Towards Combining Dense Linear Order with Random Graph

Jiamou Liu[1], Ting Zhang[2]

[1] University of Auckland, Department of Computer Science
jliu036@ec.auckland.ac.nz
[2] Microsoft Research Asia, Theory Group
tingz@microsoft.com

Abstract. In this paper we present our work in progress towards obtaining a Nelson-Oppen style combination for combining quantified theories, where each individual component theory admits quantifier elimination. We introduces the notion of *good model* for union theories, for which there exists a simple quantifier elimination scheme that uses the elimination procedures for individual component theories as black boxes. We show that a good model exists for the union theory of dense linear order and random graph, and it coincides with the Fraïssé limit of the class of finite graph with ordered vertices.

1 Introduction

In 1979, Nelson and Oppen [1] proposed a framework for combining decision procedures on quantifier-free formulas: if theories T_1 and T_2 are stably infinite, over disjoint signatures and stably infinite, then one can obtain a decision procedure for the quantifier-free fragment of the union theory $T_1 \cup T_2$, using the decision procedures for T_1 and T_2 as modules. Ever since the foundational work of Nelson and Oppen, researchers have been asking the general question: *under what condition do we have a combination method for arbitrary first-order (not necessarily quantifier-free) theories?* Recently, a lot of progresses have been made to relax the conditions on component theories to be combined [2, 3, 4], as well as to obtain Nelson-Oppen like results for other combination problems such as many-sorted logic [5], modal systems [6], and abstract interpreters [7].

In this paper we consider a restricted version of the question: providing that two theories T_1 and T_2 both admit quantifier elimination, does the union theory $T_1 \cup T_2$ also admit quantifier elimination? If it does, can we find an elimination procedure for $T_1 \cup T_2$, using the elimination procedures for T_1 and T_2 as modules?

Suppose for $i \in \{1, 2\}$, T_i is an L_i theory and φ_i is a conjunction of L_i-literals. To eliminate the existential quantifier in $\exists y(\varphi_1(\bar{x}, y) \wedge \varphi_2(\bar{x}, y))$, it is desirable to have

$$T_1 \cup T_2 \models \forall \bar{x}\ [(\exists y\, \varphi_1(\bar{x}, y)\ \wedge\ \exists\, \varphi_2(\bar{x}, y)) \leftrightarrow \exists y\ (\varphi_1(\bar{x}, y))\ \wedge\ \varphi_2(\bar{x}, y))]\,. \quad (1)$$

However, Condition (1) does not hold in general, as shown by the following example.

Example 1 (Incompatible Dense Linear Orders) *For $i \in \{1, 2\}$, let L_i be the signature $\{<_i\}$, and let T_i be the L_i-theory of dense linear orders. Consider $\mathcal{A} = \langle A, <_1, <_2 \rangle$ where A is the set of rational numbers, and $<_1$, $<_2$ are such that for any $u, v \in A$, $u <_1 v$ iff $v <_2 u$ iff $u < v$. Obviously $\mathcal{A}$ is a model of $T_1 \cup T_2$. However, for any $a \in A$, $\mathcal{A} \models \exists x\,(x <_1 a)$, $\mathcal{A} \models \exists x\,(x <_2 a)$, but $\mathcal{A} \not\models \exists x\,(x <_1 a \wedge x <_2 a)$.*

In this paper we consider only the models of $T_1 \cup T_2$ that satisfy Condition (1). Using a priority argument, we show that such models do exist for the union theory of dense linear order and random graph, and hence we obtain a decision procedure for the union theory restricted to those models.

Linear orders and graphs are fundamental objects in computer science and mathematics. Dense linear order (countable, without endpoints) is essentially the structure of rational numbers under natural order. Random graph captures the almost sure theory of finite graphs [8], i.e., $\{\phi \mid \lim_{n\to\infty} p_n(\phi) = 1\}$, where $p_n(\phi)$ is the probability of a graph with n vertices satisfying sentence ϕ. Also dense linear order (without endpoints) is the Fraïssé limit of the class of all finite linear orders and random graph is the Fraïssé limit of the class of all finite graphs [9]. We show that by combining the theory of these two Fraïssé limits, one can obtain the Fraïssé limit for all finite models of the combined theory of linear orders and graphs. This itself is an interesting phenomenon in theory combination and we believe that it deserves further investigations.

