

Computing with Gödel's completeness theorem

Hugo Herbelin

(based on joint work with Danko Ilik)

IMJ-PRG, Paris, 3 December 2018

Introduction

The completeness of classical first-order logic w.r.t. Tarskian models

First proof by Gödel [1929]

- reasoning on the prenex form + induction on the number of alternation of quantifiers + reasoning by contradiction

Standard proof by Henkin [1949]

- reasoning by contradiction + construction of a counter-model by enumeration of the formulas over a language extended with Henkin constants coming from the skolemisation of the drinkers' paradox ($\exists x(P(x) \Rightarrow \forall y P(y))$).

Tableaux-based proofs by Beth [1955], Hintikka [1955], Schütte [1956], Kanger [1957]

- building a tableau + reasoning by contradiction to show it has no infinite branch

Excerpt of alternative proofs

- proofs by Mostowski [1948], Rasiowa-Sikorski [1950], relying on the ultrafilter theorem
- a generic abstract proof by Joyal [1978] (with Tarskian completeness for coherent logic behind the scene?)

Different formulations of completeness

One of the following three classically equivalent statements

S_1 . Formula A true in all models of theory T implies A provable from (a finite subset) of T

S_2 . Theory T consistent implies T has a model *(model-theoretic view)*

S_3 . Theory T either is inconsistent or has a model *(proof search view)*

Constructivisability of the different formulations of completeness

It happens that S_2 is constructive (the model can be “constructed” as a particular predicate and proved to be a model when the object language has only negative connectives and the language is countable).

S_3 is strongly classical as the disjunction is not decidable. However, this does not exclude computing with, since classical logic is computational: one could compute with it when completeness is used as a lemma in the proof of a Σ_1^0 formula.

S_1 is the statement for which we are looking for a computational content.

Logical strength of completeness

- Kreisel [1962], after Gödel [1957]: S_1 for an empty theory and the object language of negative connectives is equivalent to Markov's principle over intuitionistic second-order arithmetic
- Generalised by McCarty [2008]: S_1 for recursively enumerable theories over the language of negative connectives is equivalent to Markov's principle over intuitionistic second-order arithmetic
- McCarty [2008]: using non-decidable theories, S_1 implies classical logic
- Simpson [1999]: strong completeness for a countable language is classically equivalent to weak König's lemma over RCA_0
- Henkin [1999]: strong completeness for an uncountable theory (hence for uncountable language) classically implies the Boolean Prime Ideal axiom

Avoiding the need for Markov's principle

Krivine's proof of completeness for an empty theory [1996]

- restricted to minimal classical logic (no $\perp \Rightarrow A$) so that negation does not have to be interpreted; Friedman's A-translation [1978] is then applicable to get rid of Markov's principle
- analysed by Berardi and Valentini [2001]: Krivine adds one extra (degenerated) model, the always-true model (similar to Friedman's fallible models and Veldman's exploding nodes in intuitionistic logic semantics)
- the modified statement is *classically equivalent* to the original one but does not need Markov's principle
- formalised in the PhoX proof assistant and later in Coq

The statement of completeness

(empty theory, countable language, restricted to the \rightarrow - \perp - \forall fragment)

$$t \in \mathcal{Term} ::= x \mid f(t_1, \dots, t_{ar_f})$$

$$A, B \in \mathcal{F} ::= P(t_1, \dots, t_{ar_P}) \mid \perp \mid A \rightarrow B \mid \forall x A$$

A model is a quadruple $(\mathcal{M}_D, \mathcal{F}_M(f) \in \mathcal{M}_D^{ar_f} \rightarrow \mathcal{M}_D, \mathcal{F}_M(P) \in \mathcal{P}(\mathcal{M}_D^{ar_P}), \perp_M \in \mathcal{P}(\{\emptyset\}))$.

Truth in \mathcal{M} is defined recursively:

$$\begin{aligned} \llbracket x \rrbracket_{\mathcal{M}}^{\sigma} &\triangleq \sigma(x) \\ \llbracket f(t_1, \dots, t_{ar_f}) \rrbracket_{\mathcal{M}}^{\sigma} &\triangleq \mathcal{F}_M(f)(\llbracket t_1 \rrbracket_{\mathcal{M}}^{\sigma}, \dots, \llbracket t_{ar_f} \rrbracket_{\mathcal{M}}^{\sigma}) \\ \llbracket P(t_1, \dots, t_{ar_P}) \rrbracket_{\mathcal{M}}^{\sigma} &\triangleq \mathcal{P}_M(P)(\llbracket t_1 \rrbracket_{\mathcal{M}}^{\sigma}, \dots, \llbracket t_{ar_P} \rrbracket_{\mathcal{M}}^{\sigma}) \\ \llbracket \perp \rrbracket_{\mathcal{M}}^{\sigma} &\triangleq \perp_M \\ \llbracket A \rightarrow B \rrbracket_{\mathcal{M}}^{\sigma} &\triangleq \llbracket A \rrbracket_{\mathcal{M}}^{\sigma} \Rightarrow \llbracket B \rrbracket_{\mathcal{M}}^{\sigma} \\ \llbracket \forall x A \rrbracket_{\mathcal{M}}^{\sigma} &\triangleq \forall t \in \mathcal{M}_D \llbracket A \rrbracket_{\mathcal{M}}^{\sigma[x \leftarrow t]} \end{aligned}$$

A model is classical on σ , written $Class(\mathcal{M})$ if for each A , $\llbracket \neg \neg A \rrbracket_{\mathcal{M}}^{\sigma} \Rightarrow \llbracket A \rrbracket_{\mathcal{M}}^{\sigma}$ (and in particular, it is exploding: $\perp_M \Rightarrow \llbracket A \rrbracket_{\mathcal{M}}^{\sigma}$).

