On the conservativity of type theories with classical logic over arithmetic

Pietro Sabelli
j.w.w. Michele Contente and Maria Emilia Maietti

Department of Logic
Institute of Philosophy of the Czech Academy of Sciences

EuroProofNet WG6 meeting
Genoa, April 2025

Conservativity

A key notion in the foundations of mathematics is the following.

Definition

Let T be a theory. We say that an extension T^+ of T is *conservative over* T if every statement expressible in the language of T and provable in T^+, is already provable in T.

Conservativity

A key notion in the foundations of mathematics is the following.

Definition

Let T be a theory. We say that an extension T^+ of T is *conservative over* T if every statement expressible in the language of T and provable in T^+, is already provable in T.

Observation

Conservativity implies equiconsistency.

Conservativity of Type Theory over Arithmetic

Theorem (Beeson, 1985)

The first-order fragment of Martin-Löf's type theory* $\mathbf{ML}_0$ *is conservative over Heyting Arithmetic* **HA**.

Daniël Otten and Benno van den Berg. "Conservativity of Type Theory over Higher-Order Arithmetic". In: 32nd EACSL Annual Conference on Computer Science Logic (CSL 2024)

Conservativity of Type Theory over Arithmetic

Theorem (Beeson, 1985)

The first-order fragment of Martin-Löf's type theory* $\mathbf{ML}_0$ *is conservative over Heyting Arithmetic* $\mathbf{HA}$.

*meaning without universes but with the axiom $0 \neq 1$.

Daniël Otten and Benno van den Berg. "Conservativity of Type Theory over Higher-Order Arithmetic". In: 32nd EACSL Annual Conference on Computer Science Logic (CSL 2024)

Conservativity of Type Theory over Arithmetic

Theorem (Beeson, 1985)

The first-order fragment of Martin-Löf's type theory* $\mathbf{ML}_0$ *is conservative over Heyting Arithmetic* **HA**.

*meaning without universes but with the axiom $0 \neq 1$.

Theorem (Otten and van den Berg, 2024)

(An extensional version of) the Calculus of Inductive Constructions **CIC** *is conservative over Higher Order Heyting Arithmetic* **HAH**.

Daniël Otten and Benno van den Berg. "Conservativity of Type Theory over Higher-Order Arithmetic". In: 32nd EACSL Annual Conference on Computer Science Logic (CSL 2024)

Conservativity of Type Theory over Arithmetic

Theorem (Beeson, 1985)

The first-order fragment of Martin-Löf's type theory* $\mathbf{ML}_0$ *is conservative over Heyting Arithmetic* **HA**.

*meaning without universes but with the axiom $0 \neq 1$.

Theorem (Otten and van den Berg, 2024)

(An extensional version of) the Calculus of Inductive Constructions **CIC** *is conservative over Higher Order Heyting Arithmetic* **HAH**.

Goal

Transfer these results to the case of classical logic – in particular replacing **HA** with Peano Arithmetic **PA**.

Daniël Otten and Benno van den Berg. "Conservativity of Type Theory over Higher-Order Arithmetic". In: 32nd EACSL Annual Conference on Computer Science Logic (CSL 2024)

Classical logic in Predicative Foundations

Issue

The classical version $\mathbf{ML}_0^c$ of $\mathbf{ML}_0$ is stronger than **PA** (in fact, even of **PAH**!)

Classical logic in Predicative Foundations

Issue

The classical version $\mathbf{ML}_0^c$ of $\mathbf{ML}_0$ is stronger than **PA** (in fact, even of **PAH**!)

From a proof-theoretic perspective, classical logic interacts poorly with most predicative foundations, e.g.

- Homotopy Type Theory
- Constructive Zermelo-Fraenkel set theory

Classical logic in Predicative Foundations

Issue

The classical version $\mathbf{ML}_0^c$ of $\mathbf{ML}_0$ is stronger than **PA** (in fact, even of **PAH**!)

From a proof-theoretic perspective, classical logic interacts poorly with most predicative foundations, e.g.

- Homotopy Type Theory
- Constructive Zermelo-Fraenkel set theory

If we want to obtain a classical version of Beeson's theorem, we need to replace Martin-Löf's type theory with something more appropriate...

The Minimalist Foundation

The *Minimalist Foundation* **MF** is a type theory *compatible* with the most relevant foundations of mathematics.

