Proof mining: Lecture 2

A closer look at the functional interpretation

Thomas Powell

Technische Universität Darmstadt

Autumn School on Proof and Computation

Fischbachau

20 September 2018

The main result of Lecture 1

We went very quickly through the definition of the functional interpretation, in order to Gödel's main result: [CHANGE TO HA^ω AND MAKE CONSISTENT]

Theorem (K. Gödel, 1958)

Suppose that

$$\mathrm{HA} \vdash A$$

Then there exists a term t of System T such that

$$\text{System T} \vdash A_D(t, y)$$

and moreover, we can formally extract t from the proof of A.

In this lecture, we have three main objectives:

- Take a step back and examine the functional interpretation in more detail.
- Explain how to extend Gödel's result to classical logic.
- Give lots of examples!

Outline

Back to prime numbers

Theorem

There are infinitely many prime numbers.

Theorem (Formal version)

For all $m \in \mathbb{N}$ there exists some $p > m$ such that p is prime.

Theorem (Very formal version)

$\forall m \exists p (p > m \wedge \mathtt{Prime}(p))$

How does the functional interpretation treat this formula? It's quite simple:

$$\underbrace{\forall m \exists p (p > m \wedge \mathtt{Prime}(p))}_{A} \mapsto \exists f \forall m (\underbrace{f(m) > m \wedge \mathtt{Prime}(f(m))}_{A_D(f,m)})$$

The soundness proof tells us: Given any proof

$$\mathrm{HA} \vdash \forall m \exists p (p > m \wedge \mathtt{Prime}(p))$$

we can extract a term $t : \mathbb{N} \to \mathbb{N}$ of System T satisfying

$$\forall m (t(m) > p \wedge \mathtt{Prime}(t(m)))$$

Euclid's elementary proof

Let's first consider the ancient Greek proof, which we mentioned in Lecture 1.

Proof (Euclid).

Fix $m \in \mathbb{N}$ and consider the number

$$N := 1 + p_1 \cdots p_k$$

where $p_1, \ldots, p_k$ are all the prime numbers $\leq m$. Then N cannot be divisible by any prime number $\leq m$. But N contains at least one prime factor p, which must therefore be greater than m. □

A formal analysis of this proof would lead to an extracted term t which looks something like

$$t(m) := \text{least } p \leq 1 + p_1 \cdots p_k \text{ such that } p \text{ prime}.$$

We can now use this to bound the size of the mth prime number p_m. We have

$$p_m \leq 1 + p_1 \ldots p_{m-1}$$

and therefore (by induction)

$$p_m < 2^{2^m}$$

A simple instance of 'proof mining'

Theorem (Original)

There are infinitely many prime numbers.

Theorem (Stronger)

There are infinitely many prime numbers, and the mth prime p_m is bounded by 2^{2^m}.

The latter may look like a genuinely new result, but it was actually already 'hidden' in the proof of the original theorem. We just had to make it explicit!

Outline

The formal definition

Let's take another look at the functional interpretation.

- If A is atomic then $A \mapsto A$ i.e. x, y are empty and $A_D := A$.

Suppose that $A \mapsto \exists x \forall y A_D(x, y)$ and $B \mapsto \exists u \forall v B_D(u, v)$. Then

- $A \wedge B \mapsto \exists x, u \forall y, v(A_D(x, y) \wedge B_D(u, v))$
- $A \vee B \mapsto \exists b^0, x, u \forall y, v((b = 0 \to A_D(x, y)) \wedge (b \neq 0 \to B_D(u, v)))$
- $A \to B \mapsto \exists U, Y \forall x, v(A_D(x, Yxv) \to B_D(Ux, v))$
- $\exists z A(z) \mapsto \exists z, x \forall y A_D(x, y, z)$
- $\forall z A(z) \mapsto \exists X \forall z, y A_D(X(z), y, z)$

This all looks reasonable, except perhaps for the interpretation of implication...

The functional interpretation of implication I

Suppose that $A \mapsto \exists x \forall y A_D(x, y)$ and $B \mapsto \exists u \forall v B_D(u, v)$, and consider the implication

$$\exists x \forall y A_D(x, y) \to \exists u \forall v B_D(u, v).$$

We want to bring the quantifiers to the front in the least non-constructive way possible.

