

Computing with Gödel's completeness theorem: Weak Fan Theorem, Markov's Principle and Double Negation Shift in action

(with some typos fixed)

Hugo Herbelin

Thierry's 60th birthday

Göteborg, 25 August 2022

Computing with completeness proofs

1990-1995: PhD thesis in Paris, remotely with Thierry in Göteborg, visiting Chalmers from time to time.

One day, in the “old” “historiska” Chalmers CS building, Thierry shared with me some of his thoughts on logical completeness: if you compose soundness and (a special normalising form of) completeness, you may prove cut-elimination. What’s the computational content of that?



Computing with completeness proofs

In LICS 1991, Ulrich Berger and Helmut Schwichtenberg initiate the idea of Normalisation-by-Evaluation.

In TLCA 1993, Ulrich Berger presents a paper on Program Extraction from Normalisation Proofs.

In CSL 1993, Catarina Coquand presents a formalised normalisation proof of simply-typed λ -calculus by transiting to Kripke semantics.

In 1996, Jean-Louis Krivine gives a detailed purely constructive proof of Gödel's completeness theorem (i.e. wrt Tarski semantics) obtained by double-negation translation.

Then, a lot of a lot of works on constructive proofs of completeness wrt diverse “informative-enough” semantics: Kripke, Beth (Thierry, Jan Smith), point-free topology (Giovanni Sambin), phase semantics (Mitsuhiro Okada), Heyting algebras (Jim Lipton, Olivier Hermant), “glued” semantics (Thierry, Peter Dybjer), ... with applications to normalisation, with various connectives or features (Thorsten Altenkirch, Martin Hofmann, Philip Scott, Andreas Abel, Christian Sattler, ...). Works also on the analysis of reducibility/realisability/logical-relation proofs as “adequacy lemma \circ escape lemma”, as well as on type-directed partial evaluation (Olivier Danvy).

The case of Tarski semantics

Compared to “informative-enough” semantics such as Kripke, Beth, phase semantics (we can consider them as “effectful”, in the sense of “carrying a state”), Tarski semantics is minimalistic. It simply replicates object syntax in the meta-language level:

$$\begin{aligned} \llbracket P(t_1, \dots, t_{a_f}) \rrbracket_{\mathcal{M}}^{\sigma} &\triangleq ([t_1]_{\mathcal{M}}^{\sigma}, \dots, [t_{a_f}]_{\mathcal{M}}^{\sigma}) \in \mathcal{M}_{\mathcal{P}}(P) \\ \llbracket \perp \rrbracket_{\mathcal{M}}^{\sigma} &\triangleq \perp \\ \llbracket A \dot{\rightarrow} B \rrbracket_{\mathcal{M}}^{\sigma} &\triangleq \llbracket A \rrbracket_{\mathcal{M}}^{\sigma} \Rightarrow \llbracket B \rrbracket_{\mathcal{M}}^{\sigma} \\ \llbracket \forall x A \rrbracket_{\mathcal{M}}^{\sigma} &\triangleq \forall v \in \mathcal{M}_{\mathcal{D}} \llbracket A \rrbracket_{\mathcal{M}}^{\sigma \cup [x \leftarrow v]} \\ \mathcal{T} \models A &\triangleq \forall \mathcal{M} \forall \sigma (\llbracket \mathcal{T} \rrbracket_{\mathcal{M}}^{\sigma} \Rightarrow \llbracket A \rrbracket_{\mathcal{M}}^{\sigma}) \end{aligned}$$

In spite of this sobriety (or weakness, depending on taste), Gödel, followed by many others (Henkin, Hasenjaeger, Beth, Hintikka, Kangers, Schütte , ...) could prove (in a classical metalanguage):

$$\mathcal{T} \models A \Rightarrow \mathcal{T} \vdash_{class} A$$

But while the proofs wrt “rich” semantics are mostly structural and are constructively mapping validity proofs into (normalised) object-syntax proofs reifying the validity proofs (thus supporting normalisation-by-evaluation), completeness proofs wrt Tarski semantics look more complicated...

First, a restriction for the rest of the talk

Completeness is commonly expressed under one of these forms:

- proof existence: $\mathcal{T} \models A \Rightarrow \mathcal{T} \vdash_{class} A$
- model existence: $\mathcal{T} \not\vdash_{class} \perp \Rightarrow \exists \mathcal{M} \exists \sigma \llbracket \mathcal{T} \rrbracket_{\mathcal{M}}^{\sigma}$
- $\mathcal{T} \vdash_{class} A \vee \exists \mathcal{M} \exists \sigma \llbracket \mathcal{T} \rrbracket_{\mathcal{M}}^{\sigma} \wedge \llbracket \neg A \rrbracket_{\mathcal{M}}^{\sigma}$

From the constructive point of view, each statement has its own specificities. We focus for the talk on the first form.

