A categorical approach to Gödel's incompleteness via arithmetic universes

Yuto Ikeda

Graduate School of Mathematical Sciences, The University of Tokyo

CSCAT 2025 at Sojo University, Kumamoto
2025/03/11

This work is supported by FoPM, WINGS Program, The University of Tokyo.

Backgrounds

- In 1973, André Joyal gave a categorical interpretation of Gödel's incompleteness theorems in his lecture, introducing certain categories called **arithmetic universes (AUs)** (Joyal 2005).
- However, his work has not been published.
- Although the notion of AUs has been studied by several authors, a precise description of its applications to incompleteness has mostly not been publicly available until now, as far as I know.
- In this talk, I present my own arrangement of Joyal's idea based on a recent literature (Dijk and Oldenziel 2020). I hope it is helpful for reconstruction, refinement and expansion of his insight.

Key idea

logic	category
meta theory formalize ⇣ object theory	category internalize ⇣ internal category

Outline

Outline

Basic observation in categorical logic

logic / type theory
(logical theory / specification)

syntactic category →
$\perp$
← internal language

category

logic	higher order logic	topos
	regular logic	regular category
	⋮	⋮
↓ "propositions as some types" ↓		
type theory	extensional dependent type theory	finitely complete (1-)category

Type constructors and structures on categories

There are correspondences between them as (Maietti 2005):

extensional DTT	category
dependent sum Σ, unit 1, extensional identity $=$	finite limits
exists $\exists$ / prop. truncation $\|-\|$	stable image
or $\vee$	stable union of subobjects
disjoint sum $+$, empty 0	stable disjoint coproduct
effective quotient	stable effective quotient
natural numbers type $\mathbb{N}$	(parameterized) NNO
list type $\mathrm{List}(A)$	(parameterized) list object
dependent product Π ($\forall$)	exponential in slice categories
type of propositions Prop	subobject classifier

Type constructors and structures on categories

For example, a **topos** (with an NNO) has all of the structures in this table:

extensional DTT	category
dependent sum Σ, unit 1, extensional identity $=$	**finite limits**
exists $\exists$ / prop. truncation $\lVert - \rVert$	**stable image**
or $\vee$	**stable union of subobjects**
disjoint sum $+$, empty 0	**stable disjoint coproduct**
effective quotient	**stable effective quotient**
natural numbers type $\mathbb{N}$	**(parameterized) NNO**
list type $\mathrm{List}(A)$	**(parameterized) list object**
dependent product Π ($\forall$)	**exponential in slice categories**
type of propositions Prop	**subobject classifier**

Type constructors and structures on categories

In contrast, a **pretopos** has only those structures:

extensional DTT	category
dependent sum Σ, unit 1, extensional identity $=$	**finite limits**
exists $\exists$ / prop. truncation $\|-\|$	**stable image**
or $\vee$	**stable union of subobjects**
disjoint sum $+$, empty 0	**stable disjoint coproduct**
effective quotient	**stable effective quotient**
natural numbers type $\mathbb{N}$	(parameterized) NNO
list type $\mathrm{List}(A)$	(parameterized) list object
dependent product Π ($\forall$)	exponential in slice categories
type of propositions Prop	subobject classifier

Type constructors and structures on categories

… and, an **arithmetic universe** is a category with here.

extensional DTT	category
dependent sum Σ, unit 1, extensional identity $=$	**finite limits**
exists $\exists$ / prop. truncation $\|-\|$	**stable image**
or $\vee$	**stable union of subobjects**
disjoint sum $+$, empty 0	**stable disjoint coproduct**
effective quotient	**stable effective quotient**
natural numbers type $\mathbb{N}$	**(parameterized) NNO**
list type $\mathrm{List}(A)$	**(parameterized) list object**
dependent product Π ($\forall$)	exponential in slice categories
type of propositions Prop	subobject classifier

Definition of AU

Definition (Maietti 2003)

An **arithmetic universe (AU)** is a pretopos which has parameterized list objects for any objects.

