Regular families of forests, antichains and duality pairs of relational structures

Péter L. Erdős[a], Dömötör Pálvölgyi[b,1], Claude Tardif[c], Gábor Tardos[a,3]

[a]*Alfréd Rényi Institute of Mathematics, Reáltanoda u 13-15 Budapest, 1053 Hungary*
email*: <erdos.peter,tardos.gabor>@renyi.mta.hu*
[b]*Institute of Mathematics, L. Eötvös University,*
Pázmány Péter sétány 1/C Budapest, 1117, Hungary
email*: dom@cs.elte.hu*
[c]*Royal Military College of Canada, PO Box 17000 Station "Forces"*
Kingston, Ontario, Canada, K7K 7B4
email*: Claude.Tardif@rmc.ca*

Abstract

Homomorphism duality pairs play a crucial role in the theory of relational structures and in the Constraint Satisfaction Problem. The case where both classes are finite is fully characterized. The case when both side are infinite seems to be very complex. It is also known that no finite-infinite duality pair is possible if we make the additional restriction that both classes are antichains. In this paper we characterize the infinite-finite antichain dualities and infinite-finite dualities with trees or forest on the left hand side. This work builds on our earlier papers [6] that gave several examples of infinite-finite antichain duality pairs of directed graphs and [7] giving a complete characterization for caterpillar dualities.

Keywords: graph homomorphism; duality pairs; general relational structures; constraint satisfaction problems; regular languages

1. Introduction

A *homomorphism duality pair* is a couple $(\mathcal{O}, \mathcal{D})$ where $\mathcal{O}$ and $\mathcal{D}$ are families of relational structures of the same type, such that the following holds.

> For any given relational structure $\mathbf{A}$, there exists a homomorphism of $\mathbf{A}$ to some member $\mathbf{D}$ of $\mathcal{D}$ if and only if there is no homomorphism of any member $\mathbf{T}$ of $\mathcal{O}$ to $\mathbf{A}$.

We refer the reader to Section 2 for basic terminology on relational structures and homomorphisms. Note that the simplest example is that of directed graphs and arc-preserving maps.

Homomorphism duality pairs (duality pairs, for short) $(\mathcal{O}, \mathcal{D})$ where $\mathcal{D}$ consists of a single structure play a major role in the study of constraint satisfaction problems (see [1]). Given a structure $\mathbf{D}$, conditions are sought that imply the existence of a duality pair $(\mathcal{O}, \{\mathbf{D}\})$, where $\mathcal{O}$ satisfies given structural properties. The structural properties considered in the literature imply that the constraint satisfaction problem for $\mathbf{D}$, that is, the problem of determining whether an input structure admits a homomorphism to $\mathbf{D}$, can be solved by an efficient algorithm. In some cases the type of duality considered gives further clues about the precise descriptive complexity of the constraint satisfaction problem (see [1]).

A duality pair $(\mathcal{O}, \{\mathbf{D}\})$ is called a *tree duality* if $\mathcal{O}$ consists of (relational) trees, and a *finite duality* if $\mathcal{O}$ is finite. The structures $\mathbf{D}$ such that there exists a tree duality $(\mathcal{O}, \{\mathbf{D}\})$ are characterized in [8], and the structures $\mathbf{D}$ such that there exists a finite duality $(\mathcal{O}, \{\mathbf{D}\})$ are characterized in [10]. In particular any finite duality $(\mathcal{O}, \{\mathbf{D}\})$ is essentially a tree duality (see [12]).

Finite dualities can be interpreted from an order theoretic point of view in a category of relational structures preordered by the existence of homomorphisms. Any finite duality can be reduced to a form $(\mathcal{O}, \{\mathbf{D}\})$, where $|\mathcal{O}|$ is minimal (with respect to $\mathbf{D}$). In most cases $\mathbf{D}$ does not admit a homomorphism to any member of $\mathcal{O}$. The set $\mathcal{A} = \mathcal{O} \cup \{\mathbf{D}\}$ is then a "maximal antichain" in the sense that there exists no homomorphism between any two of its members, and for any other structure $\mathbf{B}$ of the same type, there is a structure $\mathbf{A}$ in $\mathcal{A}$ such that there exists a homomorphism of $\mathbf{A}$ to $\mathbf{B}$ or a homomorphism of $\mathbf{B}$ to $\mathbf{A}$. More precisely, for any structure $\mathbf{B}$ of the same type, there is a structure $\mathbf{A}$ in $\mathcal{O}$ such that there exists a homomorphism of $\mathbf{A}$ to $\mathbf{B}$, or there is a structure $\mathbf{A}$ in $\mathcal{D} = \{\mathbf{D}\}$ such that there exists a homomorphism of $\mathbf{B}$ to $\mathbf{A}$. An antichain $\mathcal{A}$ in an ordered set $\mathcal{P}$ is said to have the "splitting property" if it admits a partition into sets $\mathcal{O}$ and $\mathcal{D}$ such that for any element $\mathbf{B}$, there exists $\mathbf{A} \in \mathcal{O}$ such that $\mathbf{A} \leq \mathbf{B}$, or there exists $\mathbf{A} \in \mathcal{D}$ such that $\mathbf{B} \leq \mathbf{A}$. (An antichain with the splitting property is necessarily maximal.) When $\mathcal{P}$ is a category of relational structures and $\leq$ denotes the existence of a homomorphism, the antichains with the splitting property correspond to the dualities $(\mathcal{O}, \mathcal{D})$ where $\mathcal{O} \cup \mathcal{D}$ is an antichain. Thus viewing dualities from an order-theoretic point of view leads to removing the emphasis on the case where $\mathcal{D}$ is a singleton, and putting it instead on the case where $\mathcal{O} \cup \mathcal{D}$ is a maximal antichain.

In [9] it is shown how the finite duality pairs $(\mathcal{O}, \mathcal{D})$ are built up from finite duality pairs $(\mathcal{O}', \mathcal{D}')$ with $|\mathcal{D}'| = 1$. Furthermore the finite maximal antichains are shown to correspond to finite dualities, at least in the case where there is only one relation (e.g., the case of directed graphs); it is possible that this correspondence extends to all categories of relational structures.

On the other hand the infinite maximal antichains are essentially unclassifiable, because any antichain can be greedily extended to a maximal antichain (see [3, 4]). However the antichains with the splitting property have a lot of structure. In [5] it is shown that in the case of directed graphs, such an antichain $\mathcal{A} = \mathcal{O} \cup \mathcal{D}$ cannot have $\mathcal{O}$ finite and $\mathcal{D}$ infinite. The question of the existence of antichains with the splitting property $\mathcal{A} = \mathcal{O} \cup \mathcal{D}$ with $\mathcal{O}$ infinite and $\mathcal{D}$ finite is answered positively in [6]. The context is again that of directed graphs, and it is shown that in any such infinite-finite antichain $\mathcal{A} = \mathcal{O} \cup \mathcal{D}$, $\mathcal{O}$ must consist of forests. Thus the question arises as to which properties must an antichain $\mathcal{O}$ of forests satisfy for the existence of a family $\mathcal{D}$ such that $\mathcal{A} = \mathcal{O} \cup \mathcal{D}$ is an infinite-finite antichain. Note that this reverses the original question from the context of constraint satisfaction problems, where conditions were sought on $\mathcal{D} = \{\mathbf{D}\}$ for the existence of $\mathcal{O}$ such that $(\mathcal{O}, \mathcal{D})$ is a duality pair (with prescribed properties). In [6] it is shown that if $\mathcal{O}$ is an antichain of digraph paths, then there can exist a finite $\mathcal{D}$ such that $(\mathcal{O}, \mathcal{D})$ is a duality only if $\mathcal{O}$ is "regular" in the sense of automata theory.

In [7], it is shown that regular languages can be used to characterize the "caterpillar dualities" in general relational structures. Caterpillars are generalizations of paths, and the caterpillar dualities $(\mathcal{O}, \{\mathbf{D}\})$ are of interest in the constraint satisfaction community as the dualities for which $\mathcal{O}$ can be described in the "smallest natural recursive fragment of Datalog" (see [2]).

In the present paper we extend the context of [6] from digraphs to general relational structures, and the context of [7] from caterpillar dualities to general forest dualities. The criterion of regularity remains relevant, but in the context of forest dualities it is necessary to generalize it. Our notion of regular families of forests is similar (but not identical) to other logics for trees surveyed in [11].

The paper is structured as follows. The next section is a brief introduction to relational structures and homomorphisms. In Section 3 we present regular families of forests and prove that the family of forests that do not admit a homomorphism to a given structure is regular. In Section 4 we prove that for any regular family $\mathcal{O}$ of forests, there exists a finite family $\mathcal{D}$ of structures such that $(\mathcal{O}, \mathcal{D})$ is a duality pair. In Sections 5 and 6 we deal with antichain dualities $(\mathcal{O}, \mathcal{D})$ with $\mathcal{D}$ finite, and prove that these are essentially forest dualities. In Section 5 we prove that if $(\mathcal{O}, \mathcal{D})$ is a duality pair such that $\mathcal{D}$ is finite, $\mathcal{O}$ consists of cores and $\mathcal{O} \cup \mathcal{D}$ is an antichain, then $\mathcal{O}$ consists of forests. Then Theorem 5.3 states that if $\mathcal{O}$ is an antichain of core forests, then there exists a finite family $\mathcal{D}$ such that $(\mathcal{O}, \mathcal{D})$ is a duality pair if and only if $\mathcal{O}$ is regular. This uses Theorem 5.4 which states that the cores of the minimal elements of a regular set of forests form a regular set. Section 6 is dedicated to the (quite technical) proof of this theorem. In our results, specifying that $\mathcal{O}$ consists of trees rather than forests is equivalent to specifying that $\mathcal{D}$ consists of a single structure.

