

Discrete Linear Ensemble Logic: Decidability, Expressiveness, and Axiomatization

Manfred Droste 

Institute of Computer Science, Leipzig University, Germany

Guo-Qiang Zhang¹ 

The University of Texas Health Science Center at Houston, TX, USA

Abstract

We study the discrete point-based fragment of Ensemble Logic $\text{EL}(\mathbb{N})$ over the natural numbers, a logic combining displacement φ_u , term-dependent metric modalities $\Box_t$ and $\Diamond_t$ with respect to an additive term t , Boolean connectives, and first-order quantification over $\mathbb{N}$. Motivated by the need for a unified symbolic representation of biomedical knowledge with temporal, spatial, and multimodal metric content, we develop the foundational discrete theory of the formalism. We give syntax and semantics, and prove a forward embedding of $\text{EL}(\mathbb{N})$ (over a finite proposition set $\mathcal{P}$) into first-order monadic Presburger arithmetic $\text{FO}(\mathbb{N}, <, +; \mathcal{P})$. This embedding yields the analytical upper bounds, while a reduction from nondeterministic two-counter Turing machines with recurring control states proves that satisfiability is Σ_1^1 -complete and validity is dually Π_1^1 -complete. Expressively, $\text{EL}(\mathbb{N})$ strictly extends the star-free ω -languages and is incomparable with the ω -regular languages: it defines the non- ω -regular counting language $\{a^m b^m c^m d^m \mid m \geq 1\} \cdot \Sigma^\omega$, whereas a delimited parity language remains outside the logic by a quantifier-bounding argument combined with classical circuit lower bounds for parity. On the proof-theoretic side, we present a sound Hilbert system $\mathcal{H}_{\text{EL}}$ and establish completeness relative to monadic Presburger validity as oracle, noting that completeness relative to plain Presburger arithmetic is impossible. We also prove that the existential fragment $\exists \text{EL}(\mathbb{N})$ has NP-complete satisfiability and coNP-complete unsatisfiability. Finally, we show that finite active-domain model-checking has PTIME data complexity and PSPACE-complete combined complexity. Together, these results give a precise decidability, expressiveness, proof-theoretic, and model-checking baseline for subsequent algorithmic applications of Ensemble Logic in biomedicine. This work is a part of the Symbolic Biomedicine program championed by the corresponding author.

2012 ACM Subject Classification Theory of computation $\rightarrow$ Modal and temporal logics; Theory of computation $\rightarrow$ Logic and verification; Theory of computation $\rightarrow$ Algebraic language theory; Applied computing $\rightarrow$ Health informatics

Keywords and phrases Temporal logic, monadic Presburger arithmetic, descriptive complexity, electronic health records

Digital Object Identifier 10.4230/OASICS.TIME.2026.4

Funding *Guo-Qiang Zhang*: This work has been supported in part by the U.S. National Science Foundation Award IIS2500624 and the National Institutes of Health grants R01AG084236 and U24AG098157. The views of the paper are those of the authors and do not reflect those of the funding agencies.

Acknowledgements Part of this research was done during a research stay of MD in April 2026 at the University of Texas Health Science Center at Houston. He would like to thank his colleagues for a fruitful and most enjoyable time.

¹ Corresponding author.



© Manfred Droste and Guo-Qiang Zhang;

licensed under Creative Commons License CC-BY 4.0

33rd International Symposium on Temporal Representation and Reasoning (TIME 2026).

Editors: AndreA Orlandini and Sophie Pinchinat; Article No. 4; pp. 4:1–4:18

OpenAccess Series in Informatics



OASICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

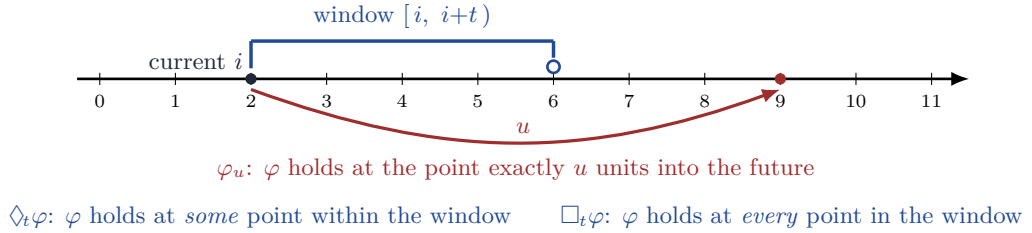
1 Introduction

Modern biomedicine has a formalism gap. Taxonomical ontologies such as SNOMED CT and UMLS [4] capture terminological structure but not metric time; pathway languages such as SBML [17] and BioPAX [8] capture biochemical network structure but not spatial containment; standard temporal logics [24] capture traces but not anatomical geometry; mathematical morphology captures shapes but not protocol-level quantifier scope [26]. A clinical constraint such as an HbA1c normalization condition that must hold continuously for 24 weeks, or a lymph-node tumour deposit large enough to contain a 1 mm spherical core, has a clear mathematical truth condition. Yet neither condition maps naturally to a single formal representation without external scripting in varied programming settings.

The central observation behind Ensemble Logic (EL) [28] is that biomedical reasoning repeatedly uses three non-Boolean, *term-dependent operations*: *exact displacement*, *bounded existence*, and *bounded universality*. In the discrete setting these are written φ_u , $\Diamond_t\varphi$, and $\Box_t\varphi$ (**Fig. 1**). Together with Boolean connectives and first-order quantifiers over an additive index domain, they yield the syntax

$$\varphi, \psi ::= p \mid \varphi_u \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \Diamond_t\varphi \mid \Box_t\varphi \mid \exists x\varphi \mid \forall x\varphi.$$

The interpretation domain of EL accommodates discrete time, dense continuous time, Euclidean space, discrete spatial coordinates, or mixed product spaces depending on the underlying monoid interpretation of the additive terms u and t [28], which in general come from different monoids. Over the discrete point-based domain $\mathbb{N}$, the focus of this paper, the interpretation of the three variable-dependent operations are illustrated below.



■ **Figure 1** An illustration of how the three operators are interpreted over the integer line.

Contributions

In this paper, we focus the mathematical investigation on the foundational discrete case $\text{EL}(\mathbb{N})$, originally introduced in [28], where $\mathbb{N}$ is interpreted temporally (modeling disease trajectories) or spatially (such as base coordinates in genomics). After brief motivational overview and comparison with existing temporal logic formalisms, we define the syntax and semantics of $\text{EL}(\mathbb{N})$ (§ 3). We show that satisfiability of $\text{EL}(\mathbb{N})$ formulas is undecidable, and in fact Σ_1^1 -hard, by a reduction from the recurring problem for nondeterministic two-counter machines. Building on the foundation laid out in [28], this paper establishes results in the directions of decidability, expressiveness, and axiomatization, with specific topics that lie at the heart of any new logical formalism:

1. We translate $\text{EL}(\mathbb{N})$ into monadic Presburger arithmetic, obtaining the analytical upper bounds Σ_1^1/Π_1^1 (§ 4).
2. We further prove Σ_1^1 -hardness by a two-counter-machine reduction tailored to EL's uniformly positive modal bounds (§ 5).

3. We place $\text{EL}(\mathbb{N})$ in the descriptive hierarchy: it contains all $\text{FO}(\mathbb{N}, <; \mathcal{P})$ languages, defines non- ω -regular counting languages, but does not contain all ω -regular languages (§ 6).
4. We give a sound Hilbert system and prove $\text{MONFO}_{\mathbb{N},+}$ -relative completeness (§ 7).
5. We prove NP-completeness of existential-fragment satisfiability, coNP-completeness of unsatisfiability, and give a complete proof system for satisfiability (§ 8).
6. We prove finite active-domain model checking is in PTIME for fixed formulas and PSPACE-complete (§ 9).

2 Background

2.1 Motivations from Biomedicine

Modern biomedicine depends on knowledge expressed across free-text discovery claims, longitudinal records, trial protocols, physiological signals, imaging, and omics. Data in different modalities are still analyzed with fragmented formalisms and ad hoc scripts. The same patient, lesion, molecular locus, or protocol event is represented differently across literature mining, EHR phenotyping, trial eligibility, image analysis, and genomics, even when the underlying scientific observation has a similar structure.

In prior work [20, 16], we demonstrated that the tri-modality of φ_u , $\Diamond_t\varphi$, and $\Box_t\varphi$ (together with standard Boolean and first-order quantifiers) suffices to articulate the entire temporal structure of clinical-trial protocols in a single closed formula, such as one for Alzheimer’s-disease trial NCT01266525 [27]. We illustrate the key modeling idea behind these works using a deliberately compact monitoring protocol example.

Whenever therapy starts on day x , a baseline measurement must occur during the four-day window beginning at x ; a dose (of medication) must be taken on every day of the fourteen-day window beginning one day later; and an outcome review must occur exactly twenty-eight days after the start.

