Two-Variable Logic with Counting and Data-Trees

Witold Charatonik¹, Ian Pratt-Hartmann², and Piotr Witkowski¹

¹Institute of Computer Science, University of Wrocław, Poland
{wch, pwit}@cs.uni.wroc.pl
²Wydział Matematyki Informatyki i Mechaniki, Uniwersytet Warszawski
Instytut Informatyki, Uniwersytet Opolski
School of Computer Science, University of Manchester
ipratt@cs.man.ac.uk

Abstract

We show NExpTime-completeness of the finite satisfiability problem for two-variable logic with counting quantifiers and data-trees. The logic allows access to trees by the mother-daughter relation; data-value comparison is modelled by an equivalence relation. Apart from these relations, the logic allows for arbitrarily many other binary and unary relations, and thus properly extends two logics considered earlier: two-variable logic with counting and trees (without equivalence) and two-variable logic with counting and equivalence (without trees). Our decision procedure is based on a novel reduction to the latter of these.

Keywords: logic, complexity, two-variable fragment, counting quantifier, tree, equivalence.

1 Introduction

1.1 Two-variable logics

The two-variable fragment of first-order logic, \( \mathcal{FO}^2 \), is the collection of first-order formulas (with equality) involving only two variables, say \( x \) and \( y \). Vocabularies of these formulas may contain relational symbols of arbitrary arity; but without losing expressive power one may confine attention to signatures of unary and binary predicates only. Thus, the considered models are are simply node- and edge- coloured directed (multi)graphs. The two-variable logic with counting, \( C^2 \), extends \( \mathcal{FO}^2 \) with counting quantifiers of the form \( \exists_{\leq C} \), \( \exists_{= C} \) and \( \exists_{\geq C} \), where \( C \) is a natural number. Variables may be re-used so that formulas express properties concerning whole ensembles of objects. For example,

\[
\forall x (\exists y (\text{red}(x, y) \land \exists x \text{ red}(y, x)) \rightarrow \exists_{\leq 3} y \text{ blue}(x, y))
\]

says that every vertex connected to another vertex by a chain of two red edges is connected to at most three vertices by a blue edge. The logic \( \mathcal{FO}^2 \) has the finite model property, and its satisfiability (= finite satisfiability) problem is
Although $C^2$ lacks the finite model property, its satisfiability and finite satisfiability problems remain in NExpTime [8, 18, 19].

Both $FO^2$ and $C^2$ can be extended by imposing restrictions on the interpretations of certain distinguished binary predicates. Of particular interest are extensions of $FO^2$ and $C^2$ in which certain binary predicates are constrained to be interpreted as the graphs of words (finite, linear orders) or of trees. The finite satisfiability problem for $FO^2$ remains decidable when up to two (but not three) binary predicates are interpreted as linear orders; moreover, of these two linear orders, at most one may have a successor relation [7, 14, 15, 17, 21, 23].

Corresponding extensions of $C^2$ are also possible, though, as might be expected, decidability of finite satisfiability is harder to retain. When no additional binary relations are present, the finite satisfiability problem for $C^2$ with one linear order and its induced successor is NExpTime-complete; however, this increases to VASS-complete (hence non-elementary) when additional binary predicates are allowed [5]. Regarding trees, the available results are more fragmentary. The satisfiability problem for $FO^2$ interpreted over domains assumed to be (finite) trees, and in which the only available binary predicates are the ‘navigational’ relations (mother-daughter, next-sister, descendant, younger-sister) is investigated in [1]: the problem is decidable in all cases, with complexity depending on the available palette of navigational predicates. It was shown in [4] that the finite satisfiability problem for the full logic $C^2$ (with any number of unary and binary predicates) remains decidable—and indeed in NExpTime—when up to two distinguished predicates are constrained to be interpreted as the daughter relation in a pair of trees. It is not known whether this problem remains decidable for three or more trees.

When using logics to describe words and trees, it is useful to model the situation where the vertices of these data-structures contain data values from some (potentially infinite) domain. We speak in this case of data words and data-trees, respectively. If the only available mode of comparison is equality, we may model the comparison operation by means of a binary predicate interpreted as an equivalence relation: vertices are equivalent when they contain the same data value. Thus, for example, $FO^2$ interpreted over data-words (with just three binary predicates denoting the next relation, the right-of relation and data-equivalence) has a decidable satisfiability problem [3]. It is also known that $FO^2$ interpreted over data-trees (with just three binary predicates denoting the daughter and next-sister and data-equivalence) has satisfiability problem in 3-NExpTime; however, when the full palette of tree-navigation predicates is available, the corresponding problem becomes at least as hard as reachability in Branching Vector Addition Systems (BVASS), the decidability status of which is unknown.

The results of the previous paragraphs raise the question of what happens when the full logic $C^2$, is interpreted over data-trees, but with a limited range of navigational predicates. More precisely, we denote by $C^{2\text{1D1E}}$ the logic whose formulas are the same as $C^2$, but in which one distinguished binary predicate is required to be interpreted as the mother-daughter relation in a forest (a disjoint union of trees), and another distinguished binary predicate is required to be interpreted as an equivalence relation. When considering this logic, we assume all structures to be finite. Since a satisfiable formula of $C^{2\text{1D1E}}$ is by definition satisfiable over a finite structure we may drop the qualifier “finite”, and simply speak of the satisfiability problem for $C^{2\text{1D1E}}$. 

2
We show that the satisfiability problem for $C^21D1E$ remains NExpTime-complete. This is closely related to (but does not quite properly extend) earlier results [4, 6], which establish that the (finite) satisfiability for $C^2$ with two trees (without equivalence) is in NExpTime. It is also related to, but again does not properly extend, the result in [2] mentioned above, which establishes that the satisfiability problem for first-order monadic two-variable logic interpreted over a finite tree with one equivalence relation and with predicates for both the mother-daughter and next-sister relations is in 3-NExpTime.

We mention in passing that the decidability and complexity of the satisfiability and finite satisfiability problems for extensions of $FO^2$ and $C^2$ with distinguished predicates interpreted as equivalence relations over general node- and edge-coloured directed multigraphs (not words or trees) is known in all cases. Thus, the satisfiability and finite satisfiability problems for $FO^2$ remain decidable when up to two binary predicates are required to be interpreted as equivalence relations (or indeed, as equivalence-closures) [11, 12, 13], but become undecidable in the presence of three equivalence relations [10]. By contrast, denote by $C^21E$ the logic $C^2$ in which a single distinguished binary predicate is required to be interpreted as an equivalence relation. The satisfiability and finite satisfiability problems for $C^2$ remain decidable (in NExpTime); however the corresponding problems for $C^2$ in the presence of two equivalence relations are both undecidable [20].

1.2 Overview of the decision procedure

Our approach to $C^21D1E$ parallels that adopted in [4, 6] for the logic $C^21D$ (i.e. the same logic, but without the equivalence relation). There we proceeded as follows. Suppose we wish to know whether a given $C^2$-formula $\varphi$ has a finite model in which the predicate $t$ is interpreted as a forest. (More precisely: $t$ denotes the relation holding between any non-root vertex in some forest and its mother.) We cannot express this latter condition in first-order logic; however, we can write down a $C^2$-formula saying that $t$ is irreflexive and anti-symmetric, and that every element is related to at most one other element by $t$. Any finite model of this formula is a graph each of whose connected components is either a tree or a single $t$-cycle with some number of trees attached. It is shown in [4] that, under certain technical conditions (that are $C^2$-expressible), each of the cyclic components of the graph of $t$ may be prized apart and spliced into one of the tree-components; moreover, this model-surgery does not affect the truth of the formula $\varphi$. In this way, from our original formula $\varphi$, we compute a $C^2$-formula $\varphi^*$ such that $\varphi$ has a (finite) model in which $t$ is interpreted as a forest if and only if $\varphi^*$ has any finite model at all. Since finite satisfiability in $C^2$ is decidable, this solves our problem.

In the present paper, we reduce the satisfiability problem for $C^21D1E$ to the finite satisfiability problem for $C^21E$, i.e. $C^2$ with a single equivalence relation. From a given $C^21D1E$-formula $\varphi$, we compute a $C^21E$-formula $\varphi^*$ such that $\varphi$ has a (finite) model in which $t$ is interpreted as a forest and $e$ as an equivalence relation if and only if $\varphi^*$ has a finite model in which $e$ is interpreted as an equivalence relation. Since finite satisfiability in $C^21E$ is decidable, this solves our problem. The presence of an equivalence relation presents us with a series of challenges, however. The first is to ensure that the process of cycle-removal does not disturb the equivalence classes defined by $e$. We meet this challenge
in Sec. 3 by defining the properties of *galactic validity* and *cosmic validity*; the former allows us to eliminate t-cycles lying entirely within a single equivalence class; the latter allows us to remove those spanning several equivalence classes. The second challenge is to show how both galactic and cosmic validity can be secured by the realization of certain configurations of elements occurring in the structure being considered. We meet this challenge in Sec. 4 by introducing the notions of *galactic shrubbery* (for galactic validity) and *cosmic shrubbery* (for cosmic validity). The essential difficulty here is to bound the sizes of these configurations; we overcome it by—in effect—running the cycle-removal process in *reverse*. The third challenge is to write appropriate \( C^{21}E \)-sentences guaranteeing the realization of such shrubberies (galactic and cosmic) within structures. We meet this challenge in Sec. 6 in a series of technical lemmas. The principal difficulty here is to write a succinct formula ensuring that every equivalence class realizes an (exponential-sized) galactic shrubbery. Overcoming this difficulty requires a re-analysis of the finite satisfiability procedure for \( C^{21}E \) that yields new, parametrized, complexity bounds for that logic.

2 Preliminaries

2.1 Structures, arboreal structures and dendral structures

A *forest* is a finite, simple, acyclic, directed graph \( G = (V, E) \) with out-degree at most 1. If \((u, v) \in E\) is an edge, we call \( u \) a daughter of \( v \) and \( v \) the mother of \( u \). (Thus, in this paper, edges point from daughter to mother.) That \( G \) is acyclic is taken to imply in particular that the relation \( E \) is irreflexive and antisymmetric. A *root* of \( G \) is any vertex which has no mother (i.e. has out-degree 0); if \( V \) is non-empty, there is at least one root, and indeed every connected component of \( G \) contains a unique root. A *tree* is a connected forest, i.e. a forest with at most one root. Note that \( G \) may contain isolated vertices, lying on no edges at all; these are, by definition, roots. If \( G = (V, E) \) is a forest and \( U \subseteq V \), then the restriction of \( G \) to \( U \) is also a forest. Where \( U \) is clear from context, a vertex \( u \in U \) is a local root if it is a root of the restriction of \( G \) to \( U \), that is, if it is either a root of \( G \), or its mother is not in \( U \).

In the sequel, we consider signatures consisting only of unary and binary predicates that contain the distinguished binary predicates \( e \), \( t \), and the distinguished unary predicate \( r \). We fix some conventions and terminology regarding these predicates. We consider exclusively *finite* structures in which \( e \) is interpreted as an *equivalence relation*; henceforth, then, the word “structure” will always be silently subject to these restrictions. Say that an *arboreal structure*, or simply an *a-structure*, is a structure in which, in addition, \( t \) is interpreted as an irreflexive, asymmetric relation in which every element has at most one successor, and \( r \) is interpreted as holding of precisely those elements which have no \( t \)-successors; we call any element of an a-structure satisfying \( r \) a root.

If \( \mathfrak{A} \) is an a-structure, it is not necessarily true that the graph of \( t \) is a forest. (Indeed, we cannot define the class of such structures with any first-order sentence.) However, a moment’s thought shows that, in any a-structure, every component of the graph of \( t \) is either a tree or contains a (directed) cycle. The elements contained in tree-components are precisely those elements which have a directed path to a root; we call such elements *dendral*. If an element
does not lie on any t-edge, it is by definition a root and hence dendral. We call an element cyclic if it lies on a (directed) cycle in the graph of t. Say that a dendral structure, or simply d-structure, is an a-structure in which the graph of t contains no cyclic elements, or equivalently, one in which the graph of t is a forest, or equivalently again, one in which every element is dendral. Obviously, if A is a d-structure, the elements satisfying r are exactly the roots (in the usual sense) of the trees formed by the components of the graph of t. Finally, suppose A is an a-structure and U ⊆ A. If U is clear from context, we say that an element u ∈ U is a locally dendral if it is dendral in ΣU—that is, if there is a directed t-path in U to a local root.

2.2 The logics C^21E and C^21D1E

The two-variable fragment with counting is the set of first-order formulas over a signature of unary and binary predicates featuring only the logical variables x and y, but with the counting quantifiers ∃≥C, ∃≤C and ∃=C allowed, where C is any (binary string encoding a) non-negative integer. We generally equivocate between integers and the binary strings encoding them, it being understood that, when measuring the size |ϕ| of a formula ϕ, a counting subscript C > 0 contributes at least ⌈log C⌉ + 1 symbols. The semantics is as expected. The set of formulas comprising this logic is standardly denoted C^2; however, given our assumption that t is always interpreted as an equivalence relation, we shall prefer the designation C^21E. By C^21D1E we mean the logic whose formulas are the same as those of C^21E, but subject to the additional restriction that t is interpreted as a forest, and r as the set of its roots. In other words, C^21D1E is C^21E restricted to d-structures.