The internal languages of an AU have

- type constructors $1, \times, 0, +, \Sigma, A/R, \mathbb{N}, \mathrm{List}(A)$,
- logical connectives $=, \top, \wedge, \bot, \vee, \exists$,

but it **lacks**

- type constructors $\to, \Pi, \mathrm{Prop}$,
- logical connectives $\neg, \to, \forall$.

Mathematics in AU

Although AUs have weaker structures than topoi,
some extent of (strongly predicative) mathematics can be done in AUs:

- (Π_2-fragments of) $I\Sigma_1$ arithmetic (Ye, 2022)
  - Since AUs do not have $\forall$, only Σ_1-formulas are expressible.
- Construction of algebras from generators and relations
  - It makes essential use of lists and quotients.
- Free construction of categories and presheaves from graphs (Maietti, 2003)

Outline

Initial AU

An **AU-functor** is a functor preserving the structures of an AU.

Proposition

There is the **initial AU** $\mathcal{A}_0$ in the sense that, for any AU $\mathcal{A}$, there is an AU-functor $F : \mathcal{A}_0 \to \mathcal{A}$ unique up to natural isomorphism.

In fact, we have a **specific construction** of $\mathcal{A}_0$:
it is just the **syntactic category** of a certain dependent type theory.

(Joyal gave a more "down-to-earth" construction of $\mathcal{A}_0$.)

Internal initial AU

Key observation: the construction of a syntactic category only involves basic operations to **finite lists** of symbols and taking **quotients**.

Proposition–Definition (Morrison 1996), (Maietti 2003), (Vickers 2019)

The construction of the initial AU can be performed within internal languages of AUs, specifically in $\mathcal{A}_0$.

It gives rise to the internal category (or the "internal AU") $\mathbb{A}_0$ in $\mathcal{A}_0$, called the **internal initial AU**.

$\mathcal{A}_0$ and $\mathbb{A}_0$ correspond to a **meta theory** and an **object theory** respectively.

Externalization

Let $\Gamma := \mathrm{Hom}(1, -) : \mathcal{C} \to \mathbf{Set}$, the global sections functor of $\mathcal{C}$.

Definition

For an internal category $\mathbb{C}$ in $\mathcal{C}$, its **externalization** $\mathrm{Ext}(\mathbb{C})$ is the small category which is the image of $\mathbb{C}$ by $\Gamma : \mathcal{C} \to \mathbf{Set}$.

$\mathrm{Ext}(\mathbb{C})$ corresponds to a theory consisting of **closed terms** in the meta theory $\mathcal{C}$ denoting syntactic entities (terms, formulas, ...) of the object theory $\mathbb{C}$.

Internalization of global sections and subobjects

Using $\mathrm{Ext}(\mathbb{C})$, we can define ordinary functors for $\mathbb{C}$ in $\mathcal{C}$, such as

the global sections functor $\Gamma_{\mathbb{C}} : \mathrm{Ext}(\mathbb{C}) \to \mathcal{C}$,

the subobject functor $\mathrm{Sub}_{\mathbb{C}} : \mathrm{Ext}(\mathbb{C})^{\mathrm{op}} \to \mathcal{C}$,

pullbacks $\mathrm{ev}_a : \Gamma_{\mathbb{C}}(a) \times \mathrm{Sub}_{\mathbb{C}}(a) \to \mathrm{Sub}_{\mathbb{C}}(1_{\mathbb{C}})$ for $a \in \mathrm{Ext}(\mathbb{C})$,

… if we assume $\mathcal{C}$ or $\mathbb{C}$ has sufficient structures.

Gödel coding functor

Proposition

$\mathrm{Ext}(\mathbb{A}_0)$ is an AU.

Definition

The unique AU-functor $\ulcorner - \urcorner : \mathcal{A}_0 \to \mathrm{Ext}(\mathbb{A}_0)$ is called **Gödel coding functor**.

This functor sends entities of the meta theory $\mathcal{A}_0$ into closed terms denoting corresponding entities of the object theory $\mathbb{A}_0$.
It corresponds to taking **numerals of Gödel numbers** $\varphi \mapsto \overline{\ulcorner \varphi \urcorner}$.