A model satisfies theory T on σ , written $\llbracket T \rrbracket_{\mathcal{M}}^{\sigma}$ if $\llbracket B \rrbracket_{\mathcal{M}}^{\sigma}$ for all $B \in T$.

The completeness statement : $\forall T \forall A (\forall \mathcal{M} \forall \sigma (Class(\mathcal{M}) \wedge \llbracket T \rrbracket_{\mathcal{M}}^{\sigma} \Rightarrow \llbracket A \rrbracket_{\mathcal{M}}^{\sigma}) \Rightarrow T \vdash A)$

Remarks on the formulation

We placed ourselves in intuitionistic second-order arithmetic, interpreting predicates by predicates, defining truth recursively.

Some authors reason instead in the arithmetic of types of rank 2 and define $\mathcal{P}_{\mathcal{M}}(P)$ as a Boolean function in $\mathcal{M}_{\mathcal{D}}^{arp} \Rightarrow \text{bool}$. The completeness proof then needs a reification axiom $\forall P \exists f. \forall x (P(x) \iff f(x) = \text{true})$. This can be obtained from the axiom of unique choice (AC!) and excluded-middle and makes the metatheory actually equivalent to second-order arithmetic. To avoid having to computationally interpret reification, what should be doable in dPA^ω (see LICS 2012), we prefer to directly reason in second-order arithmetic.

It is also common to replace $\mathcal{P}_{\mathcal{M}}$ by a set of formulas enriched over \mathcal{D} such that:

$$\begin{aligned} \perp \in \mathcal{P}_{\mathcal{M}} &\leftrightarrow \perp \\ A \dot{\rightarrow} B \in \mathcal{P}_{\mathcal{M}} &\leftrightarrow A \in \mathcal{P}_{\mathcal{M}} \Rightarrow B \in \mathcal{P}_{\mathcal{M}} \\ \dot{\forall} x A \in \mathcal{P}_{\mathcal{M}} &\leftrightarrow \forall t A[t/x] \in \mathcal{P}_{\mathcal{M}} \\ A \in \mathcal{P}_{\mathcal{M}} &\leftrightarrow \neg\neg A \in \mathcal{P}_{\mathcal{M}} \end{aligned}$$

Our approach has both the advantage of avoiding to consider formulas enriched over \mathcal{D} and to make the connection with intuitionistic models (e.g. Kripke) closer.

Part I
Analysis of Henkin's proof

Henkin's proof (usual presentation)

To prove $T \vdash A_0$, prove instead $\neg A_0, T \vdash \perp$, using the abbreviation $\neg B \triangleq B \rightarrow \perp$.

Reason by contradiction (Markov's principle) and assume $(\neg A_0, T \vdash \perp) \Rightarrow \perp$, i.e. that the context $\Gamma_0 \cup T$ where $\Gamma_0 \triangleq \neg A_0$ is consistent.

For an enumeration $\phi(0) \triangleq \forall x B_0, \phi(2) \triangleq \forall x B_2, \dots$ of all universal formulas and $\phi(1) \triangleq A_1 \rightarrow B_1, \phi(3) \triangleq A_3 \rightarrow B_3, \dots$ of all implicative formulas, classically build:

- $\Gamma_{2n+1} \triangleq \Gamma_{2n}, (B_{2n}[x_n/x] \rightarrow \forall x B_{2n})$
- $\Gamma_{2n+2} \triangleq \Gamma_{2n+1}$ if $\Gamma_{2n+1}, A_{2n+1} \rightarrow B_{2n+1}, T \vdash \perp$
- $\Gamma_{2n+2} \triangleq \Gamma_{2n+1}, A_{2n+1} \rightarrow B_{2n+1}$ otherwise

where the formulas $B_{2n}[x_n/x] \rightarrow \forall x B_{2n}$, for x_n taken fresh in all $\phi(i)$ for $i < 2n$ are Henkin axioms (no need for fresh constants, fresh variables are enough).

This construction propagates consistency from $\Gamma_0 \cup T$ to $\Gamma_n \cup T$.

The proof (usual presentation), continued

Build the infinite theory $\Gamma_\omega \triangleq \cup_n(\Gamma_n \cup T)$.

Under the initial assumption that $T \vdash A_0$ is contradictory, one gets that Γ_ω is consistent.

Define a syntactic model \mathcal{M}_0 by

$$\begin{aligned}\mathcal{D} & \triangleq \mathcal{T}erm \\ \mathcal{F}_{\mathcal{M}}(f)(t_1, \dots, t_{ar_f}) & \triangleq f(t_1, \dots, t_{ar_f}) \\ \mathcal{P}_{\mathcal{M}}(P)(t_1, \dots, t_{ar_P}) & \triangleq P(t_1, \dots, t_{ar_P}) \in \Gamma_\omega \\ \perp_{\mathcal{M}} & \triangleq \perp\end{aligned}$$

Using the converse $\lceil A \rceil$ of the Gödel's numbering of formulas, one proves by induction on A that $\llbracket A \rrbracket_{\mathcal{M}_0}^{id}$ iff $A \in \Gamma_\omega$.

The model is complete in the sense that $\neg A \notin \Gamma_\omega$ implies $A \in \Gamma_\omega$. Hence it satisfies $Class(\mathcal{M}_0)$.

The model satisfies T since $T \subset \Gamma_\omega$.

By validity of A_0 , get $\llbracket A_0 \rrbracket_{\mathcal{M}_0}^{id}$, hence $A_0 \in \Gamma_\omega$, hence $\Gamma_\omega \vdash \perp$, a contradiction.

What is the computational meaning of this proof?

Finally, for A_0 provable in T , is the model built consistent or not?

Obviously no!

If we turn the proof positively, what it shows is that $\Gamma_\omega \vdash \perp$ implies $\neg A_0, T \vdash \perp$.

