This is a ``preproof'' accepted article for *The Review of Symbolic Logic*.
This version may be subject to change during the production process.
DOI: 10.1017/S1755020326101439

LEARNING THE INTENDED MODEL OF ARITHMETIC

LEON HORSTEN, LUCA SAN MAURO, AND GIORGIO VENTURI

Abstract. We study epistemic games between a teacher and a student. The student fixes a presentation of a model of true arithmetic and truthfully answers atomic questions about it; after each answer the teacher conjectures whether the presentation is standard, and she wins if her conjectures eventually stabilise on the correct verdict. In the game $\mathbb{G}_1$, where the student may choose any ω-presentation of a model of $\mathrm{Th}(\mathcal{N})$, the teacher has no winning strategy: even if the student's answers are governed by a fixed, determinate model, finite arithmetical interaction does not guarantee, even in the limit, recognition of whether that model is standard. In $\mathbb{G}_2$, standard presentations must have computable addition (equivalently, they must be computable); the teacher then has a winning strategy, and we leave open whether she has a computable one. In $\mathbb{G}_3$, standard presentations must have primitive-recursive addition; the teacher then has a computable winning strategy. The results show how constraints on the presentations of arithmetic bear on whether another agent can recognise its standardness.

1. Introduction

Many mathematical theories are taken to have an *intended model*: the structure that our mathematical practice is about. Arithmetic is the paradigm case. When we reason about the natural numbers, we take ourselves to speak about a single structure $\mathcal{N}$, in which claims such as the associativity of addition or the uniqueness of prime factorisation are determinately true or false.[1]

Classical model theory makes this picture hard to vindicate. By the Löwenheim–Skolem theorem and compactness, first-order Peano Arithmetic (PA) does not characterise $\mathcal{N}$ up to isomorphism; indeed, every consistent completion of PA has continuum many pairwise non-isomorphic countable models.[2] Yet mathematicians take their arithmetical discourse to concern $\mathcal{N}$, not a nonstandard model, even one that satisfies exactly the same first-order sentences. We call this tension the *intended-model puzzle* P for arithmetic.[3]

2020 *Mathematics Subject Classification.* 03A05, 03C62, 03C57.

Key words and phrases. intended model of arithmetic, computable structure theory, learning in the limit, Tennenbaum's theorem.

[1]For classic discussion of reference and determinacy in arithmetic, see e.g. Benacerraf (1965); Putnam (1980); Shapiro (1991); Resnik (1997); Hellman (1989).

[2]See, for example, Chang and Keisler (1990); Kaye (1991).

[3]The puzzle is often discussed alongside broader "model-theoretic" concerns about reference and determinacy; see Putnam (1980); Shapiro (1991).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 20:34:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1755020326101439

A natural response is to move to second-order logic: under full semantics, the second-order Peano axioms are categorical (Dedekind, 1888). This move is controversial, because full semantics presupposes a determinate notion of *all* subsets of ω, which many regard as no less problematic than the notion of natural number itself.[4]

A different line of response appeals to computability theory, and in particular to a classical theorem of Tennenbaum.[5]

Theorem (Tennenbaum)**.** *No countable nonstandard model* $\mathcal{M}$ *of* PA *has computable addition or computable multiplication.*

Some philosophers have taken this theorem to supply the missing ingredient for solving P: our grasp of the natural numbers incorporates an effectiveness requirement, and among the countable models of PA the standard one is, up to isomorphism, the only one with a computable presentation.[6] Others have been sceptical. They question whether computability conditions can be built into the content of arithmetical concepts, rather than being facts about our epistemic access, and whether the non-computability of nonstandard models suffices to exclude them as candidates for idealised agents.[7] A further point is that the philosophical appeal to Tennenbaum's theorem presupposes a background notion of effective computation. Its use in addressing the intended-model puzzle therefore depends on how that background notion is understood and justified.[8]

Computability constraints of this kind can also be asked to do epistemic work. Suppose that a speaker's arithmetical practice is governed by a fully determinate model. Can another agent, through communication, recognise whether that model is standard? And what difference do such constraints make to the answer? These are the questions we address.

Our set-up is inspired by Gold-style identification in the limit.[9] A student **S** and a teacher **T** interact. The student's answers are governed by a fixed ω-presentation $\mathcal{M}$ of a model of true arithmetic (in particular, a countable structure with domain ω), subject to background constraints; the teacher is not told which one. At each round the teacher asks an atomic question about $\mathcal{M}$—for instance, whether $a + b = c$ or $a < b$ holds for the elements carrying given labels—and, after receiving the answer,

[4]Full second-order semantics is not required: arithmetic also admits a categorical axiomatisation in weak second-order logic, with second-order quantifiers ranging over externally finite subsets of the domain; see Shapiro (1991). On the controversy over full second-order semantics, see Shapiro (1991); Button (2013); McGee (1997). Other responses to the puzzle appeal to schematic resources, internal categoricity, generalised quantifiers, and computability constraints; for a survey and taxonomy, see Button and Walsh (2018, Chs. 7, 10).

[5]Tennenbaum (1959); for a textbook proof, see Kaye (1991).

[6]For versions of this idea, and for the label "computational structuralism", see Halbach and Horsten (2005). For an early argument from Tennenbaum's theorem using cognitive and computational assumptions, see Quinon and Zdanowski (2007). For critical assessment and refinements, see Button and Smith (2012); Dean (2014); Button and Walsh (2018).

[7]See Button and Smith (2012) for a map of the immediate dialectic around Tennenbaum's theorem, and Dean (2014) for a detailed discussion of models, presentations, and computability.

[8]For a related point, see Picollo and Waxman (2025).

[9]Gold (1967) initiated the area of inductive inference, which in turn helped stimulate the broader field of computational learning theory (cf. Angluin and Smith (1983); Kearns and Vazirani (1994); Kelly (1996)). Recently, the same general paradigm has been applied to questions about the learnability of algebraic and computable structures; see, for example, Bazhenov et al. (2020), Bazhenov et al. (2023), Bazhenov et al. (2026). Some of our non-learnability arguments echo construction patterns from this structural learnability literature.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 20:34:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1755020326101439

outputs a conjecture: **st** ($\mathcal{M}$ is a copy of $\mathcal{N}$) or **nst** ($\mathcal{M}$ is nonstandard). The teacher *validates* $\mathcal{M}$ if her conjectures eventually stabilise on the correct verdict; she need not know when this has happened.

For instance, the teacher may ask whether $+_{\mathcal{M}}(2,3,5)$ holds, that is, whether the student's addition, applied to the elements labelled 2 and 3, yields the element labelled 5. A positive answer reflects the chosen presentation as much as the underlying model: some presentations of $\mathcal{N}$ falsify this atomic fact (permute the labels), and suitable presentations of nonstandard models satisfy it.