Specificities of completeness wrt Tarski semantics (proof existence form)

Apparently requires non-constructive features:

- requires Markov's principle (MP, i.e. $\neg\neg A \Rightarrow A$ for $A \in \Sigma_1^0$) according to Kurt Gödel (1957) and Georg Kreisel (1962)
- requires the law of excluded-Middle (LEM) according to Charles McCarty (2004) or Christian Espíndola (2016)
- requires the Ultrafilter Theorem according to Leon Henkin or more recently Christian Espíndola (2016)
- requires Weak König's Lemma (WKL) according to Stephen Simpson's textbook on the classical reverse mathematics of the subsystems of second order arithmetic
- requires the Fan Theorem as suggested by Harvey Friedman (1975) and Wim Veldman (1976) and shown by Victor Krivtsov (2014)
- but actually requires no classical reasoning at all according to Jean-Louis Krivine (1996)!

Also apparently crucially rely on an enumeration of formulas (or to appeal to strong principles such as the ultrafilter theorem) and to not preserve the structure of the validity proof given as input. Why is it so? How crucial is it?

How to sort this out?

A first clarification by Stefano Berardi (1999) and Berardi-Valentini (2001):
connectives have an impact on the logical strength

In the presence of falsity

If object-level $\dot{\perp}$ is interpreted as meta-level \perp , Markov's principle is required (formalised e.g. in Kirst-Forster-Wehr, 2020)

\hookrightarrow This explains how Krivine bypasses the need for Markov's principle (no $\dot{\perp}$ in his language)

\hookrightarrow The observation holds also for Beth and Kripke semantics and explains how Friedman (wrt Beth models) and Veldman (wrt Kripke models) bypass the need for Markov's principle

In the presence of disjunction

If object-level $\dot{\vee}$ is interpreted as meta-level \vee , the proof cannot be constructive

Based on Wim Veldman's completeness proof wrt Kripke semantics (whose interpretation of $\dot{\vee}$ is Tarskian), the weakly classical part (called L_{fan} in Josef Berger) of the Weak Fan Theorem will be required.

Eventually, it will happen that it can be characterised as a generalised form of Double Negation Shift (DNS_T).

Note: The above assumes either the presence of \forall , \exists , or of infinitely many atoms (otherwise, we are in propositional logic which is decidable).

More on the need for Markov's principle

Remind from linear logic the existence of two *falsity* connectives:

- 0 is the neutral element of the \oplus (additive, positive) disjunction (i.e. the intuitive disjunction):

$$\frac{}{\Gamma, 0 \vdash \Delta} 0_L \quad \text{no right introduction rule}$$

its semantics is the *falsity* connective of the metalanguage

- \perp is the neutral element of the \wp (multiplicative, negative) disjunction:

$$\frac{}{\perp \vdash} \perp_L \quad \frac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, \perp} \perp_R$$

its semantics is an arbitrary formula of the metalanguage (leading to so-called “exploding” models)

In particular: it is the presence of 0 in the language (the positive form of *falsity* which forces completeness to imply Markov's principle.

This suggests the slogan: it is the *positive* connectives which impact the logical strength.

(Incidentally, this also suggests that the same analysis should be done for a language highlighting the decomposition of connectives in linear logic, and within a meta-language that also makes explicit the decomposition of connectives given by linear logic.)

Second clarification: the need for excluded-middle

How can it be that McCarty and Espíndola derive **LEM** from completeness?

Answer: This is about the logical complexity of the theory under consideration.

If the theory \mathcal{T} is allowed to be described by a formula of complexity S (typically Σ_1^0 , but possibly another complexity), and *falsity* is interpreted like 0, then $\neg\neg A \Rightarrow A$ can be derived for any A of complexity S .

For instance, completeness for a recursively enumerable theory when *falsity* is interpreted in the standard (positive) way cannot prove more than Markov's principle (Kreisel).

To fix things, we now restrict ourselves to recursively enumerable theories.

Third clarification: Ultrafilter Theorem vs Weak König's Lemma vs Weak Fan Theorem

Stefano Berardi and Silvio Valentini (2001) again give an answer: exploiting Krivine's fully constructive proof, they were able to give a constructive proof of the Ultrafilter Theorem on countable domains.

Otherwise said: crossing with Stephen Simpson's result, Weak König's Lemma can (classically) be reinterpreted as the countable restriction of the Ultrafilter Theorem.