Paper Organization Section 2 provides basic notions and terminology in model theory, and introduces some notations in our presentation. Section 3 proves the existence of good models for the union theory of dense linear order and random graphs. Section 4 provides a further discussion on the properties of good models. Section 5 concludes with a discussion of complexity and future work.

2 Preliminary

In this section we introduce notions and terminology used in this paper. We assume the first-order syntactic notions of variables, parameters and quantifiers, and semantic notions of structures, satisfiability and validity as in [10].

Basic Notations. We use $\mathbb{N}$ to denote the set of natural numbers, and $\mathbb{Q}$ the set of rational numbers. We use $\overline{u}$ to denote the sequence $u_1, \ldots, u_n$ (for some $n > 0$). We abuse notation a bit by also using $\overline{u}$ to denote the set that consists of elements in the sequence. For example, by $\overline{u} \in S$ we mean that all elements in $\overline{u}$ are contained in S. The meaning should be clear from the context. Also by $(u_i)_{i<\omega}$ we mean an infinite enumeration of the form $u_0, u_1, \ldots$.

By default we use calligraphic letters $\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$, to denote structures and the capital letters $A, B, C, \ldots$, to denote the corresponding domains. For example, a model of graph is denoted by $\mathcal{G} = \langle G, E^{\mathcal{G}} \rangle$. When there is no confusion, we drop superscripts on function symbols and predicate symbols.

We use $\mathcal{A} \cong \mathcal{B}$ to mean that $\mathcal{A}$ and $\mathcal{B}$ are isomorphic. We use $\mathcal{A} \subset \mathcal{B}$ to mean that $\mathcal{A}$ properly embeds into $\mathcal{B}$, i.e., $\mathcal{A}$ is isomorphic to a proper substructure of $\mathcal{B}$. For a structure $\mathcal{A}$ and a tuple $\overline{a} \in A$, whenever we use $\overline{a}$ in variable substitution, it should be understood that the underlying language is extended with constants $\overline{a}$, each of which names itself in the extended structure $\mathcal{A}' = (\mathcal{A}, \overline{a})$.

Dense Linear Order. A *dense linear order* (*DLO*) without endpoints is a linear order $\mathcal{D} = \langle D, < \rangle$ such that there is no minimal or maximal element and

$$\forall x, y \in D\ (x < y \rightarrow \exists z\,(x < z \wedge z < y))\,. \tag{2}$$

Let $L_{\mathcal{D}}$ denote the language of $\mathcal{D}$ and $T_{\mathcal{D}}$ the theory of $\mathcal{D}$. It is well-known that $T_{\mathcal{D}}$ is ω-categorical, complete and decidable, and it admits quantifier elimination. [9]. In particular, the linear order on rational numbers, denoted by $\mathcal{Q} = \langle \mathbb{Q}, <^{\mathcal{Q}} \rangle$, is the unique countable model of $T_{\mathcal{D}}$ up to isomorphism. In the paper we identify $\mathcal{Q}$ with $\mathcal{D}$.

Lemma 1. *For any conjunction of* positive *$L_{\mathcal{Q}}$-literals $\Phi(\overline{x}, y)$, where y does not appear in equalities, for any $\overline{a} \in \mathbb{Q}$, if $\mathcal{Q} \models \exists y \Phi(\overline{a}, y)$, then there are infinitely many $b \in \mathbb{Q}$ such that $\mathcal{Q} \models \Phi(\overline{a}, b)$.*

Proof. Let $\Phi(\overline{x}, y)$ be a conjunction of *positive* $L_{\mathcal{Q}}$-literals, where y does not appear in equalities, and let $\overline{a}$ be any tuple in $\mathcal{Q}$. Note that $\Phi(\overline{a}, y)$ states that y is contained in the intersection of finitely many open intervals whose boundaries are elements in $\overline{a}$. Since the intersection of finitely many open intervals is an open interval, if there is a solution of $\Phi(\overline{a}, y)$, then by the denseness property of $\mathcal{Q}$, there exist infinitely many such solutions. $\square$

Random Graph. A *Random Graph* (*RG*) is a countable graph $\mathcal{G} = \langle G, E \rangle$ such that for any $n, m > 0$,

$$\forall x_1 \ldots \forall x_n \forall y_1 \ldots \forall y_m \left(\bigwedge_{i=1}^{n} \bigwedge_{j=1}^{m} x_i \neq y_j \;\rightarrow\; \exists z \left(\bigwedge_{i=1}^{n} E(x_i, z) \;\wedge\; \bigwedge_{j=1}^{m} \neg E(y_j, z) \right) \right)\,. \tag{3}$$

Let $L_{\mathcal{G}}$ denote the language of $\mathcal{G}$ and $T_{\mathcal{G}}$ the theory of random graph. Like $T_{\mathcal{D}}$, $T_{\mathcal{G}}$ is ω-categorical, complete and decidable, and it admits quantifier elimination [9]. The definition is in line with the standard construction of a random graph whose edges are defined independently on pairs of vertices with probability $\frac{1}{2}$.