Interactions of this kind, in which an agent's answers are tested case by case and verdicts may be revised, are familiar from the rule-following discussions after Wittgenstein and Kripke, where finite behavioural evidence underdetermines which rule a pupil is following.[10]

We study three infinite games, $\mathbb{G}_1$, $\mathbb{G}_2$, and $\mathbb{G}_3$, which differ only in the constraints imposed on the presentation that the student selects before the first round. We distinguish *in-principle* learnability, i.e. the existence of some winning strategy, possibly relying on non-effective information such as a non-computable list of candidates, from *effective* learnability, i.e. the existence of a computable winning strategy. Our main results, proved in Sections 3–5 and summarised in Table 1, are the following:

- In $\mathbb{G}_1$ there is no effectiveness constraint: the student may choose any copy of a model of $\mathrm{Th}(\mathcal{N})$. The teacher has no winning strategy (Theorem 3.8).
- In $\mathbb{G}_2$, if the student chooses a copy of $\mathcal{N}$, its addition must be computable (equivalently, the whole presentation must be computable). The teacher has a winning strategy (Theorem 4.2); whether she has a computable one is left open (Question 4.3).
- In $\mathbb{G}_3$, if the student chooses a copy of $\mathcal{N}$, its addition must be primitive recursive. The teacher has a computable winning strategy (Theorem 5.3), which exploits the effective enumeration of the primitive recursive binary functions.

Game	Constraint on standard presentations	Winning-strategy status
$\mathbb{G}_1$	none	no winning strategy exists
$\mathbb{G}_2$	computable addition	winning strategy exists; computability open
$\mathbb{G}_3$	primitive-recursive addition	computable winning strategy exists

TABLE 1. The three games and the main results.

Technically, we combine computable structure theory with learning in the limit: we study what growing finite fragments of the atomic diagram of a model of true arithmetic allow a teacher to infer, in the limit, about its isomorphism type, and we use this to construct, or rule out, winning teacher strategies.

Categoricity, determinacy, and recognition. Because the computability proposals were advanced as answers to P, our results are easily misread. The negative result for $\mathbb{G}_1$ may look like evidence that arithmetical reference is indeterminate, and the positive results for $\mathbb{G}_2$ and $\mathbb{G}_3$ like a vindication of computational structuralism.

[10]See Wittgenstein (1953); Kripke (1982); McDowell (1984); Wright (1980). Unlike those discussions, we take it as given that there is a fact about the student's practice, namely the model that governs her answers.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 20:34:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1755020326101439

To block both readings, we separate three questions. A *categoricity* question asks whether a theory, relative to a logic, a semantics, and a class of admissible models, determines a single isomorphism type. A *metasemantic* question asks whether a speaker's use, theory, and practice determine what her arithmetical vocabulary refers to and which truth values her arithmetical sentences have. An *intersubjective epistemic* question asks what one agent can establish about the arithmetical system that governs another agent's judgements. These questions are related, but distinct. Our games concern only the third question: they grant determinacy and ask about recognition.

The intersubjective dimension is already present in the debate on the intended model, notably in work on internal categoricity (Parsons, 1990; Väänänen, 2012; Väänänen and Wang, 2015; Maddy and Väänänen, 2023). In Parsons's scenario, Kurt and Michael use distinct vocabularies for number, zero, and successor. If each can use the other's vocabulary and apply induction to predicates of the combined language, they can prove that the map sending zero to zero and successors to successors is an isomorphism between their number systems. Since the internal approach largely dispenses with model-theoretic semantics, what matters philosophically is the consequence that the two speakers cannot disagree on the truth value of any arithmetical sentence (sometimes called *intolerance*).[11] Our games differ in two respects. First, they are formulated in terms of models and their presentations, which the internal approach sets aside, and they assume intolerance rather than establish it: every admissible student model satisfies $\mathrm{Th}(\mathcal{N})$. Second, the teacher does not compare two systems represented in a shared language; she has access to a single presentation, and only through successive finite answers. The question is whether she can nevertheless recognise, in the limit, whether that presentation is standard.

Scope and idealisations. Our framework is metatheoretic: we work with ω-presentations and indices in the metalanguage. Atomic queries model elementary, publicly checkable uses of arithmetical vocabulary, such as calculations and comparisons on specified inputs, and give the teacher no direct access to global notions such as standardness, well-foundedness, or satisfaction. The positive results use only such atomic access; the negative result for $\mathbb{G}_1$ does not depend on it, since it persists when the teacher may ask arbitrary first-order questions with finitely many named parameters (Remark 3.9). Convergence in the limit idealises eventual coordination: stable correction after finitely many revisions, not recognition in finite time. Finally, the notions of computability, primitive recursiveness, and standardness that occur in our results are, like those in Tennenbaum's theorem, relative to a background notion of computation fixed in the metatheory.

Organisation of the paper. Section 2 fixes notation for ω-presentations and atomic diagrams; readers mainly interested in the philosophical upshot may treat it as a glossary. Section 3 introduces $\mathbb{G}_1$ and proves that the teacher has no winning strategy. Section 4 introduces $\mathbb{G}_2$, establishes the existence of a winning strategy, and formulates the open effectiveness question. Section 5 introduces $\mathbb{G}_3$ and gives a

[11] See Parsons (1990), Button and Walsh (2016), Button and Walsh (2018, Ch. 10, esp. §10.a), and Maddy and Väänänen (2023). On the limits of this reading, see Button and Walsh (2018, §10.3 and §10.a).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 20:34:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1755020326101439

computable winning strategy. We conclude with a discussion of the philosophical significance of these results.

2. Formal preliminaries

In this section we fix notation and recall some basic notions that will be used throughout the paper. We assume that the reader is familiar with the standard foundations of first-order logic, including elementary model theory, as well as with the basic concepts of computability theory. Suitable background can be found, for instance, in standard textbooks such as Hodges (1993) for model theory and Soare (2016) or Odifreddi (1989) for computability theory.

We let $2^{<\omega}$ denote the set of all finite binary strings and 2^{ω} the set of all infinite binary sequences. If $\sigma, \tau \in 2^{<\omega}$, we write $\sigma \preceq \tau$ to mean that σ is an initial segment of τ. Concatenation of strings is denoted by $\sigma^\frown\tau$. We denote by λ the empty string. A *tree* on $2^{<\omega}$ is a set $T \subseteq 2^{<\omega}$ that is closed under initial segments: if $\tau \in T$ and $\sigma \preceq \tau$, then $\sigma \in T$. An *infinite path* through T is an element $\pi \in 2^{\omega}$ such that $\pi \upharpoonright n \in T$ for all $n \in \omega$, where $\pi \upharpoonright n$ denotes the restriction of π to its first n bits.

In Sections 4 and 5 we schedule countably many finite checks while asking only one atomic query per round. We therefore fix standard computable codings once and for all.

Fact 2.1. *There is a computable bijection*

$$\rho_2 \colon \omega \longrightarrow \omega^2$$

with computable coordinate projections. Moreover, there is a computable enumeration $(\eta_u)_{u\in\omega}$ *of all pairs* $\eta_u = (n, \bar{a})$ *of the following forms:*

$$(n, \bar{a}) = (0, \langle\rangle), \quad \text{or} \quad n \geq 1 \text{ and } \bar{a} = (a_1, \ldots, a_{n-1}) \in \omega^{n-1}.$$

If $\rho_2(v) = (x_v, y_v)$, iterating over v visits every input pair. The pair $\eta_u = (n, \bar{a})$ will code a candidate finite chain witnessing that the element carrying a given external label is the n-fold successor of 0 in the student's model. The separate clause for $n = 0$ avoids treating ω^{-1} as a tuple space.