M. E. Maietti, G. Sambin. "Toward a minimalist foundation for constructive mathematics". 2005

M. E. Maietti. "A minimalist two-level foundation for constructive mathematics". 2009

The Minimalist Foundation

The *Minimalist Foundation* **MF** is a type theory *compatible* with the most relevant foundations of mathematics.

Example

- Martin-Löf's type theory
- Homotopy Type Theory
- Calculus of Inductive Constructions
- Constructive Zermelo-Fraenkel set theory
- Internal language of toposes

M. E. Maietti, G. Sambin. "Toward a minimalist foundation for constructive mathematics". 2005

M. E. Maietti. "A minimalist two-level foundation for constructive mathematics". 2009

The Minimalist Foundation

The *Minimalist Foundation* **MF** is a type theory *compatible* with the most relevant foundations of mathematics.

Example

- Martin-Löf's type theory
- Homotopy Type Theory
- Calculus of Inductive Constructions
- Constructive Zermelo-Fraenkel set theory
- Internal language of toposes

For what concerns us here, **MF** can be thought of as a *predicative version* of **CIC**.

M. E. Maietti, G. Sambin. "Toward a minimalist foundation for constructive mathematics". 2005

M. E. Maietti. "A minimalist two-level foundation for constructive mathematics". 2009

First-order fragment of the Minimalist Foundation

We consider on the *first-order fragment* $\mathbf{MF}_0$ of the Minimalist Foundation – you can think of it as **CIC** without the universe **Prop**.

First-order fragment of the Minimalist Foundation

We consider on the *first-order fragment* $\mathbf{MF}_0$ of the Minimalist Foundation – you can think of it as **CIC** without the universe **Prop**.

Warning!

In both systems there is *cumulativity* of propositions into types *prop* $\hookrightarrow$ *type*, however...

First-order fragment of the Minimalist Foundation

We consider on the *first-order fragment* $\mathbf{MF}_0$ of the Minimalist Foundation – you can think of it as **CIC** without the universe **Prop**.

Warning!

In both systems there is *cumulativity* of propositions into types
prop $\hookrightarrow$ *type*, however...
... a *prop* is not any *type* (as in **ML**),

First-order fragment of the Minimalist Foundation

We consider on the *first-order fragment* $\mathbf{MF}_0$ of the Minimalist Foundation – you can think of it as **CIC** without the universe **Prop**.

Warning!

In both systems there is *cumulativity* of propositions into types
prop $\hookrightarrow$ *type*, however...
... a *prop* is not any *type* (as in **ML**),
is not any mono-type (as in toposes),

First-order fragment of the Minimalist Foundation

We consider on the *first-order fragment* $\mathbf{MF}_0$ of the Minimalist Foundation – you can think of it as **CIC** without the universe **Prop**.

Warning!

In both systems there is *cumulativity* of propositions into types *prop* $\hookrightarrow$ *type*, however...

... a *prop* is not any *type* (as in **ML**),

is not any mono-type (as in toposes),

is not any h-prop (as in **HoTT**).

First-order fragment of the Minimalist Foundation

We consider on the *first-order fragment* $\mathbf{MF}_0$ of the Minimalist Foundation – you can think of it as **CIC** without the universe **Prop**.

Warning!

In both systems there is *cumulativity* of propositions into types *prop* $\hookrightarrow$ *type*, however...
... a *prop* is not any *type* (as in **ML**),
is not any mono-type (as in toposes),
is not any h-prop (as in **HoTT**).
A *prop* is a *prop* is a *prop* is a *prop*.

First-order fragment of the Minimalist Foundation

We consider on the *first-order fragment* $\mathbf{MF}_0$ of the Minimalist Foundation – you can think of it as **CIC** without the universe **Prop**.

Warning!

In both systems there is *cumulativity* of propositions into types *prop* $\hookrightarrow$ *type*, however...

... a *prop* is not any *type* (as in **ML**),

is not any mono-type (as in toposes),

is not any h-prop (as in **HoTT**).

A *prop* is a *prop* is a *prop* is a *prop*.

Corollary (to Beeson's theorem)

$\mathbf{MF}_0$ *is conservative over* **HA**.