Lemma 2. *For any conjunction of $L_{\mathcal{G}}$-literals $\Phi(\overline{x}, y)$, where y does not appear in equalities, for any $\overline{a} \in G$, if $\mathcal{G} \models \exists y \Phi(\overline{a}, y)$, then there are infinitely many $b \in G$ such that $\mathcal{G} \models \Phi(\overline{a}, b)$.*

Proof. Let $\Phi(\overline{x}, y)$ be a conjunction of $L_{\mathcal{G}}$-literals, where y does not appear in equalities, and let $\overline{a} = a_1, \ldots, a_n$ be any tuple in G. Then $\Phi(\overline{a}, y)$ is of the form

$$\bigwedge_{i=1}^{s} E(a_i, y) \ \wedge \ \bigwedge_{j=s+1}^{n} \neg E(a_j, y) \ \wedge \ \bigwedge_{b \in P} y \neq b \ \wedge \ \Phi'(\overline{a}) \ , \tag{4}$$

where $P \subseteq \overline{a}$, y does not appear in Φ' and $s \leq n$. Since $\mathcal{G} \models \exists y \Phi(\overline{a}, y)$, we have $\bigwedge_{i=1}^{s} \bigwedge_{j=s+1}^{n} a_i \neq a_j$ and $\mathcal{G} \models \Phi'(\overline{a})$. Now take a finite set $S \subseteq G$ such that $S \cap \overline{a} = \emptyset$. Then by (3) we have, for any $S' \subseteq S$,

$$\mathcal{G} \models \exists y \left(\bigwedge_{i=1}^{s} (E(a_i, y) \wedge \bigwedge_{b \in S'} E(b, y) \wedge \bigwedge_{j=s+1}^{n} \neg E(a_j, y) \wedge \bigwedge_{b \in S \setminus S'} \neg E(b, y) \right) \ , \tag{5}$$

which tells us that there are at least $2^{|S|}$ distinct witnesses to $\bigwedge_{i=1}^{s} E(a_i, y) \wedge \bigwedge_{j=s+1}^{n} \neg E(a_j, y)$, and hence at least $2^{|S|} - |P|$ solutions to (4). In fact there must be infinitely many solutions to (4) because S can be arbitrarily large. ◻

3 Combining Dense Linear Order with Random Graph

In this section we present a model of $T_{\mathcal{D}} \cup T_{\mathcal{G}}$ which admits quantifier elimination.

Lemma 3. *There exists a model $\mathcal{A} = \langle A, <^{\mathcal{A}}, E^{\mathcal{A}} \rangle$ of $T_{\mathcal{D}} \cup T_{\mathcal{G}}$ such that for any conjunction of* positive *$L_{\mathcal{D}}$-literals $\Phi(\overline{x}, y)$, and for any conjunction of $L_{\mathcal{G}}$-literals $\Psi(\overline{x}, y)$, if y does not occur in equality in either Φ or Ψ, then*

$$\mathcal{A} \models \forall \overline{x} \left[(\exists y \Phi(\overline{x}, y) \wedge \exists y \Psi(\overline{x}, y)) \leftrightarrow \exists y \left(\Phi(\overline{x}, y) \wedge \Psi(\overline{x}, y) \right) \right] \ . \tag{6}$$

Proof. We first outline our construction idea for $\mathcal{A}$. Then we present the detailed construction. Finally we prove that $\mathcal{A}$ is our desired model.