We work in the usual first-order language of arithmetic

$$L_{\text{arit}} = \{0, 1, +, \times, <\},$$

where $+$ and $\times$ are understood (for technical convenience) as ternary relation symbols encoding the graphs of addition and multiplication.

Let $\mathcal{N}$ denote the standard model of arithmetic, i.e., the L_{arit}-structure with domain ω and the usual interpretations of the symbols. We write

$$\text{Th}(\mathcal{N}) \;=\; \{\, \varphi : \mathcal{N} \models \varphi \,\}$$

for its complete first-order theory. An *ω-presentation* of a model of $\text{Th}(\mathcal{N})$ is any L_{arit}-structure $\mathcal{S}$ with domain ω such that $\mathcal{S} \models \text{Th}(\mathcal{N})$. Thus, an ω-presentation is simply a coding of a model of true arithmetic with underlying set ω.

Definition 2.2. *We denote by* $\mathfrak{N}$ *the class of all* ω*-presentations* $\mathcal{M} \models \text{Th}(\mathcal{N})$*. A model* $\mathcal{M} \in \mathfrak{N}$ *is called* standard *if* $\mathcal{M} \cong \mathcal{N}$*, and* nonstandard *otherwise.*

Remark 2.3. The fixed domain ω is only a presentation convention. An external number $j \in \omega$ serves as a label for an element of the presented structure; it need not coincide with the interpretation $\bar{j}^{\mathcal{S}}$ of the corresponding arithmetical numeral, nor

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 20:34:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1755020326101439

need the labelled element be definable in the arithmetical language. Thus the coding through which the teacher accesses a structure is kept distinct from the arithmetic internal to that structure.

Notice that all elements of $\mathfrak{N}$ are, by definition, elementarily equivalent.

We use the complete theory $\mathrm{Th}(\mathcal{N})$, rather than merely PA, deliberately. This grants the student every parameter-free first-order truth of arithmetic, so the negative result below cannot be attributed to a disagreement about which arithmetical sentences are true. The remaining issue is the global organisation of a presentation and its isomorphism type. Since $\mathrm{Th}(\mathcal{N})$ extends PA, Tennenbaum's theorem applies to all nonstandard members of $\mathfrak{N}$.

Whenever two models are isomorphic, we shall simply say they are *copies* of each other.

Definition 2.4. *An ω-presentation $\mathcal{S}$ is* computable *if the relations $+_{\mathcal{S}}, \times_{\mathcal{S}}, <_{\mathcal{S}}$ are computable.*

Remark 2.5. For a standard presentation $\mathcal{S} \cong \mathcal{N}$, computability of the graph of $+_{\mathcal{S}}$ is equivalent to computability of the whole presentation. To see the nontrivial direction, start with the interpretations $0_{\mathcal{S}}$ and $1_{\mathcal{S}}$ of the language constants and compute successively

$$s(0) = 0_{\mathcal{S}}, \qquad s(n+1) = \text{the unique } z \text{ such that } +_{\mathcal{S}}(s(n), 1_{\mathcal{S}}, z).$$

Because $\mathcal{S}$ is standard, $s\colon \omega \to \omega$ is a computable bijection. Given a presentation label a, a search for the unique n with $s(n) = a$ therefore terminates. Transporting the usual algorithms for order and multiplication through this computable bijection makes $<_{\mathcal{S}}$ and $\times_{\mathcal{S}}$ computable. Accordingly, in the definition of $\mathbb{G}_2$ it is enough to require computable addition.

Let $(\varphi_i^{\mathrm{at}})_{i\in\omega}$ be a fixed effective enumeration of all atomic L_{arit}-formulas with free variables among $\{x_0, x_1, x_2, \dots\}$. Given an ω-presentation $\mathcal{S}$, its *atomic diagram* $D(\mathcal{S}) \in 2^{\omega}$ is defined by:

$$D(\mathcal{S})(i) = \begin{cases} 1 & \text{if } \mathcal{S} \models \varphi_i^{\mathrm{at}}[x_j \mapsto j : j \in \omega], \\ 0 & \text{otherwise.} \end{cases}$$

Thus, $D(\mathcal{S})$ encodes the truth-values of all atomic formulas under the canonical assignment $x_j \mapsto j$ to elements of the domain.

Remark 2.6. Equivalently, one may expand the language L_{arit} by adding constant symbols $(c_j)_{j\in\omega}$, interpret each c_j as the element j in any ω-presentation, and view $D(\mathcal{S})$ as the (complete) atomic diagram of that expanded structure in the usual model-theoretic sense. We will occasionally use this equivalent "constants" viewpoint when applying compactness arguments.

For a finite set $F \subseteq \omega$ and a function $f \in 2^F$, we say that $\mathcal{S}$ *extends* f, and write $\mathcal{S} \succeq f$, if

$$\big(\forall i \in F\big)\big(D(\mathcal{S})(i) = f(i)\big).$$

The following lemma expresses the key fact that finite atomic information never suffices to force standardness or nonstandardness, even when we restrict to models of $\mathrm{Th}(\mathcal{N})$.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 20:34:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1755020326101439

Lemma 2.7. *Let $\mathcal{M} \in \mathfrak{N}$ be nonstandard, $F \subseteq \omega$ be finite, and $f \in 2^F$. Then*

$$(\exists\, \mathcal{S}_0 \cong \mathcal{N})(\mathcal{S}_0 \succeq f) \qquad \Longleftrightarrow \qquad (\exists\, \mathcal{S}_1 \cong \mathcal{M})(\mathcal{S}_1 \succeq f).$$

Proof. We prove the left-to-right direction; the converse is symmetric.

Assume that $\mathcal{S}_0 \cong \mathcal{N}$ and $\mathcal{S}_0 \succeq f$, where $F = \{i_0, \dots, i_{m-1}\}$. For each $k < m$, the atomic formula $\varphi^{\mathrm{at}}_{i_k}$ mentions only finitely many variables; let r be large enough so that every variable occurring in some $\varphi^{\mathrm{at}}_{i_k}$ lies among $\{x_0, \dots, x_r\}$. Writing $\bar{x} = (x_0, \dots, x_r)$, define the quantifier-free formula

$$\theta(\bar{x}) \;:=\; \Big(\bigwedge_{k<m} \big(\varphi^{\mathrm{at}}_{i_k}\big)^{f(i_k)} \Big) \;\wedge\; \Big(\bigwedge_{0 \le i < j \le r} x_i \neq x_j \Big),$$

where $\psi^1 = \psi$ and $\psi^0 = \neg\psi$.

Since $\mathcal{S}_0 \succeq f$, the canonical assignment $x_j \mapsto j$ witnesses that $\mathcal{S}_0 \models \exists \bar{x}\, \theta(\bar{x})$. As $\mathcal{S}_0 \cong \mathcal{N}$, we have $\mathcal{N} \models \exists \bar{x}\, \theta(\bar{x})$. Because $\mathcal{N} \equiv \mathcal{M}$, also $\mathcal{M} \models \exists \bar{x}\, \theta(\bar{x})$. Let $\bar{b} = (b_0, \dots, b_r)$ be witnesses in $\mathcal{M}$; by the distinctness conjuncts, the b_j are pairwise distinct.