That some $\phi(2n+1)$ has been added to the context reduces to have $(\Gamma_{2n}, \phi(2n+1) \vdash \perp) \Rightarrow \perp$ under the assumption that $(\neg A_0, T \vdash \perp) \Rightarrow \perp$.

Turned positively, this means that Γ_{2n} can be extended as soon as we know how to get rid of the extension.

Otherwise said, the model construction collects continuations.

The proof (an optimisation)

In practice, we do not need to characterize the elements of Γ_n and Γ_ω but only the provability predicate that Γ_ω generates. This means that we only need to know when a given finite context Γ is in Γ_n , which can simply be defined by the following (overlapping) clauses:

- $\Gamma \subset \Gamma_{n+1}$ whenever $\Gamma \subset \Gamma_n$
- $\Gamma, B_{2n}[x_n/x] \dot{\rightarrow} \forall x B_{2n} \subset \Gamma_{2n+1}$ whenever $\Gamma \subset \Gamma_{2n}$
- $\Gamma, A_{2n+1} \dot{\rightarrow} B_{2n+1} \subset \Gamma_{2n+2}$ whenever $\Gamma \subset \Gamma_{2n+1}$ and $\Gamma_{2n+1}, A_{2n+1} \dot{\rightarrow} B_{2n+1} \vdash \perp$ implies $\dot{\rightarrow} A_0 \vdash \perp$.
The condition $\Gamma_{2n+1}, A_{2n+1} \dot{\rightarrow} B_{2n+1} \vdash \perp$ itself reduces to the existence of $\Gamma \subset \Gamma_{2n+1}$ such that $\Gamma, A_{2n+1} \dot{\rightarrow} B_{2n+1} \vdash \perp$

We can then define $A \in \Gamma_\omega$ to mean $\exists n \exists \Gamma \subset \Gamma_n (\Gamma, T \vdash A)$ (“ A gets provable at some step of the construction of a context $\Gamma_n \cup T$ equiconsistent to $\dot{\rightarrow} A_0 \cup T$ ”).

The proof (bypassing the need for Markov's principle)

We take the following definition of the syntactic model \mathcal{M}_0 with exploding nodes:

$$\begin{aligned}\mathcal{D} &\triangleq \mathcal{T}erm \\ \mathcal{F}_{\mathcal{M}}(f)(t_1, \dots, t_{ar_f}) &\triangleq f(t_1, \dots, t_{ar_f}) \\ \mathcal{P}_{\mathcal{M}}(P)(t_1, \dots, t_{ar_P}) &\triangleq P(t_1, \dots, t_{ar_P}) \in \Gamma_{\omega} \\ \perp_{\mathcal{M}} &\triangleq \dot{\perp} \in \Gamma_{\omega}\end{aligned}$$

Giving notations to express the computational contents

We reformulate $\Gamma \subset \Gamma_n$ as an “inductive predicate” so as to be able to manipulate proof constructors as data:

$$\begin{array}{c}
 \frac{}{\neg A_0 \subset \Gamma_0} I_0 \qquad \frac{\Gamma \subset \Gamma_n}{\Gamma \subset \Gamma_{n+1}} I_S \qquad \frac{\Gamma \subset \Gamma_{2n}}{\Gamma, A(x_n) \dot{\rightarrow} \forall x A(x) \subset \Gamma_{2n+1}} I_V \\
 \\
 \frac{\Gamma \subset \Gamma_{2n+1} \quad \exists \Gamma' (\Gamma' \subset \Gamma_{2n+1} \wedge \Gamma', A \dot{\rightarrow} B, T \vdash \perp) \Rightarrow (\neg A_0, T \vdash \perp)}{\Gamma, A \dot{\rightarrow} B \subset \Gamma_{2n+2}} I_{\Rightarrow}
 \end{array}$$

where $\phi(2n) \equiv \forall x A(x)$ in I_V and $\phi(2n + 1) \equiv A \dot{\rightarrow} B$ in I_{\Rightarrow} .

The object language

We assume given a (non-minimal) set of appropriate object language constructions, parametrized by a recursively enumerable theory T :

$$\text{ax}^+ : A \in T \longrightarrow [\Gamma, T \vdash A]$$

$$\text{ax}_i : [\Gamma, A, \Gamma', T \vdash A] \quad (\text{for } \Gamma' \text{ of length } i)$$

$$\text{ax}'_i : [\Gamma, A, \Gamma', T \vdash A] \quad (\text{for } \Gamma \text{ of length } i)$$

$$\text{dn} : [\Gamma, T \vdash \neg\neg A] \longrightarrow [\Gamma, T \vdash A]$$

$$\text{abs} : [\Gamma, A, T \vdash B] \longrightarrow [\Gamma, T \vdash A \rightarrow B]$$

$$\text{app}^\Rightarrow : [\Gamma, T \vdash A \rightarrow B] \longrightarrow [\Gamma', T \vdash A] \longrightarrow [\Gamma \cup \Gamma', T \vdash B]$$

$$\text{drinker}_n : [B_{2n}[x_n/x] \rightarrow \forall x B_{2n}, \Gamma, T \vdash \perp] \longrightarrow [\Gamma, T \vdash \perp] \quad \text{where } \phi(2n) = \forall x B_{2n} \text{ and } x_n \text{ as before}$$

$$\text{app}^\forall : [\Gamma, T \vdash \forall x A(x)] \longrightarrow \forall t \in \mathcal{T}erm [\Gamma, T \vdash A(t)]$$

$$\pi_1^{\rightarrow} : [\Gamma, A \rightarrow B, T \vdash \perp] \longrightarrow [\Gamma, T \vdash A]$$

$$\pi_2^{\rightarrow} : [\Gamma, A \rightarrow B, T \vdash \perp] \longrightarrow [\Gamma, T \vdash \neg B]$$