With propositions **Start**, **Baseline**, **Dose**, and **Review**, the protocol is expressed as

$$\forall x \left(\text{Start}_x \rightarrow \left((\Diamond_4 \text{Baseline})_x \wedge (\Box_{14} \text{Dose})_{x+1} \wedge \text{Review}_{x+28} \right) \right). \quad (1)$$

The half-open-window convention ($[i, i+t)$ in **Fig. 1**) makes the three clauses precise:

Fragment	Plain reading	Truth condition
Review_{x+28}	exact displacement	Review takes place on day $x + 28$;
$(\Diamond_4 \text{Baseline})_x$	bounded existence	Baseline holds on day $x, x + 1, x + 2$, or day $x + 3$;
$(\Box_{14} \text{Dose})_{x+1}$	bounded universality	Dose is taken on every day in $[x + 1, x + 15)$.

The variable x refers to (a lapse of) x units from the time at which the closed formula is evaluated. Thus $\forall x$ checks all shifts into the future (i.e. displacement) from a given point for evaluation.

Quantification over metric parameters and logical inference become useful when the same unknown length must be reused. For example, the schematic rule

$$\forall x \left(\text{Start}_x \rightarrow \exists n \left[\Box_{n+1} \text{Stable} \wedge (\Box_{n+1} \text{Treatment})_{n+1} \wedge (\Diamond_{n+1} \text{Outcome})_{2(n+1)} \right]_x \right) \quad (2)$$

asks for a positive, record-dependent phase length $n + 1$ and reuses exactly that length for a stabilization phase, a treatment phase, and an outcome window. This reuse of a semantically chosen additive parameter is the most distinctive syntactic feature of **EL**.

A complete clinical-trial protocol is naturally a conjunction of many clauses of the form illustrated by Formula (1) above, together with domain-specific eligibility predicates and data-quality assumptions. Prior work shows how such clauses can be composed for trial protocols [20]; the present paper is concerned with the underlying discrete logic, not with validating any particular medical protocol.

The same approach extends to other medical domains. In sleep medicine [9], the formula φ_u places a finding φ at a specific offset from a reference time point; for example, snore occurs 6 seconds later can be formalized as Snore_6 . Similarly, $\Box_t \varphi$ states that the condition φ holds continuously for duration t ; for example, airflow reduction greater than 30% persisting for 10 seconds can be formalized as $\Box_{10} \text{FlowDrop30}$. The formula $\Diamond_t \varphi$ states that property φ holds somewhere within a bounded window of length t , so that a spindle occurring within a 3-second window can be formalized as $\Diamond_3 \text{Spindle}$.

The scientific content with respect to how long, how far, how soon, how continuously, is exactly the metric structure unaccounted-for by mainstream approaches. The *EL thesis* [28] is that three *variable-dependent operations*: *exact displacement*, *bounded existence*, and *bounded universality* are parts of a stable “lingua franca” for biomedical knowledge. The novelty is not that any one operator is new. Linear temporal logic and metric temporal logic already contain important special cases (see Section 2.2). The novelty is the cross-modal stability of the operator basis: the same tri-modality can be interpreted on sequences, claim windows, EHR days [20], dense waveforms [9, 15], Euclidean image regions, stranded genomic coordinates, and product spaces for spatial-temporal interactions.

2.2 Comparison with neighboring formalisms

The formalisms in Table 1 are selected according to three principled axes: the semantic carrier (points, intervals, or signals), the way a metric parameter is supplied (fixed constant, frozen timestamp, or quantified offset), and the domain over which first-order variables range. They are not intended as an exhaustive survey, but as an illustration of how EL thesis [28] is genuinely distinct from existing formalisms.

Propositional linear temporal logic (LTL) is the classical point-based logic of linear executions: formulas are evaluated at positions of a trace, and temporal operators such as next X and until U describe how truth evolves along that trace [18, 24]. LTL is qualitative: it can express that an event eventually occurs, but not that it occurs within a specified numerical delay.

Timed and metric temporal logics add quantitative time in different ways. Metric temporal logic (MTL [19, 22]) decorates temporal operators with fixed intervals, as in $\varphi U_{[a,b]} \psi$. Timed propositional temporal logic (TPTL [3]) instead introduces freeze variables: a formula may store the current timestamp and later compare it with other timestamps by arithmetic constraints [3]. First-order LTL combines temporal operators with quantification over an external data domain, while Halpern-Shoham logic evaluates formulas over intervals rather than points, using Allen-style interval relations [2, 12].

The logic EL is closest in spirit to metric temporal logics, but its metric parameters are not merely constants written into the syntax. Instead, EL quantifies over additive displacements of the timeline. Thus a formula can choose an offset n and then reason about positions such as $v+n$ or $v+n+1$. In discrete time, a bounded metric-until formula has the EL-style rendering $\varphi U_{[a,b]} \psi \equiv \bigvee_{r=a}^b (\psi_r \wedge \Box_r \varphi)$, up to the usual endpoint convention for until. Conversely, EL can quantify over the bound itself, as in formulas of the form $\exists n (\Box_{n+1} \varphi \wedge \dots)$, which is outside ordinary fixed-interval MTL. This ability to quantify over an unbounded domain of indices is also what conceptually brings EL closer to highly undecidable interval logics such as Halpern-Shoham logic over discrete linear orders [5, 12].

■ **Table 1** Comparison along semantic carrier, metric mechanism, and quantified domain.

Logic	Semantic carrier	Metric mechanism	Quantified domain
EL($\mathbb{N}$)	point of an ω -word	exact shifts and bounded windows whose additive parameters may contain variables	numerical offsets in $\mathbb{N}$
LTL	point of an infinite trace	next and qualitative until; no numerical parameter in the standard syntax	none in propositional LTL
MTL	timed position or time point	syntactically fixed metric intervals I on temporal operators	normally no quantification over interval bounds
TPTL	timed-word position with timestamp	freeze binder plus clock constraints	frozen timestamps
FO-LTL	trace position carrying a first-order structure	qualitative temporal operators plus first-order atoms	external objects, not temporal offsets

The most distinct feature of EL is that it is not one single logical formalism. EL represents a class of logics when its semantic carriers are instantiated in unlimited ways: discrete time, dense continuous time, Euclidean space, discrete spatial coordinates, or mixed product spaces, in which the underlying monoid interpretation of the additive terms u and t are fixed [28]. Each such instantiation results in a distinct logic (such as the focus of this paper – Discrete Linear Ensemble Logic), with likely distinct theoretical properties regarding decidability, expressiveness, complexity, and axiomatization.

Attributes of EL: Natural and Minimalist

An important distinction between discrete linear EL and the other formalisms mentioned above lies in its natural and minimalist character. A property of an infinite word is a yes/no question one can ask about the sequence $\alpha(0), \alpha(1), \alpha(2), \dots$. Classical temporal logic LTL answers such questions by stepping through the word one position at a time, using operators such as “next” and “until;” it can look arbitrarily far ahead, but it can only *count* by chaining single steps.

Discrete Linear Ensemble Logic keeps the temporal intuition of a point moving along the word, but it hands that point a ruler (**Fig. 1**). Instead of “sometime later” it can say “somewhere within the next t positions,” and instead of “the next position” it can say “the position exactly u steps further from here” (also called “displacement”) where t and u are *additive terms* built from numerical variables (integer variables in this paper). The single design decision that gives the logic its character is that these lengths are first-class objects: one variable may fix a distance that is then reused, combined, and referred to elsewhere in the same formula.

This is also the reason that complexity results based on LTL do not directly apply in the discrete linear EL context. Whereas the operator X encodes future positions in unary in LTL, the displacement φ_u refers to them through the underlying length measurement captured by u , typically yielding an encoding of logarithmic size. Moreover, nested applications of “until” are even less aligned with the natural-language patterns commonly found in biomedical discourse.

Natural and minimalist are also reflected in EL's general treatment of biomedical data records directly as semantic models (with existing annotated features). Other formalisms usually require a more elaborate semantic structure to work with. A full exposition, however, of how biomedical data records directly correspond to semantic models for EL, is beyond the scope of the current paper. For specific EL instantiations and applications, please see [9, 16, 15, 20].

3 Syntax and Semantics

We first recall basic definitions from [28]. Throughout, $\mathbb{N} = \{0, 1, 2, \dots\}$ and $\mathbb{N}_{>0} = \{1, 2, \dots\}$. We adopt a uniform convention: all first-order variables range over $\mathbb{N}$, while modal lengths must evaluate to numbers ≥ 1 .