Henceforth, formula means a formula of C^21D1E (equivalently, of C^21E or of C^2). A formula ϕ is in normal form if it conforms to the pattern:

∀x∀y(x = y ∨ α) ∧ \bigwedge_{h=1}^{m} ∀x\exists[◁_hC_h]y(β_h ∧ x ≠ y),  \tag{1}

where α and the β_h are quantifier-free, equality-free formulas, m is a positive integer, the ◁_h are symbols chosen from {=, ≤}, and the C_h are non-negative integers. If all of the comparisons ◁_h are equalities, we say that the formula is in strict normal form. The quantifier-free formula θ given by β_1 ∨ ⋯ ∨ β_m is called the modulus of ϕ and the number C = C_1 + ⋯ + C_m its amplitude. We shall sometimes have occasion to refer to the parameter m as the multiplicity of ϕ. We assume without essential loss of generality that C ≥ 1.

Lemma 1. Let ϕ be a C^21D1E-formula. We can compute, in time bounded by a polynomial function of |ϕ|, a strict normal-form C^21D1E-sentence ϕ′ of the form (1), such that: (i) ⊨ ϕ′ → ϕ; and (ii) if ϕ is satisfiable over a domain A of cardinality greater than C_{max} = \max_{h=1}^{m} C_h, then so is ϕ′.

Lemma 1 allows us to confine attention to (strict) normal-form C^21D1E-sentences. We take as our starting point a known decision procedure for the logic C^21E. The following result is shown in [20].

Theorem 1. There is a non-deterministic procedure which, given a normal-form C^21E-formula ϕ, has a successful run if and only if ϕ is finitely satisfiable, and which terminates in time bounded by a fixed exponential function of |ϕ|. 

5
Actually [20] assumes that \( \phi \) is in strict normal form, i.e., all of the comparisons \( \varphi_i \) are equalities. However, the proof given there works unchanged if some (or even all) of the \( \varphi_i \) are allowed to be \( \leq \). In this article, we require the slightly more general formulation given here. We shall strengthen Theorem 1 further in the sequel.

2.3 Local configurations in structures

We make extensive use of the notions of (atomic) 1- and 2-types. Let \( \Sigma \) be a signature of unary and binary predicates. A 1-type is a maximal consistent set of literals over \( \Sigma \) involving only the variable \( x \). Likewise, a 2-type is a maximal consistent set of literals over \( \Sigma \) involving only the variables \( x \) and \( y \) and containing the literal \( x \neq y \). In both cases, we take the notion of consistency to incorporate the constraint that the distinguished predicate \( \epsilon \) is interpreted as a reflexive, symmetric and transitive relation. If \( \tau \) is a 2-type, we denote by \( \tau^{-1} \) the 2-type obtained by exchanging the variables \( x \) and \( y \) in \( \tau \), and call \( \tau^{-1} \) the inverse of \( \tau \). We denote by \( \operatorname{tp}_1(\tau) \) the 1-type obtained by removing from \( \tau \) any literals containing \( y \); and we denote by \( \operatorname{tp}_2(\tau) \) the 1-type obtained by first removing from \( \tau \) any literals containing \( x \), and then replacing all occurrences of \( y \) by \( x \). Evidently, \( \operatorname{tp}_2(\tau) = \operatorname{tp}_1(\tau^{-1}) \). We equivocate freely between finite sets of formulas and their conjunctions; thus, we treat 1-types and 2-types as formulas, where convenient. Let \( \mathfrak{A} \) be any structure interpreting \( \Sigma \). If \( a \in A \), then there exists a unique 1-type \( \pi \) such that \( \mathfrak{A} \models \pi[a] \); we denote \( \pi \) by \( \operatorname{tp}^\mathfrak{A}[a] \) and say that \( \pi \) realizes \( \pi \). If, in addition, \( b \in A \setminus \{a\} \), then there exists a unique 2-type \( \tau \) such that \( \mathfrak{A} \models \tau[a,b] \); we denote \( \tau \) by \( \operatorname{tp}^\mathfrak{A}[a,b] \) and say that the pair \( a,b \) realizes \( \tau \). Evidently, in that case, \( \tau^{-1} = \operatorname{tp}^\mathfrak{A}[b,a] \); \( \operatorname{tp}_1(\tau) = \operatorname{tp}^\mathfrak{A}[a] \); and \( \operatorname{tp}_2(\tau) = \operatorname{tp}^\mathfrak{A}[b] \) (see Fig. 1a).

Let \( \theta \) be a quantifier-free formula over \( \Sigma \). A \( \theta \)-ray-type is a 2-type \( \rho \) such that \( \models \rho \rightarrow \theta \). If \( \mathfrak{A} \models \rho[a,b] \) for distinct elements \( a, b \), then we say that the pair \( \langle a,b \rangle \) is a \( \theta \)-ray. We call a \( \theta \)-ray-type \( \rho \) invertible if \( \rho^{-1} \) is also a \( \theta \)-ray-type; and we call any 2-type that is not a \( \theta \)-ray-type, \( \theta \)-dark. Additionally, if \( \models \rho \rightarrow \epsilon(x,y) \), we call \( \rho \) galactic; otherwise, cosmic. If \( \mathfrak{A} \) is a structure interpreting \( \Sigma \), we say that the \( \theta \)-luminosity of \( \mathfrak{A} \) is the supremum of the cardinalities of the sets \( \{b \in A \setminus \{a\} : \mathfrak{A} \models \theta[a,b]\} \) as \( a \) ranges over \( A \). That is, the \( \theta \)-luminosity of \( \mathfrak{A} \) is the supremum of the number of \( \theta \)-rays sent by the elements of \( \mathfrak{A} \). If \( \varphi \) is a normal-form \( C^2\epsilon \)-formula with modulus \( \theta \) and amplitude \( C \), then since all the comparisons \( \varphi_i \) are either \( = \) or \( \leq \), \( \mathfrak{A} \models \varphi \) entails that the \( \theta \)-luminosity of \( \mathfrak{A} \) is bounded by \( C \).

We now construct apparatus for describing the ‘local environment’ of elements in structures. Let \( \theta \) be a quantifier-free, formula over \( \Sigma \), and let the \( \theta \)-ray-types be listed in some fixed order (depending on \( \Sigma \) and \( \theta \)) as \( \rho_1, \ldots, \rho_J \). A \( \theta \)-star-type is a \( (J+1) \)-tuple \( \sigma = (\pi, v_1, \ldots, v_J) \), where \( \pi \) is a 1-type over \( \Sigma \) and the \( v_j \) are non-negative integers such that \( v_j \neq 0 \) implies \( \operatorname{tp}_1(v_j) = \pi \) for all \( j \) (\( 1 \leq j \leq J \)). We denote the 1-type \( \pi \) by \( \operatorname{tp}(\sigma) \). To motivate this terminology, suppose \( \mathfrak{A} \) is a structure interpreting \( \Sigma \). For any \( a \in A \), we define

\[
\operatorname{st}_\theta^\mathfrak{A}[a] = \langle \operatorname{tp}^\mathfrak{A}[a], v_1, \ldots, v_J \rangle,
\]

where \( v_j = |\{b \in A : b \neq a \text{ and } \operatorname{tp}^\mathfrak{A}[a,b] = \rho_j\}| \). Evidently, \( \operatorname{st}_\theta^\mathfrak{A}[a] \) is a \( \theta \)-star-type; we call it the \( \theta \)-star-type of \( a \) in \( \mathfrak{A} \), and say that \( a \) realizes \( \operatorname{st}_\theta^\mathfrak{A}[a] \).
\[ \text{tp}_A[a] = \text{tp}_1(\tau) \]

\[ \text{tp}_A[b] = \text{tp}_2(\tau) \]

(a) 2-type

\[ (a) \text{ 2-type} \]

\[ \text{tp}_A[a] = \text{tp}_1(\tau) \]

\[ \text{tp}_A[b] = \text{tp}_2(\tau) \]

(b) Star type

\[ (b) \text{ Star type} \]

Figure 1: An element \( a \) connected by a 2-type \( \tau \) to an element \( b \) in a structure \( A \); and a star type \( \langle \pi, v_1, v_2, \ldots, v_J \rangle \), emitting \( v_j \) rays of type \( \rho_j \) for all \( j \) \((1 \leq j \leq J)\).

Intuitively, the \( \theta \)-star-type of an element records the number of \( \theta \)-rays of each type emitted by that element. It helps to think, informally, of a \( \theta \)-star-type \( \sigma \) as emitting a collection of \( \theta \)-rays of various types (see Fig. 1b).

To understand the significance of \( \theta \)-star-types, consider the formula \( \phi \) given in (1), and again let \( \theta := \beta_1 \lor \cdots \lor \beta_m \). Evidently, if \( \mathfrak{A} \models \phi \) and \( \mathfrak{B} \) is a structure interpreting the same signature, and realizing the same set of 2-types and the same set of \( \theta \)-star-types as \( \mathfrak{A} \), then \( \mathfrak{B} \models \phi \). More formally, we say that a 2-type \( \tau \) is compatible with \( \phi \) if \( \tau \land \alpha \land \alpha(y, x) \) is consistent; and we say that a \( \theta \)-star-type \( \sigma \) given in (2) is compatible with \( \phi \) if (i) each of the ray-types \( \sigma \) emits is compatible with \( \phi \) and, (ii) for all \( h \) \((1 \leq h \leq m)\), \( \sigma \) emits either exactly or at most (depending on \( \triangleleft h \) being either \( = \) or \( \leq \)) \( C_h \) rays whose types entail \( \beta_h \), i.e.

\[
\sum_{j=1}^{m} \{ v_j \mid 1 \leq j \leq J \text{ and } \models \rho_j \rightarrow \beta_h \} \triangleleft h \ C_h.
\]

Thus, \( \mathfrak{A} \models \phi \) just in case all the \( \theta \)-star-types and all the 2-types realized in \( \mathfrak{A} \) are compatible with \( \phi \). Equivalently, \( \mathfrak{A} \models \phi \) just in case all the \( \theta \)-star-types and all the \( \theta \)-dark 2-types realized in \( \mathfrak{A} \) are compatible with \( \phi \).

In the above definitions, \( \theta \) may be any quantifier-free formula. For the most part, however, we shall be interested the case where \( \theta \) has \( t(x, y) \) as a disjunct, or, more generally, where \( t(x, y) \rightarrow \theta \). For this reason, we shall always refer to any 2-type containing the atom \( t(x, y) \) as a t-ray-type, and to any pair of (distinct) elements \( (a, b) \) satisfying \( t(x, y) \) in some structure \( \mathfrak{A} \) as a t-ray. Thus, for the most part, t-rays will be \( \theta \)-rays. A t-ray is dendral if the elements it involves are dendral, and cyclic if the elements it involves are cyclic, i.e. if it lies on a t-cycle. A non-dendral t-ray need not be cyclic; on the other hand, it must lie in a component of \( \mathfrak{C}^s \) which contains a cycle.

### 2.4 Weak normal form

We now come to the strengthening of Theorem 1 promised earlier. The strengthening is a routine extraction of a parametrized complexity bound from the original proof in [20], and requires only a very high-level re-analysis of that proof.
To establish Theorem 1, one starts with a $C^2\text{E}$-formula $\varphi$ in normal-form with modulus $\theta$ and amplitude $C$, and non-deterministically computes from $\varphi$ a system of linear Diophantine equations $E$, in variables, say, $z_1, \ldots, z_L$. The non-determinism here involves guessing a collection $\Delta$ of $\theta$-dark 2-types (i.e. 2-types that are not $\theta$-ray-types). The total size of $E$ is bounded by a function of the form $p\left(2^{2^{|\Sigma|}} + C\right)$, where $p$ is a polynomial; and $|\Delta|$ is certainly bounded by $2^{4^{|\Sigma|}}$. Intuitively, each variable $z_\ell$ ($1 \leq \ell \leq L$) corresponds to a finite multiset $M_\ell$ of $\theta$-star-types over $\Sigma$, itself involving at most $p\left(2^{2^{|\Sigma|}} + C\right)$ different $\theta$-star-types. Suppose that $z = (z_1, \ldots, z_L)$ is a solution of $E$; we say that a $\theta$-star-type $\sigma$ appears in $z$ if, for some $\ell$ ($1 \leq \ell \leq L$), $z_\ell > 0$ and $\sigma$ occurs with non-zero multiplicity in the multiset $M_\ell$. Clearly, we may list the $\theta$-star-types appearing in $z$ in some way, say $\sigma_1, \ldots, \sigma_K$. The core of the proof of Theorem 1 is an argument showing that, if $z$ is a solution of $E$, then we can take $z_\ell$ copies of the multiset $M_\ell$ of $\theta$-star-types, for each $\ell$ ($1 \leq \ell \leq L$), and assemble all the copies of the $\theta$-star-types involved into a well-defined finite structure $\mathfrak{A}$ such that: (i) these are the only $\theta$-star-types realized in $\mathfrak{A}$; and (ii) the only $\theta$-dark 2-types realized in $\mathfrak{A}$ are in $\Delta$. To check that $\mathfrak{A} \models \varphi$, we simply check that: (i) for all $k$ ($1 \leq k \leq K$), $\sigma_k$ is compatible with $\varphi$; and (ii) for every $\tau \in \Delta$, $\tau$ is compatible with $\mathfrak{A}$. We can determine whether $E$ has a solution $z$ in non-deterministic time bounded by a polynomial function of the size of $E$; we can check that all of the $\theta$-star-types $\sigma_1, \ldots, \sigma_K$ appearing in $z$ are compatible with $\varphi$ in time polynomial in $|\varphi| + K$; and we can check that all of the $\theta$-dark 2-types in $\Delta$ are compatible with $\varphi$ in time polynomial in $|\varphi| + |\Delta|$. This yields the following, more refined version of Theorem 1.

**Corollary 1.** There is a non-deterministic procedure which, given a normal-form $C^2\text{E}$-formula $\varphi$ over signature $\Sigma$ with modulus $\theta$ and amplitude $C$, has a successful run if and only if $\varphi$ is finitely satisfiable, and which terminates in time bounded by $p\left(2^{2^{|\Sigma|}} + C + |\varphi|\right)$, where $p$ is a fixed polynomial. Moreover, if $\varphi$ is finitely satisfiable, then it has a finite model in which the number of $\theta$-star-types is bounded by $f\left(2^{2^{|\Sigma|}} + C\right)$, where $f$ is a fixed polynomial.