3

2. Preliminaries

2.1. Relational structures

A *type* is a finite set $\sigma = \{R_1, \ldots, R_m\}$ of *relation symbols*, each with an *arity* r_i assigned to it. A σ-structure is a relational structure $\mathbf{A} = \langle A; R_1(\mathbf{A}), \ldots, R_m(\mathbf{A}) \rangle$ where A is a finite set called the *universe* of $\mathbf{A}$, and $R_i(\mathbf{A})$ is an r_i-ary relation on A for each i. The elements of $R_i(\mathbf{A})$, $1 \leq i \leq m$ will be called *hyperedges* of $\mathbf{A}$. By analogy with the graph theoretic setting, the universe A of $\mathbf{A}$ will also be called its vertex-set, denoted $V(\mathbf{A})$.

A σ-structure $\mathbf{A}$ may be described by its bipartite *incidence multigraph* $\mathrm{Inc}(\mathbf{A})$ defined as follows. The two parts of $\mathrm{Inc}(\mathbf{A})$ are $V(\mathbf{A})$ and $\mathrm{Block}(\mathbf{A})$, where

$$\mathrm{Block}(\mathbf{A}) = \{(R, (x_1, \ldots, x_r)) : R \in \sigma \text{ has arity } r \text{ and } (x_1, \ldots, x_r) \in R(\mathbf{A})\},$$

and with edges $e_{a,i,B}$ joining $a \in V(\mathbf{A})$ to $B = (R, (x_1, \ldots, x_r)) \in \mathrm{Block}(\mathbf{A})$ when $x_i = a$. Thus, the degree of $B = (R, (x_1, \ldots, x_r))$ in $\mathrm{Inc}(\mathbf{A})$ is precisely r. Here "degree" means number of incident edges rather than number of neighbors because parallel edges are possible: If $x_i = x_j = a \in V(\mathbf{A})$, then $e_{a,i,B}$ and $e_{a,j,B}$ both join a and B. A σ-structure $\mathbf{A}$ is called a σ-*tree* (or *tree* for short) if $\mathrm{Inc}(\mathbf{A})$ is a (graph-theoretic) tree, that is, it is connected and has no cycles or parallel edges. Similarly, $\mathbf{A}$ is called a σ-*forest* (or *forest* for short) if $\mathrm{Inc}(\mathbf{T})$ is a (graph-theoretic) forest, that is, it has no cycles or parallel edges.

A σ-structure $\mathbf{A}$ is a *substructure* of $\mathbf{B}$ (in notation: $\mathbf{A} \subseteq \mathbf{B}$) if the universe and relations of $\mathbf{A}$ are subsets of the corresponding families for $\mathbf{B}$. In this case $\mathrm{Inc}(\mathbf{A})$ is a subgraph of $\mathrm{Inc}(\mathbf{B})$. In particular the components of $\mathrm{Inc}(\mathbf{B})$ determine substructures that we call the *components* of $\mathbf{B}$.

2.2. Homomorphisms

For σ-structures $\mathbf{A}$ and $\mathbf{B}$, a *homomorphism* from $\mathbf{A}$ to $\mathbf{B}$ is a map $f : V(\mathbf{A}) \to V(\mathbf{B})$ such that $f(R_i(\mathbf{A})) \subseteq R_i(\mathbf{B})$ for all $i = 1, \ldots, m$, where for any r-ary relation $R \in \sigma$ we have

$$f(R) = \{(f(x_1), \ldots, f(x_r)) : (x_1, \ldots, x_r) \in R\}.$$

We write $\mathbf{A} \to \mathbf{B}$ if there exists a homomorphism from $\mathbf{A}$ to $\mathbf{B}$, and $\mathbf{A} \nrightarrow \mathbf{B}$ otherwise. We write $\mathbf{A} \leftrightarrow \mathbf{B}$ when $\mathbf{A} \to \mathbf{B}$ and $\mathbf{B} \to \mathbf{A}$; $\mathbf{A}$ and $\mathbf{B}$ are then called *homomorphically equivalent*. For a finite structure $\mathbf{A}$, we can always find a structure $\mathbf{B}$ such that $\mathbf{A} \leftrightarrow \mathbf{B}$ and the cardinality of $V(\mathbf{B})$ is minimal with respect to this property. It is well known (see [12]) that such $\mathbf{B}$ is unique up to isomorphism. We then call $\mathbf{B}$ the *core* of $\mathbf{A}$.

Recall the definition of homomorphism duality pairs from the beginning of the Introduction. In our definition of σ-structure we insisted that it

must be finite but in this paragraph we comment on the natural extension to infinite structures. The families $\mathcal{O}$ and $\mathcal{D}$ in the duality pairs $(\mathcal{O}, \mathcal{D})$ we consider in this paper consist of finite structures and their defining condition is required to hold for any finite structure $\mathbf{A}$. For general duality pairs this does not imply that the same condition also holds for infinite structures $\mathbf{A}$. But in this paper we consider duality pairs $(\mathcal{O}, \mathcal{D})$ with $\mathcal{D}$ finite. For a finite family $\mathcal{D}$ of finite structures, a compactness argument shows that if an infinite structure $\mathbf{A}$ has no homomorphism to any member of $\mathcal{D}$, then this also holds for a finite substructure of $\mathbf{A}$. This implies that the duality pairs we consider in this paper are duality pairs even if considered in the context of arbitrary (not necessarily finite) structures.

3. Regular families of forests

Let $\mathbb{T}$ be the set of σ-trees and $\mathbb{F}$ the set of σ-forests. A *rooted σ-structure* $(\mathbf{A}, a)$ is a σ-structure $\mathbf{A}$ with an arbitrary element $a \in V(\mathbf{A})$ designated as the root. Let $\mathbb{F}_{\mathbf{r}}$ and $\mathbb{T}_{\mathbf{r}}$ denote respectively the set of rooted σ-forests and of rooted σ-trees. We do not distinguish isomorphic structures, so more precisely $\mathbb{T}$, $\mathbb{F}$, $\mathbb{T}_{\mathbf{r}}$ and $\mathbb{F}_{\mathbf{r}}$ are the set of *isomorphism classes* of trees, forests and rooted trees and rooted forests, respectively. We introduce two operations:

- **unrooting** $[\cdot] : \mathbb{F}_{\mathbf{r}} \to \mathbb{F}$ forgets the root;

- **combining** $+ : \mathbb{F}_{\mathbf{r}} \times \mathbb{F}_{\mathbf{r}} \to \mathbb{F}_{\mathbf{r}}$ takes the disjoint union of two rooted forests and identifies their roots.

Note that the single-vertex rooted tree with empty relations is the identity of the combining operation. We denote it by $\mathbf{T}_0$.

Definition 3.1. Let $\mathcal{O}$ be a subset of $\mathbb{F}$.

- For $(\mathbf{A}, a) \in \mathbb{F}_{\mathbf{r}}$, the set $\mathcal{O} - (\mathbf{A}, a) \subseteq \mathbb{F}_{\mathbf{r}}$ is defined by

$$\mathcal{O} - (\mathbf{A}, a) = \{(\mathbf{B}, b) \in \mathbb{F}_{\mathbf{r}} \mid [(\mathbf{A}, a) + (\mathbf{B}, b)] \in \mathcal{O}\}.$$

- The equivalence $\sim_{\mathcal{O}}$ on $\mathbb{F}_{\mathbf{r}}$ is defined by $(\mathbf{A}, a) \sim_{\mathcal{O}} (\mathbf{A}', a')$ if $\mathcal{O} - (\mathbf{A}, a) = \mathcal{O} - (\mathbf{A}', a')$.

- $\mathcal{O}$ is called *regular* if $\mathbb{F}_{\mathbf{r}}$ has only a finite number of equivalence classes under $\sim_{\mathcal{O}}$.

Given a set $\mathcal{O} \subseteq \mathbb{F}$, we can define a graph G whose vertex-set is $\mathbb{F}_{\mathbf{r}}$ and whose edges join pairs $(\mathbf{A}, a)$, $(\mathbf{B}, b)$ such that $[(\mathbf{A}, a) + (\mathbf{B}, b)] \in \mathcal{O}$. Then $\mathcal{O} - (\mathbf{A}, a)$ is the neighborhood of $(\mathbf{A}, a)$ in G. Thus $\mathcal{O}$ is regular if and only if this graph can be obtained from a finite graph by blowing up its vertices.