► **Definition 1 (Terms).** We fix a countably infinite set of variables $V = \{x, y, z, \dots\}$. Additive terms are defined as $u, v, t, s ::= n \mid x \mid u + v$, where $n \in \mathbb{N}$ is a natural-number numeral and $x \in V$. Numerals are encoded in binary. The syntactic size of a numeral n is bounded by $|n| = \Theta(\log(n + 1))$, and the size $|u|$ of a term is given by the size of its syntax tree under this convention. Throughout the paper, $|\varphi|$ denotes the binary-encoded length of φ .

An assignment $\eta : V \rightarrow \mathbb{N}$ naturally extends to terms by homomorphic evaluation: $\eta(n) = n$ and $\eta(u + v) = \eta(u) + \eta(v)$. Whenever a term t appears as the metric subscript of a modality, its evaluation must be strictly positive in the current assignment ($\eta(t) \geq 1$). Syntactically, this can be ensured by writing modality bounds as $t + 1$.

► **Definition 2 (Formulas).** For a finite set $\mathcal{P}$ of atomic propositions, the formulas of $\text{EL}(\mathbb{N})$ are generated by the grammar

$$\varphi, \psi ::= p \mid \varphi_u \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \Diamond_t \varphi \mid \Box_t \varphi \mid \exists x \varphi \mid \forall x \varphi,$$

subject to the condition that every closed instance of a modal length evaluates to ≥ 1 . Standard abbreviations for $\top, \perp, \rightarrow, \leftrightarrow$ are derived as usual.

An interpretation is defined as $\alpha : \mathbb{N} \rightarrow 2^{\mathcal{P}}$, equivalently an ω -word over the alphabet $2^{\mathcal{P}}$. We denote the support of a proposition p by $P_p^\alpha = \{i \in \mathbb{N} \mid p \in \alpha(i)\}$.

► **Definition 3 (Satisfaction Relation).** For a given interpretation α , variable assignment $\bullet\eta$, and temporal/spatial origin point $i \in \mathbb{N}$, satisfaction is defined inductively by standard Boolean clauses paired with the metric primitives:

$$\begin{aligned} (\alpha, \eta, i) \models p &\iff p \in \alpha(i), \\ (\alpha, \eta, i) \models \varphi_u &\iff (\alpha, \eta, i + \eta(u)) \models \varphi, \\ (\alpha, \eta, i) \models \Diamond_t \varphi &\iff \exists j [i \leq j < i + \eta(t) \wedge (\alpha, \eta, j) \models \varphi], \\ (\alpha, \eta, i) \models \Box_t \varphi &\iff \forall j [i \leq j < i + \eta(t) \rightarrow (\alpha, \eta, j) \models \varphi], \\ (\alpha, \eta, i) \models \exists x \varphi &\iff \exists n \in \mathbb{N} (\alpha, \eta[x \mapsto n], i) \models \varphi, \\ (\alpha, \eta, i) \models \forall x \varphi &\iff \forall n \in \mathbb{N} (\alpha, \eta[x \mapsto n], i) \models \varphi. \end{aligned}$$

A formula is closed if it has no free first-order variables. Variables occurring inside additive subscripts are free unless they occur in the syntactic scope of an enclosing first-order quantifier; displacement and metric modalities do not bind variables. For a closed formula φ , $\alpha \models \varphi$ is a syntactic shorthand for $(\alpha, \emptyset, 0) \models \varphi$, and the corresponding language is defined as $L(\varphi) = \{\alpha \in (2^{\mathcal{P}})^\omega \mid \alpha \models \varphi\}$.

► **Proposition 4** ([28]). *For all formulas φ, ψ and additive terms u, v, s, t evaluating positively at modal lengths, the following semantic equivalences hold:*

- (i) $(\neg\varphi)_u \leftrightarrow \neg(\varphi_u)$, $(\varphi \wedge \psi)_u \leftrightarrow \varphi_u \wedge \psi_u$, and dually for $\vee$;
- (ii) $(\varphi_u)_v \leftrightarrow \varphi_{u+v}$;
- (iii) $\Box_t$ distributes over $\wedge$ and dually $\Diamond_t$ over $\vee$;
- (iv) $\neg\Box_t\varphi \leftrightarrow \Diamond_t\neg\varphi$ and dually for $\neg\Diamond_t$;
- (v) *Monotonicity: if $\eta(s) \leq \eta(t)$, then $\Box_t\varphi \rightarrow \Box_s\varphi$ and $\Diamond_s\varphi \rightarrow \Diamond_t\varphi$.*

As an immediate structural corollary, every $\text{EL}(\mathbb{N})$ formula is equivalent to a strict negation-normal form (NNF) where $\neg$ occurs exclusively as a terminal operator on propositional atoms.

4 Forward Embedding into Monadic Arithmetic

Let $\text{FO}(\mathbb{N}, <, +; \mathcal{P})$ represent first-order logic over the standard structure $(\mathbb{N}, <, +)$ expanded by unary predicates P_p for each $p \in \mathcal{P}$. Atomic formulas consist of $P_p(x)$ and standard equalities and inequalities for Presburger terms.

► **Theorem 5** (Forward embedding). *There exists a linear-time translation T assigning to every $\text{EL}(\mathbb{N})$ formula φ and every arithmetic term z denoting the current evaluation point an arithmetical first-order formula $T_z(\varphi)$ such that, for every interpretation α , assignment η , and point $m \in \mathbb{N}$,*

$$(\alpha, \eta, m) \models \varphi \iff (\mathbb{N}, <, +, (P_p^\alpha)_{p \in \mathcal{P}}), \eta[z \mapsto m] \models T_z(\varphi),$$

where bound variables are alpha-renamed to avoid capture of the distinguished variable z and of the fresh variables introduced by the modal clauses.

Proof. Define T_z by structural recursion:

$$\begin{aligned} T_z(p) &:= P_p(z), \\ T_z(\neg\varphi) &:= \neg T_z(\varphi), \\ T_z(\varphi \text{ op } \psi) &:= T_z(\varphi) \text{ op } T_z(\psi), \quad (\text{op} \in \{\wedge, \vee\}), \\ T_z(\varphi_u) &:= T_{z+u}(\varphi), \\ T_z(Qx\varphi) &:= Qx T_z(\varphi), \quad (Q \in \{\exists, \forall\}), \\ T_z(\Diamond_t\varphi) &:= \exists j (z \leq j \wedge j < z + t \wedge T_j(\varphi)), \\ T_z(\Box_t\varphi) &:= \forall j ((z \leq j \wedge j < z + t) \rightarrow T_j(\varphi)), \end{aligned}$$

with j fresh and with the harmless alpha-renaming convention just stated. Correctness is by structural induction on φ . The atomic, Boolean, and first-order quantifier cases are immediate from the semantics and the induction hypothesis. The displacement clause uses $(\alpha, \eta, m) \models \varphi_u$ iff $(\alpha, \eta, m + \eta(u)) \models \varphi$, which is exactly the induction hypothesis for $T_{z+u}(\varphi)$. The modal clauses match the half-open interval $[m, m + \eta(t))$ by quantifying over the fresh variable j with $z \leq j < z + t$. Each constructor contributes a constant-size translation node when the output is represented as a shared directed acyclic graph, so the translation is linear in $|\varphi|$ under the binary encoding convention of Definition 1. ◀

► **Corollary 6** (Analytical upper bound). *Satisfiability of arbitrary closed $\text{EL}(\mathbb{N})$ formulas resides in Σ_1^1 , and algorithmic validity resides in Π_1^1 .*

Proof. By Theorem 5, the satisfiability of an EL formula φ is equivalent to the satisfiability of the formula $T_0(\varphi)$. Hence deciding satisfiability means existentially quantifying the continuous array of possible monadic subset interpretations P_p^α . This can be expressed by a second-order existential prefix $\exists(P_p)_{p \in \mathcal{P}}$ appended to an arithmetical first-order formula, bounding the procedure within Σ_1^1 . Validity is obtained by the dual quantification $\forall(P_p)_{p \in \mathcal{P}}$, bounded in Π_1^1 . $\blacktriangleleft$

5 Σ_1^1 -Hardness via Two-Counter Machines

By Corollary 6, the EL($\mathbb{N}$)-satisfiability problem is in Σ_1^1 . Here we use a reduction to the recurring problem for nondeterministic two-counter machines to show its Σ_1^1 -hardness. A nondeterministic two-counter machine (2CM) [21] comprises a finite control (Q, q_0, q^*, I) alongside absolute counters $C_1, C_2 \in \mathbb{N}$ driving functional increment and test-and-decrement branches. The recurring problem asks whether there is an infinite computational sequence starting from $(q_0, 0, 0)$ and containing the loop state q^* infinitely often.