Construction Plan. The direction "$\leftarrow$" is obvious as it holds for any models. The other direction is considerably involved. We construct an infinite ascending chain of finite structures, $\mathcal{A}_0 \subset \mathcal{A}_1 \subset \ldots$, whose limit is our desired $\mathcal{A}$, i.e., $\mathcal{A} = \bigcup_{i \in \mathbb{N}} \mathcal{A}_i$. The domain A of $\mathcal{A}$ consists of tuples of the form (u, v) where $u \in \mathbb{Q}$ and $v \in G$. Moreover, every $u \in \mathbb{Q}$ and every $v \in G$ appear in exactly one tuple in A. Essentially we construct an infinite ascending chain of functions $f_0 \subset f_1 \subset \ldots$, where each f_i is a $1-1$ partial function from $\mathbb{Q}$ to G. Let $dom(f_i)$ denote the effective domain of f_i. For each $i \in \mathbb{N}$, f_i induces $\mathcal{A}_i = \langle A_i, <^{\mathcal{A}_i}, E^{\mathcal{A}_i} \rangle$ as follows.

$$A_i = \{ (u, f_i(u)) \in \mathbb{Q} \times G \mid u \in dom(f_i) \} \tag{7}$$

$$<^{\mathcal{A}_i} = \{ ((u, v), (u', v')) \in A_i \times A_i \mid u <^{\mathcal{Q}} u' \} \tag{8}$$

$$E^{\mathcal{A}_i} = \{ ((u, v), (u', v')) \in A_i \times A_i \mid E^{\mathcal{G}}(v, v') \} \tag{9}$$

Note that the limit of this chain is a bijective function $f : \mathbb{Q} \to G$ which induces $\mathcal{A} = \langle A, <^{\mathcal{A}}, E^{\mathcal{A}} \rangle$ in the same way as defined above.

The essential construction from stage i to stage $i+1$ is to, for each tuple $\bar{a} \in \mathcal{A}_i$, find witnesses for formulas of the form $\exists y(\Phi(\bar{a}, y) \wedge \Psi(\bar{a}, y))$, providing that both $\exists y \Phi(\bar{a}, y)$ and $\exists y \Psi(\bar{a}, y)$ hold separately in $\mathcal{A}_i$. Obviously, at a single stage we might not find witnesses for all pairs of formulas of the form $(\Phi(\bar{x}, y), \Psi(\bar{x}, y))$ as there are infinitely many such pairs. However, by a standard encoding technique we make sure that witnesses for every such pair will eventually be discovered at a certain stage. We present the detailed construction as follows, which is essentially a priority argument.

Construction. Let $(\Phi_i)_{i<\omega}$ be an enumeration of all finite conjunctions of *positive* $L_{\mathcal{D}}$-literals of the form $\varphi(\bar{x}, y)$ where y does not appear in equalities, and $(\Psi_i)_{i<\omega}$ an enumeration of all finite conjunctions of $L_{\mathcal{G}}$-literals of the form $\psi(\bar{x}, y)$ where y does not appear in equalities. Note that such enumerations exist since both $L_{\mathcal{D}}$ and $L_{\mathcal{G}}$ are countable languages. Let $\langle \cdot, \cdot \rangle : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ be a pairing function (i.e., a bijection from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$), and $l : \mathbb{N} \to \mathbb{N}$, $r : \mathbb{N} \to \mathbb{N}$ be the corresponding projection functions such that for any $n \in \mathbb{N}$, $\langle l(n), r(n) \rangle = n$. This pairing function is used to enumerate $\{(\Phi_i, \Psi_j) \mid i, j \in \mathbb{N}\}$. Also let $(u_i)_{i<\omega}$ be an enumeration of $\mathbb{Q}$, and $(v_i)_{i<\omega}$ an enumeration of G.

Let $f_0 = \emptyset$ and hence $\mathcal{A}_0$ be an empty structure. Suppose f_i and $\mathcal{A}_i$ have been obtained. We run Algorithm 1 to obtain f_{i+1} and $\mathcal{A}_{i+1}$.

Algorithm 1 Construction of $\mathcal{A}_{i+1}$.

1: Set $f_{i+1} = f_i$.
2: Find the first *unused* element $u \in (u_i)_{i<\omega}$ and the first *unused* element $v \in (v_i)_{i<\omega}$. Mark u, v as *used*. Set $f_{i+1} = f_{i+1} \cup (u, v)$.
3: **for all** $\bar{a} \in \mathcal{A}_i$ and $j \leq i$ **do**
4:     **if** $\mathcal{A}_i \models \exists y \Phi_{l(j)}(\bar{a}, y) \wedge \mathcal{A}_i \models \exists y \Psi_{r(j)}(\bar{a}, y)$ **then**
5:         Find the first *unused* element $u \in (u_i)_{i<\omega}$ such that $\mathcal{Q} \models \Phi_{l(j)}(\bar{a}, u)$ and the first *unused* element $v \in (v_i)_{i<\omega}$ such that $\mathcal{G} \models \Psi_{r(j)}(\bar{a}, v)$. Mark u, v as *used*. Set $f_{i+1} = f_{i+1} \cup (u, v)$.
6:     **end if**
7: **end for**