The core of the proof: $\llbracket A \rrbracket_{\mathcal{M}_0}^{id}$ iff $A \in \Gamma_\omega$

$$\downarrow_A : \llbracket A \rrbracket_{\mathcal{M}_0}^{id} \rightarrow A \in \Gamma_\omega$$

$$\downarrow_{P(\vec{t})} m \triangleq m$$

$$\downarrow_{\perp} m \triangleq m$$

$$\downarrow_{A \rightarrow B} m \triangleq (n, (\dot{\rightarrow} A_0, A \rightarrow B),$$

$$I_n(\text{inj}_n, (\Gamma, f, p) \mapsto \text{dest } \downarrow_B (m(\uparrow_A (n, \Gamma, f, \pi_1^{\rightarrow} p))) \text{ as } (n', \Gamma', f', p') \\ \text{in flush}_{\max(n, n')}^{\Gamma \cup \Gamma'} (\text{join}_{nn'}^{\Gamma'} (f, f'), \text{app}^{\Rightarrow} (\pi_2^{\rightarrow} p, p'))), \\ \text{ax}_1) \quad \text{where } n = \llbracket A \rightarrow B \rrbracket$$

$$\downarrow_{\dot{\forall} x A} m \triangleq \text{dest } \downarrow_{A[x_n/x]} (m x_n) \text{ as } (n', \Gamma', f', p')$$

$$\text{in } (\max(n, n'), \Gamma', \text{join}_{nn'}^{(\dot{\rightarrow} A_0)\Gamma'} (\text{inj}_n, f'), \text{app}^{\Rightarrow} (\text{ax}'_0, p'))$$

$$\text{where } n = \llbracket \dot{\forall} x A \rrbracket$$

$$\uparrow_A : A \in \Gamma_\omega \rightarrow \llbracket A \rrbracket_{\mathcal{M}_0}^{id}$$

$$\uparrow_{P(\vec{t})} (n, \Gamma, f, p) \triangleq (n, \Gamma, f, p)$$

$$\uparrow_{\perp} (n, \Gamma, f, p) \triangleq (n, \Gamma, f, p)$$

$$\uparrow_{A \rightarrow B} (n, \Gamma, f, p) \triangleq m \mapsto \text{dest } \downarrow_A m \text{ as } (n', \Gamma', f', p') \\ \text{in } \uparrow_B (\max(n, n'), \Gamma \cup \Gamma', \text{join}_{nn'}^{\Gamma'} (f, f'), \text{app}^{\Rightarrow} (p, p'))$$

$$\uparrow_{\dot{\forall} x A} (n, \Gamma, f, p) \triangleq t \mapsto \uparrow_{A[t/x]} (n, \Gamma, f, \text{app}^{\forall} (p, t))$$

Auxiliary lemma: propagating an inconsistency at level n to level 0

$$\begin{aligned} \text{flush}_n^\Gamma & : \Gamma \subset \Gamma_n \wedge [\Gamma, T \vdash \perp] \longrightarrow \neg A_0, T \vdash \perp \\ \text{flush}_0^\Gamma & (I_\emptyset, p) \quad \triangleq \quad p \\ \text{flush}_{n+1}^\Gamma & (I_S f, p) \quad \triangleq \quad \text{flush}_n^\Gamma(f, p) \\ \text{flush}_{2n+1}^{\Gamma, A} & (I_\forall f, p) \quad \triangleq \quad \text{flush}_{2n}(f, \text{drinker}_{x_n} p) \\ \text{flush}_{2n+2}^{\Gamma, A} & (I_\Rightarrow(f, k), p) \quad \triangleq \quad k(f, p) \end{aligned}$$