Provability functor

Definition (Dijk and Oldenziel 2020)

The **provability functor** $\Box$ for $\mathcal{A}_0$ is an endofunctor

$$\mathcal{A}_0 \xrightarrow{\ulcorner - \urcorner} \mathrm{Ext}(\mathbb{A}_0) \xrightarrow{\Gamma_{\mathbb{A}_0}} \mathcal{A}_0 .$$

In fact, $\Gamma_{\mathbb{C}}$ acts on subterminals in $\mathbb{C}$ as taking an equalizer with $\top_{\mathbb{C}}$, so it serves as a **provability predicate** on subterminals.

The functor $\Box$ generalizes classical provability $\Box\varphi :\equiv \mathrm{Pr}\left(\ulcorner\varphi\urcorner\right)$ to the whole syntactic category (without non-canonical coding)!

Remark. $\Gamma_{\mathbb{A}_0}$ is NOT an AU-functor.

Outline

Construction of an undecidable sentence

Since the Gödel sentence is a Π_1-sentence, which is not expressible in AUs, we construct **the Jeroslow sentence** $J \leftrightarrow \Box\neg J$ instead.

Lemma

We can construct a subterminal $J \rightarrowtail 1$ in $\mathcal{A}_0$ such that J is an equalizer of $\ulcorner J \urcorner, \ulcorner \bot \urcorner : 1 \rightrightarrows \mathrm{Sub}_{\mathbb{A}_0}(\ulcorner 1 \urcorner)$.

Proof sketch. The diagonal argument. $\Box$

Construction of an undecidable sentence

Proof. Let N be the NNO in $\mathcal{A}_0$. We can construct:

enumeration of formulas point-surjective $e : N \twoheadrightarrow \mathrm{Sub}_{\mathbb{A}_0}(\ulcorner N \urcorner)$,
numeral function $\eta_N : N \to \Gamma_{\mathbb{A}_0}(\ulcorner N \urcorner)$, which expresses $n \mapsto \overline{\ulcorner n \urcorner}$.

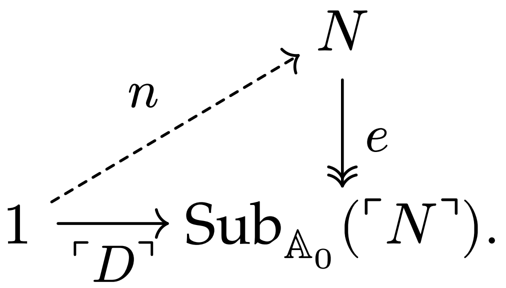
Using them, make the pullback

$$\begin{array}{ccccccc}
D & & & \longrightarrow & & & 1 \\
\downarrow & & & & & & \downarrow{\scriptstyle \ulcorner \bot \urcorner} \\
N & \xrightarrow{\Delta_N} & N \times N & \xrightarrow{\eta_N \times e} & \Gamma_{\mathbb{A}_0}(\ulcorner N \urcorner) \times \mathrm{Sub}_{\mathbb{A}_0}(\ulcorner N \urcorner) & \xrightarrow{\mathrm{ev}_{\ulcorner N \urcorner}} & \mathrm{Sub}_{\mathbb{A}_0}(\ulcorner 1 \urcorner).
\end{array}$$

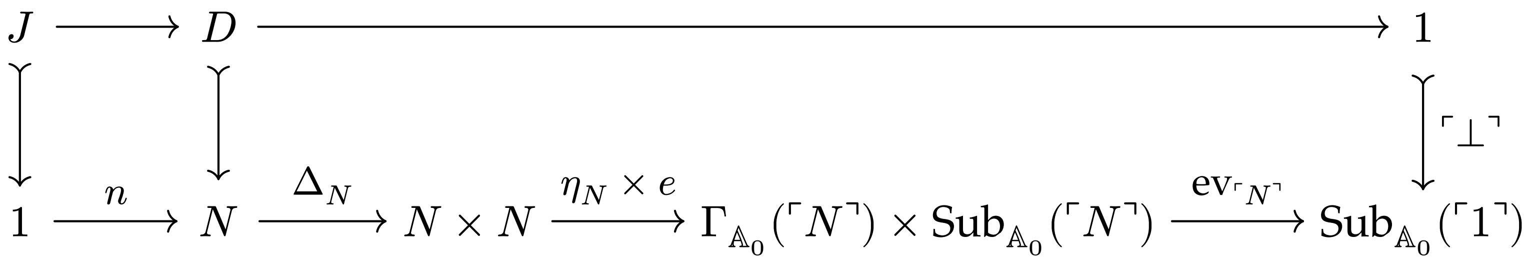
$D \rightarrowtail N$ expresses the predicate of n which states "$\varphi_n(n)$ is provably equivalent to $\bot$ in $\mathbb{A}_0$," or concisely, $\Box \neg \varphi_n(n)$.