► **Lemma 7** (Harel-Pnueli-Stavi [13]). *The recurring problem for nondeterministic two-counter machines is Σ_1^1 -complete.*

Now we can show:

► **Theorem 8** (Σ_1^1 -hardness). *The satisfiability problem for closed EL($\mathbb{N}$) formulas is Σ_1^1 -hard.*

Proof. Fix a nondeterministic two-counter machine $M = (Q, q_0, q^*, I)$. We use propositions $Q \cup \{c_1, c_2, e\}$ and first define exact letter formulas

$$H_q := q \wedge \bigwedge_{r \in Q \setminus \{q\}} \neg r \wedge \neg e \wedge \neg c_1 \wedge \neg c_2,$$

$$E := e \wedge \neg c_1 \wedge \neg c_2 \wedge \bigwedge_{r \in Q} \neg r, \quad C_i := c_i \wedge \neg e \wedge \neg c_{3-i} \wedge \bigwedge_{r \in Q} \neg r \quad (i = 1, 2).$$

A configuration with state q and counter values (x, y) begins at position p and is encoded by

$$B_q(p, x, y) := (H_q)_p \wedge E_{p+1} \wedge (\Box_{x+1} C_1)_{p+2} \wedge E_{p+x+3} \\ \wedge (\Box_{y+1} C_2)_{p+x+4} \wedge E_{p+x+y+5}.$$

Thus a counter value x is represented by a C_1 -block of length $x + 1$, and similarly for y ; all modal lengths are positive. The next configuration begins immediately at

$$N(p, x, y) := p + x + y + 6.$$

Let $\Phi_{\text{init}} := B_{q_0}(0, 0, 0)$. For the transition formula, put the machine instructions into the usual finite normal form: increments are always enabled; a decrement/test instruction for counter C_i has one branch for value 0 and one branch for a positive value. For each state q and each zero/positive case $\chi \in \{0, +\}^2$, write $\bar{x}_0 = 0$, $\bar{x}_+ = x' + 1$, $\bar{y}_0 = 0$, and $\bar{y}_+ = y' + 1$. The clause for q, χ is

$$\forall p \forall x' \forall y' \left(B_q(p, \bar{x}_\chi, \bar{y}_\chi) \rightarrow \bigvee_{\iota \in I(q, \chi)} \text{Succ}_\iota(p, \bar{x}_\chi, \bar{y}_\chi) \right),$$

where $I(q, \chi)$ is the finite set of instructions enabled in that case; an empty disjunction is $\perp$. The successor formula is obtained by writing the corresponding next block at $N(p, \bar{x}_\chi, \bar{y}_\chi)$: for example,

$$\text{Succ}_{q: C_1 := C_1 + 1; q'}(p, x, y) = B_{q'}(N(p, x, y), x + 1, y),$$

and the positive decrement branch for C_1 is written with the current value $x' + 1$ in the antecedent and successor value x' in the consequent,

$$B_q(p, x' + 1, y) \rightarrow B_{q'}(N(p, x' + 1, y), x', y).$$

The zero branch is the analogous clause with current value 0. No arithmetic equality atom is needed; the same variables are reused in the block positions and lengths. Let Φ_{trans} be the conjunction of these finitely many clauses. Finally, recurrence of q^* is expressed by

$$\Phi_{\text{rec}} := \forall r \exists s \exists x \exists y B_{q^*}(r + s, x, y).$$

Set $\varphi_M := \Phi_{\text{init}} \wedge \Phi_{\text{trans}} \wedge \Phi_{\text{rec}}$.

If M has a recurring run, concatenate the corresponding exact blocks. The initial, transition, and recurrence clauses are then immediate. Conversely, suppose $\alpha \models \varphi_M$. From $B_{q_0}(0, 0, 0)$ and Φ_{trans} , induction on successor blocks yields an infinite sequence of adjacent blocks whose states and counter values follow an enabled run of M . Because the formulas H_q, E, C_1, C_2 are mutually exclusive and every successor starts exactly at the end of the current block, no state marker can occur inside a block of this chain; the adjacent blocks therefore tile $\mathbb{N}$. Hence every block satisfying $B_q(p, x, y)$ is a boundary of this run. The formula Φ_{rec} says that such q^* -boundaries occur beyond every cutoff, so the extracted run visits q^* infinitely often. Thus φ_M is satisfiable iff M has a recurring run. The construction uses one constant-size family of clauses per state, instruction, and zero/positive case, and all numerals are fixed small constants; hence $|\varphi_M| = \text{poly}(|M|)$. ◀

► **Corollary 9** (Analytical classification). *The satisfiability problem for closed $\text{EL}(\mathbb{N})$ formulas is Σ_1^1 -complete; the validity problem is Π_1^1 -complete.*

Proof. Upper limits are given by Corollary 6. The hardness of the satisfiability problem is given by Theorem 8. The hardness of the validity problems follows by considering $\neg\varphi_M$. ◀

► **Remark.** A reduction from the Post Correspondence Problem to the satisfiability problem for closed $\text{EL}(\mathbb{N})$ formulas was given in [28] earlier.

6 Position in the Descriptive Hierarchy

In this section we wish to compare the class of EL -definable languages with standard classes of ω -languages. As is well-known, the $\text{MSO}(\mathbb{N}, <; \mathcal{P})$ -definable languages coincide with the ω -regular languages, and the $\text{FO}(\mathbb{N}, <; \mathcal{P})$ -definable languages coincide with the ω -LTL-definable and the ω -star-free rational languages.

► **Example 10** (A non- ω -regular language in EL). We consider the standard alphabet of letters $\Sigma = \{a, b, c, d\}$. Since interpretations are subsets of $\mathcal{P}$ rather than singleton alphabet symbols, using a, b, c, d alone would also accept positions where several propositions hold simultaneously. Therefore, we let $\ell_a := a \wedge \neg b \wedge \neg c \wedge \neg d$ and define ℓ_b, ℓ_c, ℓ_d analogously. The closed formula

$$\exists n \left[\Box_{n+1} \ell_a \wedge (\Box_{n+1} \ell_b)_{n+1} \wedge (\Box_{n+1} \ell_c)_{(n+1)+(n+1)} \wedge (\Box_{n+1} \ell_d)_{(n+1)+(n+1)+(n+1)} \right]$$

defines, under the standard singleton-letter encoding over $\{a, b, c, d\}$, the language $L_{\text{abcd}} = \{a^m b^m c^m d^m \mid m \geq 1\} \cdot \Sigma^\omega$ (choose target index $m = n + 1$). By the pumping lemma, $L_{\text{abcd}} \notin \omega\text{-Reg}$.

► **Example 11** (An ω -regular language outside EL). Let $\Sigma := \{a, b, c, d\}$, and L_{par} over Σ be defined as $L_{\text{par}} := \{wcd^\omega \mid w \in \{a, b\}^*, |w|_a \equiv 0 \pmod{2}\}$. This language is ω -regular: a two-state automaton records the parity of the number of a 's before the unique delimiter c and then checks the d^ω tail. If L_{par} were definable in $\text{EL}(\mathbb{N})$, Theorem 5 would make it definable in $\text{FO}(\mathbb{N}, <, +; \mathcal{P})$. Restricting such a definition to ultimately constant words of the form wcd^ω would yield an $\text{FO}[<, +]$ definition of finite-word parity. This contradicts the classical lower bound that parity is not definable in first-order logic with addition on finite ordered words, equivalently parity is not in uniform AC^0 [10, 1, 14].

Now we obtain:

► **Theorem 12** (Hierarchy classification matrix). *We have:*

- (i) $L(\text{FO}(\mathbb{N}, <; \mathcal{P})) \subsetneq L(\text{EL}(\mathbb{N})) \subseteq L(\text{FO}(\mathbb{N}, <, +; \mathcal{P}))$.
- (ii) $L(\text{EL}(\mathbb{N})) \setminus \omega\text{-Reg} \neq \emptyset$ and $\omega\text{-Reg} \setminus L(\text{EL}(\mathbb{N})) \neq \emptyset$.

Proof. (i) By Kamp's theorem [7, 18], $\text{FO}(\mathbb{N}, <; \mathcal{P})$ -definability of ω -languages coincides with LTL-definability. It is straightforward to translate LTL-formulas into equivalent EL-formulas (i.e., preserving the languages): For the *Next* operator, replace $X\varphi$ by φ_1 , and for *strict Until*, replace $\varphi U_{>0} \psi$ by $\psi_1 \vee \exists x(\psi_{x+2} \wedge (\Box_{x+1} \varphi)_1)$. By Example 10, the inclusion is proper. For the second inclusion, we apply Theorem 5. (ii) Immediate by Examples 10 and 11. ◀

7 A Hilbert System and Oracle-Relative Completeness

In this section, we wish to obtain a Hilbert system for EL-logic and to consider completeness questions. Absolute completeness is impossible by Corollary 9; completeness relative only to Presburger arithmetic is also impossible, since Presburger arithmetic is decidable (Cooper [6], Presburger [25]) and any r.e. system using only $\text{Th}(\mathbb{N}, <, +)$ as oracle would have an r.e. proof relation, contradicting the Π_1^1 -completeness of EL-validity. Therefore we use the monadic Presburger validity theory as oracle.