Proof Continued. We show by induction that for each $i \in \mathbb{N}$, Algorithm 1 is sound and terminates, and for each $i \in \mathbb{N}$, $\mathcal{A}_i \subset \mathcal{A}_{i+1}$, $\mathcal{A}_i^{L_{\mathcal{D}}} \subset \mathcal{Q}$, $\mathcal{A}_i^{L_{\mathcal{G}}} \subset \mathcal{G}$, and

$$\forall j \leq i \, \forall \bar{a} \in A_i \Big[\Big(\mathcal{A}_i \models \exists y \Phi_{l(j)}(\bar{a}, y) \ \wedge \ \mathcal{A}_i \models \exists y \Psi_{r(j)}(\bar{a}, y) \Big) \Rightarrow \mathcal{A}_{i+1} \models \exists z \left(\Phi_{l(j)}(\bar{a}, z) \ \wedge \ \Psi_{r(j)}(\bar{a}, z) \right) \Big]. \tag{10}$$

The case $i = 0$ is trivial. By Step (2), $f_i \subset f_{i+1}$ and hence $A_i \subset A_{i+1}$. By (7)-(9), we have $\mathcal{A}_i^{L_{\mathcal{D}}} \subset \mathcal{Q}$ and $\mathcal{A}_i^{L_{\mathcal{G}}} \subset \mathcal{G}$. Now $\mathcal{A}_i \models \exists y \Phi_{l(j)}(\bar{a}, y)$ implies $\mathcal{A}_i^{\mathcal{D}} \models \exists y \Phi_{l(j)}(\bar{a}, y)$, which in turn implies $\mathcal{Q} \models \exists y \Phi_{l(j)}(\bar{a}, y)$. Similarly, we have $\mathcal{A}_i^{\mathcal{G}} \models \exists y \Psi_{r(j)}(\bar{a}, y)$ implies $\mathcal{G} \models \exists y \Psi_{r(j)}(\bar{a}, y)$. Therefore, Step (5) can be realized due to Lemma 1 and Lemma 2. The termination of Algorithm 1 follows because there are only finitely many $\bar{a} \in A_i$ and $j \leq i$. Property (10) holds obviously thanks to Step (5).

Since Step (2) pairs elements in Q with elements in G according to the enumerations $(u_i)_{i<\omega}$ and $(v_i)_{i<\omega}$, eventually every element in $\mathbb{Q}$ is paired with one element in G, and vice versa. Therefore, we have $\mathcal{A}^{L_{\mathcal{D}}} \cong \mathcal{Q}$ and $\mathcal{A}^{L_{\mathcal{G}}} \cong \mathcal{G}$, and hence $\mathcal{A}$ is a model of $T_{\mathcal{D}} \cup T_{\mathcal{G}}$.

Let $\Phi \equiv \Phi_i$ and $\Psi \equiv \Psi_j$ for some $i, j \in \mathbb{N}$, and $\bar{a}$ be an arbitrary tuple in A. Suppose that $\mathcal{A} \models \Phi(\bar{a}, u) \wedge \Psi(\bar{a}, v)$ for some $u, v \in A$. Take $k \in \mathbb{N}$ such that $k > \langle i, j \rangle$, and $\bar{a}, u, v \in A_k$. We have

$$\begin{aligned}
\mathcal{A} \models \Phi(\bar{a}, u) \wedge \Psi(\bar{a}, v) \quad &\Rightarrow \quad \mathcal{A}_k \models \Phi(\bar{a}, u) \wedge \Psi(\bar{a}, v) \\
&\Rightarrow \quad \mathcal{A}_k \models \exists y \Phi(\bar{a}, y) \wedge \exists y \Psi(\bar{a}, y) \\
&\Rightarrow \quad \mathcal{A}_{k+1} \models \exists y\, (\Phi(\bar{a}, y) \wedge \Psi(\bar{a}, y)) \\
&\Rightarrow \quad \mathcal{A} \models \exists y\, (\Phi(\bar{a}, y) \wedge \Psi(\bar{a}, y)) \qquad \square
\end{aligned}$$

We call the models that satisfy Lemma 3 *good models* of $T_{\mathcal{D}} \cup T_{\mathcal{G}}$. Let $(T_{\mathcal{D}} \cup T_{\mathcal{G}})_{GOOD}$ be the theory of all good models of $T_{\mathcal{D}} \cup T_{\mathcal{G}}$.