Now choose a bijection $h \colon \mathrm{dom}(\mathcal{M}) \to \omega$ such that $h(b_j) = j$ for all $j \le r$, and let $\mathcal{S}_1$ be the L_{arit}-structure on ω obtained by transporting the interpretation of symbols along h. Then h is an isomorphism $\mathcal{M} \cong \mathcal{S}_1$. Moreover, since $\mathcal{M} \models \theta(\bar{b})$, we have $\mathcal{S}_1 \models \big(\varphi^{\mathrm{at}}_{i_k}\big)^{f(i_k)}[x_j \mapsto j]$ for each $k < m$. Equivalently, $\mathcal{S}_1 \succeq f$, as required. □

3. Teachers, students, and the game $\mathbb{G}_1$

The game $\mathbb{G}_1$. We now formalise the scenario outlined in the introduction. A student **S** secretly chooses $\mathcal{M} \in \mathfrak{N}$. A teacher **T** has no direct access to $\mathcal{M}$, but may ask atomic questions about it and, after each round, must guess whether $\mathcal{M}$ is standard or not.

Definition 3.1. *The game $\mathbb{G}_1$ proceeds as follows.*

- *Before round 0, the student selects an ω-presentation $\mathcal{M} \models \mathrm{Th}(\mathcal{N})$.*
- *At every round $n \in \omega$:*
  1. *on the basis of the first n answers, the teacher chooses an atomic formula $\varphi^{\mathrm{at}}_{i_n}$ and asks whether*
  $$\mathcal{M} \models \varphi^{\mathrm{at}}_{i_n}[x_j \mapsto j];$$
  2. *the student answers 'yes' exactly when the displayed statement holds;*
  3. *on the basis of the first $n+1$ answers, the teacher outputs a conjecture $c_n \in \{\mathbf{st}, \mathbf{nst}\}$.*

The teacher **T** wins *the play with student model $\mathcal{M}$ if the sequence $(c_n)_{n\in\omega}$ converges and*

$$\lim_{n\to\infty} c_n = \mathbf{st} \quad \textit{iff} \quad \mathcal{M} \cong \mathcal{N}.$$

A strategy *for* **T** *is a rule for choosing the next query from the answers received so far and the next conjecture after the new answer; it is* winning *if it wins against every admissible choice of $\mathcal{M}$.*

Thus $\mathbb{G}_1$ is an infinite game of imperfect information: the student's initial choice of $\mathcal{M}$ is never revealed to the teacher, who only sees a growing finite fragment of its atomic diagram.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 20:34:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1755020326101439

Remark 3.2. Saying that the student "selects" a model is metatheoretic shorthand. We do not assume that the student can survey the structure, specify it by an index, know whether it is standard, or compute its atomic diagram. A play is evaluated relative to a fixed ω-presentation $\mathcal{M}$, so that every query has a determinate answer and all answers come from a single structure. The set-up thus takes the determinacy of the student's practice as given, without explaining it, and asks what another agent can recognise about its isomorphism type from successive finite answers.

It is convenient to represent teacher strategies as labelled trees on $2^{<\omega}$. We interpret finite binary strings as finite answer histories: if $\sigma \in 2^{<\omega}$ has length n, then $\sigma(k) = 1$ (resp. $\sigma(k) = 0$) records that at round k the student answered *yes* (resp. *no*).

Definition 3.3. *A* teacher strategy *is any function*

$$\mathfrak{S}\colon 2^{<\omega} \to \big(\{\mathbf{st}, \mathbf{nst}\} \times \{\varphi_i^{\mathrm{at}} : i \in \omega\}\big),$$

written

$$\mathfrak{S}(\sigma) = \big(\mathbf{conj}_{\mathfrak{S}}(\sigma),\, \mathbf{ques}_{\mathfrak{S}}(\sigma)\big)$$

for $\sigma \in 2^{<\omega}$. Here $\mathbf{conj}_{\mathfrak{S}}(\sigma)$ is the conjecture that the teacher would output after seeing the answer history σ, and $\mathbf{ques}_{\mathfrak{S}}(\sigma)$ is the atomic formula that she will ask at the next *stage. The value $\mathbf{conj}_{\mathfrak{S}}(\lambda)$ is a harmless dummy label for the root; no conjecture is output before the first answer.*

We call a teacher strategy $\mathfrak{S}$ *computable* if there is an algorithm which, given a finite answer history $\sigma \in 2^{<\omega}$, computes both the conjecture $\mathbf{conj}_{\mathfrak{S}}(\sigma)$ and an index for the next query $\mathbf{ques}_{\mathfrak{S}}(\sigma)$ in the fixed enumeration of atomic formulas.

Given $\mathcal{S} \in \mathfrak{N}$ and a teacher strategy $\mathfrak{S}$, we define the associated *answer path* $\pi_{\mathcal{S}} \in 2^{\omega}$ recursively by $\pi_{\mathcal{S}} \restriction 0 = \lambda$, and, for each $n \in \omega$,

$$\pi_{\mathcal{S}}(n) = \begin{cases} 1 & \text{if } \mathcal{S} \models \mathbf{ques}_{\mathfrak{S}}(\pi_{\mathcal{S}} \restriction n)[x_j \mapsto j], \\ 0 & \text{otherwise.} \end{cases}$$

The corresponding *conjecture sequence* along $\mathcal{S}$ is

$$\mathbf{conj}_{\mathfrak{S}}(\pi_{\mathcal{S}}) \;=\; \big(\mathbf{conj}_{\mathfrak{S}}(\pi_{\mathcal{S}} \restriction (n+1))\big)_{n\in\omega},$$

since at round n the teacher outputs her conjecture after receiving the $(n+1)$-st answer.

Definition 3.4. *Let $\mathfrak{S}$ be a teacher strategy and let $\mathcal{M} \in \mathfrak{N}$. We say that $\mathfrak{S}$ is* winning for $\mathcal{M}$ *if the sequence $\mathbf{conj}_{\mathfrak{S}}(\pi_{\mathcal{M}})$ converges and its limit is $\mathbf{st}$ iff $\mathcal{M} \cong \mathcal{N}$. We say that $\mathfrak{S}$ is a* winning strategy *in $\mathbb{G}_1$ if it is winning for every $\mathcal{M} \in \mathfrak{N}$.*

Definition 3.5. *Let $\mathfrak{S}$ be a teacher strategy. A string $\sigma \in 2^{<\omega}$ is an* admissible node *(for $\mathfrak{S}$) if it is compatible with at least one model in $\mathfrak{N}$, i.e. if there exists some $\mathcal{S} \in \mathfrak{N}$ such that $\sigma \preceq \pi_{\mathcal{S}}$. In this case we say that σ* supports *$\mathcal{S}$.*

Intuitively, σ is admissible if it could arise as the initial segment of answers given by *some* student model when the teacher follows $\mathfrak{S}$. Non-admissible strings encode patterns of answers that no model of $\mathrm{Th}(\mathcal{N})$ can produce.