The system $\mathcal{H}_{\text{EL}}$

The axioms below are given as schemas; rules are modus ponens, universal generalisation, and necessitation for $\varphi_u, \Box_t, \Diamond_t$ from valid formulas.

Group I (propositional and first-order):

- (A1) propositional tautologies;
- (A2) $\forall x \varphi(x) \rightarrow \varphi[u/x]$ when u free for x ;
- (A3) $\forall x(\varphi \rightarrow \psi) \rightarrow (\varphi \rightarrow \forall x\psi)$ when $x \notin \text{free}(\varphi)$.

Group II (displacement):

- (B1) displacement distributes over $\wedge, \vee, \neg$;
- (B2) $(\varphi_u)_v \leftrightarrow \varphi_{u+v}$;
- (B3) if $u = v$ is arithmetically true, $\varphi_u \leftrightarrow \varphi_v$;
- (B4) $(Qx \varphi)_u \leftrightarrow Qx(\varphi_u)$ for $Q \in \{\exists, \forall\}$, $x \notin \text{free}(u)$.

Group III (metric modalities):

- (C1) $\neg \Box_t \varphi \leftrightarrow \Diamond_t \neg \varphi$ (dually);
- (C2) $\Box_t(\varphi \rightarrow \psi) \rightarrow (\Box_t \varphi \rightarrow \Box_t \psi)$;
- (C3) distribution of $\Box_t / \Diamond_t$ over $\wedge / \vee$;
- (C4) monotonicity: if $s \leq t$ holds in $(\mathbb{N}, <, +)$, then $\Box_t \varphi \rightarrow \Box_s \varphi$ and $\Diamond_s \varphi \rightarrow \Diamond_t \varphi$;
- (C5) $\Box_1 \varphi \leftrightarrow \varphi, \Diamond_1 \varphi \leftrightarrow \varphi$;

- (C6) $(\Box_t \varphi)_u \leftrightarrow \Box_t(\varphi_u)$ (dually for $\Diamond$);
- (C7) for s, t positive under every assignment, nested modalities compose: $\Diamond_s \Diamond_{t+1} \varphi \leftrightarrow \Diamond_{s+t} \varphi$ (dually for $\Box$);
- (C8) *additive decomposition*: for u, v positive under every assignment: $\Diamond_{u+v} \varphi \leftrightarrow \Diamond_u \varphi \vee (\Diamond_v \varphi)_u$ (dually for $\Box$).

Group IV (induction over displacement variables):

- (D1) $\exists x \varphi(x) \leftrightarrow \varphi(0) \vee \exists y \varphi(y+1)$; (dually for $\forall$);
- (D2) induction $\varphi(0) \wedge \forall x(\varphi(x) \rightarrow \varphi(x+1)) \rightarrow \forall x \varphi(x)$.

► **Theorem 13 (Soundness).** *If $\mathcal{H}_{EL} \vdash \varphi$ then $\models \varphi$.*

Proof. We verify the non-trivial axioms. **C7:** $(\alpha, \eta, i) \models \Diamond_s(\Diamond_t \varphi)$ iff $\exists r < s, q < t. (\alpha, \eta, i + r + q) \models \varphi$; the attainable sums $r + q$ form $[0, s + t - 1)$, giving $\Diamond_{s+t-1} \varphi$.

C8: for $a = \eta(u), b = \eta(v) \geq 1$, use the partition

$$[i, i + a + b) = [i, i + a) \cup [i + a, (i + a) + b).$$

D1: from $\mathbb{N} = \{0\} \cup \{y + 1 \mid y \geq 0\}$; **D2** is mathematical induction. Other axioms follow from Proposition 4 and direct inspection. ◀

The oracle

$$\text{MONFO}_{\mathbb{N},+} := \{ \Theta \in \text{FO}(\mathbb{N}, <, +; \mathcal{P}) \mid (\mathbb{N}, <, +, (P_p)) \models \Theta \text{ for every interpretation of the } P_p \}.$$

This is the standard-validity theory of monadic Presburger arithmetic.

► **Theorem 14 (Oracle-relative completeness).** *Extend $\mathcal{H}_{EL}$ with the rule: from $T_0(\varphi) \in \text{MONFO}_{\mathbb{N},+}$ infer φ (for closed φ). Then $\models \varphi \Leftrightarrow \mathcal{H}_{EL} + \text{MONFO}_{\mathbb{N},+} \vdash \varphi$.*

Proof. The proof is provided in the Appendix. ◀

By Halpern [11], $\text{MONFO}_{\mathbb{N},+}$ is itself Π_1^1 -complete; this is the correct level of relative completeness given Corollary 9. Axiom C8 (additive decomposition) is the central modal-arithmetic primitive: it states that inspection of a window of length $u + v$ factors through inspection of a window of length u followed by a displaced window of length v .

As an example for the power of this calculus, we just note that Barcan-type interaction laws can be derived. With the usual assumption of always-positive t , and $x \notin \text{free}(t)$, we have: $\vdash \forall x \Box_t \varphi \leftrightarrow \Box_t \forall x \varphi$; $\vdash \exists x \Diamond_t \varphi \leftrightarrow \Diamond_t \exists x \varphi$; $\vdash \exists x \Box_t \varphi \rightarrow \Box_t \exists x \varphi$; and $\vdash \Diamond_t \forall x \varphi \rightarrow \forall x \Diamond_t \varphi$.

8 The Existential Fragment $\exists \text{EL}(\mathbb{N})$

► **Definition 15 (Existential fragment).** *The existential fragment $\exists \text{EL}(\mathbb{N})$ consists of the closed formulas in negation normal form generated by*

$$\varphi, \psi ::= p \mid \neg p \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi_u \mid \Diamond_t \varphi \mid \exists x \varphi.$$

Thus $\forall$ and $\Box$ are not primitive in the fragment, and negation is allowed only in front of atomic propositions. As throughout the paper, metric bounds are required to be positive on every branch where the modality is used; in applications this is enforced syntactically by writing bounds as $t + 1$.

An explicit translation to existential Presburger arithmetic

We compile $\exists\text{EL}(\mathbb{N})$ -formulas into existential Presburger arithmetic with the standard binary numeral encoding. Up to inessential bookkeeping, the construction is determined by the following idea: the only semantic freedom in satisfiability is the choice of the monadic predicates $P_p^\alpha \subseteq \mathbb{N}$. This freedom can be eliminated as follows: each diamond chooses one witness inside its window, and each literal occurrence inspects *one* truth value of *one* proposition at *one* arithmetically computed position. If two such inspection positions happen to collide, the two truth values must agree. Once those constraints are written down, the satisfiability question reduces to a pure existential Presburger question.

Fix a closed formula $\varphi \in \exists\text{EL}(\mathbb{N})$. First alpha-rewrite all bound variables and assign a fresh variable h_δ to every occurrence δ of a modality $\Diamond_t$. Let $\bar{v}_\varphi$ be the tuple of all renamed first-order variables and all modal-witness variables h_δ . For every literal occurrence ρ introduce a fresh Boolean Presburger variable b_ρ . We define simultaneously a quantifier-free Presburger formula $B_\tau(\psi)$ and, for every literal occurrence reached during the recursion, its proposition p_ρ and absolute position term a_ρ :

$$\begin{aligned} B_\tau(p^\rho) &:= b_\rho = 1, & a_\rho &:= \tau, & p_\rho &:= p, \\ B_\tau((\neg p)^\rho) &:= b_\rho = 0, & a_\rho &:= \tau, & p_\rho &:= p, \\ B_\tau(\psi \wedge \chi) &:= B_\tau(\psi) \wedge B_\tau(\chi), & B_\tau(\psi \vee \chi) &:= B_\tau(\psi) \vee B_\tau(\chi), \\ B_\tau(\psi_u) &:= B_{\tau+u}(\psi), & B_\tau(\exists x \psi) &:= B_\tau(\psi), \\ B_\tau(\Diamond_t^\delta \psi) &:= (0 \leq h_\delta \wedge h_\delta < t) \wedge B_{\tau+h_\delta}(\psi). \end{aligned}$$

The clause for $\exists x$ is legitimate because all existential variables have already been alpha-renamed and are placed in the global existential prefix. Moving them to the front preserves truth, since only existential first-order quantifiers are present.