Theorem 2. *$(T_{\mathcal{D}} \cup T_{\mathcal{G}})_{GOOD}$ admits quantifier elimination.*

Proof. It suffices to show that one can eliminate $\exists y$ from formulas of the form $\exists y \varphi(\bar{x}, y)$ where $\varphi(\bar{x}, y)$ is a conjunction of literals. Since $L_{\mathcal{D}} \cup L_{\mathcal{G}}$ contains no function symbols, any such $\exists y \varphi(\bar{x}, y)$ can be rewritten as

$$\exists y \ (\Phi(\bar{x}, y) \wedge \Psi(\bar{x}, y)) \tag{11}$$

where $\Phi(\bar{x}, y)$ is conjunction of $L_{\mathcal{D}}$-literals, $\Psi(\bar{x}, y)$ is a conjunction of $L_{\mathcal{G}}$-literals. We further assume that $\Phi(\bar{x}, y)$ contains only positive literals as $\neg(x < y)$ can be replaced by $x = y \vee y < x$. We also assume that y does not appear in equalities (otherwise the elimination of $\exists y$ is trivial). Now $\Phi(\bar{x}, y)$ and $\Psi(\bar{x}, y)$ satisfy the requirements in Lemma 3. So (11) can be rewritten as

$$\exists y\, \Phi(\bar{x}, y) \wedge \exists y \Psi(\bar{x}, y) \tag{12}$$

Now $\exists y \Phi(\bar{x}, y)$ is a pure $L_{\mathcal{D}}$-formula and $\exists y \Psi(\bar{x}, y)$ is a pure $L_{\mathcal{G}}$-formula. We can carry out the elimination using the elimination procedure for $\mathcal{Q}$ and the elimination procedure for $\mathcal{G}$. $\square$

Corollary 1. *The decision problem for$(L_{\mathcal{D}} \cup L_{\mathcal{G}})$-formulas in good models of $T_{\mathcal{D}} \cup T_{\mathcal{G}}$ is decidable .*

Proof. Using the quantifier elimination described in Theorem 2, one can transform an arbitrary closed first-order $(L_{\mathcal{D}} \cup L_{\mathcal{G}})$-formula into an equivalent quantifier-free formula, which must be either *false* or *true* as $(L_{\mathcal{D}} \cup L_{\mathcal{G}})$ has no constants. $\square$

4 Properties of Good Models

In this section we further explore the properties of good models. Let L be a finite signature without function symbols. We say an L-structure $\mathcal{A}$ is *homogeneous* if every isomorphism between finite substructures of $\mathcal{A}$ extends to an automorphism of $\mathcal{A}$. Let $\mathfrak{K}$ be a class of finite L-structures and $\mathcal{A}$ be a countable L-structure. We say that $\mathcal{A}$ is the *Fraïssé limit* of $\mathfrak{K}$ if $\mathcal{A}$ is homogeneous and $\mathfrak{K}$ is precisely the class of finite structures that can be embedded into $\mathcal{A}$. The Fraïssé limit is sometimes referred to as the *universal homogeneous structure of age* $\mathfrak{K}$ and it is unique up to isomorphism [9]. Now let $T^0_{\mathcal{D}}$ be the theory of linear orders and $T^0_{\mathcal{G}}$ be the theory of graphs. Let $\mathfrak{G}$ be the class of all finite models of the combined theory $T^0_{\mathcal{D}} \cup T^0_{\mathcal{G}}$. Note that each model $\mathcal{B}$ in $\mathfrak{G}$ is a finite graph $(B, E, <)$ with vertex set B, edge relation E, and linear order $<$ on B.

Theorem 3. *Let $\mathcal{A}$ be a good model of $T_{\mathcal{D}} \cup T_{\mathcal{G}}$. Then*

1. *$\mathcal{A}$ is the Fraïssé limit of $\mathfrak{G}$, and*
2. *$(T_{\mathcal{D}} \cup T_{\mathcal{G}})_{GOOD}$ is ω-categorical and complete.*

The rest of the section is devoted to the proof of the above theorem. We say an L-structure $\mathcal{H}$ is *weakly homogeneous* if for any finite L-structures $\mathcal{B}, \mathcal{C}$ such that $\mathcal{B} \subseteq \mathcal{C}$ and $|C| = |B| + 1$, any embedding $f : B \to H$ can be extended to an embedding $g : C \to H$. An L-structure is homogeneous if and only if it is weakly homogeneous [9].