The next lemma ensures that all admissible nodes support both standard and nonstandard models. To streamline later proofs, we fix a specific nonstandard ω-presentation $\mathcal{N}^* \in \mathfrak{N}$.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 20:34:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1755020326101439

Lemma 3.6. *Let $\mathfrak{S}$ be a teacher strategy and let $\sigma \in 2^{<\omega}$ be admissible. Then there exist ω-presentations $\mathcal{M}_0 \cong \mathcal{N}$ and $\mathcal{M}_1 \cong \mathcal{N}^*$ such that*

$$\sigma \preceq \pi_{\mathcal{M}_0} \quad \text{and} \quad \sigma \preceq \pi_{\mathcal{M}_1}.$$

Proof. Because σ is admissible, there is some $\mathcal{S} \in \mathfrak{N}$ with $\sigma \preceq \pi_{\mathcal{S}}$. Let $F \subseteq \omega$ be the finite set of indices of queries actually asked along σ, and let $f \in 2^F$ be the partial atomic diagram determined by the answers.

By Lemma 2.7, there exists a copy $\mathcal{M}_0 \cong \mathcal{N}$ with $\mathcal{M}_0 \succeq f$ and a copy $\mathcal{M}_1 \cong \mathcal{N}^*$ with $\mathcal{M}_1 \succeq f$. Then playing the strategy $\mathfrak{S}$ against $\mathcal{M}_0$ or $\mathcal{M}_1$ induces the answer prefix σ. Thus $\sigma \preceq \pi_{\mathcal{M}_0}$ and $\sigma \preceq \pi_{\mathcal{M}_1}$. □

We now establish the key fact underlying the failure of any putative winning strategy for $\mathbb{G}_1$: along admissible nodes, one can always find extensions that force the teacher's conjecture to flip.

Lemma 3.7. *Let $\mathfrak{S}$ be a winning strategy in $\mathbb{G}_1$. Then every admissible node $\sigma \in 2^{<\omega}$ has an admissible extension $\tau \succeq \sigma$ such that*

$$\mathbf{conj}_{\mathfrak{S}}(\tau) \neq \mathbf{conj}_{\mathfrak{S}}(\sigma).$$

Proof. Suppose not. Then some admissible σ satisfies:

$$\mathbf{conj}_{\mathfrak{S}}(\tau) = \mathbf{conj}_{\mathfrak{S}}(\sigma) \quad \text{for all admissible } \tau \succeq \sigma.$$

By Lemma 3.6, there exist $\mathcal{M}_0 \cong \mathcal{N}$ and $\mathcal{M}_1 \cong \mathcal{N}^*$ such that $\sigma \preceq \pi_{\mathcal{M}_0}$ and $\sigma \preceq \pi_{\mathcal{M}_1}$. For each large enough n, $\pi_{\mathcal{M}_j} \upharpoonright n$ is an admissible extension of σ. Hence

$$\mathbf{conj}_{\mathfrak{S}}(\pi_{\mathcal{M}_j} \upharpoonright (n+1)) = \mathbf{conj}_{\mathfrak{S}}(\sigma) \quad \text{for } j = 0, 1.$$

Thus the limit conjecture along both $\mathcal{M}_0$ and $\mathcal{M}_1$ is the same. But since $\mathcal{M}_0 \cong \mathcal{N}$ and $\mathcal{M}_1 \cong \mathcal{N}^*$, a winning strategy must give different limiting conjectures. Contradiction. □

Theorem 3.8. *In the game $\mathbb{G}_1$, the teacher has no winning strategy.*

Proof. Assume, for contradiction, that $\mathfrak{S}$ is a winning strategy. Use the fixed nonstandard ω-presentation $\mathcal{N}^* \in \mathfrak{N}$. We build an infinite answer path along which the teacher changes her conjecture infinitely often, while simultaneously arranging that the path is realised by a genuine ω-presentation.

Work in the expanded language

$$L^* = L_{\text{arit}} \cup \{c_j : j \in \omega\}.$$

For $\sigma \in 2^{<\omega}$, let A_σ be the finite set of signed L^*-sentences recording the answers in σ. More explicitly, for $k < |\sigma|$ put $q_k = \mathbf{ques}_{\mathfrak{S}}(\sigma \upharpoonright k)$, let $\widehat{q}_k$ be obtained from q_k by replacing every x_j by c_j, and include $\widehat{q}_k$ or $\neg\widehat{q}_k$ according as $\sigma(k) = 1$ or 0.

Call a finite set Γ of L^*-sentences *canonically realisable* if some ω-presentation $\mathcal{P} \in \mathfrak{N}$, in its canonical expansion interpreting c_j by the presentation label j, satisfies Γ. We use the following finite-transfer observation. If Γ is canonically realisable, then it is canonically realisable both in a standard presentation and in a copy of $\mathcal{N}^*$. Indeed, replace the finitely many constants occurring in Γ by variables, conjoin the pairwise inequalities saying that distinct labels receive distinct values, and existentially close the conjunction. The resulting L_{arit}-sentence is true in one model of $\mathrm{Th}(\mathcal{N})$, hence in every such model. Choosing distinct witnesses in $\mathcal{N}$ or $\mathcal{N}^*$ and transporting the structure to domain ω sends those witnesses back to the prescribed labels.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 20:34:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1755020326101439

We construct increasing answer histories σ_s and increasing finite side conditions Δ_s so that $A_{\sigma_s} \cup \Delta_s$ is canonically realisable. Start with $\sigma_0 = \lambda$ and $\Delta_0 = \varnothing$. Suppose the pair has been constructed. If $\mathbf{conj}_{\mathfrak{S}}(\sigma_s) = \mathbf{st}$, use finite transfer to choose a canonical realisation of $A_{\sigma_s} \cup \Delta_s$ in a copy of $\mathcal{N}^*$; if the current conjecture is **nst**, choose instead a standard realisation. Follow $\mathfrak{S}$ against that presentation. Since $\mathfrak{S}$ is winning there, at some finite later history $\tau_s \succeq \sigma_s$ it must output the opposite conjecture. The chosen presentation canonically realises $A_{\tau_s} \cup \Delta_s$, and

$$\mathbf{conj}_{\mathfrak{S}}(\tau_s) \neq \mathbf{conj}_{\mathfrak{S}}(\sigma_s).$$

In particular, τ_s is a proper extension of σ_s.

It remains to add enough side conditions to ensure that the limiting constants exhaust the eventual domain. Enumerate all existential L^*-sentences as $(\exists x\, \theta_s(x))_{s\in\omega}$. Choose a canonical realisation $\mathcal{P}_s^*$ of $A_{\tau_s} \cup \Delta_s$. If $\mathcal{P}_s^* \models \exists x\, \theta_s(x)$, choose a witness $a \in \omega$ and put $j = a$; otherwise choose any $j \in \omega$. In either case the canonical expansion satisfies the Henkin implication

$$H_s : \qquad (\exists x\, \theta_s(x)) \longrightarrow \theta_s(c_j).$$

Let C_s be a finite set of indices containing $0, \dots, s$ and every index of a constant occurring in A_{τ_s}, Δ_s, or H_s, and put

$$D_s = \{c_i \neq c_j : i, j \in C_s \text{ and } i < j\}.$$

Because $\mathcal{P}_s^*$ is canonical, it satisfies $\{H_s\} \cup D_s$. Set

$$\sigma_{s+1} = \tau_s, \qquad \Delta_{s+1} = \Delta_s \cup \{H_s\} \cup D_s.$$

This preserves the invariant.