First-order fragment of the Minimalist Foundation

We consider on the *first-order fragment* $\mathbf{MF}_0$ of the Minimalist Foundation – you can think of it as **CIC** without the universe **Prop**.

Warning!

In both systems there is *cumulativity* of propositions into types *prop* $\hookrightarrow$ *type*, however...

... a *prop* is not any *type* (as in **ML**),

is not any mono-type (as in toposes),

is not any h-prop (as in **HoTT**).

A *prop* is a *prop* is a *prop* is a *prop*.

Corollary (to Beeson's theorem)

$\mathbf{MF}_0$ *is conservative over* **HA**.

Idea

We claim that *this* result can be extended to classical logic.

A cube of theories

$$
\begin{tikzcd}
 & \mathbf{PAH} \arrow[rr] & & \mathbf{CIC}^c \\
\mathbf{PA} \arrow[ur] \arrow[rr] & & \mathbf{MF}_0^c \arrow[ur] & \\
 & \mathbf{HAH} \arrow[uu] \arrow[rr, "cons."] & & \mathbf{CIC} \arrow[uu] \\
\mathbf{HA} \arrow[uu] \arrow[ur] \arrow[rr, "cons."] & & \mathbf{MF}_0 \arrow[uu] \arrow[ur] &
\end{tikzcd}
$$

- x-axis ($\rightarrow$): add type theory
- y-axis ($\uparrow$): add classical logic
- z-axis ($\nearrow$): add impredicativity

A cube of theories

$$
\begin{tikzcd}
 & \mathbf{PAH} \arrow[rr, "cons?"] & & \mathbf{CIC}^c \\
\mathbf{PA} \arrow[ur] \arrow[rr, "cons?"] & & \mathbf{MF}_0^c \arrow[ur] & \\
 & \mathbf{HAH} \arrow[uu] \arrow[rr, "cons."] & & \mathbf{CIC} \arrow[uu] \\
\mathbf{HA} \arrow[uu] \arrow[ur] \arrow[rr, "cons."] & & \mathbf{MF}_0 \arrow[uu] \arrow[ur] &
\end{tikzcd}
$$

- x-axis ($\rightarrow$): add type theory
- y-axis ($\uparrow$): add classical logic
- z-axis ($\nearrow$): add impredicativity

The Double-Negation Translation

If φ is an arithmetic formula, let $\varphi^{\mathcal{N}}$ be the formula obtained by prefixing a double-negation $\neg\neg$ in front of each existential quantifier and each disjunction appearing in φ.

The Double-Negation Translation

If φ is an arithmetic formula, let $\varphi^{\mathcal{N}}$ be the formula obtained by prefixing a double-negation $\neg\neg$ in front of each existential quantifier and each disjunction appearing in φ.

Theorem (Gödel, 1933)

$\mathbf{PA} \vdash \varphi$ *if and only if* $\mathbf{HA} \vdash \varphi^{\mathcal{N}}$.

The Double-Negation Translation

If φ is an arithmetic formula, let $\varphi^{\mathcal{N}}$ be the formula obtained by prefixing a double-negation $\neg\neg$ in front of each existential quantifier and each disjunction appearing in φ.

Theorem (Gödel, 1933)

PA $\vdash \varphi$ *if and only if* **HA** $\vdash \varphi^{\mathcal{N}}$.

The result is readily extended to higher sorts.

Theorem (Kreisel, 1968)

PAH $\vdash \varphi$ *if and only if* **HAH** $\vdash \varphi^{\mathcal{N}}$.

A cube of theories

$$
\begin{tikzcd}
& \mathbf{PAH} \arrow[rr] & & \mathbf{CIC}^c \\
\mathbf{PA} \arrow[ur] \arrow[rr] & & \mathbf{MF}_0^c \arrow[ur] & \\
& \mathbf{HAH} \arrow[uu] \arrow[rr, "cons."] & & \mathbf{CIC} \arrow[uu] \\
\mathbf{HA} \arrow[uu] \arrow[ur] \arrow[rr, "cons."] & & \mathbf{MF}_0 \arrow[uu] \arrow[ur] &
\end{tikzcd}
$$

- x-axis ($\rightarrow$): add type theory
- y-axis ($\uparrow$): add classical logic
- z-axis ($\nearrow$): add impredicativity