Digression on the logical structure of choice axioms

Exploiting the property that **WKL** lies at the intersection of the Ultrafilter Theorem and the axiom of Dependent Choices, Brede-Herbelin (2021) developed a generic view at choice and bar induction axioms connecting an “intensional” (or “operational”, “finite”, “actual”, “effective”) view and an “extensional” (or “observational”, “ideal”, “potential”) view. They develop a “generalised dependent choice” scheme GDC_{ABT} over functions from A to B satisfying some filter T :

$$\underbrace{T \text{ coinductively } A\text{-}B\text{-approximable}}_{\text{effective}} \implies \underbrace{T \text{ has an } A\text{-}B\text{-choice function}}_{\text{observational}}$$

such that

$\text{GDC}_{\mathbb{N}BT}$ is logically equivalent to Dependent Choices in B

$\text{GDC}_{A\mathbb{B}oolT}$ is logically equivalent to the Ultrafilter Theorem on A

$\text{GDC}_{\mathbb{N}\mathbb{B}oolT}$ is logically equivalent to Weak König's Lemma

GDC_{ABT} for “unary” filter T is logically equivalent to the full Axiom of Choice from A to B

and, a dual (asynchronous) “generalised bar induction” scheme GBI_{ABT} over functions from A to B satisfying some filter T :

$$\underbrace{T \text{ } A\text{-}B\text{-barred}}_{\text{observational}} \implies \underbrace{T \text{ } A\text{-}B\text{-inductively barred}}_{\text{effective}}$$

such that

$\text{GBI}_{\mathbb{N}BT}$ is logically equivalent to Bar Induction and to countable Zorn's Lemma over B

$\text{GBI}_{\mathbb{N}\mathbb{B}oolT}$ is logically equivalent to Weak Fan Theorem

Towards a fourth clarification: **WKL** vs **WFT** and why incompatible equivalences with **WFT**

A glitch remains:

- Jean-Louis Krivine proved completeness wrt Tarski semantics fully constructively (reasoning in PA_2)
- after Harvey Friedman and Wim Veldman, Victor Krivtsov showed that the Weak Fan Theorem is needed (reasoning in a variant of Weak Kleene-Wesley system)
- but after Josef Berger, the Weak Fan Theorem is known to include a bit of classical logic.

How is this possible? It is that not everyone is talking about the same formulation of Weak König's Lemma and Weak Fan Theorem.

- **WKL** and **WFT** can be formulated in PA_2 in which case they are *intuitionistically* provable¹
- when paths are functions to \mathbb{Bool} , **WKL** and **WFT** include a bit of classical logic (respectively LLPO, i.e. de Morgan's law for Σ_1^0 -formulas, and DNS)

¹ for decidable trees in the WKL case

Propositional, Decidably Propositional, and Functional Weak Fan Theorem

There are **three distinct** possible definitions of a **set** of natural numbers in (second-order) constructive logic:

- a **subset**:

$$P : \omega \rightarrow Prop$$

- a **functional relation** mapping formulas to Booleans

$$R : \omega \times Bool \rightarrow Prop \text{ such that } \forall n \exists! b R(n, b)$$

or, equivalently, a **decidable subset** of formulas:

$$P : \omega \rightarrow Prop \text{ such that } \forall n (n \in P \vee n \notin P)$$

- a **function to Bool**

$$f : \omega \rightarrow Bool$$

This gives in turn **three different** kinds of **comprehension axioms**.

Contrastingly, in classical logic, the first two formulations cannot be distinguished.

How the three representations of sets relate?

Obviously:

$$\omega \rightarrow \mathbb{Bool}$$

$$\Downarrow$$

$$R : \omega \times \mathbb{Bool} \rightarrow Prop \text{ such that } \forall n \exists! b R(n, b)$$

$$\Downarrow$$

$$\omega \rightarrow Prop$$

Map $f : \omega \rightarrow \mathbb{Bool}$ to $R(n, b) \triangleq (f(n) = b)$ which is trivially functional

Map $R : \omega \times \mathbb{Bool} \rightarrow Prop$ to $X(n) \triangleq R(n, \text{true})$

How the three representations of sets relate?

And also:

$$\omega \rightarrow \mathbb{Bool}$$

$$\mathbf{AC}^!_{\mathbb{N}, \mathbb{Bool}} \uparrow$$

$$R : \omega \times \mathbb{Bool} \rightarrow Prop \text{ such that } \forall n \exists! b R(n, b)$$

$$\mathbf{LEM} \uparrow$$

$$\omega \rightarrow Prop$$

Map $X : \omega \rightarrow Prop$ to $R(n, b) \triangleq (b = \mathbf{true} \Leftrightarrow X(n))$, this is functional by **LEM**

Map $\forall n \exists! b R(n, b)$ to a function by unique choice.