Let $\pi = \bigcup_s \sigma_s$ and

$$\Theta = \mathrm{Th}(\mathcal{N}) \cup \bigcup_s A_{\sigma_s} \cup \bigcup_s \Delta_s.$$

Every finite subset of Θ is canonically realisable and hence consistent. The theory contains $c_i \neq c_j$ for all $i < j$, and it contains a witness implication for every existential L^*-sentence. Extend it to a complete consistent theory Θ^+. The witness implications make Θ^+ a Henkin theory, so its term model is a model of $\mathrm{Th}(\mathcal{N})$.

There are no nonconstant function symbols in L^*: addition and multiplication are relational. Thus every closed term is one of $0, 1, c_0, c_1, \dots$. The witness implications for $\exists x(x = 0)$ and $\exists x(x = 1)$ ensure that the interpretations of 0 and 1 coincide with interpretations of some c_j, while the sentences $c_i \neq c_j$ make the c_j's pairwise distinct. Consequently $j \mapsto c_j^{\Theta^+}$ is a bijection from ω onto the domain of the term model. Transporting along this bijection yields an ω-presentation $\mathcal{S} \in \mathfrak{N}$ whose canonical expansion satisfies every sentence in Θ. Its answer path is therefore exactly π.

For every s, the conjectures at σ_s and σ_{s+1} differ. Since the histories have strictly increasing lengths, the conjecture sequence along π changes infinitely often and does not converge. This contradicts the assumption that $\mathfrak{S}$ wins against every member of $\mathfrak{N}$. □

Remark 3.9. Theorem 3.8 is not an artefact of restricting the teacher to atomic questions. The proof uses atomicity only to describe a finite transcript as a finite fragment of an atomic diagram. Suppose instead that the teacher may query arbitrary first-order L_{arit}-formulas with finitely many named presentation labels as parameters. A finite transcript is then a finite set Γ of signed L^*-sentences. Let $c_{j_0}, \dots, c_{j_{r-1}}$ be the presentation constants occurring in Γ. Replace them by

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 20:34:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1755020326101439

variables $x_0, \dots, x_{r-1}$, conjoin the inequalities $x_\ell \neq x_t$ for $\ell < t < r$, and existentially close the result. If Γ is canonically realised in one model of $\mathrm{Th}(\mathcal{N})$, the resulting parameter-free sentence is true in that model and hence, by completeness of $\mathrm{Th}(\mathcal{N})$, in every model of true arithmetic. Choosing distinct witnesses and transporting them to the prescribed labels therefore canonically realises Γ both in a standard copy and in a copy of any fixed nonstandard model. The flip and Henkin steps above then go through exactly as before.

For example, no single first-order query can say that a named label denotes a nonstandard element: this would require the infinite family of assertions that it lies above the n-th finite successor of 0, one for each $n \in \omega$. Every finite subfamily is realised in $\mathcal{N}$ by a sufficiently large standard element. Atomic access remains philosophically useful as a minimal model of publicly checkable calculation, but it is not responsible for the negative result.

> **Interim moral ($\mathbb{G}_1$).** Truthful answers from a model of true arithmetic do not guarantee that another agent will recognise, even in the limit, whether the model is standard. Theorem 3.8 thus separates the determinacy of a practice from its intersubjective recognisability; it does not show that arithmetical reference is indeterminate.

4. The second game: adding a computability constraint

We now refine the admissible standard presentations. The motivating thought is that ordinary arithmetical practice is closely tied to step-by-step methods for calculating sums, and that Turing computability gives a precise idealisation of such methods. We impose the condition on addition alone; by Remark 2.5, for a standard copy this is equivalent to requiring the whole presentation to be computable.[12]

The game $\mathbb{G}_2$.

Definition 4.1. *The game $\mathbb{G}_2$ has the same rounds, queries, conjectures, and winning condition as $\mathbb{G}_1$. Before round 0, however, the student must select $\mathcal{M} \in \mathfrak{N}$ satisfying*

$$\mathcal{M} \cong \mathcal{N} \quad \Longrightarrow \quad +_{\mathcal{M}} \text{ is computable.}$$

Thus every nonstandard model of true arithmetic remains admissible, while only standard copies with computable addition are admitted on the standard side. Tennenbaum's theorem guarantees that no nonstandard member of $\mathfrak{N}$ has computable addition.

The strategy we use alternates between two local tests. In **st**-mode it tests a candidate total computable function against the student's addition. If a mismatch is found, it enters **nst**-mode and tests whether the current external label denotes a member of the standard cut. A successful finite successor-chain moves the strategy to the next label; failure to find such a chain leaves it in **nst**-mode.

Theorem 4.2. *In the game $\mathbb{G}_2$, the teacher has a winning strategy.*

Proof. Fix, externally, an enumeration $(f_e)_{e\in\omega}$ of all total computable binary functions $\omega^2 \to \omega$. No effective procedure taking e, x, y to $f_e(x, y)$ is assumed. Write $\rho_2(v) = (x_v, y_v)$, and use the witness codes (η_u) from Fact 2.1.

[12] For the surrounding philosophical debate, see Halbach and Horsten (2005); Dean (2014); Grabmayr (2026).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 20:34:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1755020326101439

For an external label e and a code $\eta_u = (n, (a_1, \ldots, a_{n-1}))$, let $W(e, u)$ be the following finite list of atomic formulas:

- if $n = 0$, the one-term list $x_e = 0$;
- if $n = 1$, the one-term list $+(0, 1, x_e)$;
- if $n \geq 2$, the list

$$+(0, 1, x_{a_1}), \qquad +(x_{a_i}, 1, x_{a_{i+1}}) \ (1 \leq i < n-1), \qquad +(x_{a_{n-1}}, 1, x_e).$$

(The middle block is empty when $n = 2$.)

The entire list is true exactly when the displayed labels form a chain witnessing that the element carrying label e is the n-fold successor of 0.

The teacher maintains a state

$$X_s = (\chi, e, u, v), \qquad \chi \in \{\mathbf{st}, \mathbf{nst}\}, \quad e, u, v \in \omega,$$

with $X_0 = (\mathbf{st}, 0, 0, 0)$. The state X_s determines the query at round s. After receiving the answer, the teacher updates to X_{s+1} by the following rules and outputs the mode of X_{s+1}.

st-*mode.* If $X_s = (\mathbf{st}, e, u, v)$, let $z = f_e(x_v, y_v)$ and ask

$$+(x_{x_v}, x_{y_v}, x_z).$$

A 'yes' answer gives $X_{s+1} = (\mathbf{st}, e, 0, v + 1)$; a 'no' answer proves that f_e is not the student's addition and gives $X_{s+1} = (\mathbf{nst}, e, 0, 0)$.

nst-*mode.* If $X_s = (\mathbf{nst}, e, u, v)$, ask the v-th formula in $W(e, u)$. On a 'no' answer, refute that candidate chain and set $X_{s+1} = (\mathbf{nst}, e, u + 1, 0)$. On a 'yes' answer before the end of the list, set $X_{s+1} = (\mathbf{nst}, e, u, v + 1)$. On a 'yes' answer to the last formula, a complete chain has been verified; set $X_{s+1} = (\mathbf{st}, e + 1, 0, 0)$.

