

Two-Variable Logic for Hierarchically Partitioned and Ordered Data (Extended Abstract)

Oskar Fiuk¹, Emanuel Kieroński¹ and Vincent Michielini²

¹University of Wrocław, Poland

²University of Bordeaux, France

Abstract

We study the Two-Variable fragment of First-Order Logic, FO^2 , under semantic constraints that model hierarchically structured data. Our first logic extends FO^2 with a linear order $<$ and a chain of increasingly coarser equivalence relations $E_1 \subseteq E_2 \subseteq \dots$: we show that its finite satisfiability problem is $NEXPTIME$ -complete. We also demonstrate that a weaker variant of this logic, without the linear order, enjoys the exponential model property. Our second logic extends FO^2 with a chain of nested total preorders $\preceq_1 \subseteq \preceq_2 \subseteq \dots$: we prove that its finite satisfiability problem is also $NEXPTIME$ -complete. However, we show that the complexity increases to $EXPSpace$ -complete once access to the successor relations of the preorders is allowed. Our last result is the undecidability of FO^2 with two independent chains of nested equivalence relations.

Keywords

satisfiability, finite model property, hierarchical partitions, two-variable logic

1. Introduction

Hierarchically partitioned data are pervasive in modern computer systems: Geographical Information Services often organise geospatial information using progressively more detailed fields: country, region, state, and city. Similar hierarchies appear in numerous contexts: Data Storage (organised into folders, subfolders, and files), Network Management (addressing schemes like IPv4/IPv6 with subnet hierarchies), Dependency Maintenance (tools for tracking dependencies between modules, libraries, and services).

To model such hierarchical data, we consider domains in which elements are annotated with data values drawn from potentially infinite or very large domains. A key aspect is that these data values can be tested for equality at multiple levels of precision. This is naturally captured by a family of increasingly coarser equivalence relations: two elements are related by the k -th equivalence relation if the k -th level equality test holds between their associated data values.

In this work, we establish results on the decidability and complexity of the satisfiability problems for several variants of the Two-Variable Fragment of First-Order Logic, FO^2 , supporting such increasingly coarser equivalence relations. Our goal is to provide a logical framework that is expressive enough to model complex multi-level relationships while retaining desirable computational properties for reasoning. As most of the investigated logics do not enjoy the finite model property, their general and finite satisfiability problems differ. Our primary focus is on finite satisfiability.

The motivation for studying FO^2 stems from its good algorithmic and model-theoretic properties. FO^2 combines an $NEXPTIME$ -complete satisfiability problem and the exponential model property [1] with a reasonable expressive power. In particular, it embeds (via the so-called *standard translation*) many modal, temporal, and description logics (up to $\mathcal{ALC}IOH^{\cap, \neq}$). Also, it is the maximal fragment of First-Order Logic, in terms of number of variables, whose satisfiability problem is decidable, as already the Three-Variable Fragment is undecidable [2]. In the last few decades, FO^2 together with its variations have been extensively studied, and plenty of results have been obtained in various scenarios. All of this makes FO^2 often the first-choice option for various reasoning tasks.

 DL 2026: 39th International Workshop on Description Logics, July 17–19, 2026, Lisbon, Portugal

 contact.oskarfiuk@gmail.com (O. Fiuk); emanuel.kieronski@cs.uni.wroc.pl (E. Kieroński); vincent.michielini@u-bordeaux.fr (V. Michielini)

 0009-0006-1312-4899 (O. Fiuk); 0000-0002-8538-8221 (E. Kieroński); 0000-0002-1413-9316 (V. Michielini)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

2. Logics of Interest and Our Results

Formulas of Two-Variable Fragment of First-Order Logic, FO^2 , use only the variables x and y ; any number of unary and binary *common* relation symbols (i.e., with unconstrained interpretations); the equality symbol; and constant symbols, but no function symbols of positive arity. Extensions of FO^2 are denoted by listing the *special* symbols (i.e., those with constrained interpretations) in brackets, e.g., $\text{FO}^2[<, \mathcal{EQ}^\subseteq]$. This notation makes it explicit which semantic constraints are imposed.