Propositional, Decidably Propositional, and Functional Weak Fan Theorem

Let T be an arbitrary predicate on $\mathbb{B}ool^*$ (finite sequences of Booleans)

$$\mathbf{WFT}_{fun} \triangleq \forall f \exists n T(f|_n) \Rightarrow T^*$$

$$\mathbf{WFT}_{fun-rel} \triangleq \forall R \text{ functional } \exists l (l \approx R \wedge T(l)) \Rightarrow T^*$$

$$\mathbf{WFT}_{pred} \triangleq \forall X \exists l (l \approx X \wedge T(l)) \Rightarrow T^*$$

where $l \approx_n X$ (resp. $l \approx_n R$) expresses that l approximates the n first “values” of X (resp R):

$$\frac{}{\epsilon \approx X} \quad \frac{l \approx X \quad X(|l|)}{l \cdot \text{true} \approx X} \quad \frac{l \approx X \quad \neg X(|l|)}{l \cdot \text{false} \approx X}$$

$$\frac{}{\epsilon \approx R} \quad \frac{l \approx R \quad R(|l|, b)}{l \cdot b \approx R} \quad \begin{array}{l} f|_0 \triangleq \epsilon \\ f|_{n+1} \triangleq f|_n \cdot f(n) \end{array}$$

and

$$T^* \triangleq \exists N \forall l (|l| = N \Rightarrow \exists l' \subset l T(l')) \quad (= \text{uniformly barred})$$

Note: We do not care here about the logical complexity of T

Three forms of Weak Fan Theorem

Thus we have:

$$\mathbf{WFT}_{fun} \Rightarrow \mathbf{WFT}_{fun-rel} \Rightarrow \mathbf{WFT}_{pred}$$

While \mathbf{WFT}_{fun} (considered in intuitionistic reverse mathematics) is told equivalent to the full Fan Theorem on finite (non-necessarily binary) “trees” (Iris Loeb 2005), $\mathbf{WFT}_{fun-rel}$ and \mathbf{WFT}_{pred} are not equivalent to the corresponding formulation of the full Fan Theorem (based on Stephen Simpson’s book).

\mathbf{WFT}_{pred} is intuitionistically provable (in \mathbf{PA}_2) and is enough to constructively prove completeness in the presence of \Rightarrow , \wedge , \forall (over recursively enumerable theories) (Jean-Louis Krivine 1996) while disjunction requires a stronger version (Wim Veldman 1976, Victor Krivtsov 2014).

What do we need for disjunction?

Summary of complementary results:

- Stefano Berardi and Silvio Valentini shows that disjunction cannot be treated purely intuitionistically.
- Wim Veldman (for Kripke semantics) and Victor Krivtsov (for Tarski semantics) have a proof using WFT_{fun} .
- WFT_{pred} is intuitionistically provable thus not sufficient.
- Josef Berger decomposes WFT_{fun} into a weak classical axiom L_{fan} and a pure choice axiom C_{fan} . Can L_{fan} be the missing piece?

Moreover:

- Herbelin-Ilik (2022), inspired by Andreas Abel and Christian Sattler (2019), handled disjunction by using DNS_T , where:

(Double Negation Shift) $\text{DNS} \triangleq \forall n \neg\neg A(n) \Rightarrow \neg\neg \forall n A(n)$

(Generalised DNS) $\text{DNS}_T \triangleq \forall n ((A(n) \Rightarrow T) \Rightarrow T) \Rightarrow ((\forall n A(n)) \Rightarrow T) \Rightarrow T \quad (T \in \Sigma_1^0)$

The purpose of Generalised DNS (Danko Ilik, 2011) is to provide DNS even in contexts where MP does not hold (otherwise $\text{MP} + \text{DNS}_T \Rightarrow \text{DNS}$).

- Dominik Kirst was also able to handle disjunction in completeness for a modal logic using the Weak Law of Excluded-Middle (WLEM) or DNS (private communication, 2022).

How to express that $\text{WFT}_{fun-rel}$ adds DNS_T to WFT_{pred} ?