Lemma 4. *$\mathcal{A}$ is weakly homogeneous.*

Proof. Let $\mathcal{B} = (B, E_{\mathcal{B}}, <_{\mathcal{B}}), \mathcal{C} = (C, E_{\mathcal{C}}, <_{\mathcal{C}})$ be finite graphs whose vertices are linearly ordered, $\mathcal{B} \subseteq \mathcal{C}$ and $C = B \cup \{c\}$. Let $B = \{b_1, ..., b_n\}$. Suppose $f : B \to A$ is an embedding of $\mathcal{B}$ into $\mathcal{A}$. Let $a_i = f(b_i)$ for $1 \le i \le n$. Let $\Phi(x, a_1, ..., a_n)$ be the conjunction of literals in the set

$$\{\, a_i < x \mid \mathcal{C} \models b_i < c \,\} \cup \{\, x < a_i \mid \mathcal{C} \models c < b_i \,\},$$

and $\Psi(x, a_1, ..., a_n)$ the conjunction of literals in

$$\{\, E(a_i, x) \mid \mathcal{C} \models E(b_i, c) \,\} \cup \{\, \neg E(a_i, x) \mid \mathcal{C} \models \neg E(b_i, c) \,\}.$$

By the denseness of $<^{\mathcal{A}}$ (Property (2)), $\mathcal{A} \models \exists x \Phi(x, a_1, ..., a_n)$, and by the homogeneity of $E^{\mathcal{A}}$ (Property (3)), $\mathcal{A} \models \exists x \Psi(x, a_1, ..., a_n)$. Thanks to the *good-model* property of $\mathcal{A}$, we then have $\mathcal{A} \models \exists x(\Phi(x, a_1, ..., a_n) \wedge \Psi(x, a_1, ..., a_n))$, which means that we can find an element $a \in A$ such that $a \notin \{a_1, ..., a_n\}$ and $\mathcal{A} \models \Phi(a, a_1, ..., a_n) \wedge \Psi(a, a_1, ..., a_n)$. Therefore the mapping $f \cup \{(c, a)\}$ is an embedding of $\mathcal{C}$ into $\mathcal{A}$. ◻

Lemma 5. *Any structure in the class $\mathfrak{G}$ is embeddable into $\mathcal{A}$. Hence $\mathfrak{G}$ is precisely the class of finite structures embeddable into $\mathcal{A}$.*

Proof. Let $\mathcal{C} = (C, E, <)$ be a finite graph with vertex set C, edge relation E, and linear order $<$ on C. We prove by induction on $|C|$ that $\mathcal{C}$ can be embedded into $\mathcal{A}$. This trivially holds for $|C| = 1$. Suppose that any structures in $\mathfrak{G}$ with n elements can be embedded into $\mathcal{A}$ and $|C| = n + 1$. Let $\mathcal{B} = (B, E_B, <_B)$ be a substructure of $\mathcal{C}$ such that $|B| = n$. By induction hypothesis there is an embedding $f : B \to A$. By Lemma 4, f can be extended to an embedding $g : C \to A$. □

Proof (Theorem 3). The first statement holds by Lemma 4 and Lemma 5. Since the Fraïssé limit of any class of finite structures is unique up to isomorphism, $(T_{\mathcal{D}} \cup T_{\mathcal{G}})_{GOOD}$ is ω-categorical, and therefore complete. □

5 Conclusion and Future Work

In this paper we introduced the notion of *good model* and showed a simple quantifier elimination scheme for good models of union theories. Using a priority argument we showed that $T_{\mathcal{D}} \cup T_{\mathcal{G}}$ has good models and hence admits quantifier elimination with respect to those good models. Furthermore, we showed that $(T_{\mathcal{D}} \cup T_{\mathcal{G}})_{GOOD}$ is ω-categorical and complete, and it has interesting implications on $T^0_{\mathcal{D}} \cup T^0_{\mathcal{G}}$, the combined theory of linear orders with graphs. We showed that $(T_{\mathcal{D}} \cup T_{\mathcal{G}})_{GOOD}$ is indeed the theory of the Fraïssé limit for all finite models of $T^0_{\mathcal{D}} \cup T^0_{\mathcal{G}}$. By quantifier elimination, the almost sure theory of all these finite models is decidable.