Order on domain elements. Let $\mathcal{EQ}^\subseteq$ denote the family of special symbols $E_1, E_2, \dots$ whose interpretations are constrained to *nested* equivalence relations, that is for every $k \in \mathbb{N}$: (i) the interpretation of E_k is an equivalence relation, and (ii) the interpretation of E_{k+1} is *coarser* than that of E_k (i.e., $x E_k y \rightarrow x E_{k+1} y$). Let $<$ be a special symbol interpreted as a *strict* linear order on domain elements.¹

Our first considered logic is $\text{FO}^2[<, \mathcal{EQ}^\subseteq]$, supporting both a linear order on elements and hierarchical data values. Potential applications of $\text{FO}^2[<, \mathcal{EQ}^\subseteq]$ can be found in, e.g., temporal verification of multi-process systems. We provide a motivational example of enforcing isolation policy in an environment of processes, containers, and events.

System execution is modeled as a linear sequence of events ordered by time using the relation $<$. Events are generated by processes, and processes are grouped into containers. The equivalence relations E_1 and E_2 capture this hierarchy: $x E_1 y$ holds when events x and y come from the same process, and $x E_2 y$ when they come from possibly distinct processes yet running in the same container. These are naturally nested: $x E_1 y \rightarrow x E_2 y$.

Importantly, the domain of our model consists only of events: processes are implicitly represented as equivalence classes of the relation E_1 . That is, all events belonging to the same process form a single equivalence class of E_1 . (We assume that every process reports at least one event, e.g., *spawn event*, to ensure that it has a non-empty class.) Likewise, containers are implicitly represented as equivalence classes of E_2 . (We assume that in every container at least one process is running, e.g., *root process*.)

Process-level properties (e.g., *sandboxed*, *privileged*) are expressed via class-wise unary predicates. For example:

$$\forall x, y. x E_1 y \rightarrow (\text{sandboxed}(x) \leftrightarrow \text{sandboxed}(y))$$

This ensures that all events from the same process share the same *sandboxed* or *not sandboxed* label, even though the logic only quantifies over events.

Using the above described encoding, we can impose the following isolation policy: “A *sandboxed* process must not communicate with events outside its container unless an explicit grant was made before from a *privileged* process.” This policy can be formalised with a sentence:

$$\forall x_1. (\text{sandboxed}(x_1) \wedge \varphi_{\text{cross-cont-message}}(x_1)) \rightarrow (\exists y_1. y_1 < x_1 \wedge x_1 E_1 y_1 \wedge \varphi_{\text{permission-grant}}(y_1)),$$

where $\varphi_{\text{cross-cont-message}}(x_1)$ stands for the formula

$$\exists y_2. x_1 < y_2 \wedge \neg x_1 E_2 y_2 \wedge \text{message}(x_1, y_2),$$

and $\varphi_{\text{permission-grant}}(y_1)$ for

$$\exists x_3. \text{privileged}(x_3) \wedge x_3 < y_1 \wedge \text{grant}(x_3, y_1).$$

In the above formulas, we annotated variables as x_1, y_1, y_2, x_3 for readability to reflect their roles in different processes. However, just x and y , with reuse across quantifiers would suffice.

As stated above, all variables refer to events: x_1 is a *dispatch message* event from Process 1 (the *sender*) that initiates the communication, and y_2 is a *deliver message* event from Process 2 (the *receiver*) that receives the message from Process 1, as indicated by $\text{message}(x_1, y_2)$. The event y_1 is an earlier event from the same process as x_1 , representing the *grant acknowledge* event, marking the point at which Process 1 becomes aware of and is authorised to act on the granted permission. Finally, x_3 is a *grant authorise* event from Process 3 (a *privileged admin*) that issues the permission and notifies Process 1, as indicated by $\text{grant}(x_3, y_1)$.

¹We use also the derived non-strict linear order $\leq$; note then in $\text{FO}^2 \leq$ is definable from $<$ and vice-versa.

The intended temporal order of events $x_3 < y_1 < x_1 < y_2$ is enforced: the permission is issued before it is acknowledged ($x_3 < y_1$); the message is sent after the permission being acknowledged ($y_1 < x_1$) and before being delivered ($x_1 < y_2$).

Notice also that, to detect that Process 1 and Process 2 are running in distinct containers, we refer to their representative events using $\neg x E_2 y$. Since E_1 is nested within E_2 , it follows that if two events are not E_2 -related, then their respective E_1 -classes (i.e., processes) must also belong to distinct E_2 -classes (i.e., containers).