Auxiliary lemma: joining contexts in binary rules

$$\begin{array}{l}
 \text{join}_{n_1 n_2}^{\Gamma_1 \Gamma_2} : \Gamma_1 \subset \Gamma_{n_1} \wedge \Gamma_2 \subset \Gamma_{n_2} \longrightarrow \Gamma_1 \cup \Gamma_2 \subset \Gamma_{\max(n_1, n_2)} \\
 \text{join}_{00}^{\dot{A}_0 \dot{A}_0} \quad \mathbb{I}_\emptyset \quad \mathbb{I}_\emptyset \quad \triangleq \quad \mathbb{I}_\emptyset \\
 \text{join}_{(2n+2)(2n+2)}^{(\Gamma_1 A)(\Gamma_2 A)} \quad \mathbb{I}_{\Rightarrow}(f_1, k_1) \quad \mathbb{I}_{\Rightarrow}(f_2, k_2) \quad \triangleq \quad \mathbb{I}_{\Rightarrow}(\text{join}_{(2n+1)(2n+1)}^{\Gamma_1 \Gamma_2} f_1 f_2, k_1) \\
 \text{join}_{(2n+2)(2n+2)}^{(\Gamma_1 A)\Gamma_2} \quad \mathbb{I}_{\Rightarrow}(f_1, k_1) \quad \mathbb{I}_S f_2 \quad \triangleq \quad \mathbb{I}_{\Rightarrow}(\text{join}_{(2n+1)(2n+1)}^{\Gamma_1 \Gamma_2} f_1 f_2, k_1) \\
 \text{join}_{(2n+2)(2n+2)}^{\Gamma_1(\Gamma_2 A)} \quad \mathbb{I}_S f_1 \quad \mathbb{I}_{\Rightarrow}(f_2, k_2) \quad \triangleq \quad \mathbb{I}_{\Rightarrow}(\text{join}_{(2n+1)(2n+1)}^{\Gamma_1 \Gamma_2} f_1 f_2, k_2) \\
 \text{join}_{(n+1)(n+1)}^{\Gamma_1 \Gamma_2} \quad \mathbb{I}_S f_1 \quad \mathbb{I}_S f_2 \quad \triangleq \quad \mathbb{I}_S(\text{join}_{nn}^{\Gamma_1 \Gamma_2} f_1 f_2) \\
 \text{join}_{(2n+1)(2n+1)}^{\Gamma_1 \Gamma_2} \quad \mathbb{I}_S f_1 \quad \mathbb{I}_{\forall} f_2 \quad \triangleq \quad \mathbb{I}_{\forall}(\text{join}_{(2n)(2n)}^{\Gamma_1 \Gamma_2} f_1 f_2) \\
 \text{join}_{(2n+1)(2n+1)}^{\Gamma_1 \Gamma_2} \quad \mathbb{I}_{\forall} f_1 \quad \mathbb{I}_S f_2 \quad \triangleq \quad \mathbb{I}_{\forall}(\text{join}_{(2n)(2n)}^{\Gamma_1 \Gamma_2} f_1 f_2) \\
 \text{join}_{(2n+1)(2n+1)}^{\Gamma_1 \Gamma_2} \quad \mathbb{I}_{\forall} f_1 \quad \mathbb{I}_{\forall} f_2 \quad \triangleq \quad \mathbb{I}_{\forall}(\text{join}_{(2n)(2n)}^{\Gamma_1 \Gamma_2} f_1 f_2) \\
 \text{join}_{n_1 n_2}^{\Gamma_1 \Gamma_2} \quad \mathbb{I}_S f_1 \quad f_2 \quad \triangleq \quad \mathbb{I}_S(\text{join}_{n'_1 n'_2}^{\Gamma_1 \Gamma_2} f_1 f_2) \quad \text{if } n_1 = n'_1 + 1 > n_2 \\
 \text{join}_{n_1 n_2}^{(\Gamma_1 A_1)\Gamma_2} \quad \mathbb{I}_{\Rightarrow}(f_1, k_1) \quad f_2 \quad \triangleq \quad \mathbb{I}_{n'_1}(\text{join}_{n'_1 n'_2}^{\Gamma_1 \Gamma_2} f_1 f_2, k_1) \quad \text{if } n_1 = n'_1 + 1 > n_2 \\
 \text{join}_{n_1 n_2}^{\Gamma_1 \Gamma_2} \quad f_1 \quad \mathbb{I}_S f_2 \quad \triangleq \quad \mathbb{I}_S(\text{join}_{n_1 n'_2}^{\Gamma_1 \Gamma_2} f_1 f_2) \quad \text{if } n_1 < n'_2 + 1 = n_2 \\
 \text{join}_{n_1 n_2}^{\Gamma_1(\Gamma_2 A_2)} \quad f_1 \quad \mathbb{I}_{n'_2}(f_2, k_2) \quad \triangleq \quad \mathbb{I}_{n'_2}(\text{join}_{n_1 n'_2}^{\Gamma_1 \Gamma_2} f_1 f_2, k_2) \quad \text{if } n_1 < n'_2 + 1 = n_2
 \end{array}$$

Auxiliary lemma: $\neg A_0$ is in all contexts

inj_n : $(\neg A_0) \subset \Gamma_n$

$\text{inj}_0 \triangleq I_\emptyset$

$\text{inj}_{n+1} \triangleq I_S(\text{inj}_n)$

Remarks about the computational content

If A_0 is provable, then the countermodel (virtually) built is actually the degenerated countermodel that contains all formulas, including \perp .

Along the computational interpretation of Markov's principle, reasoning classically by assuming $[\vdash A_0] \Rightarrow \perp$ is the same as providing an exception which returns a derivation of $[\vdash A_0]$ as soon as a contradiction is obtained. Along Friedman's A -translation, this amounts to reinterpret \perp as the formula $[\vdash A_0]$.

Computationally, the proof of a negation can be seen as a continuation. Combined with the computational interpretation of Markov's principle, this is the same as a continuation that eventually returns a derivation of $[\vdash A_0]$.

In particular, a proof that some finite section $\neg A_0, \Gamma$ of the countermodel is consistent is the same as a continuation that transforms a derivation of $[\neg A_0, \Gamma, T \vdash \perp]$, that is of $[\Gamma, T \vdash A_0]$, into a derivation of $[\vdash A_0]$.

More remarks about the computational content

The ordering of formulas has an effect on the order of application of continuations: continuations are applied in the decreasing order of the Gödel number of the formulas.

In case of branching, i.e. in the case of modus ponens, if two continuations are available at level n , the `join` function arbitrary chooses one of them. In particular, some subproofs of the initial meta-proof might be lost and replaced by an other proof of the same formula in the same original meta-proof.

Compared to the completeness proof with respect to Kripke semantics where the world is locally extended with the knowledge of A to show that $A \dot{\rightarrow} B$ is provable, here, in the completeness for two-valued semantics, one extends the (counter)knowledge Γ with $A \dot{\rightarrow} B$ but altogether with a proof that contradicting $\Gamma, A \dot{\rightarrow} B$ (in the sense of a derivation of $\Gamma, A \dot{\rightarrow} B, T \vdash \perp$) eventually reduces to a derivation of $\vdash A_0$.

It is worth noticing that the definition of $\Gamma \subset \Gamma_n$ is of high implicational complexity. Due to the contravariance in the clause I_n , the definition of $\Gamma \subset \Gamma_{n+1}$ involves implications nested at level n . Henceforth, the definition of $A \in \Gamma_\omega$ is a formula whose implication nesting depth is not finite. Logically, it is however a Σ_1^0 formula.