A cube of theories

$$
\begin{tikzcd}
 & \mathbf{PAH} \arrow[rr] \arrow[dd, bend right, red, "\mathcal{N}"'] & & \mathbf{CIC}^c \\
\mathbf{PA} \arrow[ur] \arrow[rr] \arrow[dd, bend right, red, "\mathcal{N}"'] & & \mathbf{MF}_0^c \arrow[ur] & \\
 & \mathbf{HAH} \arrow[uu] \arrow[rr, "cons."] & & \mathbf{CIC} \arrow[uu] \\
\mathbf{HA} \arrow[uu] \arrow[ur] \arrow[rr, "cons."] & & \mathbf{MF}_0 \arrow[uu] \arrow[ur] &
\end{tikzcd}
$$

- x-axis ($\rightarrow$): add type theory
- y-axis ($\uparrow$): add classical logic
- z-axis ($\nearrow$): add impredicativity

A cube of theories

$$
\begin{tikzcd}
 & \mathbf{PAH} \arrow[rr, "\textcolor{red}{cons?}"] \arrow[dd, bend right, swap, "\mathcal{N}"] & & \mathbf{CIC}^c \arrow[dd, bend right, red, swap, "\mathcal{N}"] \\
\mathbf{PA} \arrow[ur] \arrow[rr, "\textcolor{red}{cons?}"] \arrow[dd, bend right, swap, "\mathcal{N}"] & & \mathbf{MF}_0^c \arrow[ur] \arrow[dd, bend right, red, swap, "\mathcal{N}"] & \\
 & \mathbf{HAH} \arrow[uu] \arrow[rr, "cons."] & & \mathbf{CIC} \arrow[uu] \\
\mathbf{HA} \arrow[uu] \arrow[ur] \arrow[rr, "cons."] & & \mathbf{MF}_0 \arrow[uu] \arrow[ur] &
\end{tikzcd}
$$

- x-axis ($\rightarrow$): add type theory
- y-axis ($\uparrow$): add classical logic
- z-axis ($\nearrow$): add impredicativity

The Challenge

In dependent type theories, logical and set-theoretical constructors are highly intertwined:

- terms appear in formulas through equality $a = b$ (as in predicate logic)
- types appear in formulas as domains of quantification $(\exists x : A)\varphi(x)$
- formulas appear in types as in the quotient set constructor A/R
- formulas appear in terms as in the subset term constructor $\{x : A \mid \varphi(x)\} : \mathcal{P}(A)$.

...we need to extend the double-negation translation to every entity!

The double-negation translation for Type Theory

In the case of **MF** and **CIC**, the definition of the translation turns out to be surprisingly simple. The relevant cases are the following.

$$
\begin{aligned}
(\varphi \vee \psi)^{\mathcal{N}} &:\equiv \neg\neg(\varphi^{\mathcal{N}} \vee \psi^{\mathcal{N}}) \\
((\exists x : A)\varphi)^{\mathcal{N}} &:\equiv \neg\neg(\exists x : A^{\mathcal{N}})\varphi^{\mathcal{N}} \\
\mathbf{Prop}^{\mathcal{N}} &:\equiv \sum_{P\,:\,\mathbf{Prop}} \neg\neg P \Rightarrow P
\end{aligned}
$$

The double-negation translation for Type Theory

In the case of **MF** and **CIC**, the definition of the translation turns out to be surprisingly simple. The relevant cases are the following.

$$
\begin{aligned}
(\varphi \vee \psi)^{\mathcal{N}} &:\equiv \neg\neg(\varphi^{\mathcal{N}} \vee \psi^{\mathcal{N}}) \\
((\exists x : A)\varphi)^{\mathcal{N}} &:\equiv \neg\neg(\exists x : A^{\mathcal{N}})\varphi^{\mathcal{N}} \\
\mathbf{Prop}^{\mathcal{N}} &:\equiv \sum_{P : \mathbf{Prop}} \neg\neg P \Rightarrow P
\end{aligned}
$$

Lemma

For any type A we have that $\neg\neg \mathrm{Eq}_{A^{\mathcal{N}}}(x, y) \Rightarrow \mathrm{Eq}_{A^{\mathcal{N}}}(x, y)$ holds.