Our first main contribution is establishing the complexity of $\text{FO}^2[<, \mathcal{EQ}^\subseteq]$:

Theorem 1. *Finitely satisfiable sentences of $\text{FO}^2[<, \mathcal{EQ}^\subseteq]$ admit models of exponential size. The finite satisfiability problem for $\text{FO}^2[<, \mathcal{EQ}^\subseteq]$ is NEXPTIME -complete.*

Notice that $\text{FO}^2[<, \mathcal{EQ}^\subseteq]$ does not include the induced successor of $<$.² It is shown in [3] that adding it leads to undecidability (cf. Section 4).

Order on data values. We also consider a different way to incorporate a linear order into our scenario. We trade an order on domain elements for a family of nested linear orders on data values, i.e., on equivalence classes of $E_1, E_2, \dots$.

Let $\preceq^\subseteq$ denote the family of special symbols $\preceq_1, \preceq_2, \dots$ whose interpretations are constrained to *nested* total preorders, that is for each $k \in \mathbb{N}$: (i) the interpretation of $\preceq_k$ is a total preorder³, and (ii) the interpretation of $\preceq_k$ is a subrelation of the interpretation of $\preceq_{k+1}$ (i.e. $x \preceq_k y \rightarrow x \preceq_{k+1} y$).

For example, we can define (over $\mathbb{N}$): $n \preceq_k m$ iff $\lfloor \frac{n}{10^k} \rfloor \leq \lfloor \frac{m}{10^k} \rfloor$.

Our second logic is $\text{FO}^2[\preceq^\subseteq]$, supporting hierarchical data values that can be compared at multiple levels of granularity using less-than, equal, and greater-than comparison tests. Nested equivalence relations $E_1 \subseteq E_2 \subseteq \dots$ are definable in this formalism: $x E_k y \leftrightarrow x \preceq_k y \wedge y \preceq_k x$. We establish:

Theorem 2. *Finitely satisfiable sentences of $\text{FO}^2[\preceq^\subseteq]$ admit models of exponential size. The finite satisfiability problem for $\text{FO}^2[\preceq^\subseteq]$ is NEXPTIME -complete.*

Further, we consider the logic $\text{FO}^2[\preceq_{\text{succ}}^\subseteq]$ that enriches the syntax of $\text{FO}^2[\preceq^\subseteq]$ by adding, for each $\preceq_k$ -symbol, its induced successor predicate $\mathcal{S}_k$, defined as $\mathcal{S}_k(x, y) := x \prec_k y \wedge \forall z. (x \prec_k z \rightarrow y \preceq_k z)$.

The second main contribution of this paper is:

Theorem 3. *Finitely satisfiable sentences of $\text{FO}^2[\preceq_{\text{succ}}^\subseteq]$ admit models of doubly exponential size. The finite satisfiability problem for $\text{FO}^2[\preceq_{\text{succ}}^\subseteq]$ is EXPSPACE -complete.*

Absence of orders. Both $\text{FO}^2[<, \mathcal{EQ}^\subseteq]$ and $\text{FO}^2[\preceq^\subseteq]$ admit sentences that enforce infinite models (e.g., $\forall x. \exists y. x < y$ and $\forall x. \exists y. x \prec_1 y$). This contrasts with pure FO^2 which enjoys the *finite model property* (every satisfiable sentence has a finite model). A natural question arises: Can nested equivalence relations alone, without a linear order, enforce infinite models?

Let $\text{FO}^2[\mathcal{EQ}^\subseteq]$ denote FO^2 extended with the family of nested equivalence relations $E_1 \subseteq E_2 \subseteq \dots$. We answer that $\text{FO}^2[\mathcal{EQ}^\subseteq]$ does indeed enjoy the finite model property:

Theorem 4. *Satisfiable sentences of $\text{FO}^2[\mathcal{EQ}^\subseteq]$ admit models of exponential size. The satisfiability and finite satisfiability problems for $\text{FO}^2[\mathcal{EQ}^\subseteq]$ coincide and are NEXPTIME -complete.*