The crucial point here is that, while the overall complexity bound of $\text{NExpTime}$ from Theorem 1 is unaffected, for a fixed signature $\Sigma$ and amplitude $C$, our (non-deterministic) procedure runs in time polynomial in $|\varphi|$. This has some important consequences. For as long as we know we are dealing with structures of $\theta$-luminosity at most $C$, we can impose additional conditions on the realized $\theta$-star-types without essentially changing the complexity of the decision procedure. The following details formalize this idea. Let $\varphi$ be a normal-form $C^2\text{E}$-formula with modulus $\theta$ and amplitude $C$. Say that a formula $\epsilon$, with $x$ as its only free variable, is $\theta$-eclipsed if it is a Boolean combination of formulas that are either (i) quantifier-free or (ii) of any of the forms $\exists_{\leq D} y \cdot \eta$, $\exists_{= D} y \cdot \eta$ or $\exists_{\geq D} y \cdot \eta$, where $\eta$ is quantifier-free and $\models \eta \rightarrow \theta$. A $C^2\text{E}$-formula $\psi$ is in weak normal form if it conforms to the pattern $\varphi \land \forall x. \epsilon$, where $\varphi$ is in normal form with modulus $\theta$, and $\epsilon$ is $\theta$-eclipsed. We take the modulus and amplitude of $\psi$ to be those of $\varphi$. As an example, the formula $\forall x \exists_{\leq N} (\epsilon(x, y) \land x \neq y)$ is in normal form (actually, strict normal form), with modulus $\theta = \epsilon(x, y)$ and amplitude $N$. Hence, if $\eta_0, \ldots, \eta_{N-1}$ are quantifier-free formulas and $p$ a unary
predicate, the sentence

$$\forall x \exists \eta \forall y \in \mathbb{N} \exists \epsilon (e(x, y) \land x \neq y) \land \forall x \left( p(x) \rightarrow \exists \eta \exists \epsilon (e(x, y) \land \eta) \right)$$

is in weak normal form, with the same modulus and amplitude, since $p(x)$ is quantifier-free, and each of the formulas $e(x, y) \land \eta$ trivially entails $\theta$. We remark that, if $\varphi_i \land \forall x.\epsilon_i$ is in weak normal form with modulus $\theta_i$ and amplitude $C_i$, for $i = 1, 2$, then the conjunction $\varphi_1 \land \varphi_2 \land \forall (\epsilon_1 \land \epsilon_2)$ is in weak normal form with modulus $\theta_1 \lor \theta_2$ and amplitude $C_1 + C_2$.

Suppose then $\psi = \varphi \land \forall x.\epsilon$ is in weak normal form, with modulus $\theta$ and amplitude $C$. Since $\psi$ entails $\varphi$, all $\theta$-star-types appearing in any model of $\psi$ are compatible with $\varphi$ (and hence $C$-bounded); moreover, all 2-types occurring in this model are also compatible with $\varphi$. In fact, to test whether $\psi$ is finitely satisfiable, we may construct the same system $\mathcal{E}$ of Diophantine equations, and check that the realized $\theta$-star-types are compatible with $\varphi$ and satisfy the additional conditions imposed by the conjunct $\forall x.\epsilon$. Since we know exactly how many $\theta$-rays of each type are emitted by each $\theta$-star-type, verification of this latter requirement involves only a straightforward check on each realized star-type. Thus, we can strengthen Corollary 1.

**Corollary 2.** There is a non-deterministic procedure which, given a weak normal-form $C^{2\Sigma}$-formula $\psi$ over signature $\Sigma$ with modulus $\theta$ and amplitude $C$, has a successful run if and only if $\psi$ is finitely satisfiable, and which terminates in time bounded by $p(2^{2|\Sigma|} + C + \|\psi\|)$, where $p$ is a fixed polynomial. Moreover, if $\psi$ is finitely satisfiable, then it has a finite model in which the number of $\theta$-star-types is bounded by $f(2^{2|\Sigma|} + C)$, where $f$ is a fixed polynomial.

The parametrized bounds in Corollary 2 are of particular note here. For fixed $\Sigma$ and $C$, the quantity $p(2^{2|\Sigma|} + C + \|\psi\|)$ appearing in the first statement is polynomial in $\|\psi\|$, while the quantity $f(2^{2|\Sigma|} + C)$ in the second statement is constant (i.e. does not depend on $\|\psi\|$ at all). We shall exploit both these facts in Sec. 7.

## 3 Switching rays in a-structures

Recall that all structures considered in this paper (are finite and) interpret the distinguished binary predicate $\tau$ as an equivalence relation, while all a-structures additionally interpret the distinguished binary predicate $t$ as an irreflexive, antisymmetric relation with out-degree at most 1, and the distinguished unary predicate $r$ as its set of roots (elements with out-degree 0). Where a structure $\mathfrak{A}$ is clear from context, we speak of elements $a, b \in A$ as being equivalent if $\mathfrak{A} \models \epsilon[a, b]$.

Let $\mathfrak{A}$ be an a-structure interpreting signature $\Sigma$ and $\theta$ a quantifier-free formula over $\Sigma$. Define the parental 1-type of an element of $\mathfrak{A}$ to be the 1-type of that element’s mother in the graph of $t$ (undefined if it has no mother). We say that $\mathfrak{A}$ is $\theta$-parental if no element of 1-type $\pi$ sends a $\theta$-ray to any element that is not one of its daughters, and whose parental 1-type is $\pi$. Equivalently, $\mathfrak{A}$ is $\theta$-parental if there do not exist distinct elements $a, b, c \in A$ such that $\mathfrak{A} \models \theta[a, b], \mathfrak{A} \models t[b, c]$ and $\text{tp}_\mathfrak{A}[a] = \text{tp}_\mathfrak{A}[c]$. 


Lemma 2. Let $\mathfrak{A}$ be an $a$-structure interpreting a signature $\Sigma$, and let $\theta$ be a quantifier- and equality-free formula over $\Sigma$. If $\mathfrak{A}$ has finite $\theta$-luminosity $C$, then $\mathfrak{A}$ can be expanded to a $\theta$-parental structure $\mathfrak{A}'$ interpreting $\Sigma$ together with $\lceil \log(2C+1) \rceil$ fresh unary predicates.

Proof. Let $G = (A, E)$ be the directed graph whose vertices are the elements of $A$ and whose edges are given by

$$\{ (a, c) \mid a \neq c, \text{tp}^A[a] = \text{tp}^A[c] \text{ and there exists } b \in A \text{ s.t. } \mathfrak{A} \models \theta[a, b] \text{ and } \mathfrak{A} \models \theta[b, c] \}. $$

This directed graph has out-degree at most $C$, and so the undirected graph admits a $(2C + 1)$-colouring. Now interpret the $\lceil \log(2C+1) \rceil$ fresh predicates to encode these colours, and let $\mathfrak{A}'$ be the resulting expansion of $\mathfrak{A}$. $\square$

In $\theta$-parental $a$-structures, it is possible to exchange pairs of invertible $t$-rays of identical type, provided that the elements involved satisfy certain equivalence-conditions. The following remarks formalize this idea. Let $\Sigma$ be a signature, $\theta$ a quantifier-free formula over $\Sigma$, and $\rho$ a $t$-ray-type over $\Sigma$. Let $\mathfrak{A}$ be a $\theta$-parental $a$-structure interpreting $\Sigma$, and suppose $a, b, c, d$ are distinct elements of $\mathfrak{A}$ such that $\langle a, b \rangle$ and $\langle c, d \rangle$ are $t$-rays of type $\rho$. It follows from the $\theta$-parental property of $\mathfrak{A}$ that $\langle b, c \rangle$ and $\langle d, a \rangle$ are not $\theta$-rays. We say that $\langle a, b \rangle$ and $\langle c, d \rangle$ are switchable if one of the following three conditions obtains: (i) $a$ and $c$ are equivalent in $\mathfrak{A}$; (ii) $b$ and $d$ are equivalent in $\mathfrak{A}$; or (iii) $a, b, c$ and $d$ are pairwise inequivalent in $\mathfrak{A}$. If $(a, b)$ and $(c, d)$ are switchable, we define the $a$-structure $\mathfrak{A}'$ to be exactly like $\mathfrak{A}$, except that:

$$\text{tp}^{\mathfrak{A}}[a, d] = \text{tp}^{\mathfrak{A}}[a, b] \quad \text{ and } \quad \text{tp}^{\mathfrak{A}}[c, d] = \text{tp}^{\mathfrak{A}}[c, b],$$

and we denote $\mathfrak{A}'$ by $\mathfrak{A}(a, b || c, d)$. The transformation is illustrated in Fig. 2. It is easy to see from this diagram that the interpretation of $\tau$ is not disturbed, and thus remains an equivalence relation. Moreover, $\mathfrak{A}'$ realizes exactly the same set of 2-types as $\mathfrak{A}$, and every element of $\mathfrak{A}'$ emits at most one $t$-ray; thus $\mathfrak{A}(a, b || c, d)$ is an $a$-structure. In fact, slightly more follows. Suppose the ray $\langle a, b \rangle$ is dendral in $\mathfrak{A}$, and $(c, d)$ cyclic, and let $C$ be the component of the graph of $t^3$ containing $c$ and $d$. Then all of the elements which were dendral in $\mathfrak{A}$, as well as all of the elements in $C$, will be dendral in $\mathfrak{A}(a, b || c, d)$. In effect, $C$ is spliced in to the tree containing $a$ and $b$, as illustrated in Fig 3.

Lemma 3. Let $\Sigma$ be a signature, $\theta$ a quantifier-free formula over $\Sigma$ such that $\models (x, y) \to \theta$, and $\rho$ a $t$-ray-type over $\Sigma$. Let $\mathfrak{A}$ be a $\theta$-parental $a$-structure interpreting $\Sigma$, and suppose $\langle a, b \rangle$ and $\langle c, d \rangle$ are switchable realizations of $\rho$.
in $\mathfrak{A}$. Let $\mathfrak{A}' = \mathfrak{A}(a, b\parallel c, d)$. Then: (i) the sets of 2-types realized in $\mathfrak{A}$ and $\mathfrak{A}'$ are identical; (ii) every element of $\mathfrak{A}$ realizes the same $\theta$-star-type in $\mathfrak{A}$ as in $\mathfrak{A}'$; and (iii) $\mathfrak{A}'$ is $\theta$-parental.

Proof. Statement (i) is immediate. For statement (ii), since $tp^\mathfrak{A}_a = tp^\mathfrak{A}_c$, $a \neq c$, and $\mathfrak{A}$ is $\theta$-parental, it follows that $b$ sends no $\theta$-ray to $c$, and $d$ sends no $\theta$-ray to $a$. It is then immediate that $a$, $b$, $c$ and $d$ have the same $\theta$-star-types in $\mathfrak{A}$ as in $\mathfrak{A}'$, and certainly the $\theta$-star-type of no other element is changed. For statement (iii), we must show that if $a_0$ in the structure $\mathfrak{A}'$ sends a $\theta$-ray to some element $b_0$ with mother $c_0 \neq a_0$, then $a_0$ and $c_0$ have different 1-types. Observe first that the parental 1-type of every element is the same in $\mathfrak{A}'$ as in $\mathfrak{A}$. This is because the only elements which do not have the same mother in $\mathfrak{A}$ and $\mathfrak{A}'$ are $a$ and $c$, and these exchange their mothers, which are of the same 1-type. Suppose then that $a_0 \notin \{a, b, c, d\}$; it follows that $tp^\mathfrak{A}_a[a_0] = tp^\mathfrak{A}_c[a_0]$. Since the parental 1-type of $b_0$ is the same in both structures, it follows from the fact that $\mathfrak{A}$ is parental that $tp^\mathfrak{A}_a[a_0] \neq tp^\mathfrak{A}_c[a_0]$. Suppose now that $a_0$ is one of $a$, $b$, or $c$. By assumption, $a$ and $c$ have the same 1-type, say $\pi$. But the only new elements that $a_0$ sends a $\theta$-ray to in $\mathfrak{A}'$ are either of $b$ or $d$; and we already know from the fact that $\mathfrak{A}$ is $\theta$-parental and $\models t(x, y) \to \theta$ that the mothers of $b$ and $d$ do not have 1-type $\pi$. The only remaining possibility is that $a_0$ is one of $b$ or $d$. Here, the only possible new $\theta$-rays occurring in $\mathfrak{A}'$ are of type $\rho^{-1}$ (assuming $\rho$ is invertible). But $\rho$ is a $t$-ray (i.e. points from daughter to mother), whence $c_0$ is not distinct from $a_0$.

We now introduce two notions which will play a crucial role in the removal of cycles from a-structures. The first is a ‘local’ property concerning equivalence classes in a-structures. Say that an a-structure is \textit{galactically valid} if, for every equivalence class $B$ in that structure, and every $t$-ray-type $\rho$ realized in $B$, $\rho$ is realized by a pair of locally dendral elements in $B$, i.e. by a pair of elements that are dendral in the induced sub-structure $\mathfrak{A}|B$. The notion of galactic validity is defined for a-structures generally; however, for d-structures, it is trivial.

\textbf{Lemma 4.} Every d-structure is galactically valid.

\textit{Proof.} By definition, in a d-structure, all elements are dendral, and hence certainly locally dendral.