Let $\text{Occ}(\varphi)$ be the finite set of literal occurrences. Define

$$\begin{aligned} \text{Bool}_\varphi &:= \bigwedge_{\rho \in \text{Occ}(\varphi)} (b_\rho = 0 \vee b_\rho = 1), \\ \text{Cons}_\varphi &:= \bigwedge_{\substack{\rho, \sigma \in \text{Occ}(\varphi) \\ p_\rho = p_\sigma}} (a_\rho = a_\sigma \rightarrow b_\rho = b_\sigma), \\ \text{EP}(\varphi) &:= \exists \bar{v}_\varphi \exists (b_\rho)_{\rho \in \text{Occ}(\varphi)} (B_0(\varphi) \wedge \text{Bool}_\varphi \wedge \text{Cons}_\varphi). \end{aligned}$$

Here implication is read as the Presburger formula $a_\rho \neq a_\sigma \vee b_\rho = b_\sigma$.

The formula $\text{EP}(\varphi)$ is an existential Presburger formula of size polynomial in $|\varphi|$; the only quadratic blow-up is the pairwise consistency condition.

► **Lemma 16** (Correctness of the Presburger translation). *For every closed $\varphi \in \exists\text{EL}(\mathbb{N})$,*

$$\varphi \text{ is satisfiable over } \omega\text{-words} \iff (\mathbb{N}, <, +) \models \text{EP}(\varphi).$$

Proof. Suppose first that $\alpha \models \varphi$. Choose witnesses for the first-order existential quantifiers and, along one successful evaluation of each used diamond occurrence, choose h_δ to be the offset from the current point to the witnessing point. Variables belonging only to unused disjunctive branches may be assigned arbitrarily. For every literal occurrence ρ , set $b_\rho = 1$ iff $p_\rho \in \alpha(\eta(a_\rho))$. Then Bool_φ is immediate. If two occurrences ρ, σ mention the same proposition and $\eta(a_\rho) = \eta(a_\sigma)$, they inspect the same proposition at the same position of the same word, so $b_\rho = b_\sigma$; hence Cons_φ holds. A direct induction on the successful evaluation tree of φ gives $B_0(\varphi)$.

Conversely, assume $(\mathbb{N}, <, +) \models \text{EP}(\varphi)$ and fix satisfying values for $\bar{v}_\varphi$ and the Boolean variables b_ρ . Define an ω -word α as follows. For every proposition p and position n , if there exists an occurrence ρ with $p_\rho = p$ and $\eta(a_\rho) = n$, put $p \in \alpha(n)$ exactly when $b_\rho = 1$; if no such occurrence exists, choose the truth value of p at n arbitrarily, say false. The definition is independent of the chosen occurrence by Cons_φ . Now prove by structural induction on subformulas that whenever $B_\tau(\psi)$ is true under the fixed witnesses, then $(\alpha, \eta, \eta(\tau)) \models \psi$. The literal cases follow from the construction of α ; Boolean cases are immediate; displacement changes the current term from τ to $\tau + u$; the existential case uses the globally fixed value of the renamed variable; and the diamond case uses the fixed offset h_δ , whose conjunct $0 \leq h_\delta < t$ places the witnessing point inside the half-open interval required by Definition 3. Applying the induction to $B_0(\varphi)$ yields $\alpha \models \varphi$. ◀

► **Theorem 17** (Decidability and NP-completeness). *Satisfiability of closed $\exists\text{EL}(\mathbb{N})$ formulas is decidable and NP-complete.*

Proof. Membership in NP follows from Lemma 16 and the classical NP upper bound for satisfiability of existential Presburger arithmetic with binary numerals [23]. The translation $\varphi \mapsto \text{EP}(\varphi)$ is polynomial: it introduces one numeric witness for each diamond, one Boolean variable for each literal occurrence, and only a quadratic number of consistency implications.

For hardness, reduce propositional satisfiability. Given a propositional formula F over variables $X_1, \dots, X_m$, introduce propositions $p_1, \dots, p_m$ and replace each positive literal X_i by p_i and each negative literal $\neg X_i$ by $\neg p_i$. The resulting $\exists\text{EL}(\mathbb{N})$ formula contains no quantifier and no modality. A truth assignment satisfying F is realised by choosing the truth values of the propositions p_i at position 0, and conversely every satisfying word determines such a truth assignment from its letter at position 0. Thus F is satisfiable iff the translated $\exists\text{EL}(\mathbb{N})$ formula is satisfiable. ◀

► **Corollary 18** (Finite certificate and unsatisfiability complexity). *Every satisfiable closed $\exists\text{EL}(\mathbb{N})$ formula has a certificate consisting of polynomial-bit Presburger witnesses together with truth values for the finitely many literal occurrences. Consequently, unsatisfiability of closed $\exists\text{EL}(\mathbb{N})$ formulas is coNP-complete.*

Proof. The certificate is precisely a satisfying assignment for $\text{EP}(\varphi)$. The polynomial bit bound is the standard small-certificate property behind the NP upper bound for existential Presburger arithmetic. Since satisfiability is NP-complete by Theorem 17, its complement, unsatisfiability, is coNP-complete. ◀

A complete proof system for satisfiability of the existential fragment

Let PrA be any fixed sound and complete axiomatisation of Presburger arithmetic. For

$$\text{EP}(\varphi) = \exists \bar{y} \Theta_\varphi(\bar{y}),$$

where Θ_φ is quantifier-free, define a labelled calculus $\text{S}_{\exists\text{EL}}$ for satisfiability judgements by the two terminal rules

$$\frac{\text{PrA} \vdash \Theta_\varphi(\bar{n})}{\text{S}_{\exists\text{EL}} \vdash \text{Sat}(\varphi)} \quad \frac{\text{PrA} \vdash \neg \exists \bar{y} \Theta_\varphi(\bar{y})}{\text{S}_{\exists\text{EL}} \vdash \text{Unsat}(\varphi)},$$

where $\bar{n}$ ranges over tuples of numerals of the right arity. The first rule is a witness rule; the second is a Presburger refutation rule. In both cases Θ_φ is the matrix produced by the explicit translation above.

For standard validity proofs, let $\text{H}_{\text{rel}}^\exists$ denote the restriction of the oracle-extended Hilbert system of Theorem 14 to conclusions that belong to $\exists\text{EL}(\mathbb{N})$.

► **Theorem 19** (Complete proof systems for satisfiability of $\exists\text{EL}(\mathbb{N})$). *For every closed $\varphi \in \exists\text{EL}(\mathbb{N})$:*

- (i) $S_{\exists\text{EL}} \vdash \text{Sat}(\varphi)$ iff φ is satisfiable;
- (ii) $S_{\exists\text{EL}} \vdash \text{Unsat}(\varphi)$ iff φ is unsatisfiable;
- (iii) $H_{\text{rel}}^{\exists} \vdash \varphi$ iff $\models \varphi$.

Proof. For (i), if the witness rule derives $\text{Sat}(\varphi)$, then the Presburger proof of $\Theta_{\varphi}(\bar{n})$ makes $\text{EP}(\varphi)$ true, and Lemma 16 gives a satisfying word. Conversely, if φ is satisfiable, Lemma 16 gives a tuple $\bar{n}$ satisfying Θ_{φ} ; completeness of PrA proves the true ground Presburger sentence $\Theta_{\varphi}(\bar{n})$, so the witness rule applies.

For (ii), the refutation rule is sound because $\text{PrA} \vdash \neg \exists \bar{y} \Theta_{\varphi}$ implies that $\text{EP}(\varphi)$ is false, hence φ is unsatisfiable by Lemma 16. Conversely, if φ is unsatisfiable, then $\neg \text{EP}(\varphi)$ is a true Presburger sentence, so it is provable in PrA and the refutation rule applies.

For (iii), $\exists\text{EL}(\mathbb{N})$ is a syntactic subfragment of $\text{EL}(\mathbb{N})$. Therefore the soundness and completeness equivalence $\mathcal{H}_{\text{EL}} + \text{MONFO}_{\mathbb{N},+} \vdash \psi$ iff $\models \psi$ from Theorem 14 applies in particular to every closed existential formula $\psi = \varphi$; restricting the set of allowed conclusions does not change derivability of such conclusions. ◀

9 Finite Active-Domain Model Checking

In clinical applications, model checking is performed over a finite active domain. Padding a finite trace by a don't-care suffix would leave quantifiers ranging over an infinite tail and is a different problem.

► **Definition 20** (Active-domain semantics). *A finite trace of length N is $\sigma : D_N \rightarrow 2^{\mathcal{P}}$ for $D_N = \{0, \dots, N-1\}$. Active-domain satisfaction $\sigma, \eta, i \models_N \varphi$ is as in Definition 3, except that all positions and quantifier ranges are restricted to D_N : φ_u requires $i + \eta(u) \in D_N$; modalities range over $D_N \cap [i, i + \eta(t))$; quantifiers range over D_N .*

► **Theorem 21** (Finite model-checking complexity). *For finite trace σ of length N and closed $\varphi \in \text{EL}(\mathbb{N})$, deciding $\sigma \models_N \varphi$ has (i) PTIME data complexity for fixed φ , and (ii) PSPACE-complete combined complexity in $|\varphi| + \log N$, where $|\varphi|$ is the binary-encoded size from Definition 1 and $\log N$ is the bit-length of the active-domain bound.*

Proof.