Note. For a detailed discussion of this, and the various possibilities, see [Kohlenbach, 2008, Chapter 8].

We describe the functional interpretation of implication using the language of game semantics. The idea here is to visualise the quantifiers as representing a game between two players:

- Eloise(existential quantifier) wants to find evidence that the statement is true;
- Abelard(universal quantifier) tries to confound *Eloise* by claiming that the statement is false.

The functional interpretation of implication II

$$\exists x \forall y A_D(x, y) \to \exists u \forall v B_D(u, v)$$

Let's imagine this as a game between Eloise and Abelard, who are trying to respectively prove and disprove the implication.

- Abelard: I claim that there is a realizer x for the premise, and challenge you to find a realizer for the conclusion.
- Eloise: I accept the challenge, and give you a witness u for the conclusion.
- Abelard: I claim that there is a counterexample v to your witness u.
- Eloise: In which case, I give you a counterexample y to your original witness x.

The formula is true if Eloisehas a winning strategy against any choices from Abelard.

The functional interpretation of implication III

For any witness challenge x from Abelard

$$\forall x(\forall y A_D(x,y) \to \exists u \forall v B_D(u,v))$$

there is a witness response u from Eloise

$$\forall x \exists u(\forall y A_D(x,y) \to \forall v B_D(u,v))$$

such that for any counterexample challenge v from Abelard

$$\forall x \exists u \forall v(\forall y A_D(x,y) \to B_D(u,v))$$

there is a counterexample response y from Eloise

$$\forall x \exists u \forall v \exists y(A_D(x,y) \to B_D(u,v))$$

Now we convert these to functions:

$$\exists U, Y \forall x, v(\underbrace{A_D(x, Y(x,v)) \to B_D(U(x),v)}_{(A \to B)_D(U,Y,x,v)})$$

Outline

Euler's analytic proof I

We will now analyse a more subtle proof that there are infinitely many prime numbers, which uses basic facts from analysis.

Proof (Euler).

Suppose there are only finitely many prime numbers $p_1, \ldots, p_m$. We have (by simple combinatorics)

$$\sum_{0 \leq k_1, \ldots, k_m \leq n} \frac{1}{p_1^{k_1} \cdots p_m^{k_m}} = \left(\sum_{i=0}^{n} \frac{1}{p_1^i} \right) \cdots \left(\sum_{i=0}^{n} \frac{1}{p_m^i} \right)$$

But using

$$\sum_{i=0}^{n} \frac{1}{p^i} < \sum_{i=0}^{\infty} \frac{1}{p^i} = \frac{p}{p-1}$$

we have

$$\begin{aligned} \sum_{0 \leq k_1, \ldots, k_m \leq n} \frac{1}{p_1^{k_1} \cdots p_m^{k_m}} &< \frac{p_1}{p_1 - 1} \cdot \ldots \cdot \frac{p_m}{p_m - 1} \\ &\leq \frac{2}{1} \cdot \frac{3}{2} \cdot \ldots \cdot \frac{p_m}{p_m - 1} \\ &= p_m \end{aligned}$$

□

Euler's analytic proof II

Proof cont...

We have shown that

$$\sum_{0 \leq k_1, \ldots, k_m \leq n} \frac{1}{p_1^{k_1} \cdots p_m^{k_m}} < p_m$$

for any n. Now using the prime factorisation theorem, it follows that

$$\sum_{i=1}^{n} \frac{1}{i} \leq p_m$$

for all n, contradicting the fact that

$$\sum_{i=1}^{\infty} \frac{1}{i} = \infty$$

Therefore there are infinitely many primes! □

The structure of Euler's proof

Euler's proof resorts to a crucial **Lemma**, namely that the sequence $\sum_{i=1}^{\infty} \frac{1}{i}$ diverges. This is a well-known fact, but is technically not part of the proof itself. Rather, Euler has shown the following:

Theorem

Assuming that $\sum_{i=1}^{\infty} \frac{1}{i}$ diverges then there are infinitely many primes.