Undecidability. A natural candidate to explore next is a logic supporting two independent families of nested equivalence relations $E_1 \subseteq E_2 \subseteq \dots$ and $F_1 \subseteq F_2 \subseteq \dots$. However, in this extension, we obtained a negative result, holding already for chains of depth two:

Theorem 5. *The satisfiability and finite satisfiability problems are undecidable for the constant-free, equality-free, monadic fragment of FO^2 extended with four special symbols E_1, E_2, F_1, F_2 , interpreted as equivalence relations such that E_2 is coarser than E_1 and F_2 is coarser than F_1 .*

²Notice that the induced successor is definable in FO , yet not in FO^2 , by the formula $\varphi(x, y) := x < y \wedge \forall z. (x < z \rightarrow y \leq z)$.

³A *total preorder* is a transitive relation $\preceq$ such that, for every x, y , either $x \preceq y$ or $y \preceq x$ holds (in particular, $x \preceq x$ holds).

3. Proof techniques

As typical for two-variable logics in our proofs we work with *Scott Normal Form* [4]. An FO^2 -sentence φ is in normal form if it is in the shape of $\forall x, y. \psi_0(x, y) \wedge \bigwedge_{m=1}^M \forall x. \exists y. \psi_m(x, y)$, for some $M \in \mathbb{N}$. For $m \in [0, M]$, the formula $\psi_m(x, y)$ is quantifier-free, constant-free, and uses only relation symbols of arity 1 and 2. There is a standard polynomial-time procedure transforming any FO^2 sentence into Scott Normal Form, preserving satisfiability over the same domains; see e.g., [5]. It is readily verified that it properly handles linear orders and equivalences and hence can be used in our scenarios.

Let us sketch the proofs of our theorems.

Theorem 1. Let φ be a normal form sentence. We assume that the equivalences it mentions are $E_1, \dots, E_K$, for some $K \in \mathbb{N}$; for technical convenience we use E_0 and E_{K+1} , assuming that they are interpreted as the identity, and, respectively, the universal relation true everywhere. The main ingredient of the proof is the lemma, which allows us to replace a single E_l -equivalence class $\mathcal{C}$ ($l = 1, \dots, K + 1$) in any model $\mathfrak{A} \models \varphi$, with its small counterpart, built out of some selected E_{l-1} -equivalence classes it contains, without touching the internal structure of those selected classes and the external part of $\mathfrak{A}$. The technique we use here is somehow similar to that used in the proof of the small substructure property for FO^2 [6]. We first select some small set of elements W_I of $\mathcal{C}$ that is sufficiently rich to serve as a source of witnesses for the $\forall\exists$ -requirements of φ , for all elements in A , whenever such witnesses are needed in $\mathcal{C}$. This set W_I consists of some number (linear in M) of maximal and minimal (with respect to the linear order) realizations of every 1-type in $\mathcal{C}$. Then we select a set of witnesses W_{II} in $\mathcal{C}$ for the elements of W_I , and W_{III} for the elements of W_{II} . We retain the E_{l-1} -classes of the elements from the selected sets and remove all the other E_{l-1} -classes from $\mathcal{C}$. Finally, to complete the definition of the new structure $\mathfrak{B}$, we modify the connections of its every element $a \notin W_I \cup W_{II}$ with the elements of W_I , in such a way that a has all the required witnesses; in this process we do not change the way the elements are connected by the linear order and the equivalence relations. Our strategy ensures that the 2-type of any pair of elements in $\mathfrak{B}$ is either retained from $\mathfrak{A}$ or copied from some other pair from $\mathfrak{A}$, so the $\forall x, y. \psi_0$ holds in $\mathfrak{B}$, so $\mathfrak{B} \models \varphi$.

Now, to get a small model for a given finitely satisfiable normal form sentence φ , we start from its any finite model and successively decrease its E_l -equivalence classes, proceeding from $l = 1$ to $l = K + 1$. Careful calculations show that the resulting model is indeed of size exponential in the length of φ .