In [6] and [7], the natural description of paths and caterpillars in terms of words over an alphabet allowed a more direct correspondence between regular languages and regular families of paths or caterpillars. Definition 3.1 is in the spirit of the Myhill-Nerode theorem, which states that a language $\mathcal{L}$ over an alphabet Σ is regular if and only if the infinitely many words $a \in \Sigma^*$ define only finitely many distinct extension sets $\mathcal{L} - a = \{b \in \Sigma^* \mid ab \in \mathcal{L}\}$. In Definition 3.1, $[\cdot + \cdot]$ plays the role of concatenation and the sets $\mathcal{O} - (\mathbf{A}, x)$ play the role of extension sets.

The following result establishes basic properties of regular sets of forests.

Lemma 3.2. *Let $\mathcal{O}_1$ and $\mathcal{O}_2$ be regular subsets of $\mathbb{F}$. Then $\mathcal{O}_1 \cup \mathcal{O}_2$, $\mathcal{O}_1 \cap \mathcal{O}_2$ and $\mathbb{F} \setminus \mathcal{O}_1$ are also regular.*

Proof. This follows from the fact that for any $(\mathbf{A}, x) \in \mathbb{F}_\mathbf{r}$ we have

$$
\begin{aligned}
(\mathcal{O}_1 \cup \mathcal{O}_2) - (\mathbf{A}, x) &= (\mathcal{O}_1 - (\mathbf{A}, x)) \cup (\mathcal{O}_2 - (\mathbf{A}, x)), \\
(\mathcal{O}_1 \cap \mathcal{O}_2) - (\mathbf{A}, x) &= (\mathcal{O}_1 - (\mathbf{A}, x)) \cap (\mathcal{O}_2 - (\mathbf{A}, x)) \text{ and} \\
(\mathbb{F} \setminus \mathcal{O}_1) - (\mathbf{A}, x) &= \mathbb{F}_\mathbf{r} \setminus (\mathcal{O}_1 - (\mathbf{A}, x)).
\end{aligned}
$$

Therefore

$$
\begin{aligned}
|\mathbb{F}_\mathbf{r} / \sim_{\mathcal{O}_1 \cup \mathcal{O}_2}| &\leq |\mathbb{F}_\mathbf{r} / \sim_{\mathcal{O}_1}| \cdot |\mathbb{F}_\mathbf{r} / \sim_{\mathcal{O}_2}|, \\
|\mathbb{F}_\mathbf{r} / \sim_{\mathcal{O}_1 \cap \mathcal{O}_2}| &\leq |\mathbb{F}_\mathbf{r} / \sim_{\mathcal{O}_1}| \cdot |\mathbb{F}_\mathbf{r} / \sim_{\mathcal{O}_2}| \text{ and} \\
|\mathbb{F}_\mathbf{r} / \sim_{\mathbb{F} \setminus \mathcal{O}_1}| &= |\mathbb{F}_\mathbf{r} / \sim_{\mathcal{O}_1}|.
\end{aligned}
$$

$\square$

Lemma 3.3. *For any σ-structure $\mathbf{D}$, the family $\mathcal{H}_\mathbf{D} = \{\mathbf{A} \in \mathbb{F} \mid \mathbf{A} \to \mathbf{D}\}$ is regular.*

Proof. For any vertex $z \in V(\mathbf{D})$ let $\mathcal{S}_z$ be the set of the rooted forests $(\mathbf{B}, y)$ such that there exists a homomorphism $f : \mathbf{B} \to \mathbf{D}$ with $f(y) = z$. Then for any $(\mathbf{A}, x) \in \mathbb{F}_\mathbf{r}$, we have

$$
\mathcal{H}_\mathbf{D} - (\mathbf{A}, x) = \bigcup \{\mathcal{S}_z \mid (\mathbf{A}, x) \in \mathcal{S}_z\}.
$$

Therefore $|\mathbb{F}_\mathbf{r} / \sim_{\mathcal{H}_\mathbf{D}}| \leq 2^{|V(\mathbf{D})|}$. $\square$

Corollary 3.4. *Let $\mathcal{D}$ be a finite family of σ-structures. Then the family $\mathcal{O}_\mathcal{D}$ of forests which do not admit a homomorphism to any structure in $\mathcal{D}$ is regular.*

Proof. We have $\mathcal{O}_\mathcal{D} = \bigcap_{\mathbf{D} \in \mathcal{D}} (\mathbb{F} \setminus \mathcal{H}_\mathbf{D})$, and the latter is regular by Lemmas 3.3 and 3.2. $\square$

Corollary 3.5. *Let $\mathbf{D}$ be a σ-structure. Then the family of trees which do not admit a homomorphism to $\mathbf{D}$ is regular.*

Proof. The family $\mathbb{T}$ of σ-trees is regular, since $\mathbb{F_r}/\sim_\mathbb{T} = \{\mathbb{T_r}, \mathbb{F_r} \setminus \mathbb{T_r}\}$. Thus for any σ-structure $\mathbf{D}$, $\mathcal{H}_\mathbf{D} \cap \mathbb{T}$ and $\mathcal{O}_{\{\mathbf{D}\}} \cap \mathbb{T}$ are regular. $\qquad\square$

Note that by a "regular family of trees" we mean a "regular family of forests whose members are trees." A second interpretation is possible, obtained by replacing $\mathbb{F}$ and $\mathbb{F_r}$ by $\mathbb{T}$ and $\mathbb{T_r}$ in Definition 3.1. However the two interpretations turn out to be equivalent, since for $\mathcal{O} \subseteq \mathbb{T}$, we have $\mathcal{O} - (\mathbf{A}, x) = \emptyset$ for all $(\mathbf{A}, x) \in \mathbb{F_r} \setminus \mathbb{T_r}$. Thus there is no need for a separate definition of "regular family of trees."

How about "regular families of rooted forests"? One might be tempted to define these in a similar fashion, by defining the extension sets of $\mathcal{O}_r \in \mathbb{F_r}$ by

$$\mathcal{O}_r - (\mathbf{A}, a) = \{(\mathbf{B}, b) \in \mathbb{F_r} \mid (\mathbf{A}, a) + (\mathbf{B}, b) \in \mathcal{O}_r\}.$$

The regular families of rooted forests would then be those with finitely many extension sets. However, our regular families of forests are not simply the unrootings of such "regular families of rooted forests." Indeed, consider any family $\mathcal{O}_r$ that consists only of rooted trees whose root has degree exactly one in the incidence multigraph. With the above definition, such a family becomes regular, with $\mathcal{O}_r$, $\{\mathbf{T}_0\}$ and $\emptyset$ being the only possible extension sets. Thus this approach leads to uncountably many families of rooted forests that are regular for trivial reasons, and whose unrootings yield uncountably many families of forests.

In contrast, our last result of this section will show that just like in the case of regular languages over any given alphabet, there are countably many regular families of σ-forests over any given type σ.

Definition 3.6. Let $R \in \sigma$ be a relation of arity r, and $(\mathbf{A}_1, x_1), \ldots, (\mathbf{A}_r, x_r)$ rooted σ-structures. The *concatenation* $\mathbf{C}(R, (\mathbf{A}_1, x_1), \ldots, (\mathbf{A}_r, x_r))$ is the σ-structure $\mathbf{C}$ obtained from the disjoint union of $\mathbf{A}_1, \ldots, \mathbf{A}_r$ by adding $(x_1, \ldots, x_r)$ to $R(\mathbf{C})$.

Note that the concatenation of rooted σ-trees is a σ-tree.

Theorem 3.7. *There are countably many regular families of σ-forests.*

Proof. We associate a type $o(\sigma)$ of operations to the type σ as follows. For each $R \in \sigma$ of arity r, $o(\sigma)$ contains r operations $\mu_{R,1}, \ldots, \mu_{R,r}$ of arity r. In addition, $o(\sigma)$ contains an operation ν of arity 1. We define the algebra $(\mathbb{F_r}, o(\sigma))$ as follows.

- $\nu(\mathbf{A}, x)$ is the rooted forest obtained by adding an isolated element to $V(\mathbf{A})$, which becomes the new root.

- For $R \in \sigma$ of arity r and $i \in \{1, \ldots, r\}$, $\mu_{R,i}$ is defined by

$$\mu_{R,i}((\mathbf{A}_1, x_1), \ldots, (\mathbf{A}_r, x_r)) = (\mathbf{C}(R, (\mathbf{A}_1, x_1), \ldots, (\mathbf{A}_r, x_r)), x_i).$$

Note that the whole of $\mathbb{F}_r$ is generated by $\{\mathbf{T}_0\}$.