Construction of an undecidable sentence

Since e is point-surjective, we have $n : 1 \to N$ such that



Let the subterminal $J \in \mathrm{Sub}(1)$ be the pullback of $D \rightarrowtail N$ along n.



Since the bottom side is equal to $\ulcorner J \urcorner$, the lemma follows. $\square$

First incompleteness theorem

Lemma (repeat)

$J \rightarrowtail 1$ is an equalizer of $\ulcorner J \urcorner, \ulcorner \bot \urcorner : 1 \rightrightarrows \mathrm{Sub}_{\mathbb{A}_0}(\ulcorner 1 \urcorner)$.

Theorem (first incompleteness theorem)

$J \in \mathrm{Sub}(1)$ is not equivalent to $\bot$ nor $\top$.

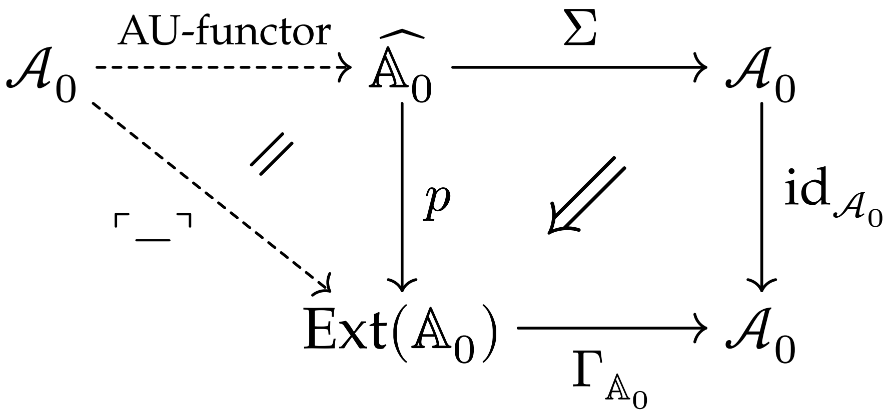
Proof. If $J \cong \bot$, then $\ulcorner J \urcorner = \ulcorner \bot \urcorner$, hence their equalizer J is equivalent to $\top$. So $\top \cong J \cong \bot$, which contradicts to non-triviality of $\mathcal{A}_0$.

If $J \cong \top$, then the lemma says $\ulcorner \top \urcorner$ and $\ulcorner \bot \urcorner$ are the same. Applying the unique AU-functor $\mathcal{A}_0 \to \mathbf{Set}$, it follows that $\top$ and $\bot$ in $\mathcal{A}_0$ are the same, which contradicts to non-triviality of $\mathcal{A}_0$ again. □

Freyd cover and Σ_1-completeness

Lemma (Freyd cover of $\mathbb{A}_0$)

The comma category $\widehat{\mathbb{A}_0} := \mathrm{id}_{\mathcal{A}_0} \downarrow \Gamma_{\mathbb{A}_0}$ is an AU. Moreover, two forgetful functors Σ, p are AU-functors.



This gives a natural transformation $\eta : \mathrm{id}_{\mathcal{A}_0} \Rightarrow \Box$.
$\eta_A : A \to \Box A$ corresponds to **formalized Σ_1-completeness** of $\mathbb{A}_0$ in $\mathcal{A}_0$.

