouble. Rather it is better than PA for this purpose since it keeps us from making silly design choices.

We can obtain G2 in theories that are in some respects weaker than S^1_2, but that requires special pleading.

Universiteit Utrecht

Interpretation Version 1

Theorem (a.o. Sam Buss)

Let U be any consistent recursively enumerable theory (in finite signature). Let α_U be a Σ_1^b- representation of the axiom set of U. Then $U \not\rhd (\mathsf{S}_2^1 + \mathsf{con}(\alpha_U))$.

Here con is a formalisation of consistency using one's favourite coding scheme. This is the Fefermanian approach. More about this later.

Suppose, for some $N : \mathsf{S}_2^1 \lhd U$, we have $U \nvdash (\mathsf{incon}(\alpha_U))^N$. Then, we can find $N_0 : \mathsf{S}_2^1 \lhd U$ and $N_1 : \mathsf{S}_2^1 \lhd U$ such that

$$U \nvdash (\mathsf{con}(\alpha_U))^{N_i} \to (\mathsf{con}(\alpha_U))^{N_{1-i}}.$$

Universiteit Utrecht

Interpretation Version 2

For a wide class of theories U like S^1_2, EA, PRA, ACA_0, GB, there is no A, such that $U \nvdash A$, but $U + (\mathrm{con}(\alpha_U))^N \vdash A$, for all $N : \mathsf{S}^1_2 \lhd U$.

So there is nothing that can pose as the weakest consistency statement. Especially, the disjunction of all $(\mathrm{con}(\alpha_U))^N$ is not generally first-order definable over U.

Alternative Versions

We can replace S^1_2 in the statement of the theorem by:

- Q. (Pavel Pudlák, using an idea of Robert Solovay.)
- PA^-.
- An appropriate weak theory of syntax. There are many variants of such theories that would fill the bill.
- An appropriate weak set theory like AS.

An insight by Fedor Pakhomov: for every true Π^0_1-sentence P we have $U \rhd (\mathsf{R} + P)$. So, *a fortiori*, $U \rhd (\mathsf{R} + \mathsf{con}(\alpha_U))$.

Fedor produced a very natural consistent theory W such that $W \vdash \mathsf{con}(\alpha_W)$.

Philosophical Remarks 1

The question *What is it for an arithmetical statement to express the consistency of a theory?* has two faces:

a. What is it for an arithmetical statement *with the standard semantics* to express consistency?
b. What is it *about a theory* that makes an arithmetical statement express consistency?

My two cents in this discussion are that (b) is a red herring. Nothing in a merely formal theory makes anything express anything. It is *our use* of the formal theory *as a meaningful theory* that makes closed formulas express statements.

Philosophical Remarks 2

The interpretation version can be viewed as stating a relation between a merely formal theory U and $\mathsf{S}^1_2 + \mathsf{con}(\alpha_U)$ *considered as a meaningful theory*. If viewed this way, we need only answer question (a).

If we replace S^1_2 in the statement of the theorem by a weak syntactical theory, we may sidestep the worry about numbers versus syntax.

But which syntax theory to choose? And, even given the syntax theory, there is a question how to codify formulas and proofs? There are still many conventional choices.

Philosophical Remarks 3

In the context of the refutation of Hilbert's Program, Mic Detlefsen has insisted that one should show that no sentence that expresses consistency be provable in the given theory.

The interpretation version we presented above quantifies over sentences that (presumably) express consistency, but is it enough?

- We restricted ourselves to Σ_1^b-axiomatisations.
- We restricted ourselves to a fixed arithmetisation of provability. Can we eliminate this? The quantifier over all reasonable arithmetisations is not a mathematical quantifier.

Overview

A Fefermanian Variation 1

Using a trivial compactness argument, we can boost the interpretation version of G2 to a result of seemingly far greater strength.

Consider a consistent theory U. Let X be the set of Gödel numbers of axioms of U. We do not put any constraints on the complexity of X. Let U_k be the theory axiomatised by X_k, the elements of X that are $\leq k$.

Suppose $N : \mathrm{S}^1_2 \lhd U$. A U-formula ξ *uniformely semi-numerates* X (w.r.t. N) iff, for every n, there is an $m \geq n$, such that U_m proves $\xi(\underline{i})$, for each $i \in X_m$.

Universiteit Utrecht

A Fefermanian Variation 2

Theorem

Suppose ξ uniformly semi-numerates the axioms of U (w.r.t. N). Then, $U \nvdash \mathrm{con}^N[\xi]$.

The square brackets emphasise that ξ is not supposed to be relativised to N.

Sol Feferman gave a celebrated example of a Π^0_1-axiomatisation π^* of PA such that PA $\vdash \mathrm{con}(\pi^*)$. The formula π^* numerates the axioms of PA in PA *but not uniformly*.

Positive examples are versions of oracle provability, like provability with an oracle for Π^0_1-truth. Such forms of provability provide alternative formulations of reflection principles. They play an important role in Beklemishev's program to do proof theory using provability logic.

Overview

Preludium

The Interpretation Version of the Theorem

A Fefermanian Variation

Coordinate Free G2?

Out of the Box Consistency

There are many design choices in the definition of prov_{α_U} for some theory U. We have to choose a treatment of syntax, a proof system, a Gödel numbering both for formulas and proofs, ... Can we somehow eliminate these conventional design choices?

We have seen that $\bigvee_{N:\mathrm{S}^1_2 \lhd U}(\mathrm{con}(\alpha_U))^N$ is not generally first-order definable over U. However, it is meaningful in all U models.

Suppose U is axiomatised by a scheme and is sequential ('has enough coding'). We can define a *coding-free* accessibility relation between U-models, thus obtaining a Kripke model $\mathbb{K}$ with the following property.

Let $\Diamond$ be the possibility operator in $\mathbb{K}$. Then $\Diamond\top$ is equivalent to $\bigvee_{N:\mathrm{S}^1_2 \lhd U}(\mathrm{con}(\alpha_U))^N$. $\Diamond\top$ is dependent on the chosen scheme.

Coordinate Free G2

We have $U \not\models \Diamond\top$.

Remarkably, it seems that the result also works for pair theories, where we need not have interpretations of S^1_2. This uses an idea of Fedor Pakhomov. Many details were provided by Max Bonnet.

So we have a G2-like result where the original G2 cannot go.

The scope of the result is larger than you would think: by a theorem of Vaught, every recursively enumerable pair theory can be axiomatised by a scheme.

Hopefully, the above is a step in the direction of a solution of the *mathematical* problem of abstracting away from implementation details. Does it also have *philosophical* bite? I suspect that it is still a long way to go to find an informally rigorous argument that $\Diamond\top$ is implied by every sentence that expresses consistency.

Open Ended End

We end open-endedly. The philosophical problem of the representation of syntax is studied intensely. There is current work by Balthasar Grabmayer on Gödel numberings. There is work by Volker Halbach amd Graham Leigh on syntax theory … We are clearly still facing many loose ends.

Thank You