Let $\mathcal{O}$ be a family of forests. We show that $\sim_{\mathcal{O}}$ is a congruence on $(\mathbb{F}_r, o(\sigma))$. Indeed if $(\mathbf{A}, x) \sim_{\mathcal{O}} (\mathbf{A}', x')$ and $[\nu(\mathbf{A}, x) + (\mathbf{B}, y)] \in \mathcal{O}$, then since $[\nu(\mathbf{A}, x) + (\mathbf{B}, y)] = [(\mathbf{A}, x) + \nu(\mathbf{B}, y)]$, we have $[\nu(\mathbf{A}', x') + (\mathbf{B}, y)] = [(\mathbf{A}', x') + \nu(\mathbf{B}, y)] \in \mathcal{O}$. Now suppose that $R \in \sigma$ has arity r, $i, j \in \{1, \ldots r\}$ and $(\mathbf{A}_i, x_i) \sim_{\mathcal{O}} (\mathbf{A}'_i, x'_i)$. We show that for any $(\mathbf{A}_k, x_k) = (\mathbf{A}'_k, x'_k)$ and $k \in \{1, \ldots, r\} \setminus \{i\}$, we have $\mu_{R,j}((\mathbf{A}_1, x_1), \ldots, (\mathbf{A}_r, x_r)) \sim_{\mathcal{O}} \mu_{R,j}((\mathbf{A}'_1, x'_1), \ldots, (\mathbf{A}'_r, x'_r))$. Suppose that $[\mu_{R,j}((\mathbf{A}_1, x_1), \ldots, (\mathbf{A}_r, x_r)) + (\mathbf{B}, y)] \in \mathcal{O}$. We define $(\mathbf{C}_1, z_1), \ldots, (\mathbf{C}_r, z_r)$ by $\mathbf{C}_k = (\mathbf{A}_k, x_k)$ if $k \notin \{i, j\}$; $\mathbf{C}_i = \{\mathbf{T}_0\}$, $\mathbf{C}_j = (\mathbf{A}_j, x_j) + (\mathbf{B}, y)$ if $j \neq i$; finally $\mathbf{C}_i = (\mathbf{B}, y)$ if $j = i$. Then $[\mu_{R,j}((\mathbf{A}_1, x_1), \ldots, (\mathbf{A}_r, x_r)) + (\mathbf{B}, y)] = [(\mathbf{A}_i, x_i) + \mu_{R,i}((\mathbf{C}_1, x_1), \ldots, (\mathbf{C}_r, x_r))]$. As $(\mathbf{A}_i, x_i) \sim_{\mathcal{O}} (\mathbf{A}'_i, x'_i)$, we then have $[(\mathbf{A}'_i, x'_i) + \mu_{R,i}((\mathbf{C}_1, x_1), \ldots, (\mathbf{C}_r, x_r))] = [\mu_{R,j}((\mathbf{A}'_1, x'_1), \ldots, (\mathbf{A}'_r, x'_r)) + (\mathbf{B}, y)] \in \mathcal{O}$. Thus $\sim_{\mathcal{O}}$ is a congruence on $(\mathbb{F}_r, o(\sigma))$.

If $\mathcal{O}$ is regular, then $(\mathbb{F}_r, o(\sigma))/\sim_{\mathcal{O}}$ is finite, thus $\sim_{\mathcal{O}}$ is the kernel of a homomorphism $\phi(\mathbb{F}_r, o(\sigma)) \to (X, o(\sigma))$, where $(X, o(\sigma))$ is a finite $o(\sigma)$-algebra. Since $\mathbb{F}_r$ is generated by $\{\mathbf{T}_0\}$, ϕ is completely determined by $\phi(\mathbf{T}_0)$. Take the unrootings of rooted forests in an $\sim_{\mathcal{O}}$ equivalence class. We have $\mathbf{A} = [(\mathbf{A}, x) + \mathbf{T}_0]$, showing that either all of these forests belong to $\mathcal{O}$ or none of them. To emphasize the similarity with the finite automaton characterization of regular languages, we call the elements of X *states*, $\phi(\mathbf{T}_0)$ the *initial state* and the states corresponding to the classes of rooted forests whose unrootings are in $\mathcal{O}$ the *terminal states*.

There are finitely many $o(\sigma)$-algebras on any finite set, finitely many choices for the initial state and for the terminal states. The family $\mathcal{O}$ is determined by these choices (except whether the empty structure belongs to it), thus the number of regular families of forests is countable. $\square$

4. Constructions of duals

4.1. Duals of families of trees

Definition 4.1. Let $\mathcal{O}$ be a regular family of trees. We define the structure $\mathbf{D}(\mathcal{O})$ as follows.

The vertices of $\mathbf{D}(\mathcal{O})$ are the sets $\mathcal{V} \subseteq \mathbb{T}_r$ satisfying the following properties.

(i) $\mathbf{T}_0 \in \mathcal{V}$.

(ii) If $(\mathbf{A}, a) \in \mathcal{V}$, then $\mathbf{A} \notin \mathcal{O}$.

(iii) $\mathcal{V} = \mathbb{T}_r \setminus \bigcup_{(\mathbf{A}, a) \in \mathcal{H}} (\mathcal{O} - (\mathbf{A}, a))$ for some family $\mathcal{H}$ of rooted trees.

For $R \in \sigma$ of arity r, we have $(\mathcal{V}_1, \ldots, \mathcal{V}_r) \in R(\mathbf{D}(\mathcal{O}))$ if for each $(\mathbf{A}_i, a_i) \in \mathcal{V}_i$, $i = 1, \ldots, r$ and $\mathbf{T} = \mathbf{C}(R, ((\mathbf{A}_1, a_1), \ldots, (\mathbf{A}_r, a_r)))$, we have $(\mathbf{T}, a_j) \in \mathcal{V}_j$, $j = 1, \ldots, r$, where $\mathbf{C}$ is the concatenation of Definition 3.6.

Note that since $\mathcal{O}$ is regular, there are finitely many sets of the form $\mathcal{O} - (\mathbf{A}, a)$, hence $V(\mathbf{D}(\mathcal{O}))$ is finite as required.

Lemma 4.2. *Let $\mathbf{A}$ be a σ-tree and $\phi : \mathbf{A} \to \mathbf{D}(\mathcal{O})$ a homomorphism. Then for every $a \in V(\mathbf{A})$, $(\mathbf{A}, a) \in \phi(a)$.*

Proof. We use induction on the number $|\bigcup_{R \in \sigma} R(\mathbf{A})|$ of hyperedges of $\mathbf{A}$. If $\mathbf{A}$ has no hyperedges, the result follows from item (i) of Definition 4.1. Now suppose that the result is valid for any tree with fewer hyperedges than $\mathbf{A}$. Let $(x_1, \ldots, x_r) \in R(\mathbf{A})$ be a hyperedge such that $a \in \{x_1, \ldots, x_r\}$. Then $\mathbf{A} = \mathbf{C}(R, (\mathbf{A}_1, x_1), \ldots, (\mathbf{A}_r, x_r))$, where $\mathbf{A}_1, \ldots, \mathbf{A}_r$ are the components of the forest obtained from removing $(x_1, \ldots, x_r)$ from $R(\mathbf{A})$. The restriction of ϕ to each $\mathbf{A}_i$ is a homomorphism, so by the induction hypothesis, $(\mathbf{A}_i, x_i) \in \phi(x_i)$ for $i = 1, \ldots, r$. Since ϕ is a homomorphism and $(x_1, \ldots, x_r) \in R(\mathbf{A})$, we then have $(\mathbf{A}, x_j) \in \phi(x_j)$ for $j = 1, \ldots, r$, and in particular, $(\mathbf{A}, a) \in \phi(a)$. $\qquad\square$

Theorem 4.3. *Let $\mathcal{O}$ be a regular family of trees. Then for any σ-structure $\mathbf{B}$, there exists a homomorphism of $\mathbf{B}$ to $\mathbf{D}(\mathcal{O})$ if and only if no tree in $\mathcal{O}$ admits a homomorphism to $\mathbf{B}$.*

Proof. Let $\mathbf{B}$ be a σ-structure that admits a homomorphism ϕ to $\mathbf{D}(\mathcal{O})$. Suppose that there exists a tree $\mathbf{A}$ in $\mathcal{O}$ that admits a homomorphism ψ to $\mathbf{B}$. Then $\phi \circ \psi : \mathbf{A} \to \mathbf{D}(\mathcal{O})$ is a homomorphism. Therefore by Lemma 4.2, for any $a \in V(\mathbf{A})$, $(\mathbf{A}, a) \in \phi(\psi(a))$. However this contradicts item (ii) in Definition 4.1. Thus no element of $\mathcal{O}$ admits a homomorphism to $\mathbf{B}$.

Conversely, suppose that no element of $\mathcal{O}$ admits a homomorphism to $\mathbf{B}$. For a vertex b of $\mathbf{B}$, we define the following sets

- $S(b) = \{(\mathbf{A}, a) \in \mathbb{T}_r \mid$ there exists a homomorphism $\phi : \mathbf{A} \to \mathbf{B}$ such that $\phi(a) = b\}$,

- $\phi(b) = \mathbb{T}_r \setminus \bigcup(\mathcal{O} - (\mathbf{A}, a))$, where the union is taken for all $(\mathbf{A}, a) \in \mathbb{T}_r$ with $S(b) \cap (\mathcal{O} - (\mathbf{A}, a)) = \emptyset$.