The second of our two notions, which is more elaborate than the first, concerns larger-scale structures than mere galaxies. If $\mathfrak{A}$ is an a-structure interpreting $\Sigma$, say that a $t$-ray-type $\rho$ over $\Sigma$ is \textit{monotelic} in $\mathfrak{A}$ if there exists a pair of dendral elements $\langle a, b \rangle$ realizing $\rho$ in $\mathfrak{A}$, and either every realization $\langle c, d \rangle$ of $\rho$...
in $\mathfrak{A}$ satisfies $\mathfrak{A} \models \varphi[a,c]$, or every realization $(c,d)$ of $\rho$ in $\mathfrak{A}$ satisfies $\mathfrak{A} \models \varphi[b,d]$. In other words, a monotonic ray-type is one which has at least one dendral realization and for which either there exists a single equivalence class that emits all instances of $\rho$, or there exists a single equivalence class that absorbs all instances of $\rho$. We say that $\rho$ is $(3,2)$-hexatopic in $\mathfrak{A}$ if there exist distinct equivalence classes $A_1, \ldots, A_6$ of $\mathfrak{A}$ and dendral elements $a_1, \ldots, a_6$ such that $A_1$ is connected to $A_{i+3}$ by the $\rho$-ray $\langle a_i, a_{i+3} \rangle$ $(1 \leq i \leq 3)$, as illustrated in Fig. 4(a).

We say that $\rho$ is $(2,3)$-hexatopic in $\mathfrak{A}$ if there exist distinct equivalence classes $A_1, \ldots, A_6$ of $\mathfrak{A}$ and dendral elements $a_1, \ldots, a_6$ such that $A_1$ and $A_2$ are connected by the $\rho$-ray $\langle a_1, a_2 \rangle$, $A_2$ and $A_3$ are connected by the $\rho$-ray $\langle a_3, a_4 \rangle$, and similarly for $A_3, \ldots, A_6$ and $a_5, \ldots, a_8$, as illustrated in Fig. 4(b). We say that $\rho$ is hexatopic in $\mathfrak{A}$ if it is either $(3,2)$-hexatopic or $(2,3)$-hexatopic in $\mathfrak{A}$. We call an $\alpha$-structure $(3,2)$-cosmically valid if every realized cosmic $t$-ray-type is either monotonic or $(3,2)$-hexatopic, and cosmically valid if every realized cosmic $t$-ray-type is either monotonic or hexatopic.

The notion of cosmic validity is defined for $\alpha$-structures generally; however, in the special case of $d$-structures, it is easy to secure, as the next lemma shows.

**Lemma 5.** Let $D$ be a $d$-structure interpreting signature $\Sigma$. Then there is a $(3,2)$-cosmically valid expansion $D'$ of $D$ interpreting $\Sigma$ together with at most 8 fresh binary predicates. If $D$ is $\theta$-parental for some quantifier-free formula $\theta$ over $\Sigma$, then so is $D'$.

**Proof.** Enumerate the cosmic $t$-ray-types over $\Sigma$ realized in $D$. If any of these are invertible, remove one of the directions from this list so that only one member of any pair $\{\rho, \rho^{-1}\}$ appears. Let the resulting list be $\rho_1, \rho_2, \ldots$. Let $D_0 = D$.

We begin by setting the extensions $r_1, \ldots, r_8$ to be empty, enlarging these sets as we consider the $\rho_i$ in turn. Assume that $\rho_1, \ldots, \rho_{i-1}$ have been processed for some $i \geq 1$, yielding the d-structure $D_{i-1}$; we consider $\rho_i$.

Let $\langle a_1, b_1 \rangle$ be an instance of $\rho_i$. Let $A_1, B_1$ be the equivalence classes of $D_{i-1}$ such that $a_1 \in A_1$ and $b_1 \in B_1$. (Hence $A_1 \neq B_1$.) Suppose first that every instance $\langle a', b' \rangle$ of $\rho_i$ in $D_{i-1}$ is either absorbed or emitted by some element of $A_1 \cup B_1$. Then we let $D_i$ be the same as $D_{i-1}$ except that, for any $t$-ray $\langle a', b' \rangle$ of type $\rho_i$, we set:

- $\langle a', b' \rangle \in r_1^{D_i} \iff a' \in A_1$
- $\langle a', b' \rangle \in r_2^{D_i} \iff a' \in B_1$
- $\langle a', b' \rangle \in r_3^{D_i} \iff b' \in A_1$
- $\langle a', b' \rangle \in r_4^{D_i} \iff b' \in B_1$

Figure 4: Witnesses for hexatopic structures
Thus, the instances of $\rho_i$ are re-distributed among six ray-types (four satisfying exactly one of the new predicates and two satisfying either $r_1$ and $r_4$ or $r_2$ and $r_3$), each of which is evidently monotonic.

We may therefore assume that there exists an instance $\langle a_2, b_2 \rangle$ of $\rho_i$, and equivalence classes $A_2, B_2$ of $\mathcal{D}_{i-1}$ such that $a_2 \in A_2$ and $b_2 \in B_2$, with $A_1, A_2, B_1, B_2$ distinct. Suppose now that every instance $\langle a', b' \rangle$ of $\rho_i$ in $\mathcal{D}_{i-1}$ is either absorbed or emitted by some element of $A_1 \cup A_2 \cup B_1 \cup B_2$. Then we let $\mathcal{D}_i$ be the same as $\mathcal{D}_{i-1}$ except that, for any t-ray $\langle a', b' \rangle$ of type $\rho_i$, we set the interpretations of $r_1, \ldots, r_4$ as above, and we set the interpretations of $r_5, \ldots, r_8$ analogously, but with $A_1$ and $B_1$ replaced by $A_2$ and $B_2$, respectively. Thus, the instances of $\rho_i$ are re-distributed among several (actually, twenty) ray-types, each of which is evidently monotonic.

We may therefore assume that there exists an instance $\langle a_3, b_3 \rangle$ of $\rho_i$, and equivalence classes $A_3, B_3$ of $\mathcal{D}_{i-1}$ such that $a_3 \in A_3$ and $b_3 \in B_3$, with $A_1, A_2, A_3, B_1, B_2, B_3$ distinct. Then $\rho_i$ is $(3,2)$-hexatopic in $\mathcal{D}_{i-1}$, and we set $\mathcal{D}_i = \mathcal{D}_{i-1}$.

At the end of this process, we obtain a structure $\mathcal{D}'$ in which every realized t-ray type is either monotonic or $(3,2)$-hexatopic. A pair $\langle a, b \rangle$ is a $\theta$-ray in $\mathcal{D}$ if and only if it is a $\theta$-ray in $\mathcal{D}'$; moreover, the 1-type of every element is the same in $\mathcal{D}$ as in $\mathcal{D}'$. Thus, $\mathcal{D}'$ is $\theta$-parental if $\mathcal{D}$ is.

The most important property of cosmic validity is that, for any cosmic $t$-ray-type $\rho$, if there is a cyclic realization of $\rho$, then we can find a dendral realization of $\rho$ such that the rays in question may be switched without compromising the property of cosmic validity. The next lemma formalizes this idea.

**Lemma 6.** Let $\Sigma$ be a signature, $\theta$ a quantifier-free formula over $\Sigma$, and $\rho$ a cosmic $t$-ray-type over $\Sigma$. Let $\mathfrak{A}$ be a cosmically valid, $\theta$-parental $a$-structure interpreting $\Sigma$. If $\langle c, d \rangle$ is a cyclic realization of $\rho$ in $\mathfrak{A}$, then we can find a dendral realization $\langle a, b \rangle$ of $\rho$ in $\mathfrak{A}$ for which $\langle a, b \rangle$ and $\langle c, d \rangle$ are switchable and $\mathfrak{A}(a, b||c, d)$ is cosmically valid.

**Proof.** If $\mathfrak{A}$ has a dendral realization $\langle a, b \rangle$ of $\rho$ such that either $\mathfrak{A} \models \varepsilon[a, c]$ or $\mathfrak{A} \models \varepsilon[b, d]$, then $\langle a, b \rangle$ and $\langle c, d \rangle$ are switchable, so that $\mathfrak{A}(a, b||c, d)$ is defined. Indeed, in this case, each pair of equivalence classes is connected by exactly the same types of $\theta$-rays in both structures, whence $\mathfrak{A}(a, b||c, d)$ remains cosmically valid.

Hence we may assume that no such $\langle a, b \rangle$ exists, and thus that $\rho$ is hexatopic in $\mathfrak{A}$. Let $A_1, \ldots, A_6$ be distinct equivalence classes witnessing the hexatopicity of $\rho$. Let $A_c$ and $A_d$ be the equivalence classes of $c$ and $d$, respectively. Assume first that $\rho$ is $(3,2)$-hexatopic, and let $a_1, \ldots, a_6$ be the corresponding witness points. Notice that, by assumption, $A_c$ is distinct from $A_1, A_2, A_3$, and $A_d$ is distinct from $A_4, A_5, A_6$. We thus have five sub-cases.

(i) $A_c$ and $A_d$ are distinct from all of $A_1, \ldots, A_6$ (Fig. 5a). Pick, say $a = a_3$ and $b = a_6$, and set $\mathfrak{A}' = \mathfrak{A}(a, b||c, d)$. Then the classes $A_1, A_2, A_c, A_4, A_5, A_6$ and points $a_1, a_2, c, a_4, a_5, a_6$ witness that $\rho$ is $(3,2)$-hexatopic in $\mathfrak{A}'$. We require that the ray $\langle c, a_6 \rangle$ is dendral in $\mathfrak{A}'$; but since $a_3, a_6$ are dendral in $\mathfrak{A}$, and $c, d$ are cyclic, all of the rays involved will be dendral in $\mathfrak{A}'$.

(ii) $A_c$ is distinct from all of $A_1, \ldots, A_6$; and $A_d$ is identical to one of $A_1, A_2, A_3$—say $A_d = A_3$ (Fig. 5b). Pick, say $a = a_2$ and $b = a_5$, and set $\mathfrak{A}' = \mathfrak{A}(a, b||c, d)$.
Thus we may assume henceforth that \( \rho \) into a \((2,3)\)-hexatopic ray type.

(i) \( \rho \) is \((2,3)\)-hexatopic. Solid arrows are rays in \( \mathfrak{A} \) (with crossed arrows absent from \( \mathfrak{A}' \)); dotted arrows are rays added in \( \mathfrak{A}' \).

Then the classes \( A_1, A_2, A_c, A_4, A_3, A_5 \) and points \( a_1, a_2, c, a_4, d, a_5 \) witness that \( \rho \) is \((3,2)\)-hexatopic in \( \mathfrak{A}' \).

(iii) \( A_d \) is distinct from all of \( A_1, \ldots, A_6 \); and \( A_c \) is identical to one of \( A_4, A_5, A_6 \). Symmetric to the previous case.

(iv) \( A_c \) is identical to one of \( A_4, A_5, A_6 \) — say \( A_c = A_6 \); and \( A_d \) is identical to the corresponding class of \( A_1, A_2, A_3 \)—namely \( A_d = A_3 \) (Fig. 5c). Pick, say \( a = a_2 \) and \( b = a_5 \), and set \( \mathfrak{A}' = \mathfrak{A}(a, b|c, d) \). Then the classes \( A_1, A_2, A_6, A_4, A_3, A_5 \) and points \( a_1, a_2, c, a_4, d, a_5 \) witness that \( \rho \) is \((3,2)\)-hexatopic in \( \mathfrak{A}' \).

(v) \( A_c \) is identical to one of \( A_1, A_5, A_6 \) — say \( A_c = A_5 \); and \( A_d \) is identical to a non-corresponding class of \( A_1, A_2, A_3 \)—say \( A_d = A_3 \) (Fig. 5d). Pick \( a = a_1 \) and \( b = a_4 \), and set \( \mathfrak{A}' = \mathfrak{A}(a, b|c, d) \). Then the classes \( A_1, A_3, A_6, A_2, A_5, A_4 \) and points \( a_1, a_2, a_3, a_4, a_5, c, a_4 \) witness that \( \rho \) is \((2,3)\)-hexatopic in \( \mathfrak{A}' \). We remark that this is the only case in which a \((3,2)\)-hexatopic ray type is transformed into a \((2,3)\)-hexatopic ray type.

Thus we may assume henceforth that \( \rho \) is \((2,3)\)-hexatopic, and let \( a_1, \ldots, a_8 \) be witness points. Notice that, by hypothesis, \( A_c \) is distinct from \( A_1, A_2, A_4, A_5 \), and \( A_d \) is distinct from \( A_2, A_3, A_5, A_6 \). We thus have five sub-cases.

(i) \( A_c \) and \( A_d \) are distinct from all of \( A_1, \ldots, A_6 \). Therefore, we can pick, say \( a = a_1 \) and \( b = a_2 \), and set \( \mathfrak{A}' = \mathfrak{A}(a, b|c, d) \). Then the classes \( A_1, A_2, A_4, A_6, A_3, A_5 \) and points \( a_1, a_2, a_3, a_4, a_5, a_6 \) witness that \( \rho \) is \((3,2)\)-hexatopic in \( \mathfrak{A}' \) (Fig. 6a).

(ii) \( A_c \) is distinct from all of \( A_1, \ldots, A_6 \); and \( A_d \) is identical to one of \( A_1 \) or \( A_4 \)—say \( A_d = A_1 \). Therefore, we can pick, say \( a = a_3 \) and \( b = a_4 \),

Figure 5: Proof of Lemma 6: \( \rho \) is \((3,2)\)-hexatopic. Solid arrows are rays in \( \mathfrak{A} \) (with crossed arrows absent from \( \mathfrak{A}' \)); dotted arrows are rays added in \( \mathfrak{A}' \).
and set $\mathcal{X}' = \mathcal{X}(a, b\|c, d)$. Then the classes $A_1, A_c, A_4, A_5, A_6$ and points $a_1, a_3, a_5, a_6$ witness that $\rho$ is $(3, 2)$-hexatopic in $\mathcal{X}'$ (Fig. 6b).