Second incompleteness theorem

Theorem (second incompleteness theorem)

Let $\mathrm{Incon}(\mathbb{A}_0) \in \mathrm{Sub}(1)$ be the equalizer of $\ulcorner\bot\urcorner$ and $\ulcorner\top\urcorner$.
Then, $\mathrm{Incon}(\mathbb{A}_0) \not\cong \bot$.

Proof.

1. J is an equalizer of $\ulcorner J\urcorner$ and $\ulcorner\bot\urcorner$.
2. Since $\Box J$ is an equalizer of $\ulcorner J\urcorner$ and $\ulcorner\top\urcorner$ ($\Box$ serves as provability), the Σ_1-completeness $J \to \Box J$ implies that J equalizes $\ulcorner J\urcorner$ and $\ulcorner\top\urcorner$.
3. Therefore, J also equalizes $\ulcorner\bot\urcorner$ and $\ulcorner\top\urcorner$, i.e. $J \leq \mathrm{Incon}(\mathbb{A}_0)$.

Hence, if $\mathrm{Incon}(\mathbb{A}_0) \cong \bot$, then $J \cong \bot$, contradicting with the first incompleteness theorem. □

Outline

Conclusions

Summary:

- The idea of "formalizing a theory in itself" can be interpreted in terms of internal categories.
- It might give a structural view to meta-object interactions, which seem unique and fascinating ideas in logic (to me).

Future work:

- Establish a precise proof, its refinement, and broader applications.
- Analyze properties of the provability functor $\square$.
- Can we find other situations like "constructing itself in itself" interpreted/unified in the framework of internal/fibred categories?

References

Dijk, Joost van, and Alexander Gietelink Oldenziel. 2020. "Gödel Incompleteness Through Arithmetic Universes After A. Joyal". arXiv. April 22, 2020. https://doi.org/10.48550/arXiv.2004.10482.

Goodstein, R. L. 1954. "Logic-Free Formalisations of Recursive Arithmetic". *MATHEMATICA SCANDINAVICA* 2 (December):246–60. https://doi.org/10.7146/math.scand.a-10412.

Jacobs, Bart. 2001. *Categorical Logic and Type Theory*. Paperback ed. Studies in Logic and the Foundations of Mathematics. Amsterdam: Elsevier.

Johnstone, Peter. 2002. *Sketches of an Elephant: A Topos Theory Compendium*. Oxford Logic Guides. Oxford: Clarendon Press.

Joyal, André. 2005. "The Gödel Incompleteness Theorem, A Categorical Approach". *Cahiers De Topologie Et Géométrie Différentielle Catégoriques* 46 (3): 202. http://www.numdam.org/item/CTGDC_2005__46_3_163_0/.

Lambek, Joachim, and Philip J. Scott. 1994. *Introduction to Higher Order Categorical Logic*. Paperback ed. (with corr.), reprinted. Cambridge Studies in Advanced Mathematics. Cambridge: Cambridge Univ. Press.

References

Maietti, Maria Emilia. 2003. “Joyal's Arithmetic Universes via Type Theory”. *Electronic Notes in Theoretical Computer Science*, CTCS'02, Category Theory and Computer Science, 69 (February):272–86. https://doi.org/10.1016/S1571-0661(04)80569-3.

Maietti, Maria Emilia. 2005. “Modular Correspondence between Dependent Type Theories and Categories Including Pretopoi and Topoi”. *Mathematical Structures in Computer Science* 15 (6): 1089. https://doi.org/10.1017/S0960129505004962.

Maietti, Maria Emillia. 2010. “Joyal's Arithmetic Universe as List-Arithmetic Pretopos”. *Theory and Applications of Categories* 24 (January).

Morrison, Alan. 1996. “Reasoning in Arithmetic Universe.”

Román, Leopoldo. 1989. “Cartesian Categories with Natural Numbers Object”. *Journal of Pure and Applied Algebra* 58 (3): 267–78. https://doi.org/10.1016/0022-4049(89)90042-X.

Vickers, Steven. 2019. “Sketches for Arithmetic Universes”. *Journal of Logic and Analysis* 11 (June). https://doi.org/10.4115/jla.2019.11.FT4.

Ye, Lingyuan. 2023. “Categorical Structure in Theory of Arithmetic”. arXiv. May 2, 2023. https://doi.org/10.48550/arXiv.2304.05477.