Upper bound. Evaluate φ by DFS recursion. A configuration is (occurrence, $i, \eta \upharpoonright \text{scope}$); each value uses $O(\log N)$ bits, scope has $\leq |\varphi|$ variables, recursion depth is $\leq |\varphi|$. Quantifiers and modalities iterate over D_N (and over $D_N \cap [i, i + \eta(t))$); arithmetic evaluations $\eta(u), \eta(t)$ on binary terms are done in $\text{poly}(|\varphi| + \log N)$ time and $O(\log N)$ space. The algorithm uses polynomial space in $|\varphi| + \log N$. For fixed φ , nested iterations give time $O(N^c)$, $c = c(\varphi)$.

Hardness. Reduce QBF on traces of length 2: $\mathcal{P} = \{T\}$, $\sigma(0) = \emptyset$, $\sigma(1) = \{T\}$; translate $X_i \mapsto T_{x_i}$, $\exists X_i \mapsto \exists x_i$, $\forall X_i \mapsto \forall x_i$. Each x_i ranges over $\{0, 1\}$ and T_{x_i} is true iff $x_i = 1$. The reduction is linear in any encoding. ◀

10 Discussion: Expressive Limits

We conjecture that the EL -definable languages form a proper subclass of the $\text{FO}(\mathbb{N}, <, +; \mathcal{P})$ -definable languages; a proof could proceed using Ehrenfeucht-Fraïssé-games. It would be interesting to explore suitable oracle boundaries for sub-hierarchies in $\mathcal{H}_{\text{EL}}$. Furthermore, it is open whether bounded sub-fragments of EL might map exactly to known classes of ω -recognizable languages.

11 Conclusion

This paper develops the foundational discrete theory of Ensemble Logic over $\mathbb{N}$. Starting from the three core operations of exact displacement, bounded existence, and bounded universality, we gave a syntax and semantics for $\text{EL}(\mathbb{N})$ that is expressive enough to represent metric constraints while remaining close to the symbolic forms used in temporal, spatial, genomic, and clinical reasoning. The formal development clarifies the precise role of additive terms, positive modal bounds, and first-order quantification over the discrete index domain.

The first main contribution is the translation of $\text{EL}(\mathbb{N})$ into monadic first-order Presburger arithmetic with unary predicates. The capture-avoiding embedding provides a uniform semantic bridge from Ensemble Logic to a classical arithmetic setting and yields sharp analytical upper bounds for satisfiability and validity. Complementing this upper bound, the two-counter-machine construction shows that these bounds are tight: satisfiability is Σ_1^1 -complete and validity is Π_1^1 -complete. Thus the full logic is highly expressive, and its undecidability is not an artefact of the presentation but an intrinsic consequence of combining metric displacement, bounded modalities, and quantification.

The second contribution is the placement of $\text{EL}(\mathbb{N})$ in the descriptive hierarchy. The logic contains the star-free ω -languages and can define counting patterns such as $\{a^m b^m c^m d^m \mid m \geq 1\} \cdot \Sigma^\omega$, showing that it reaches beyond ω -regularity. At the same time, the parity-separation argument shows that $\text{EL}(\mathbb{N})$ does not subsume all ω -regular behavior. This gives a more nuanced expressiveness profile: $\text{EL}(\mathbb{N})$ is neither simply a metric variant of linear temporal logic nor merely a notational fragment of regular-language theory.

The third contribution is proof-theoretic. We introduced the Hilbert system $\mathcal{H}_{\text{EL}}$, proved its soundness, and established an oracle-relative completeness theorem with respect to monadic Presburger validity. The relative theorem identifies the exact arithmetic oracle needed for full completeness and explains why a weaker oracle for plain Presburger arithmetic cannot suffice. For the existential fragment, the Presburger-based calculus and complexity analysis give a decidable and computationally meaningful subtheory, with NP-complete satisfiability and coNP-complete unsatisfiability.

The fourth contribution is algorithmic. Finite active-domain model checking has PTIME data complexity for fixed formulas and PSPACE-complete combined complexity. This separates the cost of evaluating a fixed biomedical rule over a finite trace from the cost of allowing the rule itself to vary, and it provides the appropriate complexity baseline for implementations that target electronic health records, clinical-trial protocols, sleep-event streams, genomic coordinates, or other finite symbolic data sources.

Overall, the results establish $\text{EL}(\mathbb{N})$ as a mathematically robust discrete metric logic: expressive enough to encode non-regular counting and analytically complete computation-theoretic phenomena, yet structured enough to admit exact embeddings, proof systems for meaningful fragments, and finite-trace model-checking algorithms. The remaining expressive-limit questions discussed above point toward a sharper classification of pure EL inside monadic Presburger arithmetic and toward practically useful bounded fragments for biomedical knowledge representation.

References

- 1 Miklós Ajtai. Σ_1^1 -formulae on finite structures. *Annals of Pure and Applied Logic*, 24(1):1–48, 1983. doi:10.1016/0168-0072(83)90038-6.
- 2 James F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843, 1983. doi:10.1145/182.358434.
- 3 Rajeev Alur and Thomas A. Henzinger. A really temporal logic. *Journal of the ACM*, 41(1):181–204, 1994. doi:10.1145/174644.174651.

- 4 Olivier Bodenreider. The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Research*, 32(Database issue):D267–D270, 2004. doi:10.1093/nar/gkh061.
- 5 Davide Bresolin, Angelo Montanari, Pietro Sala, and Guido Sciavicco. What’s decidable about Halpern and Shoham’s interval logic? The maximal fragment ABBL. In *Proceedings of the 26th Annual IEEE Symposium on Logic in Computer Science (LICS 2011)*, pages 387–396. IEEE Computer Society, 2011. doi:10.1109/LICS.2011.35.
- 6 D. C. Cooper. Theorem proving in arithmetic without multiplication. In Bernard Meltzer and Donald Michie, editors, *Machine Intelligence*, volume 7, pages 91–99. Edinburgh University Press, Edinburgh, 1972.
- 7 S. Shelah D. Gabbay, A. Pnueli and J. Stavi. On the temporal analysis of fairness. In *Conference Record of the Seventh Annual ACM Symposium on Principles of Programming Languages (POPL)*, 1980, pages 163–173. ACM Press, 1980. doi:10.1145/567446.567462.
- 8 Emek Demir et al. The BioPAX community standard for pathway data sharing. *Nature Biotechnology*, 28(9):935–942, 2010. doi:10.1038/nbt.1666.
- 9 J. Fan, Xiaojin Li, Yan Huang, X. Hao, Licong Cui, and Guo-Qiang Zhang. Ensemble logic for symbolic representation of sleep medicine guidelines. In *Proceedings of the AMIA 2026 Annual Symposium*, 2026. In press, 10 pages. doi:10.48550/arXiv.2607.21331.
- 10 Merrick L. Furst, James B. Saxe, and Michael Sipser. Parity, circuits, and the polynomial-time hierarchy. *Mathematical Systems Theory*, 17(1):13–27, 1984. doi:10.1007/BF01744431.
- 11 Joseph Y. Halpern. Presburger arithmetic with unary predicates is Π_1^1 complete. *Journal of Symbolic Logic*, 56(2):637–642, 1991. doi:10.2307/2274706.
- 12 Joseph Y. Halpern and Yoav Shoham. A propositional modal logic of time intervals. *Journal of the ACM*, 38(4):935–962, 1991. doi:10.1145/115234.115351.
- 13 David Harel, Amir Pnueli, and Jonathan Stavi. Propositional dynamic logic of nonregular programs. *Journal of Computer and System Sciences*, 26(2):222–243, 1983. doi:10.1016/0022-0000(83)90014-4.
- 14 Johan Håstad. Almost optimal lower bounds for small depth circuits. In *Proceedings of the 18th Annual ACM Symposium on Theory of Computing (STOC 1986)*, pages 6–20. ACM, 1986. doi:10.1145/12130.12132.
- 15 Yan Huang, Xiaojin Li, Licong Cui, and Guo-Qiang Zhang. A logic-based temporal cohort discovery engine: algorithms, indices, and experimental results on the National Sleep Research Resource. *Journal of Biomedical Informatics*, 2026. Under review. doi:10.48550/arXiv.2607.21377.
- 16 Yan Huang, Xiaojin Li, X. Hao, Rashmie Abeysinghe, Licong Cui, and Guo-Qiang Zhang. Scaling up formal representation of clinical trial protocols in ensemble logic using LLMs: a preliminary study. In *Proceedings of the AMIA 2026 Annual Symposium*, 2026. In press, 10 pages. doi:10.48550/arXiv.2607.21307.
- 17 Michael Hucka et al. The Systems Biology Markup Language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531, 2003. doi:10.1093/bioinformatics/btg015.
- 18 Hans Kamp. *Tense Logic and the Theory of Linear Order*. PhD thesis, University of California, Los Angeles, 1968. URL: <https://philpapers.org/rec/KAMTLA-2>.
- 19 Ron Koymans. Specifying real-time properties with metric temporal logic. *Real-Time Systems*, 2(4):255–299, 1990. doi:10.1007/BF01995674.
- 20 Xiaojin Li, Yan Huang, Rashmie Abeysinghe, Zenan Sun, Hongyu Chen, Pengze Li, Xing He, Shiqiang Tao, Cui Tao, Jiang Bian, Licong Cui, and Guo-Qiang Zhang. Temporal Ensemble Logic for Integrative Representation of the Entirety of Clinical Trials. In Thierry Vidal and Przemysław Andrzej Wałęga, editors, *32nd International Symposium on Temporal Representation and Reasoning (TIME 2025)*, volume 355 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 13:1–13:16, Dagstuhl, Germany, 2025. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.TIME.2025.13.