Note that $S(b) \subseteq \phi(b)$ for all vertex b, furthermore $\phi(b)$ is the minimal set containing $S(b)$ that is in the form required by item (iii) of Definition 4.1. Item (i) is also satisfied by $\phi(b)$ as $\mathbf{T}_0 \in S(b) \subseteq \phi(b)$. Since no element of $\mathcal{O}$ admits a homomorphism to $\mathbf{B}$, we have $\mathbf{A} \notin \mathcal{O}$ whenever $(\mathbf{A}, a) \in S(b)$. This implies $S(b) \cap (\mathcal{O} - \mathbf{T}_0) = \emptyset$, thus $\phi(b) \subseteq \mathbb{T}_r \setminus (\mathcal{O} - \mathbf{T}_0)$. Therefore $\phi(b)$ also satisfies item (ii) and thus it is a vertex of $\mathbf{D}(\mathcal{O})$. We will show that the map $\phi : V(\mathbf{B}) \to V(\mathbf{D}(\mathcal{O}))$ is a homomorphism.

Let $R \in \sigma$ be a relation of arity r and $(b_1, \ldots, b_r) \in R(\mathbf{B})$. We need to show that $(\phi(b_1), \ldots, \phi(b_r)) \in R(\mathbf{D}(\mathcal{O}))$, that is, for every $(\mathbf{A}_i, a_i) \in \phi(b_i)$, $i = 1, \ldots, r$ and $\mathbf{T} = \mathbf{C}(R, (\mathbf{A}_1, a_1), \ldots, (\mathbf{A}_r, a_r))$ we have $(\mathbf{T}, a_j) \in \phi(b_j)$, $j = 1, \ldots, r$. We will proceed by contradiction, supposing that there exists an index j such that $(\mathbf{T}, a_j) \notin \phi(b_j)$. First note that we cannot then

9

have $(\mathbf{A}_i, a_i) \in S(b_i)$, $i = 1, \dots, r$. For otherwise the corresponding homomorphisms $\psi_i : \mathbf{A}_i \to \mathbf{B}$ with $\psi_i(a_i) = b_i$, $i = 1, \dots, r$ could be combined into a homomorphism $\psi : \mathbf{T} \to \mathbf{B}$ such that $\psi(a_j) = b_j$, giving $(\mathbf{T}, a_j) \in S(b_j) \subseteq \phi(b_j)$. Thus we can assume that there exists at least one index i such that $(\mathbf{A}_i, a_i) \in \phi(b_i) \setminus S(b_i)$. (We could have $i = j$.) We will suppose that the cardinality $|\{k \mid (\mathbf{A}_k, a_k) \in \phi(b_k) \setminus S(b_k)\}|$ is as small as possible, and derive a contradiction.

Since $(\mathbf{T}, a_j) \notin \phi(b_j)$, there exists $(\mathbf{A}'_j, a'_j)$ such that $S(b_j) \cap (\mathcal{O} - (\mathbf{A}'_j, a'_j)) = \emptyset$ and $(\mathbf{T}, a_j) \in (\mathcal{O} - (\mathbf{A}'_j, a'_j))$. We then have $[(\mathbf{A}'_j, a'_j) + (\mathbf{T}, a_j)] \in \mathcal{O}$. We can rewrite the latter tree as a different sum by moving the root: $[(\mathbf{A}'_j, a'_j) + (\mathbf{T}, a_j)] = [(\mathbf{A}_i, a_i) + (\mathbf{T}', a_i)]$; this implicitly characterizes $\mathbf{T}'$. We then have $(\mathbf{A}_i, a_i) \in (\mathcal{O} - (\mathbf{T}', a_i))$. Thus $(\mathbf{A}_i, a_i) \in \phi(b_i) \not\subseteq \mathbb{T}_{\mathrm{r}} \setminus (\mathcal{O} - (\mathbf{T}'', a''_i))$. By the definition of $\phi(b_i)$, this means that there exists $(\mathbf{A}''_i, a''_i) \in S(b_i) \cap (\mathcal{O} - (\mathbf{T}', a_i))$.

For $k \neq i$, put $(\mathbf{A}''_k, a''_k) = (\mathbf{A}_k, a_k)$. The tree $\mathbf{T}'' = \mathbf{C}(R, (\mathbf{A}''_1, a''_1), \dots, (\mathbf{A}''_r, a''_r))$ is obtained by replacing $(\mathbf{A}_i, a_i)$ by $(\mathbf{A}''_i, a''_i)$ in the concatenation defining $\mathbf{T}$. By the minimality of $|\{k \mid (\mathbf{A}_k, a_k) \in \phi(b_k) \setminus S(b_k)\}|$, we have $(\mathbf{T}'', a''_k) \in \phi(b_k)$, $k = 1, \dots, r$, and in particular $(\mathbf{T}'', a''_j) \in \phi(b_j)$. A second argument will also show $(\mathbf{T}'', a''_j) \notin \phi(b_j)$: Since $S(b_j) \cap (\mathcal{O} - (\mathbf{A}'_j, a'_j)) = \emptyset$, we have $\phi(b_j) \subseteq \mathbb{T}_{\mathrm{r}} \setminus (\mathcal{O} - (\mathbf{A}'_j, a'_j))$. But by moving the root in $[(\mathbf{T}'', a''_j) + (\mathbf{A}'_j, a'_j)]$ we get

$$[(\mathbf{T}'', a''_j) + (\mathbf{A}'_j, a'_j)] = [(\mathbf{T}', a_i) + (\mathbf{A}''_i, a''_i)] \in \mathcal{O},$$

since $(\mathbf{A}''_i, a''_i) \in \mathcal{O} - (\mathbf{T}', a_i)$. Thus $(\mathbf{T}'', a''_j) \in (\mathcal{O} - (\mathbf{A}'_j, a'_j))$, which is disjoint from $\phi(b_j)$. Therefore $(\mathbf{T}'', a''_j) \notin \phi(b_j)$, a contradiction. $\qquad\square$

4.2. Duals of forests

We will use the symbol $+$ to denote the disjoint union of structures including forests (without roots). Accordingly, for $\mathcal{O} \subseteq \mathbb{F}$, we put

$$\mathcal{O} - \mathbf{A} = \{\mathbf{B} \in \mathbb{F} \mid \mathbf{A} + \mathbf{B} \in \mathcal{O}\}.$$

Note that we use $+$ and $-$ also in the context of rooted forests for a related but different notion.

The equivalence $\approx_{\mathcal{O}}$ on $\mathbb{F}$ is defined by

$$\mathbf{A} \approx_{\mathcal{O}} \mathbf{A}' \text{ if } \mathcal{O} - \mathbf{A} = \mathcal{O} - \mathbf{A}'.$$

Lemma 4.4. *If $\mathcal{O} \subseteq \mathbb{F}$ is regular, then there are finitely many $\approx_{\mathcal{O}}$-equivalence classes, and these are regular.*

Proof. Given a forest $\mathbf{A}$, let $(\mathbf{A}^+, x)$ be the rooted forest obtained by adding an isolated element x to $V(\mathbf{A})$ and designating it as the root. Then

$$\mathcal{O} - \mathbf{A} = \{[(\mathbf{B}, b)] \in \mathbb{F} \mid (\mathbf{B}, b) \in \mathcal{O} - (\mathbf{A}^+, x)\}.$$

Since $\mathcal{O}$ is regular, there are finitely many sets of the form $\mathcal{O} - (\mathbf{A}^+, x)$, hence finitely many sets of the form $\mathcal{O} - \mathbf{A}$ and finitely many $\approx_{\mathcal{O}}$ equivalence classes.

Now for $\mathbf{A} \in \mathbb{F}$ and $(\mathbf{B}, b), (\mathbf{C}, c) \in \mathbb{F}_{\mathrm{r}}$ we have $\mathbf{A} + [(\mathbf{B}, b) + (\mathbf{C}, c)] = [(\mathbf{A} + \mathbf{B}, b) + (\mathbf{C}, c)]$, Thus $(\mathcal{O} - \mathbf{A}) - (\mathbf{B}, b) = \mathcal{O} - (\mathbf{A} + \mathbf{B}, b)$. If $\mathcal{O}$ is regular, there are finitely many sets of the form $\mathcal{O} - (\mathbf{A} + \mathbf{B}, b)$, that is, finitely many sets of the form $(\mathcal{O} - \mathbf{A}) - (\mathbf{B}, b)$. Thus $\mathcal{O} - \mathbf{A}$ is regular.

Any $\approx_{\mathcal{O}}$ equivalence class can be obtained as the set of forests $\mathbf{A}$ satisfying finitely many conditions of the form $\mathbf{B}_i \in \mathcal{O} - \mathbf{A}$ or $\mathbf{B}_i \notin \mathcal{O} - \mathbf{A}$. Since $\mathbf{B}_i \in \mathcal{O} - \mathbf{A}$ if and only if $\mathbf{A} \in \mathcal{O} - \mathbf{B}_i$ the equivalence class can be obtained as the intersection of finitely many regular sets or their complements, and by Lemma 3.2 it must be regular. $\qquad\square$

Definition 4.5. Let $\mathcal{O}$ be a regular family of forests. A family $\mathcal{T} \subset \mathbb{T}$ is called $\mathcal{O}$-*admissible* if it satisfies the following properties.

- $\mathcal{T}$ contains a connected component of each element of $\mathcal{O}$.