Remarks about the Henkin axioms

Cutting Henkin axiom $A[x_n/x] \rightarrow \dot{\forall}x A$ with drinkers' paradox $\exists y(A[y/x] \rightarrow \dot{\forall}x A)$ can be seen as a way to delegate the insurance of the (moral) freshness of x_n and the ability to go from $A[x_n/x]$ to $\dot{\forall}x A$ even when in a context where other occurrences of x_n might occur (namely the context $\Gamma' \subset \Gamma_{n'}$ in the $\dot{\forall}$ clause of \downarrow). Interestingly, eliminating a cut with $\exists y(A[y/x] \rightarrow \dot{\forall}x A)$ will rename the x_n whose occurrences are possibly non-fresh using names that are actually fresh and from which $\dot{\forall}x A$ can correctly be inferred.

Part II

A proof with side effects of completeness with respect to
Tarski semantics

Completeness w.r.t. Kripke models

Kripke models

A Kripke model \mathcal{K} is an increasing family of Tarskian models indexed over a set of worlds $\mathcal{W}_{\mathcal{K}}$ ordered by $\geq_{\mathcal{K}}$. In the absence of \forall and \exists , it is enough to take $\mathcal{D}_{\mathcal{K}}$ constant.

Truth relatively to \mathcal{K} at world w is defined by:

$$\begin{aligned}
 \llbracket x \rrbracket_{\mathcal{K}}^{\sigma} &\triangleq \sigma(x) \\
 \llbracket f t_1 \dots t_{a_f} \rrbracket_{\mathcal{K}}^{\sigma} &\triangleq \mathcal{F}_{\mathcal{K}}(f)(\llbracket t_1 \rrbracket_{\mathcal{K}}^{\sigma}, \dots, \llbracket t_{a_f} \rrbracket_{\mathcal{K}}^{\sigma}) \\
 w \Vdash_{\mathcal{K}}^{\sigma} \dot{P}(t_1 \dots t_{a_p}) &\triangleq \mathcal{P}_{\mathcal{K}}(\dot{P})(w)(\llbracket t_1 \rrbracket_{\mathcal{K}}^{\sigma}, \dots, \llbracket t_{a_p} \rrbracket_{\mathcal{K}}^{\sigma}) \\
 w \Vdash_{\mathcal{K}}^{\sigma} \perp &\triangleq \perp_{\mathcal{K}}(w) \\
 w \Vdash_{\mathcal{K}}^{\sigma} A \rightarrow B &\triangleq \forall w' \geq_{\mathcal{K}} w (w' \Vdash_{\mathcal{K}}^{\sigma} A \Rightarrow w' \Vdash_{\mathcal{K}}^{\sigma} B) \\
 w \Vdash_{\mathcal{K}}^{\sigma} \forall x A &\triangleq \forall t \in \mathcal{K}_D w \Vdash_{\mathcal{K}}^{\sigma[x \leftarrow t]} A
 \end{aligned}$$

The statement of completeness w.r.t. Kripke models for an empty theory is:

$$(\forall \mathcal{K} \forall \sigma \forall w \in \mathcal{W}_{\mathcal{K}} w \Vdash_{\mathcal{K}}^{\sigma} A) \Rightarrow \vdash_M A$$

Taking a natural deduction as object language

We take the following inference rules:

$$\begin{aligned} \mathbf{Ax}^{\Gamma, A, \Gamma'} & : (\Gamma, A \subset \Gamma') \Rightarrow (\Gamma' \vdash A) \\ \mathbf{App}_{\Rightarrow}^{\Gamma, A, B} & : (\Gamma \vdash A \dot{\rightarrow} B) \Rightarrow (\Gamma \vdash A) \Rightarrow (\Gamma \vdash B) \\ \mathbf{Abs}_{\Rightarrow}^{\Gamma, A, B} & : (\Gamma, A \vdash B) \Rightarrow (\Gamma \vdash A \dot{\rightarrow} B) \\ \mathbf{Abs}_{\forall}^{\Gamma, x, A} & : (\Gamma \vdash A) \Rightarrow (x \notin FV(\Gamma)) \Rightarrow (\Gamma \vdash \dot{\forall}x A) \\ \mathbf{App}_{\forall}^{\Gamma, x, t, A} & : (\Gamma \vdash \dot{\forall}x A) \Rightarrow (\Gamma \vdash A[t/x]) \end{aligned}$$

Moreover, the following is admissible:

$$\mathbf{weak}_{\Gamma, A}^{\Gamma'} : (\Gamma \subset \Gamma') \Rightarrow (\Gamma \vdash A) \Rightarrow (\Gamma' \vdash A)$$

We shall also write r_A^{Γ} for a proof of $\Gamma \subset (\Gamma, A)$,

Completeness w.r.t Kripke models

The “standard” proof works by building the canonical model \mathcal{K}_0 defined by taking $\mathcal{W}_{\mathcal{K}_0}$ to be the typing contexts ordered by inclusion, $\mathcal{D}_{\mathcal{K}_0}$ to be the terms, $\mathcal{F}_{\mathcal{K}_0}(f)$ to be the syntactic application of f , $\mathcal{P}_{\mathcal{K}_0}(\dot{P})(\Gamma)(t_1, \dots, t_{a_{\dot{P}}})$ to be $\Gamma \vdash_M \dot{P}(t_1, \dots, t_{a_{\dot{P}}})$, and $\perp_{\mathcal{K}_0}(\Gamma)$ to be $\Gamma \vdash_M \perp$.