- 21 Marvin L. Minsky. *Computation: Finite and Infinite Machines*. Prentice-Hall, Englewood Cliffs, NJ, 1967.
- 22 Joël Ouaknine and James Worrell. On the decidability of metric temporal logic. In *Proceedings of the 20th Annual IEEE Symposium on Logic in Computer Science (LICS 2005)*, pages 188–197. IEEE Computer Society, 2005. doi:10.1109/LICS.2005.33.
- 23 Christos H. Papadimitriou. On the complexity of integer programming. *Journal of the ACM*, 28(4):765–768, 1981. doi:10.1145/322276.322287.
- 24 Amir Pnueli. The temporal logic of programs. In *Proceedings of the 18th Annual Symposium on Foundations of Computer Science (FOCS 1977)*, pages 46–57, Providence, RI, 1977. IEEE. doi:10.1109/SFCS.1977.32.
- 25 Mojżesz Presburger. Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt. In *Comptes Rendus du I Congrès des Mathématiciens des Pays Slaves*, pages 92–101, Warsaw, 1929. URL: <https://eudml.org/doc/212367>.
- 26 Jean Serra and Pierre Soille, editors. *Mathematical Morphology and Its Applications to Image Processing*, volume 2 of *Computational Imaging and Vision*. Kluwer Academic Publishers, Dordrecht, 1994. doi:10.1007/978-94-011-1040-2.
- 27 U.S. National Library of Medicine. Effect of Different Doses of SAR110894 on Cognition in Patients with Mild to Moderate Alzheimer’s Disease on Donepezil. ClinicalTrials.gov, 2011. ClinicalTrials.gov Identifier: NCT01266525. URL: <https://clinicaltrials.gov/study/NCT01266525>.
- 28 Guo-Qiang Zhang. Temporal Ensemble Logic. *CoRR*, abs/2408.14443, 2024. doi:10.48550/arXiv.2408.14443.

A Proof of Theorem 14

► **Theorem 14** (Oracle-relative completeness). *Extend $\mathcal{H}_{EL}$ with the rule: from $T_0(\varphi) \in \text{MONFO}_{\mathbb{N},+}$ infer φ (for closed φ). Then $\models \varphi \Leftrightarrow \mathcal{H}_{EL} + \text{MONFO}_{\mathbb{N},+} \vdash \varphi$.*

Proof.

Soundness of the extended system $\mathcal{H}_{EL} + \text{MONFO}_{\mathbb{N},+}$. We must show that for every closed $\varphi \in \text{EL}(\mathbb{N})$, if $\mathcal{H}_{EL} + \text{MONFO}_{\mathbb{N},+} \vdash \varphi$ then $\models \varphi$. The system $\mathcal{H}_{EL} + \text{MONFO}_{\mathbb{N},+}$ inherits all axioms and rules of $\mathcal{H}_{EL}$ and adds a single rule: from $T_0(\varphi) \in \text{MONFO}_{\mathbb{N},+}$, infer φ . We verify soundness for each kind of derivation step.

Inherited $\mathcal{H}_{EL}$ -axioms and rules. By Theorem 13, every $\mathcal{H}_{EL}$ -axiom is valid on every ω -word, and modus ponens, generalisation, and necessitation preserve validity. The verification there proceeds axiom-by-axiom: **A1** by the classical truth-table argument, **A2–A3** by standard first-order semantics, **A4** by definition of $\text{Th}(\mathbb{N}, <, +)$, **B1–B4** by the displacement clause of Definition 3, **C1–C8** by case analysis on the satisfaction intervals (with **C7**, **C8** explicitly checked in Theorem 13), **D1** from $\mathbb{N} = \{0\} \cup \{y + 1 \mid y \geq 0\}$, and **D2** from mathematical induction.

The oracle rule. Suppose the last step of the derivation is $T_0(\varphi) \in \text{MONFO}_{\mathbb{N},+} \Rightarrow \varphi$. By Theorem 5, for every ω -word α and every assignment η ,

$$(\alpha, \eta, 0) \models \varphi \iff (\mathbb{N}, <, +, (P_p^\alpha)_{p \in \mathcal{P}}, \eta) \models T_0(\varphi).$$

Since φ is closed, η is immaterial. The condition $T_0(\varphi) \in \text{MONFO}_{\mathbb{N},+}$ asserts that $T_0(\varphi)$ holds in $(\mathbb{N}, <, +)$ under *every* interpretation (P_p) of the unary predicates. In particular, for every ω -word α , the assignment $P_p \mapsto P_p^\alpha$ is one such interpretation, so $(\mathbb{N}, <, +, (P_p^\alpha)) \models T_0(\varphi)$, hence $(\alpha, 0) \models \varphi$. As α was arbitrary, $\models \varphi$.

Completeness of $\mathcal{H}_{\text{EL}} + \text{MONFO}_{\mathbb{N},+}$. Assume $\models \varphi$. By the equivalence in Theorem 5 applied in the *contrapositive* direction, for every ω -word α , $(\mathbb{N}, <, +, (P_p^\alpha)) \models T_0(\varphi)$. Crucially, the map $\alpha \mapsto (P_p^\alpha)_{p \in \mathcal{P}}$ is a bijection between ω -words and unary-predicate interpretations $\mathcal{P} \rightarrow 2^{\mathbb{N}}$: setting $P_p^\alpha = \{i \in \mathbb{N} : p \in \alpha(i)\}$ defines a bijection with inverse $\alpha(i) = \{p \in \mathcal{P} : i \in P_p\}$. Hence every interpretation (P_p) arises as (P_p^α) for some α , and

$$T_0(\varphi) \in \text{MONFO}_{\mathbb{N},+} \iff \forall \alpha (\alpha, 0) \models \varphi \iff \models \varphi.$$

The forward direction of this chain gives $T_0(\varphi) \in \text{MONFO}_{\mathbb{N},+}$, and the oracle rule then derives φ in a single step. Thus $\mathcal{H}_{\text{EL}} + \text{MONFO}_{\mathbb{N},+} \vdash \varphi$.

Axiom-usage. The completeness derivation produced above uses *only* the oracle rule and *none* of the structural axioms or rules of $\mathcal{H}_{\text{EL}}$. This is the standard content of an *oracle-relative completeness* theorem in modal logic: the underlying calculus is sound (and the oracle rule extends this soundly), but the completeness reduction itself is delegated to the oracle. The role of the $\mathcal{H}_{\text{EL}}$ -axioms is therefore:

- **Sub-derivations leading to the oracle.** While the completeness construction above invokes the oracle in one step, practical proof search would typically use the $\mathcal{H}_{\text{EL}}$ -axioms to *normalise* φ before consulting the oracle (e.g., **C8** unfolds bounded modalities, **B1-B2** flatten nested displacements, **D1** unfolds existentials), reducing the size of the formula handed to the oracle.
- **Oracle-free completeness on restricted fragments.** The structural axioms suffice without an oracle on the *existential fragment* (Theorem 19), where the existential counterparts **A1**[∃]-**E4**[∃] give a meta-complete system without an oracle, although with **A4**[∃] retained as a Presburger side-condition oracle.

Necessity of the oracle. By Corollary 9, validity in $\text{EL}(\mathbb{N})$ is Π_1^1 -complete, hence not recursively enumerable. No recursive extension of $\mathcal{H}_{\text{EL}}$ can therefore be complete; the relativisation to $\text{MONFO}_{\mathbb{N},+}$ (also Π_1^1 -complete, by Halpern [11]) is optimal in this analytical sense. ◀