- If $\mathbf{A} \in \mathcal{T}$ and $\mathbf{A}' \approx_{\mathcal{O}} \mathbf{A}$, then $\mathbf{A}' \in \mathcal{T}$.

By Lemma 4.4, each $\approx_{\mathcal{O}}$-equivalence class $\mathbf{A}/\approx_{\mathcal{O}}$ is regular. Since $\mathbb{T}$ is also regular, $\mathbb{T} \cap (\mathbf{A}/\approx_{\mathcal{O}})$ is regular. An $\mathcal{O}$-admissible family $\mathcal{T}$ is a finite union of such regular families (since $\mathbb{F}/\approx_{\mathcal{O}}$ is finite), hence it is regular. Thus by Theorem 4.3, $\mathcal{T}$ admits a dual $\mathbf{D}(\mathcal{T})$.

Definition 4.6. Let $\mathcal{O}$ be a regular family of forests. The family $\mathcal{D}(\mathcal{O})$ is defined as the set of structures $\mathbf{D}(\mathcal{T})$ such that $\mathcal{T}$ is an $\mathcal{O}$-admissible family of trees.

Theorem 4.7. *Let $\mathcal{O}$ be a regular family of forests. Then $\mathcal{D}(\mathcal{O})$ is finite, and for any σ-structure $\mathbf{B}$, there exists a homomorphism of $\mathbf{B}$ to some member of $\mathcal{D}(\mathcal{O})$ if and only if no member of $\mathcal{O}$ admits a homomorphism to $\mathbf{B}$.*

Proof. The number of elements of $\mathcal{D}(\mathcal{O})$ is at most $2^{|\mathbb{F}/\approx_{\mathcal{O}}|}$ since every $\mathcal{O}$-admissible family is a union of sets of the form $\mathbb{T} \cap (\mathbf{A}/\approx_{\mathcal{O}})$. By Lemma 4.4, $|\mathbb{F}/\approx_{\mathcal{O}}|$ is finite, hence $|\mathcal{D}(\mathcal{O})|$ is finite.

For $\mathbf{A} \in \mathcal{O}$ and $\mathbf{D}(\mathcal{T}) \in \mathcal{D}(\mathcal{O})$, there exists a component $\mathbf{T}$ of $\mathbf{A}$ which belongs to $\mathcal{T}$. We have $\mathbf{T} \nrightarrow \mathbf{D}(\mathcal{T})$ whence $\mathbf{A} \nrightarrow \mathbf{D}(\mathcal{T})$. Thus if a σ-structure $\mathbf{B}$ admits a homomorphism from some $\mathbf{A} \in \mathcal{O}$, then there is no homomorphism of $\mathbf{B}$ to any member of $\mathcal{D}(\mathcal{O})$.

Now let $\mathbf{B}$ be a σ-structure such that no member of $\mathcal{O}$ admits a homomorphism to $\mathbf{B}$. Put $\mathcal{S}(B) = \{\mathbf{T} \in \mathbb{T} \mid \mathbf{T} \nrightarrow \mathbf{B}\}$ and

$$\mathcal{T}(B) = \{\mathbf{T} \in \mathbb{T} \mid (\mathbf{T}/\approx_{\mathcal{O}}) \subseteq \mathcal{S}(B)\}.$$

Then $\mathcal{S}(B)$ contains a component of every $\mathbf{A}$ in $\mathcal{O}$. We need to show that the same holds for $\mathcal{T}(B)$.

Suppose for a contradiction that some $\mathbf{A} \in \mathcal{O}$ has no component $\mathbf{T} \in \mathcal{S}(B)$ such that $(\mathbf{T}/\approx_{\mathcal{O}}) \subseteq \mathcal{S}(B)$. Let $\mathbf{T}_1, \ldots, \mathbf{T}_k$ be the components of $\mathbf{A}$ in $\mathcal{S}(B)$ such that there exists $\mathbf{T}'_i$ with $\mathbf{T}'_i \to \mathbf{B}$ and $\mathbf{T}'_i \approx_{\mathcal{O}} \mathbf{T}_i$, $i = 1, \ldots, k$. Put $\mathbf{A}_0 = \mathbf{A}$, and for $i = 1, \ldots, k$, let $\mathbf{A}_i$ be the forest obtained from $\mathbf{A}_{i-1}$ by replacing the component $\mathbf{T}_i$ by $\mathbf{T}'_i$. Then by definition of $\approx_{\mathcal{O}}$, we have $\mathbf{A}_i \in \mathcal{O}$ for all i, hence $\mathbf{A}_k \in \mathcal{O}$. However, $\mathbf{A}_k \to \mathbf{B}$, which contradicts the fact that no member of $\mathcal{O}$ admits a homomorphism to $\mathbf{B}$.

Thus $\mathcal{T}(B)$ is a union of $\approx_{\mathcal{O}}$ classes which contains a component of every $\mathbf{A}$ in $\mathcal{O}$. Thus by definition, $\mathcal{T}(B)$ is $\mathcal{O}$-admissible, and $\mathbf{D}(\mathcal{T}(B)) \in \mathcal{D}(\mathcal{O})$. No member of $\mathcal{T}(B)$ admits a homomorphism to $\mathbf{B}$, thus $\mathbf{B}$ admits a homomorphism to $\mathbf{D}(\mathcal{T}(B))$. $\qquad\square$

5. Antichain dualities and forests

A duality pair $(\mathcal{O}, \mathcal{D})$ is called an *antichain duality* if $\mathcal{O} \cup \mathcal{D}$ is an antichain. We will suppose that $\mathcal{O}$ consists of cores, since any element of $\mathcal{O}$ can be retracted to its core.

Recall that we consider the the relation $\mathbf{A} \to \mathbf{B}$ as a preorder on the σ-structures. Keeping with this view we call a structure $\mathbf{D}$ *minimal* in a family $\mathcal{D}$ of structures if $\mathbf{D} \in \mathcal{D}$ and for any $\mathbf{C} \in \mathcal{D}$ with $\mathbf{C} \to \mathbf{D}$ we have $\mathbf{D} \to \mathbf{C}$. We use *maximal* for the converse notion.

Lemma 5.1. *Let $(\mathcal{O}, \mathcal{D})$ be a duality pair such that $\mathcal{D}$ is finite and the members of $\mathcal{O}$ are cores. Then the minimal members of $\mathcal{O}$ are forests.*

Proof. The proof uses the concept of direct product of structures and its property that for any three structures $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ we have $\mathbf{A} \to \mathbf{B} \times \mathbf{C}$ if and only if $\mathbf{A} \to \mathbf{B}$ and $\mathbf{A} \to \mathbf{C}$. See e.g., [12]. Note that this characterization is the dual of the characterization of disjoint union: $\mathbf{A} + \mathbf{B}$ admits a homomorphism to $\mathbf{C}$ if and only if $\mathbf{A} \to \mathbf{C}$ and $\mathbf{B} \to \mathbf{C}$.

Suppose that $\mathbf{A}$ is a minimal element of $\mathcal{O}$ and it is no forest. Let $\mathbf{A}_1$ be a non-tree component of $\mathbf{A}$ and let $\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_0$. Here $\mathbf{A}_0$ is empty if $\mathbf{A}$ is connected. Note that $\mathbf{A}_1$, as a component of the core structure $\mathbf{A}$ is itself a core. We have $\mathbf{A}_0 \to \mathbf{A}$ but $\mathbf{A} \nrightarrow \mathbf{A}_0$ as $\mathbf{A}$ is core. If any member of $\mathcal{O}$ would admit a homomorphism to $\mathbf{A}_0$ this would contradict the minimality of $\mathbf{A}$ in $\mathcal{O}$. Thus, there exist a at least one structure $\mathbf{D} \in \mathcal{D}$ with $\mathbf{A}_0 \to \mathbf{D}$.

Let $\mathbf{B} = \mathbf{A}_1 \times \mathbf{D}$ be a maximal element in the non-empty finite family $\{\mathbf{A}_1 \times \mathbf{D} \mid \mathbf{D} \in \mathcal{D}, \mathbf{A}_0 \to \mathbf{D}\}$. We have $\mathbf{B} \to \mathbf{A}_1$, but $\mathbf{A}_1 \to \mathbf{B}$ would imply $\mathbf{A}_1 \to \mathbf{D}$ and together with $\mathbf{A}_0 \to \mathbf{D}$ it would also imply $\mathbf{A} \to \mathbf{D}$, a contradiction.

Corollary 3.5 of [12] can be applied here as $\mathbf{A}_1$ is connected, core and not a tree and we have $\mathbf{B} \to \mathbf{A}_1 \nrightarrow \mathbf{B}$. The Corollary states that there exists a σ-structure $\mathbf{C}_1$ such that $\mathbf{B} \to \mathbf{C}_1 \to \mathbf{A}_1$ and $\mathbf{A}_1 \nrightarrow \mathbf{C}_1 \nrightarrow \mathbf{B}$. Let us fix such a structure $\mathbf{C}_1$ and let $\mathbf{C} = \mathbf{C}_1 + \mathbf{A}_0$. We have $\mathbf{C} \to \mathbf{A} \nrightarrow \mathbf{C}$ and since $\mathbf{A}$ is minimal in $\mathcal{O}$, no member of $\mathcal{O}$ admits a homomorphism to $\mathbf{C}$.