Exactly one atomic formula is queried at each round, so this is a legitimate strategy. It need not be computable, because the fixed enumeration (f_e) is not assumed to have a uniform effective evaluator.

To verify convergence, first suppose that $\mathcal{M} \cong \mathcal{N}$ is admissible, and let e_+ be the least index with $f_{e_+} = +_{\mathcal{M}}$. For each $e < e_+$, some input pair witnesses a mismatch, so the **st**-test eventually moves to **nst**-mode. Every external label denotes an element of the standard cut in a standard model, so some candidate in the exhaustive list (η_u) eventually supplies a successful finite successor-chain, and the strategy advances to $e + 1$. Once it reaches e_+, every tested equation is true and the conjecture remains **st**.

Now suppose that $\mathcal{M}$ is nonstandard. By Tennenbaum's theorem no f_e is its addition, so every **st**-test eventually fails. Let $\hat{e}$ be the least external label whose element lies outside the standard cut of $\mathcal{M}$. For $e < \hat{e}$, a finite successor-chain is eventually verified and the strategy advances. At $e = \hat{e}$, no such chain exists; every candidate is eventually refuted, so the strategy remains in **nst**-mode forever. In both cases its conjectures converge to the correct verdict. □

Question 4.3. *Does there exist a computable winning strategy for the teacher in* $\mathbb{G}_2$*?*

At present we do not know the answer. The compactness construction for $\mathbb{G}_1$ does not immediately adapt, because any standard limit presentation would now have to retain computable addition. Conversely, the strategy above cannot simply be effectivised: the class of total computable functions has no uniform computable

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 20:34:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1755020326101439

enumeration consisting only of total functions. It remains open whether that non-effectivity is merely a feature of this strategy or an unavoidable feature of every winning strategy.

> **Interim moral ($\mathbb{G}_2$).** If standard presentations have computable addition, some teacher recognises standardness in the limit. Whether a computable teacher can do so is open.

5. The third game: primitive-recursive addition

The obstacle to effectivising the strategy just given for $\mathbb{G}_2$ is not the evaluation of any individual candidate procedure, but the lack of a uniform effective enumeration of all total computable binary functions. We therefore consider a standard-side constraint whose admissible addition procedures can be uniformly enumerated and evaluated.

Primitive recursive functions provide a canonical example. They are generated from the initial functions by composition and primitive recursion; every such function is total, and syntactic codes for primitive-recursive definitions can be evaluated by a uniform computable procedure. The point is effective indexability, not practical feasibility or a claim that human calculation is bounded by primitive recursion. See Odifreddi (1989) for background.

The game $\mathbb{G}_3$.

Definition 5.1. *The game $\mathbb{G}_3$ has the same interaction and winning condition as $\mathbb{G}_1$. Before round 0, the student selects $\mathcal{M} \in \mathfrak{N}$ satisfying*

$$\mathcal{M} \cong \mathcal{N} \quad \Longrightarrow \quad \text{the binary function represented by } +_{\mathcal{M}} \text{ is primitive recursive.}$$

Equivalently, in the standard case there is a primitive recursive $f\colon \omega^2 \to \omega$ such that $+_{\mathcal{M}}(x, y, z)$ iff $z = f(x, y)$.

This assumption is strictly about the presentation of addition. Requiring primitive-recursive multiplication as well would be stronger than the proof needs.

Fact 5.2. *There is a total computable function $U\colon \omega^3 \to \omega$ such that each binary section $U_e(x, y) = U(e, x, y)$ is primitive recursive and every primitive recursive binary function occurs as some U_e.*

The evaluator U is total computable, but not primitive recursive; total computability is all that the teacher requires.

Theorem 5.3. *In the game $\mathbb{G}_3$, the teacher has a computable winning strategy.*

Proof. Use the same state machine as in the proof of Theorem 4.2, with the same lists $W(e, u)$, but replace the external enumeration (f_e) by the effective family (U_e). Thus, in **st**-mode and with $\rho_2(v) = (x_v, y_v)$, the teacher computes $z = U(e, x_v, y_v)$ and asks $+(x_{x_v}, x_{y_v}, x_z)$. The **nst**-mode is unchanged. As before, X_s determines the query, the answer determines X_{s+1}, and the mode of X_{s+1} is the conjecture for that round.

This strategy is computable: from a finite answer history one computes the current state by iterating the transition rules; ρ_2, the witness enumeration, and U are all computable, so both the next atomic query and the current conjecture are computable.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 20:34:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1755020326101439

If $\mathcal{M} \cong \mathcal{N}$ is admissible, choose the least e_+ with $U_{e_+} = +_{\mathcal{M}}$. Every smaller incorrect candidate eventually fails on some input pair and is followed by a successful finite-chain test for the external label e. At e_+, all addition tests succeed, so the conjecture is eventually constantly **st**. If $\mathcal{M}$ is nonstandard, Tennenbaum's theorem implies that no U_e computes its addition. Every **st**-test therefore fails, and at the least external label outside the standard cut the finite-chain search never succeeds. The conjecture is then eventually constantly **nst**. □

Remark 5.4. Primitive recursiveness provides a clean sufficient condition; we do not claim that it is uniquely privileged or optimal. The proof works for any class $\mathcal{C}$ of total binary functions supplied with a total computable evaluator $V(e, x, y)$ whose sections enumerate $\mathcal{C}$, provided every admissible standard addition belongs to $\mathcal{C}$. Thus the operative property is uniform effective surveyability of the allowed addition procedures.

> **Interim moral ($\mathbb{G}_3$).** If the admissible standard additions are contained in a family of total functions with a uniform computable evaluator, a computable teacher recognises standardness in the limit. Primitive-recursive addition is a natural instance.

Philosophical conclusions

Our games assume that the student's answers are governed by a fixed model of true arithmetic and ask whether another agent can recognise, through an unbounded sequence of finite exchanges, whether that model is standard.

Theorem 3.8 gives a negative answer in the unconstrained case. By construction, the failure is not due to disagreement over first-order arithmetical truth: the student's model satisfies $\mathrm{Th}(\mathcal{N})$. Agreement on arithmetical truth, therefore, does not suffice for one agent to recognise, even in the limit, whether another's model is standard. What can be recognised depends on how the model is presented, and not only on which model it is; in this sense, an epistemology of structures requires an epistemology of presentations.

Theorems 4.2 and 5.3 show how constraints on presentations change the situation. If standard presentations must have computable addition, some teacher strategy wins, although we do not know whether a computable one does. If they must have primitive-recursive addition, a computable strategy wins, because every admissible standard addition belongs to a family of total candidate functions that can be uniformly enumerated (Remark 5.4). One may object that $\mathbb{G}_2$ and $\mathbb{G}_3$ simply exclude the noncomputable copies of $\mathcal{N}$. The exclusion is essential: if every standard copy were admitted alongside every nonstandard model of true arithmetic, the admissible class would be that of $\mathbb{G}_1$, and Theorem 3.8 would apply.