We don't a priori have $\text{WFT}_{pred} + \text{DNS}_T = \text{WFT}_{fun-rel}$ but we can prove:

$$\text{WFT}_{pred}^{pseudo\neg\neg} \xrightleftharpoons[\leftarrow]{\text{DNS}_T} \text{WFT}_{fun-rel}^{pseudo\neg\neg}$$

and, suspectingly (by reasoning in the theory of arithmetic), even:

$$\text{WFT}_{pred}^{pseudo\neg\neg} + \text{DNS}_T = \text{WFT}_{fun-rel}^{pseudo\neg\neg}$$

where

$$\begin{aligned} \text{WFT}_{fun-rel}^{pseudo\neg\neg} &\triangleq \forall X \text{ decidable}_{T^*} \neg_{T^*} \neg_{T^*} \exists l (l \approx X \wedge T(l)) \Rightarrow T^* \\ \text{WFT}_{pred}^{pseudo\neg\neg} &\triangleq \forall X \neg_{T^*} \neg_{T^*} \exists l (l \approx X \wedge T(l)) \Rightarrow T^* \end{aligned}$$

and

$$\begin{aligned} \neg_S A &\triangleq A \Rightarrow S \\ X \text{ decidable}_S &\triangleq \forall n (X(n) \vee (X(n) \Rightarrow S)) \end{aligned}$$

(This is to be contrasted with Josef Berger's decomposition of WFT_{fun})

Final summary of results

- When reasoning in a weak intuitionistic second-order arithmetic, we expect the bare logic (e.g. Scott's entailment relations) and the logic with \forall , \wedge , \Rightarrow , linear-logic \perp to exactly require $\mathbf{WFT}_{pred}^{pseudo\neg\neg}$ (for all of Tarski, Kripke and Beth completeness).

Currently, we only know that we require a principle classically equivalent to it.

- Linear-logic 0 additionally requires **MP** (and more generally S -**LEM** for a theory of logical complexity S)
- \mathbf{DNS}_T is enough to additionally handle \vee (for Tarski or Kripke completeness, when \forall and \exists are present, or when a non finite theory is given).

We conjecture that the full power of \mathbf{DNS}_T is required (deriving \mathbf{DNS}_T by applying completeness to a theory encoding arithmetic).

Note: Surprisingly \exists requires only \mathbf{WFT}_{pred} for Tarski completeness (but suspectingly \mathbf{DNS}_T for Kripke completeness).

Computing with completeness wrt Tarski semantics in the presence of positive falsity and disjunction

Importantly, **MP** and DNS_T preserve the disjunction and existence properties (i.e. the constructive meaning of \vee and \exists).

Also, both have known computational contents:

- unbounded search (Kleene) or exceptions (Herbelin 2010, justified by Thierry Coquand and Martin Hofmann's generalisation of Friedman's A-translation) for **MP** (see also Pierre-Marie Pédrot, 2019 for type theory)
- bar recursion (Clifford Spector 1976) or delimited control (Danko Ilik 2011) for DNS_T

Computing with completeness wrt Tarski semantics

It seems that there are basically two kinds of completeness proofs wrt Tarski models:

- Henkin style: the data that a formula A belongs to the maximal extension of a consistent context computationally corresponds to giving a continuation reducing an inconsistency in the presence of A to an inconsistency without A .

Eventually, the resulting derivation has roughly the form of a resolution proof (i.e. a tree of cuts inferring a contradiction from the axioms of the theory).

See Herbelin-Ilik (draft, 2016, lastly revised 2022) for a detailed presentation highlighting the similarity with reify/reflect-based normalisation-by-evaluation.

Two versions are given, one without **DNS** for \forall , \rightarrow , (negative) \perp , \wedge , \exists and one with **DNS** supporting also disjunction.

- Beth-Hintikka-Kanger-Schütte style, building an “universal” infinite proof of $\mathcal{T} \vdash A$ of which only the needed steps are eventually kept, according to what the proof of validity says, ending in a finite proof.

Another direction is to observe that Kripke semantics is obtained by forcing from Tarski semantics and to use a memory effect to simulate forcing in direct style: we have then the illusion to prove completeness with respect to Tarski semantics but, under the hood, the nice structural proof underlying completeness wrt Kripke semantics is used!

Recapitulation

- Following Thierry (and Catarina), an analysis of the constructive content of completeness proofs.
- Following Stefano Berardi, an analysis of the weakly classical principles needed to handle positive connectives interpreted à la Tarski in completeness proofs:
 - positive falsity requires Markov Principle,
 - disjunction requires (Generalized) Double Negation Shift.
- A clarification of an abundant and apparently contradictory literature on the topic.
- Keeping better in mind that Weak König's Lemma is the restriction of the Ultrafilter Theorem to countable domains.
- The observation of several variants of various strengths of the Weak Fan Theorem, some of them being purely intuitionistic.