Therefore, there exists an element $\mathbf{D}'$ of $\mathcal{D}$ such that $\mathbf{C} \to \mathbf{D}'$. We then have $\mathbf{B} \to \mathbf{C}_1 \to \mathbf{A}_1 \times \mathbf{D}'$ and $\mathbf{A}_1 \times \mathbf{D}' \not\to \mathbf{B}$ (since $\mathbf{C}_1 \not\to \mathbf{B}$). This contradicts the maximality of $\mathbf{B}$ as we also have $\mathbf{A}_0 \to \mathbf{D}'$. $\qquad\square$

Lemma 5.2. *Let $(\mathcal{O}, \mathcal{D})$ be an antichain duality such that the members of $\mathcal{O}$ are cores. Then $|\mathcal{D}| = 1$ if and only if the members of $\mathcal{O}$ are connected.*

Proof. Suppose that the members of $\mathcal{O}$ are connected. If $\mathcal{D}$ contains two members $\mathbf{D}_1$ and $\mathbf{D}_2$, then $\mathbf{D}_1 \cup \mathbf{D}_2$ does not admit a homomorphism to any member of $\mathcal{D}$, therefore there exists some $\mathbf{A}$ in $\mathcal{O}$ which admits a homomorphism to $\mathbf{D}_1 \cup \mathbf{D}_2$. Since $\mathbf{A}$ is connected, this means that $\mathbf{A}$ admits a homomorphism to $\mathbf{D}_1$ or $\mathbf{D}_2$, a contradiction.

Conversely, suppose that some $\mathbf{A} \in \mathcal{O}$ has connected components $\mathbf{A}_1, \ldots, \mathbf{A}_n$, where $n \geq 2$. For $i = 1, \ldots, n$, put $\mathbf{B}_i = \bigcup_{j \neq i} \mathbf{A}_j$. Then we have $\mathbf{B}_i \to \mathbf{A} \not\to \mathbf{B}_i$, $i = 1, \ldots, n$, therefore no member of $\mathcal{O}$ admits a homomorphism to any $\mathbf{B}_i$. Thus there is a function $\delta : \{1, \ldots, n\} \to \mathcal{D}$ such that $\mathbf{B}_i \to \delta(i)$. For $i \neq j$, we must have $\delta(i) \neq \delta(j)$, otherwise $\mathbf{A} \to \delta(i)$. Therefore $|\mathcal{D}| \geq n$. $\qquad\square$

Theorem 5.3. *An antichain $\mathcal{O}$ of core σ-structures has a finite dual $\mathcal{D}$ if and only if $\mathcal{O}$ is a regular family of forests.*

For the proof of Theorem 5.3 we need that whenever the upward closure of an antichain $\mathcal{O}$ of core forests is regular, so is $\mathcal{O}$ itself. This follows from the following more general observation.

Theorem 5.4. *The cores of the minimal elements of a regular set of forests form a regular set.*

This result is more subtle than it looks. There are regular families $\mathcal{O}$ such that none of (a) the minimal elements in $\mathcal{O}$, (b) the cores among the elements of $\mathcal{O}$, or (c) the cores of all elements in $\mathcal{O}$ form regular languages. To see this, consider the type of directed graphs, consider the family $\mathcal{O}$ consisting of the oriented paths $P_{ij} = p(++(+-+)^i ++--(-+-)^j --)$, where a word $x_1 \ldots x_n \in \{+, -\}^n$ describes the orientation $p(x)$ of the n edge path whose kth edge is directed forward if $x_k = +$ and directed backward if $x_k = -$. Here P_{ij} is core if and only if $i \neq j$. The minimal elements in $\mathcal{O}$ are the oriented paths P_{ii}. The core of P_{ii} is $p(++(+-+)^i ++)$ and these latter oriented paths do form a regular set as opposed to the sets mentioned in (a–c).

We start with the proof of Theorem 5.3 using Theorem 5.4 and then we will prove Theorem 5.4 in Section 6.

Proof of Theorem 5.3 using Theorem 5.4. The "if" part of the statement readily follows from Theorem 4.7.

For the "only if" part assume $\mathcal{O}$ is an antichain of core σ-structures and it has a finite dual. By Lemma 5.1 $\mathcal{O}$ must consist of forests. Applying

Corollary 3.4 we obtain that the family $\mathcal{O}_\mathcal{D}$ of structures with no homomorphism to a member of $\mathcal{D}$ is regular. But the cores of the minimal elements in this family are exactly the elements of $\mathcal{O}$, so Theorem 5.4 finishes the proof. $\qquad\square$

Note that in a similar fashion Theorem 5.4 implies the following strengthening of Lemma 5.1: Let $(\mathcal{O}, \mathcal{D})$ be a duality pair with $\mathcal{D}$ finite and $\mathcal{O}$ consisting of cores. Then the minimal elements of $\mathcal{O}$ form a regular family of forests.

6. Cores of minimals of regular set of forest are regular

This section is devoted to the proof of Theorem 5.4. Before proving Theorem 5.4 we rephrase it.

Using category theoretic conventions we call a homomorphism $f : A \to B$ a *retraction* if it has a right inverse, namely a homomorphism $g : B \to A$ with $f \circ g$ being the identity on B. For brevity we write *non-retraction* for a homomorphism that is not a retraction.

The following characterization of non-retractions will be useful.

Proposition 6.1. *A homomorphism $h : \mathbf{A} \to \mathbf{B}$ is a non-retraction if and only if there is a component $\mathbf{C}$ of $\mathbf{B}$ such that the restriction of h to no substructure $\mathbf{D} \subseteq \mathbf{A}$ gives a $\mathbf{D} \to \mathbf{C}$ isomorphism.*

Proof. If $h \circ g : \mathbf{B} \to \mathbf{B}$ is the identity and $\mathbf{C}$ is a component of $\mathbf{B}$, then the restriction of h to the substructure $\mathbf{D}$ of $\mathbf{A}$ induced by the image of $\mathbf{C}$ under g is an isomorphism $\mathbf{D} \to \mathbf{C}$.

Conversely, if there is such a substructure $\mathbf{D}$, for each component $\mathbf{C}$ of $\mathbf{B}$, then the inverses $g_{\mathbf{C}}$ of the $\mathbf{D} \to \mathbf{C}$ isomorphisms give the right inverse $g : \mathbf{B} \to \mathbf{A}$ of h as their union. $\qquad\square$

For $\mathcal{O} \subseteq \mathbb{F}$ we define

$$
\begin{aligned}
\mathrm{UP}(\mathcal{O}) &= \{\mathbf{A} \in \mathbb{F} \mid \exists \mathbf{T} \in \mathcal{O}, \mathbf{T} \to \mathbf{A}\}, \\
\mathrm{EX}(\mathcal{O}) &= \{\mathbf{A} \in \mathbb{F} \mid \exists \mathbf{T} \in \mathcal{O} \text{ and a non-retraction } h : \mathbf{T} \to \mathbf{A}\}.
\end{aligned}
$$

Proposition 6.2. *For $\mathcal{O} \subseteq \mathbb{F}$ the family $\mathrm{UP}(\mathcal{O}) \setminus \mathrm{EX}(\mathcal{O})$ is the set of the cores of the minimal elements in $\mathcal{O}$.*

Proof. Let $\mathbf{A} \in \mathbb{F}$ be the core of minimal element $\mathbf{B}$ of $\mathcal{O}$. Clearly, $\mathbf{B} \to \mathbf{A}$ ensures that $\mathbf{A} \in \mathrm{UP}(\mathcal{O})$. Let us consider a homomorphism $h : \mathbf{T} \to \mathbf{A}$ with $\mathbf{T} \in \mathcal{O}$. From the minimality of $\mathbf{B}$ and from $\mathbf{T} \to \mathbf{A} \to \mathbf{B}$, we also have $\mathbf{B} \to \mathbf{T}$ and hence also a homomorphism $g : \mathbf{A} \to \mathbf{T}$. As $\mathbf{A}$ is a core, the homomorphism $h \circ g : \mathbf{A} \to \mathbf{A}$ must be an isomorphism. But then $g \circ (h \circ g)^{-1}$ is a right inverse of h, so h is a retraction. This shows that the cores of the minimal elements of $\mathcal{O}$ are contained in $\mathrm{UP}(\mathcal{O}) \setminus \mathrm{EX}(\mathcal{O})$.