The resulting hierarchy (summarised in Table 1) is conditional. We have not shown that ordinary speakers use primitive-recursive presentations, that primitive recursion is the threshold for effective learnability, or that infinite interaction with truthful answers is psychologically realistic. The games show that modest changes in the admissible presentations affect whether stable intersubjective recognition is possible. Computability constraints can explain specific forms of recognitional access to arithmetic, provided that their role in the presentation of arithmetic is made explicit.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 20:34:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1755020326101439

Acknowledgments

Earlier versions of this paper were presented at the Konstanz–Pisa–Torino workshop (2023) and at the symposium "New Directions in Philosophy of Computability Theory" at CLMPST 2023. We thank the audiences at both events for their helpful questions, comments, and discussions. We are also grateful to the referees for their careful reading and constructive suggestions. We used AI tools (ChatGPT and Claude) to help with the wording of the manuscript; we take full responsibility for the final text.

San Mauro is a member of INDAM-GNSAGA.

Funding

San Mauro was partially funded by the National Science Centre of Poland (grant 2023/49/B/HS1/03930).

References

D. Angluin and C. H. Smith. Inductive inference: Theory and methods. *ACM Computing Surveys*, 15(3):237–269, 1983.

N. Bazhenov, E. Fokina, and L. San Mauro. Learning families of algebraic structures from informant. *Information and Computation*, 275:104590, 2020.

N. Bazhenov, V. Cipriani, and L. San Mauro. Learning algebraic structures with the help of Borel equivalence relations. *Theoretical Computer Science*, 951:113762, 2023.

N. Bazhenov, V. Cipriani, S. Jain, L. San Mauro, and F. Stephan. Classifying different criteria for learning algebraic structures. *Annals of Pure and Applied Logic*, 177(1):103648, 2026.

P. Benacerraf. What numbers could not be. *The Philosophical Review*, 74(1):47–73, 1965.

T. Button. *The Limits of Realism*. Oxford University Press, Oxford, 2013.

T. Button and P. Smith. The philosophical significance of Tennenbaum's theorem. *Philosophia Mathematica*, 20(1):114–121, 2012.

T. Button and S. Walsh. Structure and categoricity: Determinacy of reference and truth value in the philosophy of mathematics. *Philosophia Mathematica*, 24(3):283–307, 2016.

T. Button and S. Walsh. *Philosophy and Model Theory*. Oxford University Press, Oxford, 2018.

C. C. Chang and H. J. Keisler. *Model Theory*. Studies in Logic and the Foundations of Mathematics, vol. 73. North-Holland, Amsterdam, third edition, 1990.

W. Dean. Models and computability. *Philosophia Mathematica*, 22(2):143–166, 2014.

R. Dedekind. *Was sind und was sollen die Zahlen?* Vieweg, Braunschweig, 1888.

E. M. Gold. Language identification in the limit. *Information and Control*, 10(5):447–474, 1967.

B. Grabmayr. Structure and computation. *Noûs*, 60(3):593–605, 2026.

V. Halbach and L. Horsten. Computational structuralism. *Philosophia Mathematica*, 13(2):174–186, 2005.

G. Hellman. *Mathematics without Numbers: Towards a Modal-Structural Interpretation*. Clarendon Press, Oxford, 1989.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 20:34:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1755020326101439

W. Hodges. *Model Theory*. Cambridge University Press, Cambridge, 1993.
R. Kaye. *Models of Peano Arithmetic*. Oxford Logic Guides, vol. 15. Clarendon Press, Oxford, 1991.
M. J. Kearns and U. V. Vazirani. *An Introduction to Computational Learning Theory*. MIT Press, Cambridge, MA, 1994.
K. T. Kelly. *The Logic of Reliable Inquiry*. Oxford University Press, New York, 1996.
S. A. Kripke. *Wittgenstein on Rules and Private Language: An Elementary Exposition*. Harvard University Press, Cambridge, MA, 1982.
P. Maddy and J. Väänänen. *Philosophical Uses of Categoricity Arguments*. Cambridge Elements in the Philosophy of Mathematics. Cambridge University Press, Cambridge, 2023.
J. McDowell. Wittgenstein on following a rule. *Synthese*, 58(3):325–363, 1984.
V. McGee. How we learn mathematical language. *The Philosophical Review*, 106(1):35–68, 1997.
P. Odifreddi. *Classical Recursion Theory: The Theory of Functions and Sets of Natural Numbers*. Studies in Logic and the Foundations of Mathematics, vol. 125. North-Holland, Amsterdam, 1989.
C. Parsons. The uniqueness of the natural numbers. *Iyyun: The Jerusalem Philosophical Quarterly*, 39(1):13–44, 1990.
P. Quinon and K. Zdanowski. The intended model of arithmetic: An argument from Tennenbaum's theorem. In S. B. Cooper, T. F. Kent, B. Löwe, and A. Sorbi, editors, *Computation and Logic in the Real World: Third Conference on Computability in Europe, CiE 2007, Siena, Italy, June 18–23, 2007: Local Proceedings*, pages 313–317. Quaderni del Dipartimento di Scienze Matematiche e Informatiche "Roberto Magari", Technical Report no. 487, University of Siena, June 2007.
L. Picollo and D. Waxman. Arithmetical pluralism and the objectivity of syntax. *Noûs*, 59(2):372–391, 2025.
H. Putnam. Models and reality. *Journal of Symbolic Logic*, 45(3):464–482, 1980.
M. D. Resnik. *Mathematics as a Science of Patterns*. Oxford University Press, Oxford, 1997.
S. Shapiro. *Foundations without Foundationalism: A Case for Second-Order Logic*. Clarendon Press, Oxford, 1991.
R. I. Soare. *Turing Computability: Theory and Applications*. Theory and Applications of Computability. Springer, Berlin, 2016.
S. Tennenbaum. Non-Archimedean models for arithmetic. *Notices of the American Mathematical Society*, 6:270, 1959.
J. Väänänen. Second order logic or set theory? *Bulletin of Symbolic Logic*, 18(1):91–121, 2012.
J. Väänänen and T. Wang. Internal categoricity in arithmetic and set theory. *Notre Dame Journal of Formal Logic*, 56(1):121–134, 2015.
L. Wittgenstein. *Philosophical Investigations*. Translated by G. E. M. Anscombe. Blackwell, Oxford, 1953.
C. Wright. *Wittgenstein on the Foundations of Mathematics*. Harvard University Press, Cambridge, MA, 1980.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 20:34:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1755020326101439

Fachbereich Philosophie, Universität Konstanz, Universitätsstrasse 10, 78464 Konstanz, Germany
Email address: leon.horsten@uni-konstanz.de
ORCID: 0000-0003-3610-9318

Dipartimento di Ricerca e Innovazione Umanistica, Università di Bari, Piazza Umberto I, 1, 70121 Bari, Italy
Email address: luca.sanmauro@gmail.com
ORCID: 0000-0002-3156-6870

Dipartimento di Civiltà e Forme del Sapere, Università di Pisa, Via Pasquale Paoli, 15, 56126 Pisa, Italy
Email address: gio.venturi@gmail.com
ORCID: 0000-0001-9089-4340

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 16 Nov 2026 at 20:34:55, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1755020326101439