(iii) $A_d$ is distinct from all of $A_1, \ldots, A_6$; and $A_c$ is identical to one of $A_3, A_6$. Symmetric to the previous case.

(iv) $A_c$ is identical to one of $A_1, \ldots, A_6$; and $A_d$ is identical to $A_7-i$. Suppose $A_c = A_3$ and $A_d = A_4$. Therefore, we can pick, say $a = a_1$ and $b = a_2$, and set $\mathcal{X}' = \mathcal{X}(a, b\|c, d)$. Then the classes $A_1, A_2, A_5, A_3, A_6$ and points $a_1, a_5, a_7, d, a_4, a_8$ witness that $\rho$ is $(3, 2)$-hexatopic in $\mathcal{X}'$ (Fig. 6c).

(v) $A_c$ is identical to one of $A_i$, where $i$ is either 3 or 6; and $A_d$ is identical to $A_{7-i}$. Suppose $A_c = A_3$ and $A_d = A_1$. Therefore, we can pick, say $a = a_5$ and $b = a_6$, and set $\mathcal{X}' = \mathcal{X}(a, b\|c, d)$. Then the classes $A_2, A_5, A_3, A_6, A_1$ and points $a_3, a_5, a_4, a_8, d$ witness that $\rho$ is $(3, 2)$-hexatopic in $\mathcal{X}'$ (Fig. 6d).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{Proof of Lemma 6: $\rho$ is $(2, 3)$-hexatopic. Solid arrows are rays in $\mathfrak{A}$ (with crossed arrows absent from $\mathfrak{A}'$); dotted arrows are rays added in $\mathfrak{A}'$.}
\end{figure}

\section{Removing cycles from structures}

At this point, we can present the core idea of the proof. If $\mathfrak{A}$ is an $a$-structure, we refer to any cycle in the graph $(\mathfrak{A}, t^\mathfrak{A})$, simply, as a cycle in $\mathfrak{A}$. We say that the cycle is galactic if all of the $t$-rays involved are galactic, and cosmic, if any is cosmic. Cosmically and galactically valid $a$-structures enable us to eliminate cycles by merging them into other components of the graph of $t$ one by one.

\begin{lemma}
Let $\Sigma$ be a signature, $\theta$ a quantifier-free formula over $\Sigma$, and $\mathfrak{A}$ a $\theta$-parental $a$-structure interpreting $\Sigma$ such that $\mathfrak{A}$ is galactically and cosmically
\end{lemma}
valid. Suppose $A$ contains at least one cycle. Then there exists an $a$-structure $A'$ interpreting $\Sigma$ over the same domain, $A$, also galactically and cosmically valid, such that: (i) $A$ and $A'$ realize the same 2-types; (ii) every element of $A$ has the same $\theta$-star-type in $A'$ as in $A$; (iii) $A'$ is $\theta$-parental; and (iv) $A'$ contains strictly fewer cycles than $A$. Hence, there is a $d$-structure $D$ interpreting $\Sigma$ over the domain $A$, such that: (i) $D$ and $A$ realize the same 2-types; and (ii) every element of $A$ has the same $\theta$-star-type in $D$ as in $A$.

Proof. Suppose first of all that $A$ contains a galactic cycle. Let $B$ be the equivalence class containing this cycle, and pick some $t$-ray $\langle c,d \rangle$ of type $\rho$ in the cycle. Since $A$ is galactically valid, we may find a locally dendral instance $\langle a,b \rangle$ of $\rho$ in $B$. Since $a$, $b$, $c$, $d$ are all equivalent, $\langle a,b \rangle$ and $\langle c,d \rangle$ are certainly switchable, and so $A' = A(a,b||c,d)$ is an $a$-structure. Since all elements that were locally dendral in $B$ remain locally dendral in $B$ following this change, $A'$ is galactically valid. And since all elements that were dendral in $A$ remain dendral in $A'$, and cosmic $t$-rays are unaffected, $A'$ is also cosmically valid.

Otherwise, $A$ contains at least one cosmic cycle. Pick some $t$-ray $\langle c,d \rangle$ of cosmic type $\rho$ in that cycle. Since $A$ is cosmically valid, by Lemma 6, we can find a dendral realization $\langle a,b \rangle$ of $\rho$ for which $\langle a,b \rangle$ and $\langle c,d \rangle$ are switchable, with $A' = A(a,b||c,d)$ a cosmically valid $a$-structure. Since no galactic 2-types have changed at all, $A'$ is also galactically valid.

We need to show that $A' = A(a,b||c,d)$ has the properties required by the lemma. Properties (i)–(iii) are guaranteed by Lemma 3. Since the cycle containing $\langle c,d \rangle$ has been merged either into one of the dendral components or into a distinct cycle, we have Property (iv).

The second statement of the lemma follows by repeated application of the first.

\[\square\]

5 Shrubberies

In this section, we define two properties of $a$-structures sufficient for galactic and cosmic validity, respectively. We begin with the former.

As usual, we take forests to be directed graphs with edges from daughter to mother. Let $\Sigma$ be a signature. A galactic shrubbery over $\Sigma$ is a triple $(V,E,L)$, where $(V,E)$ is a non-empty forest with $V = \{1, \ldots, N\}$, and $L$ is a vertex- and edge-labelling satisfying the following conditions:

1. for all $i \in V$, $L(i)$ is a 1-type over $\Sigma$;
2. for all $e \in E$, $L(e)$ is a $t$-ray-type over $\Sigma$ such that $\tp_1(L(u,v)) = L(u)$ and $\tp_2(L(u,v)) = L(v)$.

The size of $S$ is $N = |V|$.

Let $A$ be an $a$-structure interpreting a signature $\Sigma$ and $S = (V,E,L)$ a galactic shrubbery over $\Sigma$. An equivalence class $B$ of $A$ realizes $S$ if there exists an embedding $f : V \to B$ such that:

(i) for all $i \in V$, $\tp^A[f(i)] = L(i)$;
(ii) for all $(i,j) = e \in E$, $\tp^A[f(i), f(j)] = L(e)$;
(iii) if \( i \) is a root of \((V, E)\), then \( f(i) \) is a local root in \( B \);

(iv) every galactic t-ray-type realized in \( B \) occurs as \( L(e) \) for some \( e \in E \).

If the function \( f \) is clear from context, we speak of an element \( f(i) \in B \) as realizing the vertex \( i \in V \). Essentially, \( B \) realizes \( S \) just in case the locally dendrinal components of \( B \) have a prefix isomorphic to \( S \), with the realizing vertices having 1-types and 2-types as indicated by the labelling on \( S \); in addition, roots of \( S \) must be realized by local roots in \( B \) and all arboreal 2-types occurring in \( B \) must be accounted for in the edges of \( S \). Galactic shrubberies ensure galactic validity: the following lemma is simply a matter of unpeeling the foregoing definitions.

**Lemma 8.** Suppose \( \mathfrak{A} \) is an \( a \)-structure interpreting \( \Sigma \). Then \( \mathfrak{A} \) is galactically valid if and only if every equivalence class of \( \mathfrak{A} \) realizes some galactic shrubbery.

We now carry out a corresponding construction for cosmic validity. Let \( \Sigma \) be a signature. A cosmic shrubbery over \( \Sigma \) is a tuple

\[
T = (V, E, L, \sim, R_m, R_h, M, \kappa, \lambda_1, \ldots, \lambda_6),
\]

satisfying the following conditions.

1. \( S = (V, E, L) \) satisfies the conditions of being galactic shrubbery;

2. \( \sim \) is an equivalence relation on \( V \);

3. \( \{R_m, R_h\} \) partitions the set of cosmic t-ray-types occurring in \( L(E) \) (we allow \( R_m \) and \( R_h \) to be empty);

4. \( 0 \leq M \leq 2^{4|\Sigma|} \), and \( \kappa \) is a function \( \kappa : R_m \to \{1, \ldots, M\} \);

5. for all \( i \) (1 \( \leq i \leq 6 \)), \( \lambda_i : R_h \to V \) is a function such that, for all \( \rho \in R_h \),
   
   (i) the vertices \( \lambda_1(\rho), \ldots, \lambda_6(\rho) \) are pairwise unrelated by \( \sim \), and
   
   (ii) \( e_1 = (\lambda_1(\rho), \lambda_2(\rho)) \), \( e_2 = (\lambda_3(\rho), \lambda_4(\rho)) \) and \( e_3 = (\lambda_5(\rho), \lambda_6(\rho)) \) are edges in \( E \) satisfying \( L(e_1) = L(e_2) = L(e_3) = \rho \).

The size of \( T \) is \( N = |V| \).

Thus a cosmic shrubbery is a galactic shrubbery together with some extra structure: an equivalence relation, a partition of the realized cosmic t-ray-types into the cells \( R_m \) and \( R_h \), an assignment to each t-ray-type in \( R_m \) of some integer, and an assignment to every t-ray-type in \( R_h \) of a collection of six pairwise inequivalent vertices of \( S \) with \( \rho \)-labelled edges connecting these elements in three pairs as indicated. Condition 1 of the above definition of course means that, formally speaking, \( S = (V, E, L) \) is a galactic shrubbery. However, in this context, we are to imagine the vertices \( V \) to correspond to elements in different equivalence classes in a-structures. Intuitively, the t-ray-types in \( R_m \) are meant to be monotonic, and those in \( R_h \) (3,2)-hexatopic. For each \( \rho \in R_m \), we think of \( \kappa(\rho) \) as the index of an equivalence class which either emits all instances of \( \rho \) or absorbs all instances of \( \rho \). And for each \( \rho \in R_h \), we think of the elements \( \lambda_1(\rho), \ldots, \lambda_6(\rho) \) as witnessing the (3,2)-hexatopicity of \( \rho \).

Let \( \mathfrak{A} \) be an \( a \)-structure interpreting \( \Sigma \) and \( T \) a cosmic shrubbery as given in (3). We say that \( \mathfrak{A} \) realizes \( T \) if there exists an embedding \( g : V \to A \) and equivalence classes \( B_1, \ldots, B_M \) of \( \mathfrak{A} \) such that:

17
(i) for all \( i \in V \), \( \text{tp}^{A}[g(i)] = L(i) \);
(ii) for all \( (i, j) = e \in E \), \( \text{tp}^{A}[g(i), g(j)] = L(e) \);
(iii) if \( i \) is a root of \( (V, E) \), then \( g(i) \) is a root in \( A \);
(iv) every cosmic t-ray-type realized in \( A \) occurs as \( L(e) \) for some \( e \in E \);
(v) for all \( i, j \in V \), \( i \sim j \) if and only if \( A \models e[g(i), g(j)] \);
(vi) for every \( \rho \in R_m \), either every ray \( (a, b) \) of type \( \rho \) in \( A \) satisfies \( a \in B_{\kappa(\rho)} \),
or every ray \( (a, b) \) of type \( \rho \) in \( A \) satisfies \( b \in B_{\kappa(\rho)} \).

If the function \( g \) is clear from context, we speak of an element \( g(i) \in A \) as realizing the vertex \( i \in V \).

Cosmic shrubbery ensure \((3,2)\)-cosmic validity: the next lemma is again a matter of unpeeling definitions.

**Lemma 9.** Suppose \( A \) is an a-structure interpreting \( \Sigma \). Then \( A \) is \((3,2)\)-cosmically valid if and only if it realizes some cosmic shrubbery.

Lemmas 8 and 9 are not quite enough for our purposes: we need to show that certain galactically valid a-structures can be transformed into a-structures which realize small galactic shrubbery, and similarly for cosmic validity. Intuitively, large shrubbery contain long paths that can be shortened by switching as in Fig. 3, this time read from right to left. (Thus, whereas Lemma 7 proceeds by removing cycles from a-structures, here we introduce them.)

**Lemma 10.** Let \( \Sigma \) be a signature, \( \theta \) a quantifier-free formula over \( \Sigma \), and \( A \) a \( \theta \)-parental a-structure interpreting \( \Sigma \) such that every equivalence class of \( A \) realizes a galactic shrubbery, and \( A \) realizes a cosmic shrubbery. Define \( Z = 2^{4|\Sigma|} \). Then there exists an a-structure \( A^* \) over the same domain \( A \), such that (i) \( A \) and \( A^* \) realize the same 2-types; (ii) every element of \( A \) has the same \( \theta \)-star-type in \( A^* \) as in \( A \); (iii) every equivalence class of \( A^* \) realizes a galactic shrubbery of size at most \( M = 5Z^2 \); (iv) \( A^* \) realizes a cosmic shrubbery of size at most \( N = 500Z^4 \).

**Proof.** Observe that \( Z \) is greater than the number of 2-types over \( \Sigma \). For each equivalence class \( B \) of \( A \), let \( S_B \) be a galactic shrubbery realized by \( B \), and let \( T = (V, E, L, \sim, R_m, R_h, M, \kappa, \lambda_1, \ldots, \lambda_6) \) be a cosmic shrubbery realized by \( A \). It follows that \( R_m \) and \( R_h \) are the sets of monotonic, respectively hexatonic, cosmic t-rays realized in \( A \). For each cosmic t-ray-type in \( R_m \), choose some pair of elements realizing an edge of \( T \) labelled with \( \rho \), and mark those elements. For each cosmic t-ray-type in \( R_h \), mark the elements realizing the vertices \( \lambda_1(\rho), \ldots, \lambda_6(\rho) \) of \( T \). By assumption, these elements form triples of \( \rho \)-rays as in Fig. 4(a). Say that the elements marked in this process are cosmically marked.