Construction of the initial AU: Step 1

Let Σ_0 be the initial category with finite products and NNO. Explicitly,

Objects $N^0, N^1, N^2, \ldots$ (finite products of N).

Morphisms $N^k \to N$ are codes of primitive recursive functions with k variables, up to the congruence relation generated by the definition of each function and

$$\frac{f(\overline{x}, 0) \sim g(\overline{x}, 0) \quad f(\overline{x}, Sy) \sim h(\overline{x}, y, f(\overline{x}, y)) \quad g(\overline{x}, Sy) \sim h(\overline{x}, y, g(\overline{x}, y))}{f(\overline{x}, y) \sim g(\overline{x}, y)}$$

Construction of the initial AU: Step 2

Let $\mathrm{Pred}(\Sigma_0)$ be the category consists of:

Objects $P : N \to N$ in Σ_0 such that $P \times P = P$.

Intuitively, P indicates $P^{-1}(1) \subseteq N$ decidable by primitive recursions.

Morphisms $P \to Q$ are $f : N \to N$ in Σ_0 such that $P \leq Q \circ f$,
up to the equivalence $f \sim g$ defined by $f \times P = g \times P$.

Remark. $\mathrm{Pred}(\Sigma_0)$ satisfies the axioms of AU except for exactness.

Construction of the initial AU: Step 3

$\mathcal{A}_0$ is obtained as the exact completion of $\mathrm{Pred}(\Sigma_0)$. Explicitly:

Objects Monic equivalence relations $R \rightrightarrows X$ in $\mathrm{Pred}(\Sigma_0)$.

Morphisms $(R \rightrightarrows X) \to (S \rightrightarrows Y)$ are equivalence classes of $f : X \to Y$ such that $R \subseteq (f \times f)^{-1}(S)$.
The equivalence $f \sim g$ is defined by $R \subseteq (f \times g)^{-1}(S)$.

Global sections and provability

$\Gamma_{\mathcal{C}} : \mathcal{C} \to \mathbf{Set}$ acts on subterminals as

$$\begin{aligned} \mathrm{Sub}_{\mathcal{C}}(1_{\mathcal{C}}) &\longrightarrow \mathcal{P}(\{*\}) \\ A &\longmapsto \{x \in \{*\} \mid A \cong \text{the maximal subterminal } \top_{\mathcal{C}}\}, \end{aligned}$$

which can be read as "the closed formula A is provable in $\mathcal{C}$".
In this sense, $\Gamma_{\mathcal{C}}$ includes the notion of **provability**.

This can be internalized;
$\Gamma_{\mathbb{C}} : \mathrm{Ext}(\mathbb{C}) \to \mathcal{C}$ naturally induces $\Gamma_{\mathcal{C}}(\mathrm{Sub}_{\mathbb{C}}(1_{\mathbb{C}})) \to \mathrm{Sub}_{\mathcal{C}}(1_{\mathcal{C}})$,
which serves as **provability predicates**.
Another way to say this is taking the equalizer with $\top_{\mathbb{C}} \in \Gamma_{\mathcal{C}}(\mathrm{Sub}_{\mathbb{C}}(1_{\mathbb{C}}))$.

Global sections and provability

Proposition

Let $P : 1 \to \mathrm{Sub}_{\mathbb{C}}(1_{\mathbb{C}})$ be (a global element denoting) a subterminal in $\mathbb{C}$. Then, $\Gamma_{\mathbb{C}}(P)$ is an equalizer of P and the maximal subterminal $\top_{\mathbb{C}}$ in $\mathbb{C}$:

$$\Gamma_{\mathbb{C}}(P) \rightarrowtail 1 \underset{\top_{\mathbb{C}}}{\overset{P}{\rightrightarrows}} \mathrm{Sub}_{\mathbb{C}}(1_{\mathbb{C}})$$

In a syntactic category, coincidence with $\top \in \mathrm{Sub}(1)$ means **provability** of the sentence.

The above says $\Gamma_{\mathbb{C}}$ serves as **provability predicates** on subterminals.