Theorem (Formal version)

$$\forall m \exists k \left(\sum_{i=1}^{k} \frac{1}{i} > m \right) \rightarrow \forall n \exists p (p > n \wedge \texttt{Prime}(p))$$

The functional interpretation of this statement is as follows:

$$\exists \Phi, f \forall g, n \left(\sum_{i=1}^{g(\Phi(g,n))} \frac{1}{i} > \Phi(g,n) \rightarrow f(g,n) > n \wedge \texttt{Prime}(f(g,n)) \right)$$

An analysis of Euler's proof yields very simple computational information: We have

$$\forall g, n \left(\sum_{i=1}^{g(n)} \frac{1}{i} > n \to t(g,n) > n \wedge \texttt{Prime}(t(g,n)) \right)$$

for

$$t(g,n) := \text{least } p \leq g(n) \text{ such that } p \text{ prime}$$

In other words, the compuational content of the proof is a procedure:

$$\text{Rate of divergence of} \sum_{i=1}^{\infty} \tfrac{1}{i} \mapsto \text{Bound on the } m\text{th prime}$$

An explicit rate of divergence is given by

$$g(n) = \lceil e^{m-\gamma} \rceil$$

where $\gamma \approx 0.5772$ is the so-called Euler-Mascheroni constant.

Remark. Euler's proof provides a better bound the Euclid's! This is an important phenomenon in proof mining.

For a more detailed discussion of this example, see [Kohlenbach, 2008, Chapter 2].

Outline

The minimum principle

Theorem

Let $f : \mathbb{N} \to \mathbb{N}$ be a function. There exists some $n \in \mathbb{N}$ such that $\forall m(f(n) \leq f(m))$.

Proof.

Suppose that this were not the case. Then for any n there would exists some m with $f(n) > f(m)$.

Define the sequence (x_i) by

$$x_0 := n \text{ and } x_{i+1} \text{ satisfies } f(x_i) > f(x_{i+1})$$

Then we have an infinite decreasing sequence

$$f(x_0) > f(x_1) > f(x_2) > \ldots$$

which contradicts the wellfoundedness of $\mathbb{N}$. □ □

Some theorems are just noncomputable

Theorem

There is no computable functional $\Phi : (\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$ *which satisfies*

$$(*) \quad \exists n \leq \Phi(f) \forall m (f(n) \leq f(m)).$$

Proof.

Suppose that one did exist, and define $f = \mathbf{1}$ i.e. f is the constant 1-function. Since Φ is computable, it only looks at a finite amount of its input i.e. there exists some N such that

$$(\dagger) \quad \forall g : \mathbb{N} \to \mathbb{N} (\forall i \leq N (g(i) = 1) \to \Phi(g) = \Phi(\mathbf{1}))$$

Now define

$$h(n) := \begin{cases} 1 & \text{if } n \leq \max\{N, \Phi(\mathbf{1})\} \\ 0 & \text{otherwise} \end{cases}$$

- Then $\forall i \leq N(h(i) = 1)$ and so $\Phi(h) = \Phi(\mathbf{1})$ by $(\dagger)$.
- But by $(*)$ we have $\exists n \leq \Phi(\mathbf{1})(g(n) \leq 0)$

But $g(n) = 1$ for all $n \leq \Phi(\mathbf{1})$, a contradiction. □

There is no classical functional interpretation

It is impossible to extend the functional interpretation to classical logic.

If it were, then since PA^ω proves

$$\forall f \exists n \forall m (f(n) \leq f(m))$$

we would expect to extract a term t of System T satisfying

$$\forall f, m (f(t(f)) \leq f(m))$$

Therefore, in particular, there would be a computable functional $\Phi(f) := t(f)$ satisfying

$$\exists n \leq \Phi(f) \forall m (f(n) \leq f(m))$$

which we just demonstrated was not possible.

Outline

The minimum principle revisited

Theorem

$\forall f \exists n \forall m (f(n) \leq f(m))$.

Proof (computational version).