Fix some equivalence class \( B \) for the moment. For brevity, we refer to the realization of \( S_B \) in \( B \) simply as \( “S_B” \), and similarly for \( T \); no confusion should arise as a result. For each galactic t-ray-type \( \rho \) realized in \( B \), choose some pair of elements forming an edge of \( S_B \) labelled with \( \rho \) and mark those elements. Say that the elements marked in this process are galactically marked. The total number of elements so-far marked in \( B \) (cosmically or galactically) is certainly at most \( 2Z \), since at most one ray of each type can lead to the marking of elements.
in $B$, and only two elements can be marked as a result of considering each ray. Finally, for each pair of marked vertices in $S_B$, galactically mark their nearest common ancestor in the realization of $S_B$ (if they have a common ancestor in $B$ which is not already marked). It as easy to see that at most $2Z - 1$ additional vertices are galactically marked in this way. Thus, fewer than $4Z$ vertices of $S_B$ will have been marked in total.

Now consider the marked vertices of $S_B$ and all their local ancestors (i.e. vertices lying on paths to a local root). Clearly, we may remove any other vertices from $S_B$, since each galactic t-ray-type realized in $B$ is realized by galactically marked elements. (Note that removing elements from $S_B$ just affects the galactic shrubbery $S_B$ realized in $A$: the a-structure $A$ is not changed at all.) Once this has been done, all cosmically marked elements in $B$ are still in $S_B$, and, moreover, all the unmarked vertices of $S_B$ form a collection of disjoint linear paths in $S_B$. Now suppose one of these paths has length more than $Z$. Then there must be a t-ray of some galactic type $\rho$ that occurs at least twice on this path. Say the occurrences of $\rho$ are $\langle a_1, b_1 \rangle$ and $\langle a_2, b_2 \rangle$. Since all four elements are by assumption equivalent, the rays $\langle a_1, b_1 \rangle$ and $\langle a_2, b_2 \rangle$ are switchable, so we may form the structure $A' = A(a_1, b_1 \parallel a_2, b_2)$. Evidently, $B$ realizes a smaller galactic shrubbery, say $S'$, and $A'$ realizes a (possibly smaller) cosmic shrubbery $T'$. Hence $A'$ is galactically and cosmically valid, and, by Lemma 3, realizes the same 2-types, assigns the same $\theta$-star-type to every element, and is $\theta$-parental. Hence we may proceed until we obtain a realization of some galactic shrubbery $S^+$ in $B$ in which all paths of unmarked elements have length at most $Z$. Thus, $|S^+| \leq 4Z(Z + 1) < 5Z^2$ (since $Z > 4$). Do the same for all equivalence classes. We obtain a structure, say $A^+$, in which each equivalence class $B$ realizes a galactic shrubbery $S^+_B$ of size at most $5Z^2$, and which realizes a cosmic shrubbery $T^+$.

We need to modify $A^+$ so as to control the size of the cosmic shrubbery. Let us return to those elements that were cosmically marked. These elements were joined in pairs by cosmic rays in $A$, and these cosmic rays were not disturbed in the construction of $A^+$. In particular, they must be vertices of the cosmic shrubbery $T^+$, and there are at most $6Z$ of them in total. Take any two distinct cosmically marked vertices of $T^+$, and cosmically mark their nearest common ancestor, if any. It as easy to see that at most $6Z - 1$ additional vertices are cosmically marked in this way, making fewer than $12Z$ in total. (This is actually a silly over-estimate, but no matter.) Now consider the cosmically marked vertices of $T^+$ and all their ancestors (i.e. vertices lying on paths to a root). Clearly, we may remove any other vertices from $T^+$, since all each cosmic t-ray-type realized in $A$ is realized by cosmically marked elements. Suppose this has been done. (Note that removing elements from $T^+$ just affects the cosmic shrubbery: $A^+$ is not changed at all.) Thus, all the vertices of $T^+$ which are not cosmically marked form a collection of disjoint linear paths. Fixing any one of these linear paths, say $P$, call a maximal contiguous sub-path consisting entirely of nodes lying within the same equivalence class a segment. If there are more than $3Z + 1$ segments on $P$ then $P$ must contain at least 4 edges $\langle a_1, b_1 \rangle$, $\langle a_2, b_2 \rangle$, $\langle a_3, b_3 \rangle$ and $\langle a_4, b_4 \rangle$ all labelled with the same cosmic t-ray-type $\rho$. We claim that some pair of these rays is switchable. For if any $a_i$ is equivalent (in $A$) to $a_j$ ($i \neq j$), then $\langle a_i, b_i \rangle$ and $\langle a_j, b_j \rangle$ are switchable; similarly if any $b_i$ is equivalent to $b_j$ ($i \neq j$). Hence we may assume that the $a_i$ are pairwise inequivalent, and
similarly for the $b_k$. Therefore, $a_1$ is equivalent (in $\mathfrak{A}$) to at most one of $b_2$, $b_3$ and $b_4$: by renumbering if necessary, suppose $a_1$ is equivalent to neither $b_3$ nor $b_4$. Similarly, $b_1$ cannot be equivalent to both $a_2$ and $a_4$: by renumbering if necessary, suppose $b_1$ is not equivalent to $a_4$. Then $\langle a_1, b_1 \rangle$ and $\langle a_4, b_4 \rangle$ are switchable. Thus, we may find a pair of switchable rays of cosmic type $\rho$, say $\langle a, b \rangle$ and $\langle c, d \rangle$. Setting $\mathfrak{A}' = \mathfrak{A}(a, b|c, d)$, we see that $\mathfrak{A}'$ realizes a smaller cosmic shrubbery. Since only cosmic rays have been changed, certainly no galactic shrubberies are affected. Continuing this process until the cosmic shrubbery cannot be reduced in size any further, we see that no path $P$ of cosmically unmarked nodes in the cosmic shrubbery realized by $\mathfrak{A}'$ can have more than $3Z^2 + 1$ segments. Consider now any segment, say in an equivalence class $B$. Observe that one end of this segment is a local root in $B$. Now, if that segment intersects the galactic tree, $S^*_B$, then the common part includes the local root and forms a path in $S^*_B$, and that path certainly has length bounded by $4Z(Z + 1)$, since this is a bound on the size of $S^*_B$. The remainder of the segment can then be shortened (by switching galactic rays of the same type, as described above) so that it has length of no more than $Z$, meaning that the segment has length at most $4Z(Z + 1) + Z$. Since only rays lying outside the local galactic shrubbery have been switched, no galactic shrubberies are affected. Carrying out this process to exhaustion, we see that $P$ consists of no more than $3Z$ edges labelled with cosmic $t$-ray types and no more than $(3Z + 1) \cdot (4Z(Z + 1) + Z)$ edges labelled with galactic $t$-ray types. Thus, the total length of the path is $12Z^3 + 19Z^2 + 8Z$. Let $\mathfrak{A}^*$ denote the a-structure obtained at the end of this process, and let the cosmic shrubbery obtained be $T^*$. Thus $T^*$ consists of fewer than $12Z$ cosmically marked vertices, joined by paths of length at most $12Z^3 + 19Z^2 + 8Z$. Taking into account the fact that $Z > 4$, so that $Z^4 > 4Z^3 > 14Z^2 > 64Z$, we obtain $|T^*| \leq (12Z + 1)(12Z^3 + 19Z^2 + 8Z) < 500Z^4$. Thus, each equivalence class $B$ of $\mathfrak{A}^*$ will realize a shrubbery $S^*_B$ of size less than $3Z^2$, and $\mathfrak{A}^*$ will realize a cosmic shrubbery $T^*$ of size at most $500Z^4$, as required.

This secures Properties (iii) and (iv) of the Lemma. Properties (i) and (ii) follow by Lemma 3 (i) and (ii), respectively.

6 Encoding properties of a-structures

6.1 Definition of a-structures

Recall that an a-structure is a structure in which: (i) the binary predicate $t$ is interpreted as an irreflexive, asymmetric relation in which every element has at most one successor; and (ii) the unary predicate $r$ is satisfied by precisely those elements with no $t$-successor. Over domains of more than one element, condition (i) is expressed by the normal-form formula

$$\forall x \forall y ((\neg t(x, x) \land (t(x, y) \rightarrow \neg t(y, x))) \lor x = y) \land$$

$$\forall x \exists y [t(x, y) \land x \neq y], \quad (\Omega_1)$$

and condition (ii) is enforced by the normal-form formula

$$\forall x \exists y [r(x, y) \land x \neq y] \land$$

$$\forall x \forall y (((r(x) \rightarrow \neg t(x, y)) \land ((r(x, y) \land \neg r(x)) \rightarrow t(x, y))) \lor x = y), \quad (\Omega_2)$$

20
where \( r \) is a fresh binary predicate. Denote by \( \Omega \) the conjunction of \((\Omega_1)\) and \((\Omega_2)\).

**Lemma 11.** Any structure \( \mathfrak{A} \) such that \( \mathfrak{A} \models \Omega \) is an a-structure. Conversely, let \( \mathfrak{A} \) be an a-structure interpreting a signature that does not contain the binary predicate \( r \), with \( |A| > 1 \). Then \( \mathfrak{A} \) has an expansion \( \mathfrak{A}^+ \) such that \( \mathfrak{A}^+ \models \Omega \).

### 6.2 Parental property

The \( \theta \)-parental property for an a-structure interpreting a signature \( \Sigma \) can also be expressed by a formula in a larger signature. For each predicate \( p \) (unary or binary) of \( \Sigma \), let \( \bar{p} \) be a fresh predicate, and write \( \bar{\Sigma} = \{ \bar{p} \mid p \in \Sigma \} \). For each 1-type \( \pi \) over \( \Sigma \), denote by \( \hat{\pi}(x) \) the formula which results from replacing every predicate (unary or binary) in \( \pi \) by the corresponding predicate \( \bar{p} \). Consider then the sentence

\[
\bigwedge \{ \forall x \forall y ((t(x, y) \rightarrow (\bar{p}(x) \leftrightarrow p(y))) \vee x = y) \mid p \in \Sigma^{[1]} \} \wedge \\
\bigwedge \{ \forall x \forall y (((t(x, y) \rightarrow (\bar{p}(x, x) \leftrightarrow p(y, y))) \vee x = y) \mid p \in \Sigma^{[2]} \}, \quad (\Pi^2_0)
\]

where \( \Sigma^{[1]} \) is the set of unary predicates of \( \Sigma \) and \( \Sigma^{[2]} \) the set of binary predicates of \( \Sigma \). In any a-structure \( \mathfrak{A} \) making \((\Pi^2_0)\) true, we may read \( \hat{\pi} \) as “the parental 1-type of \( x \) over \( \Sigma \) (if defined) is \( \pi \).” For elements with no mother, nothing is fixed about the meaning of \( \hat{\pi} \). Now consider the sentence

\[
\bigwedge \{ \forall x \forall y (((\theta \land \pi \land \neg t(y, x)) \rightarrow (\bar{t}(y) \lor \neg \hat{\pi}(y))) \vee x = y) \mid \pi \text{ a } \Sigma\text{-1-type} \}. \quad (\Pi^2_0)
\]

Given the interpretations of \( \hat{\pi} \) and \( t \) suggested above, \((\Pi^2_0)\) states that the \( \Sigma \)-reduct of the a-structure in question is \( \theta \)-parental. Let \( \Pi_\theta \) be the conjunction of \((\Pi^2_0)\) and \((\Pi^2_0)\).

**Lemma 12.** Let \( \mathfrak{A}' \) be an a-structure interpreting a signature \( \Sigma' \supseteq \Sigma \) such that \( \Sigma' \cap \Sigma = \emptyset \), and let \( \mathfrak{A} \) be the \( \Sigma \)-reduct of \( \mathfrak{A}' \). Then \( \mathfrak{A} \) is \( \theta \)-parental if and only if \( \mathfrak{A}' \) has an expansion \( \mathfrak{A}^+ \models \Pi_\theta \).

### 6.3 Relativization

We now turn to the more difficult problem of encoding galactic and cosmic validity in a-structures. We make repeated use of a simple technical device to label a large number of elements with a small signature. Suppose that \( \bar{p} = p_1, \ldots, p_n \) is a sequence of unary predicates, and consider a conjunction \( \pm p_1(x) \land \cdots \land \pm p_n(x) \), where \( \pm \gamma \) denotes either \( \gamma \) or \( \neg \gamma \) for any formula \( \gamma \). Each such conjunction encodes an integer \( i \) \((0 \leq i < 2^n)\), where the \( h \)th digit of \( i \) is 1 just in case the atom \( p_h(x) \) occurs in the formula with positive polarity. Denote by \( \bar{p}(i)(x) \) the formula encoding \( i \). We may read \( \bar{p}(i)(x) \) as “\( x \) is an element with \( \bar{p} \)-index \( i \).”