The main lemma proves $\Gamma \vdash_M A \iff \Gamma \Vdash_{\mathcal{K}_0} A$ by induction on A , with $r_A^\Gamma : \Gamma \subset \Gamma, A$:

$$\begin{array}{l}
 \uparrow_A^\Gamma \quad \Gamma \vdash_M A \longrightarrow \Gamma \Vdash_{\mathcal{K}_0} A \\
 \uparrow_{\dot{P}(\vec{t})}^\Gamma \quad p \quad \triangleq \quad p \\
 \uparrow_{A \rightarrow G}^\Gamma \quad p \quad \triangleq \quad \Gamma' \mapsto h \mapsto m \mapsto \uparrow_G^{\Gamma', A, G} \mathbf{App}_{\Rightarrow}^{\Gamma', A, G}(\mathbf{weak}_{\Gamma, A}^{\Gamma'}(h, p), \downarrow_A^{\Gamma'} m) \\
 \uparrow_{\forall x A}^\Gamma \quad p \quad \triangleq \quad t \mapsto \uparrow_{A[t/x]}^\Gamma \mathbf{App}_{\forall}^{\Gamma, x, A}(p, t) \\
 \\
 \downarrow_A^\Gamma \quad \Gamma \Vdash_{\mathcal{K}_0} A \longrightarrow \Gamma \vdash_M A \\
 \downarrow_{\dot{P}(\vec{t})}^\Gamma \quad m \quad \triangleq \quad m \\
 \downarrow_{A \rightarrow B}^\Gamma \quad m \quad \triangleq \quad \mathbf{Abs}_{\Rightarrow}^{\Gamma, A, B}(\downarrow_B^{\Gamma, A} (m(\Gamma, A) r_A^\Gamma (\uparrow_A^{\Gamma, A} \mathbf{Ax}^{\Gamma, A, (\Gamma, A)}(\mathit{refl}_{\Gamma, A})))) \\
 \downarrow_{\forall x A}^\Gamma \quad m \quad \triangleq \quad \mathbf{Abs}_{\forall}^{\Gamma, x, A}(\dot{y}, \downarrow_{A[z/x]}^\Gamma (m \dot{y})) \quad \dot{y} \text{ fresh in } \Gamma
 \end{array}$$

And finally:

$$\mathbf{compl} \triangleq v \mapsto \downarrow_A^\epsilon (v \mathcal{K}_0 \emptyset \epsilon) : (\forall \mathcal{K} \forall \sigma \forall w \in \mathcal{W}_{\mathcal{K}} w \Vdash_{\mathcal{K}}^\sigma A) \Rightarrow \vdash_M A$$

Completeness w.r.t. Kripke models in direct-style

Kripke forcing translation for second-order arithmetic

We consider a second-order arithmetic meta-language, multi-sorted over first-order datatypes such as \mathbb{N} , lists, formulas, etc., and with primitive recursive atoms written $P(t_1, \dots, t_{ap})$ (morally: HA^2).

$$A, B \triangleq X(t_1, \dots, t_{ax}) \mid P(t_1, \dots, t_{ap}) \mid A \wedge B \mid A \Rightarrow B \mid \forall x A \mid \forall X A$$

Let \geq be a preorder definable over some sort W in HA^2 . We consider a (syntactic) Kripke forcing translation from HA^2 to HA^2 :

$$\begin{aligned} w \Vdash_{\geq} X(t_1, \dots, t_{ax}) &\triangleq X(w, t_1, \dots, t_{ax}) \\ w \Vdash_{\geq} P(t_1, \dots, t_{ap}) &\triangleq P(t_1, \dots, t_{ap}) \\ w \Vdash_{\geq} A \wedge B &\triangleq (w \Vdash_{\geq} A) \wedge (w \Vdash_{\geq} B) \\ w \Vdash_{\geq} A \Rightarrow B &\triangleq \forall w' \geq w [(w' \Vdash_{\geq} A) \Rightarrow (w' \Vdash_{\geq} B)] \\ w \Vdash_{\geq} \forall x A &\triangleq \forall x (w \Vdash_{\geq} A) \\ w \Vdash_{\geq} \forall X A &\triangleq \forall X (\text{mon}(X) \Rightarrow w \Vdash_{\geq} A) \end{aligned}$$

where $\text{mon}(X) \triangleq \forall w \forall w' \geq w (X(w, t_1, \dots, t_{ax}) \Rightarrow X(w', t_1, \dots, t_{ax}))$

Kripke translation of Tarskian semantics is Kripke semantics

We have the following key observation:

$$w \Vdash_{\geq} [\forall (\mathcal{D}_M, \mathcal{F}_M, \mathcal{P}_M, \perp_M) \forall \sigma (\llbracket A \rrbracket_{(\mathcal{D}_M, \mathcal{F}_M, \mathcal{P}_M, \perp_M)}^{\sigma})]$$

is the same as

$$\forall (\mathcal{D}_K, \mathcal{F}_K, \mathcal{P}_K, \perp_K) \forall \sigma w \Vdash_{(\mathcal{D}_K, \mathcal{F}_K, \mathcal{P}_K, \perp_K)}^{\sigma} A$$

Otherwise said: for a given ordered set of worlds, *the syntactic Kripke translation of validity w.r.t. Tarskian models is validity w.r.t. Kripke models over the same ordered set of worlds!*

Kripke translation of the statement of completeness

The Kripke translation of the statement of completeness w.r.t. Tarskian models:

$$w \Vdash_{\geq} [(\forall(\mathcal{D}_M, \mathcal{F}_M, \mathcal{P}_M, \perp_M) \forall \sigma (\llbracket F \rrbracket_{(\mathcal{D}_M, \mathcal{F}_M, \mathcal{P}_M, \perp_M)}^{\sigma})) \Rightarrow \vdash F]$$

is then

$$\forall w' \geq w [(\forall(\mathcal{D}_K, \mathcal{F}_K, \mathcal{P}_K, \perp_K) \forall \sigma w' \Vdash_{(\mathcal{D}_K, \mathcal{F}_K, \geq, \mathcal{P}_K)}^{\sigma} F) \Rightarrow w' \Vdash_{(\mathcal{D}_K, \mathcal{F}_K, \geq, \mathcal{P}_K, \perp_K)}^{\sigma} (\vdash F)]$$

i.e.

$$\forall w' \geq w [(\forall(\mathcal{D}_K, \mathcal{F}_K, \mathcal{P}_K, \perp_K) \forall \sigma w' \Vdash_{(\mathcal{D}_K, \mathcal{F}_K, \geq, \mathcal{P}_K, \perp_K)}^{\sigma} F) \Rightarrow \vdash F]$$

since $\vdash F$ is a Σ_1^0 formula with no second-order free variables.