Theorem 2. We proceed via a reduction to $\text{FO}^2[<, \mathcal{EQ}^\subseteq]$. Let $\mathfrak{A}_0$ be a finite $\text{FO}^2[\preceq^\subseteq]$ -model of φ . We expand it to an $\text{FO}^2[<, \mathcal{EQ}^\subseteq]$ -structure $\mathfrak{A}$ by: (i) setting $E_k^{\mathfrak{A}} := \preceq_k^{\mathfrak{A}_0} \cap (\preceq_k^{\mathfrak{A}_0})^{-1}$ for all k ; and (ii) interpreting $<^{\mathfrak{A}}$ as a linear order derived from $\preceq_1^{\mathfrak{A}_0}$ by resolving ties within E_1 -classes arbitrarily. We then apply the lemma from the proof of Thm. 1, replacing a single equivalence class $\mathcal{C}$ by its small counterpart $\mathcal{D}$ which yields a model $\mathfrak{B}$ over the domain $B = (A \setminus \mathcal{C}) \cup \mathcal{D}$ with $\mathcal{D} \subseteq \mathcal{C}$. It can be checked that $\preceq_k^{\mathfrak{B}} = \preceq_k^{\mathfrak{A} \upharpoonright B}$ for all k . Since nested preorders are preserved under taking substructures, $\mathfrak{B}$ respects the semantics of $\preceq_k$ -symbols, and thus it can be naturally transformed back into an $\text{FO}^2[\preceq^\subseteq]$ -model of φ . Since the size bound analysis from the proof of Theorem 1 still applies, we get a small model and thereby establish Theorem 2.

Theorem 3. This is technically the most involved of our proofs. The main argument is similar to the pumping lemma for regular languages. See the full version of the paper for the paper for details.

Theorem 4. The proof proceeds analogously to that of Theorem 1, but now starting from a possibly infinite model $\mathfrak{A}$. Again it relies on the Lemma decreasing a single equivalence class. In the absence of the linear order, we may choose arbitrary realisations of 1-types into the set W_I (rather than the maximal and minimal ones). The rest of the reasoning carries over directly, with the only modification being that, due to the possible infinity of $\mathfrak{A}$, it is sometimes necessary to work with natural limit structures.

Theorem 5. The undecidability is proven via a reduction from the halting problem for *two-counter machines*. This proof is similar in spirit to Pratt-Hartmann's proof of the undecidability of FO^2 with counting and two equivalence relations [7].

4. Related Work

Data Words. FO^2 with a linear order and an equivalence relation has been studied extensively in the context of *data words*. These are structures interpreting unary predicates, an equivalence relation, a linear order, and its induced successor relation. Satisfiability is decidable but non-elementary-hard [8]. When the successor relation is omitted, the problem becomes NEXPTIME -complete.

Quite close to our setting are data words with nested equivalences, that were considered in [3]. Satisfiability is decidable when the linear order is accessible only via its successor relation, and becomes undecidable as soon as both the order and its successor are accessible. The variant where only the linear order (and not its induced successor relation) is available, was not explored in that work.

Compared to our setting the main difference is that data words restrict the common part of the signature to unary symbols, whereas we allow full FO^2 with arbitrary binary relations that may freely interact with the equivalences and with one another (cf. the example in Section 2 where message and grant are common binary symbols)

Two-variable logic with equivalence relations or a linear order. FO^2 with common binary symbols and equivalences (not nested) or a linear order is well understood. With one equivalence it has the exponential model property and NEXPTIME -complete satisfiability [6]. With two equivalences (not nested), the finite model property is lost; both satisfiability and finite satisfiability are 2-NEXPTIME -complete [9]. Since FO^2 can express containment between equivalence relations, this yields decidability of FO^2 over hierarchical partitions of depth two (with no ordering). With three equivalence relations, FO^2 is undecidable [6]. The satisfiability and finite satisfiability problems for FO^2 with a linear order (without any equivalences) was shown to be NEXPTIME -complete [10].

FO^2 with total preorders. The EXPSpace -completeness of finite satisfiability for FO^2 over structures with one total preorder, its induced successor relation, a linear order, and unary relations was established in [11] (no common binary relations are permitted). When two independent total preorders are available, satisfiability becomes undecidable. The case of a single total preorder combined with arbitrary binary relations, as well as settings with nested total preorders, have not been investigated so far.