Fix a signature \( \Sigma \) for the moment, and define \( \mathbb{Z} = 2^{[\Sigma]} \), \( \mathbb{M} = 5\mathbb{Z} \) and \( \mathbb{N} = 500\mathbb{Z} \), as in Lemma 10. In the sequel, \( \mathbb{Z} \), \( \mathbb{M} \) and \( \mathbb{N} \) will always depend on \( \Sigma \) in this way. In any forest of size at most \( \mathbb{M} \) there are at most \( \mathbb{M} \) edges and isolated vertices in total. It follows that the number of galactic shruberies of size at most \( \mathbb{M} \) over a signature \( \Sigma \) is bounded by \( \mathbb{H} = \mathbb{M}^{\mathbb{M}^{-1}} \cdot \mathbb{Z}^{\mathbb{M}^{-1}} \); let
\[ h = \lceil \log([\log H]) \rceil \]. Note that \( H \leq 2^h \). Let \( c \) be a fresh unary predicate, \( d = d_1, \ldots, d_h \) a sequence of fresh unary predicates, and \( e \) a fresh binary predicate. We shall refer to elements satisfying \( c \) as visible (because they can be ‘seen’), and the remainder as invisible. We proceed to write a formula ensuring that every equivalence class containing visible elements also contains a collection of exactly \( 2^h \) invisible elements, and that, moreover, each of these has a unique \( d \)-index. Let \( \Upsilon_h \) be the formula

\[
\forall x \exists y \ (c(x, y) \land x \neq y) \land \\
\forall x \forall y \ (\forall \{e(x, y) \land c(x) \rightarrow c(x, y) \land \neg c(y) \} \forall x = y) \land \\
\forall x \forall y \ (\forall \{x(x, y) \land \neg c(x) \land \neg c(y) \rightarrow \bigvee_{1 \leq i \leq h} \neg (d_i(x) \leftrightarrow d_i(y)) \lor x = y \}). \quad (\Upsilon_h)
\]

stating that every visible element is in the same equivalence class as exactly \( 2^h \) equivalent invisible elements, no two of which have the same \( d \)-index. An a-structure is \( h \)-relativized if it satisfies \( \Upsilon_h \) (and just relativized if we do not care about \( h \)). When we say that a galactic shrubbery is realized by some equivalence class in a relativized a-structure \( \mathfrak{A} \), we shall assume that the realizing elements are all visible; and similarly for cosmic shruberies. In fact, invisible elements will be used in the sequel simply to group together equivalence classes which realize the same galactic shruberies.

Let \( \varphi \) be a \( C^{1}E \)-formula in normal form (1), with amplitude \( C \) and multiplicity \( m \). Let \( c \) be the unary predicate occurring in \( \Upsilon_h \) and let \( f_1, \ldots, f_m \) be fresh binary predicates. The relativization of \( \varphi \), denoted \( \hat{\varphi} \), is the following formula stating, in effect, that \( \varphi \) holds for the visible cosmos.

\[
\forall x \forall y (\varphi(x, y) \lor x = y) \land \\
\bigwedge_{1 \leq h \leq m} \forall x \forall y (\varphi(x, y) \lor (f_h(x, y) \leftrightarrow b_h)) \lor x = y) \land \\
\bigwedge_{1 \leq h \leq m} \forall x \forall y (\varphi(f_h(x, y), c(x)) \lor x = y) \land \\
\bigwedge_{1 \leq h \leq m} \forall x \forall y (\varphi(c(x, y) \land x \neq y) \lor x = y). \quad (\hat{\varphi})
\]

Observe that \( \hat{\varphi} \land \Upsilon_h \) is in normal form with modulus \( c(x, y) \lor f_1(x, y) \lor \cdots \lor f_m(x, y) \), amplitude \( 2^h + C \) and multiplicity \( m + 1 \).

**Lemma 13.** Let \( \varphi \) be a normal-form \( C^{1}E \)-formula over a signature \( \Sigma \) not containing the predicates \( f_1, \ldots, f_m, c, d_1, \ldots, d_h \) or \( e \), and let \( \hat{\varphi} \) and \( \Upsilon_h \) be as above. (i) If \( \mathfrak{B} \models \hat{\varphi} \), then \( \mathfrak{B} \models [\varphi^\mathfrak{B}] \models \varphi \); and (ii) if \( \mathfrak{A} \models \varphi \), then there exists a model \( \mathfrak{B} \models \hat{\varphi} \land \Upsilon_h \) such that \( \mathfrak{A} \) is the \( \Sigma \)-reduct of \( \mathfrak{B} \models [\varphi^\mathfrak{B}] \).

### 6.4 Galactic validity

We now show how galactic validity can be enforced by a sentence in weak normal form. We proceed by constructing, for any galactic shrubbery \( \mathcal{S} \), a formula \( \Gamma(\mathcal{S}) \), with free variable \( x \), which we can use to ensure that a given equivalence class in some structure realizes \( \mathcal{S} \). Write \( \mathcal{S} = (V, E, L) \), and let the size of \( \mathcal{S} \) be \( N \).

The formula in question features a sequence of unary predicates \( p = p_1, \ldots, p_n \), where \( n = \lceil \log(N + 1) \rceil \). For \( 1 \leq i \leq N \), we read \( p(i)(x) \) as “\( x \) realizes the vertex \( i \in V \), and we read \( p(0)(x) \) as “\( x \) does not realize a vertex of \( \mathcal{S} \). We build this formula conjunct by conjunct. We also assume that the signature at
contains visible elements realizing every vertex in \( V \):

\[
\bigwedge_{i=1}^{N} \exists x(y(\epsilon(y) \land \bar{p}(i)(y) \land \epsilon(x, y))).
\]  

(\( \Gamma_5^1 \))

(Recall in this regard that \( V = \{ 1, \ldots, N \} \).) Second, if \( x \) realizes a vertex of \( V \), then it has the 1-type over \( \Sigma \) mandated by the labelling \( L(i) \) (which, remember, is a formula with free variable \( x \)):

\[
\bigwedge_{i=1}^{N} (\bar{p}(i)(x) \rightarrow L(i)).
\]  

(\( \Gamma_5^2 \))

Third, if \( x \) realizes a vertex \( i \) of \( V \), and \( (i, j) \) is an edge of \( S \), then \( x \) is related to any (=the) element \( y \) in its equivalence class realizing the vertex \( j \) by the 2-type over \( \Sigma \) mandated by the labelling \( L(i, j) \) (which, remember, is a formula with free variables \( x, y \)):

\[
\bigwedge_{(i, j) \in E} (\bar{p}(i)(x) \rightarrow \exists_{i=0} y(\epsilon(x, y) \land \bar{p}(j)(y) \land \neg L(i, j))).
\]  

(\( \Gamma_5^3 \))

Fourth, if the galactic, arboreal ray \( \rho \) does not occur as the label of an edge of \( S \), then \( x \) does not send a ray of this type:

\[
\bigwedge \{ \exists_{i=0} y(\rho) \mid \rho \text{ a galactic t-ray-type not in } L(E) \}.
\]  

(\( \Gamma_5^4 \))

Fifth, if \( x \) realizes a root, \( i \), of \( (V, E) \), then \( x \) does not have a mother lying in its equivalence class:

\[
\bigwedge \{ (\bar{p}(i)(x) \rightarrow \exists_{i=0} y(t(x, y) \land \epsilon(x, y))) \mid i \text{ a root of } (V, E) \}.
\]  

(\( \Gamma_5^5 \))

We define \( \Gamma(S) \) to be the conjunction of \( (\Gamma_5^1) - (\Gamma_5^5) \). Observe that the formulas \( (\Gamma_5^1) - (\Gamma_5^3) \) are \( \theta_n \)-eclipsed, where \( \theta_n \) is the formula

\[
\bigvee_{i=1}^{2^n} (\bar{p}(i)(y) \land \epsilon(x, y)),
\]  

(\( \theta_n \))

while \( (\Gamma_5^4) \) and \( (\Gamma_5^5) \) are t-ray-type-eclipsed, since any t-ray-type \( \rho \) contains the conjunct \( t(x, y) \). Hence, the formula \( \Gamma(S) \) (with free-variable \( x \)) is \( (t(x, y) \lor \theta_n) \)-eclipsed. This observation will be used later to establish that a certain formula is in weak normal form. From the foregoing remarks, we have:

**Lemma 14.** Suppose \( \mathfrak{A} \) is a relativized a-structure interpreting a signature \( \Sigma' \supseteq \Sigma \) not containing the unary predicates \( p_1, \ldots, p_n \). Let \( \mathfrak{A} \) be the \( \Sigma \)-reduct of \( \mathfrak{A} \). Let \( S \) be a galactic shrubbery over \( \Sigma \) of size at most \( 2^n - 1 \) and let \( B \) be an equivalence class of \( \mathfrak{A} \) (and hence of \( \mathfrak{A} \)). If \( \mathfrak{A} \) has an extension \( \mathfrak{A}^+ \) such that \( \mathfrak{A}^+ \models \Gamma(S)[a] \) for every \( a \in B \), then \( B \) realizes \( S \) in \( \mathfrak{A} \). Conversely, if \( B \) realizes \( S \) in \( \mathfrak{A} \), then \( \mathfrak{A} \) has an extension \( \mathfrak{A}^+ \) such that \( \mathfrak{A}^+ \models \Gamma(S)[a] \) for every \( a \in B \); for this purpose, it does not matter how \( p_1, \ldots, p_n \) are interpreted outside \( B \).

Lemma 14 ensures that, if \( \mathfrak{A} \) is an a-structure in which every equivalence class \( B \) realizes a galactic shrubbery \( S_B \), we may expand \( \mathfrak{A} \) to a structure \( \mathfrak{A}^+ \) such that, for every such \( B \), every element of \( B \) satisfies \( \Gamma(S_B) \) in \( \mathfrak{A}^+ \). In particular, if the maximum size of any \( S_B \) is \( M \) (as promised by Lemma 10),
we need only \( m \) fresh predicates \( p_1, \ldots, p_m \), where \( m = \lceil \log(M + 1) \rceil \). Defining \( \Xi_m \) to be the sentence

\[
\bigwedge_{i=1}^{2^m} \forall x \exists i \leq |y| (\phi(i)(y) \land c(x, y) \land x \neq y),
\]

(\( \Xi_m \))

we see that \( \mathfrak{A}^+ \models \Xi_m \). Observe that \( m \), like the number \( h \) defined above, is polynomial in \( |\Sigma| \).

Now let \( \mathfrak{A} \models \Upsilon_h \) be an \( h \)-relativized a-structure also interpreting the fresh unary predicate \( d \). Every equivalence class \( B \) of \( \mathfrak{A} \) contains exactly \( 2^h \) invisible elements, say \( a_0, \ldots, a_{2^h-1} \) (numbered according to their \( d \)-indices). Since \( \mathfrak{A} \) interprets the unary predicate \( d \), we may define the invisible coordinate of \( B \) to be the number \( H (0 \leq H < 2^h) \) such that, under the standard \( 2^h \)-bit encoding, the \( z \)-th digit of \( H \) is 1 if \( a_z \) satisfies \( d \), and 0 otherwise. For all \( H \) in this range, define \( d_h(H) \) to be the formula (with free variable \( x \)) given by

\[
\bigwedge_{z=0}^{2^h-1} \exists i = 0 [y((e(x, y) \land d(z)(y)) \land \neg \delta_z(y)),
\]

(\( d_h(H) \))

where \( e \) is the predicate occurring in \( \Upsilon_h \) and, for all \( z \) \((0 \leq z < 2^h) \), \( \delta_z(y) \) is \( d(y) \) if the \( z \)-th digit in the binary representation of \( H \) is 1, and \( \delta_z(y) = \neg d(y) \) otherwise. In structures making the sentence \( \Upsilon_h \) true, and for \( x \) satisfying the predicate \( c \), we may read \( d_h(H) \) as “\( x \) lies in an equivalence class with invisible coordinate \( H \)”.

Thus, the following are \( \forall x \left( \bigvee_{H \in \mathbb{H}} d_h(H)(x) \right) \land \bigwedge_{H \in \mathbb{H}} \forall x (d_h(H)(x) \rightarrow \Gamma(S_H)(x)) \). (\( \Pi_2 \))

Thus, we may read \( \Gamma_{\mathbb{H}} \) as: “The only invisible coordinates realized are those in \( \mathbb{H} \), and any equivalence class with invisible coordinate \( H \in \mathbb{H} \) realizes the galactic shrubbery \( S_H \)”.

6.5 Cosmic validity

We now show how realization of a cosmic shrubbery \( \mathcal{T} \) can be encoded by a sentence \( \Delta_{\mathcal{T}} \) in weak normal form. The encoding is similar in character to the
encoding of galactic shrubberies. (In fact, it is slightly simpler.) Write
\[ \mathcal{T} = (V, E, L, \sim, R_m, R_h, M, \kappa, \lambda_1, \ldots, \lambda_h), \]
and let the size of \( \mathcal{T} \) be \( N \). (Thus, \( V = \{1, \ldots, N\} \).) The encoding sentence features a sequence of unary predicates \( \bar{q} = q_1, \ldots, q_n \), where \( n = \lceil \log(N+1) \rceil \).