Suppose that the statement is false, in other words

$$\exists f \forall n \exists m (f(n) > f(m)).$$

Then in particular there must exists a function $g : \mathbb{N} \to \mathbb{N}$ such that

$$\forall n (f(n) > f(g(n))).$$

But this means that

$$f(0) > f(g(0)) > f(g^{(2)}(0)) > \ldots > f(g^{(k)}(0))$$

which is a contradiction for any $k > f(0)$. Therefore we have

$$f(n) \leq f(g(n))$$

where n is one of $g(0), g^{(2)}(0), \ldots, g^{(f(0))}(0)$, and so the original statement must be true. □

A computational minimum principle

In general there is no computable functional Φ such that

$$\forall f \exists n \leq \Phi(f) \forall m (f(n) \leq f(m)).$$

However, we *can* find a functional Φ such that

$$\forall f, g \exists n \leq \Phi(f, g) (f(n) \leq f(g(n))).$$

namely:

$$\Phi(f, g) := \max\{g(0), g^{(2)}(0), \ldots, g^{(f(0))}(0)\}$$

This is a witness for the following reformulated version of the minimum priniple:

$$\forall f, g \exists n (f(n) \leq f(g(n)))$$

Reformulated principles

The original minimum principle $\forall f \exists n \forall m (f(n) \leq f(n)))$ asserts:

> *There exists an 'ideal' minimal element n, such $f(n) \leq f(m)$* $\boxed{\textit{for all } m}$

The reformulated minimum principle $\forall f, g \exists n (f(n) \leq f(g(n)))$ asserts:

> $\boxed{\textit{For any function } g}$ *there exists an approximation n to a minimum element, such that $f(n) \leq f(g(n))$.*

Key idea.

- We may not be able to compute ideal objects whose existence relies on classical logic, but we can compute approximations to those ideal objects.
- Functionals which compute these approximations can be formally extracted from the classical proof.

What is really going on here?

We are seeing the following phenomenon.

- We cannot compute direct witnesses for existential statements proven using classical logic.
- We can compute witnesses for the 'not not' version of these statements.
- The latter can be viewed as approximatations to the former.

What is going on in general?

Outline

The Gödel-Gentzen negative translation

Let A be a formula in predicate logic. We define the negative translation of A by

$$A^N :\equiv \neg\neg A^*$$

where A^* is defined inductively as

$$\begin{aligned} A^* &:\equiv A \text{ if } A \text{ is a prime formula} \\ (A \square B)^* &:\equiv A^* \square B^* \text{ if } \square \in \{\wedge, \vee, \to\} \\ (\exists x A)^* &:\equiv \exists x A^* \\ (\forall x A)^* &\equiv \forall x \neg\neg A^* \end{aligned}$$

Soundness of the negative translation

The negative translation obeys the following general pattern: Suppose that

$$\mathcal{P}_{\text{class}} \vdash A$$

for some classical theory $\mathcal{P}_{\text{class}}$. Then

$$\mathcal{P} \vdash A^N$$

where $\mathcal{P}$ is the intuitionistic version of that theory.

In particular, this is true for Peano/Heyting arithmetic.

Theorem
If $\text{PA}^\omega \vdash A$ *then* $\text{HA}^\omega \vdash A^N$.

Proof.
Induction over the structure of derivations in PA^ω. □

The negative translation of $\forall\exists\forall$ formulas

Suppose that $A :\equiv \forall k \exists n \forall m P(k,n,m)$ for $P(k,n,m)$ quantifier-free. Then

$$\begin{aligned} A^N &\equiv \neg\neg A^* \\ &\equiv \neg\neg(\forall k \exists n \forall m P(k,n,m))^* \\ &\equiv \neg\neg\forall k \neg\neg(\exists n \forall m P(k,n,m))^* \\ &\equiv \neg\neg\forall k \neg\neg\exists n \forall m \neg\neg P(k,n,m). \end{aligned}$$

This looks complicated, but in arithmetic we have

$$\neg\neg Q \leftrightarrow Q$$

for all quantifier-free formulas, and

$$\neg\neg\forall k \neg\neg B \leftrightarrow \forall k \neg\neg B$$

is provable intuitionistically. Therefore

$$A^N \leftrightarrow \forall k \neg\neg\exists n \forall m P(k,n,m).$$

and so

$$\mathrm{PA} \vdash \forall k \exists n \forall m P(k,n,m) \Rightarrow \mathrm{HA} \vdash \forall k \neg\neg\exists n \forall m P(k,n,m).$$

The classical functional interpretation

We cannot give a direct computational interpretation to classical arithmetic i.e. it is *not* the case that

$$\text{if } \mathrm{PA}^\omega \vdash A \text{ then } \mathrm{HA}^\omega \vdash \forall y A_D(t, y)$$

for some $t \in \mathsf{T}$. However, what we do have is:

- A computational interpretation of Heyting arithmetic
- An embedding of Peano arithmetic into Heyting arithmetic

So why not combine them? I.e.

$$\mathrm{PA}^\omega \mapsto \mathrm{HA}^\omega \mapsto \text{System } \mathsf{T}$$

Gödel's main theorem (second part)

Gödel's soundness theorem for *classical logic* says that we can translate a proof of A to a program witnessing $\exists x \forall y (A^N)_D(x, y)$.