Description logics. Nested equivalence relations can be simulated in expressive description logics with transitive roles ($\mathcal{S}$), inverse roles ($\mathcal{I}$), role hierarchies ($\mathcal{H}$), and the operator $\top$. Indeed, one can then express that a transitive role T is symmetric and reflexive, with the axioms $T \sqsubseteq T^{-1}$ and $\top \sqsubseteq \exists T.\top$, respectively. Role hierarchies can then specify a hierarchy of equivalences. While common relation symbols are available, interactions between binary relations are limited to role hierarchies, meaning one can only express containment between relations. Moreover, these logics do not include linear orderings.

Logics with many variables. Among decidable first-order fragments, the Unary Negation Fragment, UNFO [12], is particularly worth to mention in the context of our work. This logic restricts negation to subformulas with at most one free variable. Its extension capturing the description logic $\mathcal{SHOI}$ with the *Self* operator, denoted $\text{UNFO}+\mathcal{SOH}$, is decidable and 2-ExpTime -complete [13]. Remarkably, $\text{UNFO}+\mathcal{SOH}$ enables reasoning over arbitrarily many independent families of nested equivalence relations. Recently, decidability of the Guarded Fragment with a hierarchy of equivalences has been shown [14]; its complexity turns out to be non-elementary.

Acknowledgments

The first and second authors were supported by the Polish National Science Center grant No. 2021/41/B/ST6/00996; the third author was supported by the ANR AI Chair INTENDED (ANR-19-CHIA-0014). This paper is an extended abstract of the paper presented at KR 2025 [15].

Declaration on Generative AI

The author(s) did not employ any Generative AI tools.

References

- [1] E. Grädel, P. Kolaitis, M. Y. Vardi, On the decision problem for two-variable first-order logic, *Bulletin of Symbolic Logic* 3 (1997) 53–69.
- [2] A. Kahr, E. Moore, H. Wang, Entscheidungsproblem reduced to the $\forall\exists\forall$ case, *Proc. Nat. Acad. Sci. U.S.A.* 48 (1962) 365–377.
- [3] H. Björklund, M. Bojańczyk, Shuffle expressions and words with nested data, in: *Proceedings of the 32nd International Conference on Mathematical Foundations of Computer Science, MFCS’07*, 2007, p. 750–761.
- [4] D. Scott, A decision method for validity of sentences in two variables, *Journal Symbolic Logic* 27 (1962) 477.
- [5] I. Pratt-Hartmann, *Fragments of First-Order Logic*, Oxford Logic Guides, Oxford University Press, United Kingdom, 2023.
- [6] E. Kieroński, M. Otto, Small substructures and decidability issues for first-order logic with two variables., *Journal of Symbolic Logic* 77 (2012) 729–765.
- [7] I. Pratt-Hartmann, Logics with counting and equivalence, in: *Proceedings of the Joint Meeting of CSL and LICS, CSL-LICS ’14*, ACM, 2014, pp. 76:1–76:10.
- [8] M. Bojańczyk, C. David, A. Muscholl, T. Schwentick, L. Segoufin, Two-variable logic on data words, *ACM Trans. Comput. Log.* 12 (2011) 27.
- [9] E. Kieroński, J. Michaliszyn, I. Pratt-Hartmann, L. Tendera, Two-variable first-order logic with equivalence closure, *SIAM Journal of Computing* 43 (2014) 1012–1063.
- [10] M. Otto, Two-variable first-order logic over ordered domains, *Journal of Symbolic Logic* 66 (2001) 685–702.
- [11] T. Schwentick, T. Zeume, Two-variable logic with two order relations, *Logical Methods in Computer Science* 8 (2012).
- [12] B. ten Cate, L. Segoufin, Unary negation, *Logical Methods in Comp. Sc.* 9 (2013).
- [13] D. Danielski, E. Kieroński, Finite satisfiability of unary negation fragment with transitivity, in: *44th International Symposium on Mathematical Foundations of Computer Science, MFCS*, volume 138 of *LIPICs*, 2019, pp. 17:1–17:15.
- [14] O. Fiuk, The guarded fragment with nested equivalences, in: *Proceedings of the 41th Annual Symposium on Logic in Computer Science, LICS 2026*, 2026.
- [15] O. Fiuk, E. Kieronski, V. Michielini, Two-variable logic for hierarchically partitioned and ordered data, in: *Proceedings of the 22nd International Conference on Principles of Knowledge Representation and Reasoning, KR 2025*, 2025.