We conclude with a remark on the complexity of our decision procedure for $(T_{\mathcal{D}} \cup T_{\mathcal{G}})_{GOOD}$. The algorithm proposed in Corollary 1 is essentially a quantifier elimination procedure on any $(L_{\mathcal{D}} \cup L_{\mathcal{G}})$-formulas. Suppose the input formula is in prenex normal form and each $\forall$ is replaced by $\neg\exists\neg$. Let n be the size of the input formula. We eliminate all existential quantifiers one-by-one. In each iteration, we apply the elimination procedure to the inner-most sub-formula of the form $\exists y \varphi(\overline{x}, y)$ where φ is a quantifier-free formula in disjunctive normal form. Then we write each disjunct in the form $\Phi(\overline{x}, y) \wedge \Psi(\overline{x}, y)$, where Φ is a quantifier-free $L_{\mathcal{D}}$-formula and Ψ is a quantifier-free $L_{\mathcal{G}}$-formula. The above can be done in time $2^{O(n)}$ as the conversion to DNF requires exponential time. The complexity of the elimination procedure depends on the respective complexities of the elimination procedures for $T_{\mathcal{D}}$ and $T_{\mathcal{G}}$. Suppose that the quantifier eliminations for $T_{\mathcal{D}}$ and for $T_{\mathcal{G}}$ take time (space) $f(n)$ and $g(n)$, respectively. Then the complexity of each iteration in our algorithm requires $O(\max\{f(n), g(n)\} + 2^n)$. Let $h(n) = \max\{f(n), g(n), 2^n\}$. For a formula of size n, there could be at most n quantifiers, and hence at most n iterations. A crude analysis shows that the complexity of our elimination procedure is $O(h^n(n))$. See [11, 12] for discussions on the complexity of individual quantifier elimination procedures.

Future Work This is a work in progress towards generalizing Nelson-Oppen combination for combining quantified theories, providing that each individual component theory admits quantifier elimination. Although our current result is only limited to good models, we think it is a good starting point for investigating more general schemes for combining quantifier elimination procedures.

Note that our proof of the existence of good models relies on the "denseness" property of individual theories, that is, there are infinitely many witnesses to existential formulas (Lemmas 1 and 2). However, this property does not hold for many important theories in computer science, such as Presburger arithmetic and discrete orders. Therefore, we first plan to investigate the necessary conditions for the existence of good models and hope this would give us more insights on quantifier elimination schemes for the general models of union theories.

References

[1] Nelson, G., Oppen, D.C.: Simplification by cooperating decision procedures. ACM Transactions on Programming Languages and Systems **1**(2) (October 1979) 245–257

[2] Tinelli, C., Ringeissen, C.: Unions of non-disjoint theories and combinations of satisfiability procedures. Theoretical Computer Science **290**(1) (January 2003) 291–353

[3] Ghilardi, S.: Model-theoretic methods in combined constraint satisfiability. Journal of Automated Reasoning **33**(3-4) (2005) 221–249

[4] Avigad, J., Friedman, H.: Combining decision procedures for the reals. Logical Methods in Computer Sciences **2**(4:4) (2006) 1–42

[5] Tinelli, C., Zarba, C.G.: Combining decision procedures for sorted theories. In Algeres, J., Leite, J., eds.: Logic in Artificial Intelligence, 9th European Conference (JELIA). (2004)

[6] Ghilardi, S., Santocanale, L.: Algebraic and model-theoretic techniques for fusion decidability in modal logic. In Vardi, M., Voronkov, A., eds.: Logic for Programming, Artificial Intelligence and Reasoning (LPAR 03'). Volume 2850., Springer LNAI (2003) 152–166

[7] Gulwani, S., Tiwari, A.: Combining abstract interpreters. In: the 2006 ACM SIGPLAN conference on Programming language design and implementation. (2006) 376–386

[8] Erdös, P., Spencer, J.: Probabilistic Methods in Combinatorics. Academy Press, New York and London (1974)

[9] Hodges, W.: A Shorter Model Theory. Cambridge University Press (1997)

[10] Enderton, H.B.: A Mathematical Introduction to Logic. Academic Press (2001)

[11] Nipkow, T.: Linear quantifier elimination. In: IJCAR 2008. (to appear)

[12] Weispfenning, V.: Efficient decision algorithm for locally finite theories. In Calmet, J., ed.: Algebraic algorithms and error-correcting codes:3rd international conference, Grenoble, France. Volume 229 of LNCS., Springer (1986) 262–273