Theorem (K. Gödel, 1958)

Suppose that

$$\mathrm{PA}^\omega \vdash A$$

Then there exists a term t of System T such that

$$\mathsf{T} \vdash (A^N)_D(t, y)$$

and moreover, we can formally extract t from the proof of A.

Proof.

Combine the soundness theorem for intuitionistic logic with the negative translation. □

The classical functional interpretation of $\forall\exists\forall$ theorems

What is the functional interpretation of $B :\equiv \forall k \neg\neg \exists n \forall m P(k,n,m)$?

Recalling the interpretation of implication we have

$$\begin{aligned}\forall k \neg\neg \exists n \forall m P(k,n,m) &\mapsto \forall k \neg(\exists n \forall m P(k,n,m) \to \bot) \\ &\mapsto \forall k \neg \exists g \forall n \neg P(k,n,g(n)) \\ &\mapsto \forall k (\exists g \forall n \neg P(k,n,g(n)) \to \bot) \\ &\mapsto \exists \Phi \forall k, g P(k, \Phi(k,g), g(\Phi(k,g)))\end{aligned}$$

Therefore in the special case of theorems of this form, we have

$$\text{if } \mathrm{PA}^\omega \vdash \forall k \exists n \forall m P(k,n,m) \text{ then } \mathrm{HA}^\omega \vdash \forall k, g P(k, t(k,g), g(t(k,g)))$$

for some term t if System T.

We can equivalently view this as a bound i.e.

$$\mathsf{T} \vdash \forall k, g, \exists n \leq t(g,k) P(k,n,g(n))$$

We now see what was going on with the minimum principle!

Outline

The classical functional interpretation of $\forall\exists$ statements

Suppose that $\mathrm{PA}^\omega \vdash B$ where $B :\equiv \forall u \exists v Q(u,v)$. What does the classical functional interpretation do in this case?

Let's first look at the negative translation. We have

$$B^N \equiv \neg\neg(\forall u \exists v Q(u,v)) \equiv \neg\neg\forall u \neg\neg\exists v \neg\neg Q(u,v) \leftrightarrow \forall u \neg\neg\exists v Q(u,v)$$

where the equivalence $\leftrightarrow$ is possible in Heyting arithmetic. Therefore

$$\mathrm{HA} \vdash \forall u \neg\neg\exists v Q(u,v)$$

But what is the functional interpretation of this? We have

$$\begin{aligned} \forall u \neg\neg\exists v Q(u,v) &\mapsto \forall u \neg\exists v \neg Q(u,v) \\ &\mapsto \forall u \exists v \neg\neg Q(u,v) \\ &\mapsto \exists f \forall u Q(u, f(u)). \end{aligned}$$

But this is the same as the direct, intuitionistic functional interpretation!

Remark. What really going on here is that the functional interpretation admits Markov's principle $\neg\neg\exists x A_0(x) \to \exists x A_0(x)$ for any quantifier-free formula $A_0(x)$.

Program extraction theorem

Theorem

Suppose that

$$\mathrm{PA}^\omega \vdash \forall u \exists v Q(u, v).$$

Then there exists a term t of System T such that

$$\mathrm{HA}^\omega \vdash \forall u Q(u, tu)$$

and moreover, we can formally extract t from the proof of A.

In other words, for the special case of $\forall\exists$ theorems, we can extract a **direct** witness from their proof, even if their proof uses non-constructive reasoning and therefore doesn't seem to have any computational meaning.

At this point, you should have two burning questions:

- **How is the program extraction theorem even possible?**
- **Can we use it to find new quantitative results in mathematics?**

All will be revealed in **Lecture 3**.

Outline

References I

Kohlenbach, U. (2008).
Applied Proof Theory - Proof Interpretations and their Use in Mathematics.
Springer Monographs in Mathematics. Springer.