Now, if we take contexts ordered by inclusion \subset for worlds and concentrate on the empty context, we get:

$$(\forall(\mathcal{D}_K, \mathcal{F}_K, \mathcal{P}_K, \perp_K) \forall \sigma \in \Vdash_{(\mathcal{D}_K, \mathcal{F}_K, \subset, \mathcal{P}_K, \perp_K)}^{\sigma} F) \Rightarrow \vdash F$$

which happens to be exactly completeness w.r.t. to Kripke models over contexts ordered by conclusion and considered on the empty context, i.e. a statement of which we had a simple proof.

It just remains to interpret this latter proof in direct style to get a new proof with side effects of completeness w.r.t. Tarskian models.

Excerpt of our meta-language with effects

$$\frac{\Gamma \vdash p : A(y) \quad y \text{ fresh in } \Gamma}{\Gamma \vdash \lambda y. p : \forall y A(y)} \quad \forall_I$$

$$\frac{\Gamma \vdash p : \forall x A(x) \quad t \text{ updatable-variable-free or } t \text{ an updatable variable and } A(x) \text{ of type 1}}{\Gamma \vdash p t : A(t)} \quad \forall_E$$

$$\frac{\Gamma \vdash p : A(X) \quad X \text{ fresh in } \Gamma}{\Gamma \vdash p : \forall X A(X)} \quad \forall_I^2 \quad \frac{\Gamma \vdash p : \forall X A(X) \quad \Gamma \vdash q : \text{mon}_\Gamma B(\vec{y})}{\Gamma \vdash p : A(X)[B(\vec{y})/X(\vec{y})]} \quad \forall_E^2$$

$$\frac{\Gamma, [b : x \geq t] \vdash q : T(x) \quad \Gamma \vdash r : \text{refl} \geq \quad \Gamma \vdash s : \text{trans} \geq \quad x \text{ fresh in } \Gamma \text{ and } T(t)}{\Gamma \vdash \text{set } x := t \text{ as } b /_{(r,s)} \text{ in } q : T(t)} \quad \text{SETEFF}$$

$$\frac{\Gamma, [b : x \geq t(x')] \vdash q : T(x) \quad \Gamma \vdash r : t(x') \geq x' \quad [x \geq u] \in \Gamma \text{ for some } u \quad x' \text{ fresh in } \Gamma}{\Gamma \vdash \text{update } x := t(x) \text{ of } x' \text{ as } b \text{ by } r \text{ in } q : T(t(x))} \quad \text{UPDATE}$$

where C of type 1 means in the grammar $C ::= P(t_1, \dots, t_{ap}) \mid P(t_1, \dots, t_{ap}) \Rightarrow C \mid \forall x C$ and $\text{mon}_\Gamma B$ means B monotonic for all updatable variables in Γ

The completeness proof in direct-style

In direct style, \mathcal{K}_0 is the model \mathcal{M}_0 defined by $\mathcal{P}_{\mathcal{M}}(\dot{P})(t_1, \dots, t_{a_p}) \triangleq \Gamma \vdash \dot{P}(t_1, \dots, t_{a_p})$ for Γ a given updatable variable

$$\begin{aligned}
 \uparrow_F \quad \Gamma \vdash_M F &\longrightarrow \llbracket F \rrbracket'_{\mathcal{M}_0} \\
 \uparrow_{P(\vec{i})} \quad g &\triangleq g \\
 \uparrow_{F \dot{\rightarrow} G} \quad g &\triangleq m \mapsto \uparrow_G \mathbf{App}_{\Rightarrow}^{\Gamma, F, G}(g, \downarrow_F m) \\
 \uparrow_{\dot{\forall}_x F} \quad g &\triangleq t \mapsto \uparrow_{F[t/x]} \mathbf{App}_{\forall}^{\Gamma, x, F}(g, t)
 \end{aligned}$$

$$\begin{aligned}
 \downarrow_F \quad \llbracket F \rrbracket'_{\mathcal{M}_0} &\longrightarrow \Gamma \vdash_M F \\
 \downarrow_{P(\vec{i})} \quad m &\triangleq m \\
 \downarrow_{F \dot{\rightarrow} G} \quad m &\triangleq \mathbf{Abs}_{\Rightarrow}^{\Gamma, F, G}(\text{update } \Gamma := (\Gamma, F) \text{ of } \Gamma_1 \text{ as } b_F \text{ by } r_F^\Gamma \text{ in } \downarrow_G (m(\uparrow_F \mathbf{Ax}^{\Gamma_1, F, \Gamma}(b_F)))) \\
 \downarrow_{\dot{\forall}_x F} \quad m &\triangleq \mathbf{Abs}_{\forall}^{\Gamma, x, F}(\dot{y}, \downarrow_{F[z/x]} (m \dot{y}))
 \end{aligned}$$

$$\text{compl} \triangleq v \mapsto \text{set } \Gamma := \epsilon \text{ as } b /_{(r,s)} \text{ in } \downarrow_F (v \mathcal{M}_0 \emptyset)$$

Obviously, the resulting proof in the object language is a reification of the proof of validity as in Normalisation-by-Evaluation / semantic normalisation [C. Coquand 93, Danvy 96, Altenkirch-Hofmann 96, Okada 99, ...]