The double-negation translation for Type Theory

In the case of **MF** and **CIC**, the definition of the translation turns out to be surprisingly simple. The relevant cases are the following.

$$
\begin{aligned}
(\varphi \vee \psi)^{\mathcal{N}} &:\equiv \neg\neg(\varphi^{\mathcal{N}} \vee \psi^{\mathcal{N}}) \\
((\exists x : A)\varphi)^{\mathcal{N}} &:\equiv \neg\neg(\exists x : A^{\mathcal{N}})\varphi^{\mathcal{N}} \\
\mathbf{Prop}^{\mathcal{N}} &:\equiv \sum_{P : \mathbf{Prop}} \neg\neg P \Rightarrow P
\end{aligned}
$$

Lemma

For any type A we have that $\neg\neg\mathrm{Eq}_{A^{\mathcal{N}}}(x, y) \Rightarrow \mathrm{Eq}_{A^{\mathcal{N}}}(x, y)$ holds.

Theorem (Maietti, S.)

A judgment $\mathcal{J}$ is derivable in the classical version if and only if $\mathcal{J}^{\mathcal{N}}$ is derivable in the intuitionistic version.

Conservativity over Peano Arithemtic

Theorem (Contente, S.)

$\mathbf{MF}_0^c$ *is conservative over* **PA** *and* **CIC** *is conservative over* **PAH**.

Conservativity over Peano Arithemtic

Theorem (Contente, S.)

$\mathbf{MF}_0^c$ *is conservative over* **PA** *and* **CIC** *is conservative over* **PAH**.

Proof.

Let φ be an arithmetical proposition, and assume φ is true in $\mathbf{MF}_0^c$.

$$\begin{tikzcd}
\mathbf{PA} \arrow[r, "\textit{cons?}"] & \mathbf{MF}_0^c \quad \varphi \textit{ true} \\
\mathbf{HA} \arrow[u] \arrow[r, "\textit{cons.}"] & \mathbf{MF}_0 \arrow[u]
\end{tikzcd}$$

□

Conservativity over Peano Arithemtic

Theorem (Contente, S.)

$\mathbf{MF}_0^c$ *is conservative over* **PA** *and* **CIC** *is conservative over* **PAH**.

Proof.

Let φ be an arithmetical proposition, and assume φ is true in $\mathbf{MF}_0^c$.

$$\begin{tikzcd}
\mathbf{PA} \arrow[r, "\color{red}{cons?}"] & \mathbf{MF}_0^c \quad \varphi \text{ true} \\
\mathbf{HA} \arrow[u] \arrow[r, "cons."] & \mathbf{MF}_0 \arrow[u] \quad \varphi^{\mathcal{N}} \text{ true}
\end{tikzcd}$$

□

Conservativity over Peano Arithemtic

Theorem (Contente, S.)

$\mathbf{MF}_0^c$ *is conservative over* **PA** *and* **CIC** *is conservative over* **PAH**.

Proof.

Let φ be an arithmetical proposition, and assume φ is true in $\mathbf{MF}_0^c$.

$$
\begin{tikzcd}
& \mathbf{PA} \arrow[r, "\textcolor{red}{\mathit{cons?}}"] & \mathbf{MF}_0^c & \varphi\ \mathit{true} \\
& \mathbf{HA} \arrow[u] \arrow[r, "\mathit{cons.}"] & \mathbf{MF}_0 \arrow[u] & \\
\varphi^{\mathcal{N}}\ \mathit{true} & & & \varphi^{\mathcal{N}}\ \mathit{true}
\end{tikzcd}
$$

$\square$

Conservativity over Peano Arithemtic

Theorem (Contente, S.)

$\mathbf{MF}_0^c$ *is conservative over* $\mathbf{PA}$ *and* $\mathbf{CIC}$ *is conservative over* $\mathbf{PAH}$*.*

Proof.

Let φ be an arithmetical proposition, and assume φ is true in $\mathbf{MF}_0^c$.

$$\begin{tikzcd}
\varphi \textit{ true} & \mathbf{PA} \arrow[rr, "cons!"] & & \mathbf{MF}_0^c & \varphi \textit{ true} \\
\varphi^{\mathcal{N}} \textit{ true} & \mathbf{HA} \arrow[u] \arrow[rr, "cons."] & & \mathbf{MF}_0 \arrow[u] & \varphi^{\mathcal{N}} \textit{ true}
\end{tikzcd}$$

□

Thanks for your attention!