For \( 1 \leq i \leq N \), we read \( \bar{q}(i)(y) \) as "\( y \) realizes the vertex \( i \in V \), and we read \( \bar{q}(0)(y) \) as "\( y \) does not realize a vertex of \( \mathcal{T} \). Recall that we refer to elements satisfying the predicate \( c \) as the visible elements. We build \( \Delta_{\mathcal{T}} \) conjunct by conjunct. Firstly, the \( a \)-structure contains visible elements realizing every vertex in \( V \):
\[
\bigwedge_{i=1}^{N} \exists x \bar{q}(i)(x) \land c(x)). \quad (\Delta_1')
\]
Second, if an element realizes vertex \( i \) of \( V \), then it has the 1-type over \( \Sigma \) mandated by the labelling \( L(i) \):
\[
\bigwedge_{i=1}^{N} \forall x \bar{q}(i)(x) \rightarrow L(i)). \quad (\Delta_2')
\]
Third, if a pair of elements realize vertices \( i \) and \( j \) of \( V \), and \( (i, j) \) is a edge of \( \mathcal{T} \), then those elements are related as mandated by the labelling \( L(i, j) \):
\[
\bigwedge_{(i, j) \in E} \forall x \bar{q}(i)(x) \rightarrow \forall y \bar{q}(j)(y) \rightarrow L(i, j)). \quad (\Delta_3')
\]
Fourth, if the cosmic, arboreal ray \( \rho \) does not occur as the label of an edge of \( \mathcal{T} \), then it does not occur at all:
\[
\bigwedge \{ \forall x \forall y \lnot \rho \mid \rho \text{ a cosmic, arboreal ray-type not in } L(E) \}. \quad (\Delta_4')
\]
Fifth, if an element realizes a root, \( i \), of \( (V, E) \), then it is a root in the structure:
\[
\forall x \bigwedge \{ (\bar{q}(i)(x) \rightarrow v(x) \mid i \text{ a root of } (V, E) \}. \quad (\Delta_5')
\]
We define the sentence \( \Delta_{\mathcal{T}} \) to be the conjunction of \((\Delta_1')-(\Delta_5')\). Bearing in mind that \((\Delta_1')-(\Delta_5')\) is logically equivalent to \( \bigwedge_{i=1}^{N} \forall x \exists y \lnot (\bar{q}(i)(y) \land c(y)) \lor x \neq y \land \forall y \bigwedge_{i=1}^{N} \bar{q}(i)(y) \land c(y) \), we may regard \( \Delta_{\mathcal{T}} \) as a weak normal-form formula with modulus \( \forall \bar{q}(i)(y) \land c(y) \) and amplitude \( N \). From the foregoing remarks, we have:

**Lemma 16.** Let \( \mathfrak{A}' \) be an \( a \)-structure interpreting a signature \( \Sigma' \supseteq \Sigma \) not containing the unary predicates \( q_1, \ldots, q_n \), and let \( \mathfrak{A} \) be the \( \Sigma \)-reduct of \( \mathfrak{A}' \). Let \( \mathcal{T} \) be a cosmic shrubbery over \( \Sigma \) of size at most \( 2^n - 1 \). Then \( \mathfrak{A} \) realizes \( \mathcal{T} \) if and only if \( \mathfrak{A}' \) has an expansion \( \mathfrak{A}'^+ \) such that \( \mathfrak{A}'^+ \models \Delta_{\mathcal{T}} \).

In particular, if the size of \( \mathcal{T} \) is \( N \) (as promised by Lemma 10), we need only \( n \) fresh predicates \( q_1, \ldots, q_n \), where \( n = \lceil \log(N+1) \rceil \).

### 6.6 Summary

A brief summary of the results of this section will help with orientation. Let \( \phi \) be a \( C^2 \)DIE-formula with modulus \( \theta \) over a signature \( \Sigma \), let \( d, h, m, n \) be integers, let \( H \) be a set of indices of galactic shrubberies of size at most \( 2^m \) over \( \Sigma \) (using some standard enumeration \( S_0, S_1 \ldots \)), and let \( \mathcal{T} \) be a cosmic shrubbery of size at most \( 2^n \) over \( \Sigma \). We have shown how to construct a collection of seven \( C^2 \)-formulas, which, roughly speaking, we may interpret with respect to any (finite) structure \( \mathfrak{A} \) as follows.
Table 1: Formulas defined in Sec. 6 giving the additional predicate symbols introduced, modulus and amplitude; note that $\Delta_T$ is in weak normal form.

<table>
<thead>
<tr>
<th>Formula</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Omega$</td>
<td>The structure $\mathfrak{A}$ is arboreal.</td>
</tr>
<tr>
<td>$\Pi_\theta$</td>
<td>The $\Sigma$-reduct of $\mathfrak{A}$ is $\theta$-parental.</td>
</tr>
<tr>
<td>$\Upsilon_h$</td>
<td>Every visible element of $\mathfrak{A}$ is in the same equivalence class as exactly $2^h$ invisible elements, no two of which have the same $d$-index.</td>
</tr>
<tr>
<td>$\dot{\varphi}$</td>
<td>The visible elements of $\mathfrak{A}$ form a model of $\varphi$.</td>
</tr>
<tr>
<td>$\Xi_m$</td>
<td>The various Boolean combinations of the predicates $p_1, \ldots, p_m$, when realized, unambiguously index elements of galactic shruberies in equivalence classes in $\mathfrak{A}$.</td>
</tr>
<tr>
<td>$\Delta_T$</td>
<td>The structure $\mathfrak{A}$ realizes the cosmic shrubbery $T$.</td>
</tr>
<tr>
<td>$\Gamma_H$ (in conjunction with $\Upsilon_h$)</td>
<td>Every equivalence class of $\mathfrak{A}$ has an invisible coordinate in the set $H$, and any equivalence class with invisible coordinate $H \in H$ realizes the galactic shrubbery $S_H$.</td>
</tr>
</tbody>
</table>

The conjunction of the first six of these formulas, $\dot{\varphi} \land \Pi_\theta \land \Upsilon_h \land \Omega \land \Xi_m \land \Delta_T$, is a $C^2$-formula in weak normal form with modulus

$$\theta := \bigvee_{h=1}^m f_h(x, y) \lor e(x, y) \lor \bigvee_{i=1}^{2^m} (\bar{p}(i)(y) \land e(x, y)) \lor t(x, y) \lor r(x, y) \lor \bigvee_{i=1}^{2^n} \bar{q}(i)(y)$$

such that the seventh formula, $\Gamma_H$, has the form $\forall x.\epsilon$, with $\epsilon$ $\theta$-eclipsed. Thus, the conjunction of all seven formulas is in weak normal form. We need to be scrupulously careful about the sizes of the signatures of these formulas, as well as their moduli and amplitudes, so as to be able usefully to apply the parametrized complexity bounds given in Corollary 2. Table 1 summarizes these metrics. (We remark that $\Gamma_H$ features one additional unary predicate, $d$.)

7 Main result

We are now in a position to assemble the above material into a proof of the main result. Remember that we are assuming all structures to be finite.

**Theorem 2.** The satisfiability problem for the logic $C^21D1E$ is $\text{NExpTime}$-complete.
\[
\begin{align*}
\text{small galactic and} & \quad \text{few star-types} \\
\text{cosmic shrubberies} & \quad \theta\text{-parental} \\
\text{galactically valid} & \quad \text{expanded and} \\
(3,2)\text{-cosmically valid} & \quad \text{relativized}
\end{align*}
\]

Figure 7: If-direction of proof of Theorem 2.

Proof. Hardness for NExpTime is immediate from the fact that \(C^2\text{D1E} \in \text{FO}^2\). We consider only membership in NExpTime.

Let a \(C^2\text{D1E}\)-formula \(\varphi\) be given. By Lemma 1, we may assume without loss of generality that \(\varphi\) is in normal form, as given in (1). Let \(\varphi\) have amplitude \(C\) and multiplicity \(m\), and let \(\Sigma_0\) be the signature of \(\varphi\) (which we may assume contains \(e, t\) and \(r\)). Let \(\Sigma\) be \(\Sigma_0\) together with \(\lceil \log(2C+1) \rceil\) fresh unary predicates and 8 fresh binary predicates. It will help to think of \(\Sigma\) as our ‘base signature’ in what follows—the signature in which the important parts of the proof take place—even though we shall also consider larger signatures. As usual, let \(Z = 2^{|\Sigma|}, M = 5Z^2, N = 500Z^4, H = M^{M-1}Z^M, h = \lceil \log(\lceil \log H \rceil) \rceil, m = \lceil \log(M+1) \rceil\) and \(n = \lceil \log(N+1) \rceil\). Let all the galactic shrubberies over \(\Sigma\) of size no greater than \(M\) be enumerated \(S_0, \ldots, S_{H-1}\) and write \(H^* = \{0, \ldots, H-1\}\). Finally, let \(C^* = C + 2h + 2m + 2n + 2\), and \(s^* = 2|\Sigma| + m + h + m + n + 4\). Recalling the function \(f\) provided by Corollary 2, execute the following procedure.

1. Guess a subset \(\mathbb{H} \subseteq H^*\) satisfying \(|\mathbb{H}| \leq f(2s^* + C^*)\);
2. Guess a cosmic shrubbery \(T\) over \(\Sigma\) of size at most \(N\);
3. Let \(\varphi^*\) be the \(C^2\)-formula \(\dot{\varphi} \wedge \Upsilon_h \wedge \Pi_\theta \wedge \Xi_m \wedge \Gamma_{H} \wedge \Delta_T \wedge \Omega\).
4. Non-deterministically test finite satisfiability of \(\varphi^*\) and return the result.

We have already observed that \(\Omega \wedge \Upsilon_h \wedge \Xi_m \wedge \Gamma_{H}\) is in weak normal form; moreover, \(\dot{\varphi} \wedge \Pi_\theta\) is in normal form and \(\Delta_T\) is in weak normal form. Hence \(\varphi^*\) is in weak normal form. The signature of \(\varphi^*\) is of size at most \(s^*\), the modulus is \(\theta^*\) given by

\[
f_1(x, y) \lor \cdots \lor f_m(x, y) \lor t(x, y) \lor r(x, y) \lor e(x, y) \lor \bigvee_{i=1}^{m}(\tilde{p}(i)(y) \land e(x, y)) \lor \bigvee_{i=1}^{n}(\tilde{q}(i)(y)),
\]

and the amplitude is at most \(C^*\). By Corollary 2, then, the procedure runs in (non-deterministic) time bounded by an exponential function of \(|\varphi|\). We show that it has a successful run if and only if \(\varphi\) has a dendral model, thus proving the theorem. For each direction of the bi-conditional, we proceed via a series of stages illustrated in Figs. 7 (if-direction) and 8 (only-if direction).

If-direction: Suppose first that \(\mathcal{D}_0\) is a dendral model of \(\varphi\) interpreting the signature \(\Sigma_0\) of \(\varphi\). By Lemma 2 and Lemma 5, expand \(\mathcal{D}_0\) to a \(\theta\)-parental,
that $H$ has been replaced by the potentially much larger $\Gamma$ equivalence class of $A$ in $C$. $\psi$ is therefore finitely satisfiable. (Note that $\psi$ is the same as $\varphi^*$ except that $\Gamma_\Sigma$ has been replaced by the potentially much larger $\Gamma_{\Sigma'}$.) A simple check shows that $\psi^*$ has a signature $\Sigma'$ of size $s^* = 2|\Sigma| + m + h + n + 4$ and amplitude $C^* = C + 2h + 2m + 2n + 2$. Thus, $s^*$ is polynomial in $|\varphi|$, while $C^*$ is exponential in $|\varphi|$.

By the second statement of Corollary 2, $\psi^*$ therefore has a model, say $\mathfrak{A}^*$, realizing at most $f(2s^* + C^*)$ $\theta^*$-star-types (over $\Sigma'$). Since $\mathfrak{A}^* = \mathfrak{T}_h$, each equivalence class of $\mathfrak{A}^*$ has a well-defined dark coordinate. But only $f(2s^* + C^*)$ $\theta^*$-star-types (over $\Sigma'$), are realized in $\mathfrak{A}^*$. And since the invisible coordinate of any equivalence class is determined by the $\theta^*$-star-types of any of its elements, it follows that the set of visible coordinates occurring in $\mathfrak{A}^*$ has cardinality at most $f(2s^* + C^*)$, whence there is an $H \subseteq \mathbb{H}^*$ such that $|H| \leq f(2s^* + C^*)$, for which the $C^2E$-formula $\varphi^*$ is finitely satisfiable. Hence the procedure has a successful run.

Only-if direction: Suppose, conversely, that for some choices of $\mathbb{H}$ and $T$, the procedure has a successful run. Thus, the formula $\varphi^*$ has a model $\mathfrak{A}_2$. Since $\mathfrak{A}_2 \models \Omega$, $\mathfrak{A}_2$ is an $a$-structure, by Lemma 11. Let $\mathfrak{A}_1$ be the $\Sigma$-reduct of $\mathfrak{A}_2$. By Lemma 12, $\mathfrak{A}_1$ is $\theta$-parental, and by Lemmas 15 and 16, each equivalence class

Figure 8: Only-if direction of proof of Theorem 2.
in $\mathfrak{A}_1$ realizes a galactic shrubbery, and $\mathfrak{A}_1$ realizes a cosmic shrubbery. Let $\mathfrak{A}$ be the substructure of $\mathfrak{A}_1$ consisting of the visible elements (elements satisfying $c$), so that, by Lemma 13, $\mathfrak{A} \models \varphi$. Since shrubberies in relativized structures are by definition realized by visible elements, we see from Lemmas 8 and 9 that $\mathfrak{A}$ is galactically and cosmically valid. Certainly, $\mathfrak{A}$ is $\theta$-parental. By Lemma 7, there exists a $d$-structure $\mathfrak{D}$ over the same domain as $\mathfrak{A}$ such that $\mathfrak{D}$ and $\mathfrak{A}$ realize the same 2-types, and such that the $\theta$-star-type of every element is the same in both structures. Hence $\mathfrak{D} \models \varphi$. This proves the theorem.

\section{8 Conclusion}

We showed that the satisfiability problem for two-variable logic $C^21D1E$ with counting quantifiers, an equivalence relation and a tree-relation (interpreted over finite structures) is NExpTime-complete. Our decision procedure is based on a reduction to the finite satisfiability problem for two-variable logic $C^21E$ with counting and equivalence (without trees), using the latter essentially as a black-box. Such an approach significantly simplifies the earlier algorithm for finite satisfiability of $C^2$ with trees from [4], which re-uses many technical details of the original satisfiability algorithm for $C^2$ [19].

Further development of these techniques may lead to efficient decision procedures for other extensions of $C^2$. For example, we expect that $C^21D1E$ supplemented with next-sister relation can be decided in NExpTime by a reduction to $C^21D1E$ (without next-sister), thus improving 3-NExpTime upper bound by Bojańczyk et al. for $FO^2$ with mother-daughter, next-sister and data-equivalence [2].

\section*{Acknowledgements}

This work is supported by the Polish National Science Centre grants [2016/21/B/ST6/01444 to W.C and P.W.] and [2018/31/B/ST6/03662 to I.P-H.].

\section*{References}