It remains to show that each forest $\mathbf{A} \in \mathrm{UP}(\mathcal{O}) \setminus \mathrm{EX}(\mathcal{O})$ is indeed the core of a minimal element of $\mathcal{O}$. From $\mathbf{A} \in \mathrm{UP}(\mathcal{O})$ we have a homomorphism $h : \mathbf{B} \to \mathbf{A}$ with some $\mathbf{B} \in \mathcal{O}$. As $\mathbf{A} \notin \mathrm{EX}(\mathcal{O})$, the homomorphism h is a retraction and in particular $\mathbf{A}$ is homomorphically equivalent to $\mathbf{B}$. To show that $\mathbf{B}$ is minimal in $\mathcal{O}$, consider an arbitrary $\mathbf{C} \in \mathcal{O}$ with $\mathbf{C} \to \mathbf{B}$ and note that $\mathbf{C} \to \mathbf{B} \to \mathbf{A}$ implies (as above) that $\mathbf{C}$ and $\mathbf{A}$ (and hence also $\mathbf{C}$ and $\mathbf{B}$) are homomorphically equivalent. Finally let us consider an arbitrary homomorphism $\phi : \mathbf{A} \to \mathbf{A}$. The homomorphism $\phi \circ h : \mathbf{B} \to \mathbf{A}$ is a retraction, so we have $g : \mathbf{A} \to \mathbf{B}$ with $\phi \circ h \circ g : \mathbf{A} \to \mathbf{A}$ the identity. This implies that ϕ itself must be an automorphism and thus $\mathbf{A}$ is a core. This finishes the proof of the proposition. $\square$

Recall that by Lemma 3.2 the difference between regular sets is also regular. Thus Proposition 6.2 above and the following two propositions together imply Theorem 5.4.

Proposition 6.3. *For a regular set $\mathcal{O} \subseteq \mathbb{F}$ its upward closure $\mathrm{UP}(\mathcal{O})$ is also regular.*

Proof. By Theorem 4.7 $\mathcal{O}$ has a finite dual $\mathcal{D}$. By Corollary 3.4 $\mathrm{UP}(\mathcal{O}) = \mathcal{O}_{\mathcal{D}}$ is regular. $\square$

Note here that a direct proof of this result would only be simpler than the proof of Proposition 6.4 below by not having to distinguish danger points from safe points. The proofs give a doubly exponential bound on the number of equivalence classes of $\sim_{\mathrm{UP}(\mathcal{O})}$ or $\sim_{\mathrm{EX}(\mathcal{O})}$ in terms of the number of equivalence classes of $\sim_{\mathcal{O}}$ and we believe that the number of equivalence classes can indeed be that high for some regular sets $\mathcal{O}$.

Proposition 6.4. *For any regular set $\mathcal{O} \subseteq \mathbb{F}$ the set $\mathrm{EX}(\mathcal{O})$ is regular.*

Proof. For simplicity we assume that the type σ contains no unary relations. The proof works basically the same way in the presence of unary relations too, but making this mild assumption makes our presentation simpler. We start with a few definitions.

Recall that each forest has a unique decomposition as the disjoint union of trees, its components. For a rooted forest $(\mathbf{X}, v) \in \mathbb{F}_{\mathrm{r}}$ we denote its rooted component by $(\mathbf{X}, v)_+$, that is $[(\mathbf{X}, v)_+]$ is the component of $\mathbf{X}$ containing v and $(\mathbf{X}, v)_+ = ([(\mathbf{X}, v)_+], v) \in \mathbb{T}_{\mathrm{r}}$. Let $(\mathbf{X}, v)_- \in \mathbb{F}$ stand for the union of the remaining components of $\mathbf{X}$, that is $\mathbf{X} = [(\mathbf{X}, v)_+] + (\mathbf{X}, v)_-$. For a tree $\mathbf{A} \in \mathbb{T}$ and a homomorphism $h : \mathbf{A} \to [(\mathbf{X}, v)_+]$ we define $h^0 = \{w \in V(\mathbf{A}) \mid h(w) = v\}$ to be the set of *root points*, points mapped by h to the root of $(\mathbf{X}, v)$. We say that $w \in h^0$ is a *danger point of h* if a restriction of h to a suitable substructure of $\mathbf{A}$ containing w is an isomorphism to $[(\mathbf{X}, v)_+]$. Let h^1 stand for the set of danger points of h and note that $h^1 \subseteq h^0$. We write $h^2 = h^0 \setminus h^1$ stand for the *safe points of h*. Note that we slightly abuse

notation by not indicating the dependence of h^i on $(\mathbf{X}, v)_+$, but this will lead to no confusion.

Let us consider $\mathbf{A} \in \mathbb{T}$, $(\mathbf{X}, v) \in \mathbb{F_r}$ and a homomorphism $h : \mathbf{A} \to [(\mathbf{X}, v)_+]$. We define the $(\mathbf{E}, \mathbf{E}_0)$-*extension* of $\mathbf{A}$ for $\mathbf{E}_0 \in \mathbb{F}$ and $\mathbf{E} : h^0 \to \mathbb{T_r}$ to be the forest $\mathbf{B}$ obtained by gluing a copy of $\mathbf{E}(w)$ to w for each $w \in h^0$ and further adding $\mathbf{E}_0$. More precisely, we obtain $\mathbf{B}$ by taking the disjoint union of the forests $\mathbf{A}$, $\mathbf{E}_0$ and $[\mathbf{E}(w)]$ for each $w \in h^0$ and then identifying the root of $\mathbf{E}(w)$ with w for each $w \in h^0$. We will identify the forests $\mathbf{A}$, $\mathbf{E}_0$ and $[\mathbf{E}(w)]$ with the corresponding substructures of $\mathbf{B}$ as long as this leads to no confusion.

Let us fix the regular set $\mathcal{O} \subseteq \mathbb{F}$. We call a map $L : \mathbb{F_r}/\sim_{\mathcal{O}} \to \{0, 1, 2\}$ a *list*.

Let $\mathbf{A} \in \mathbb{T}$, $(\mathbf{X}, v) \in \mathbb{F_r}$, $h : \mathbf{A} \to [(\mathbf{X}, v)_+]$ and let $\mathbf{B}$ be the $(\mathbf{E}, \mathbf{E}_0)$-extension of $\mathbf{A}$. We define the *signature* of this extension to be the triple $(L_1, L_2, \mathcal{C}')$, where $\mathcal{C}' = \mathbf{E}_0/\approx_{\mathcal{O}} \in \mathbb{F}/\approx_{\mathcal{O}}$ and L_1 and L_2 are lists defined by $L_i(\mathcal{C}) = \min(2, |\{w \in h^i \mid \mathbf{E}(w) \in \mathcal{C}\}|)$ for $i = 1, 2$ and $\mathcal{C} \in \mathbb{F_r}/\sim_{\mathcal{O}}$.

We say that $(\mathbf{X}, v) \in \mathbb{F_r}$ is *compatible* with the signature T if there is $\mathbf{A} \in \mathbb{T}$, $h : \mathbf{A} \to [(\mathbf{X}, v)_+]$ and an extension $\mathbf{B}$ of $\mathbf{A}$ of signature T with $\mathbf{B} \in \mathcal{O}$. We say that $(\mathbf{X}, v) \in \mathbb{F_r}$ is *compatible* with $\mathcal{C}' \in \mathbb{F}/\approx_{\mathcal{O}}$ if there is a forest $\mathbf{A} \in \mathcal{C}'$ satisfying $\mathbf{A} \to (\mathbf{X}, v)_-$. If a non-retraction $\mathbf{A} \to (\mathbf{X}, v)_-$ also exists from such a forest $\mathbf{A} \in \mathcal{C}'$, we say that $(\mathbf{X}, v)$ is *strongly compatible* with $\mathcal{C}'$.

Note that since $\mathcal{O}$ is regular, $\mathbb{F_r}/\sim_{\mathcal{O}}$ and $\mathbb{F}/\approx_{\mathcal{O}}$ must be finite (the latter by Lemma 4.4), there is a finite number of different signatures. Let us accept Proposition 6.5 below. It implies that $\sim_{\mathrm{EX}(\mathcal{O})}$ has a finite number equivalence classes, in other words, that $\mathrm{EX}(\mathcal{O})$ is regular, finishing the proof of Proposition 6.4. $\qquad\square$

To complete the proof above it remains to prove the following:

Proposition 6.5. *If $(\mathbf{X}, v), (\mathbf{X}', v) \in \mathbb{F_r}$ are compatible with the same signatures and the same elements of $\mathbb{F}/\approx_{\mathcal{O}}$ and they are also strongly compatible with the same elements of $\mathbb{F}/\approx_{\mathcal{O}}$ and further $(\mathbf{X}, v)_+ = \mathbf{T}_0$ if and only if $(\mathbf{X}', v)_+ = \mathbf{T}_0$, then $(\mathbf{X}, v) \sim_{\mathrm{EX}(\mathcal{O})} (\mathbf{X}', v)$.*

Proof. Let $(\mathbf{X}, v), (\mathbf{X}', v)$ satisfy the condition of the proposition and $(\mathbf{Y}, v) \in \mathbb{F_r}$ satisfy $[(\mathbf{X}, v) + (\mathbf{Y}, v)] \in \mathrm{EX}(\mathcal{O})$. By symmetry it is enough to prove that $[(\mathbf{X}', v) + (\mathbf{Y}, v)] \in \mathrm{EX}(\mathcal{O})$.

By the definition of $\mathrm{EX}(\mathcal{O})$ we have a non-retraction $h : \mathbf{B} \to [(\mathbf{X}, v) + (\mathbf{Y}, v)]$ for some $\mathbf{B} \in \mathcal{O}$. Our goal is to find another non-retraction $h' : \mathbf{B}' \to [(\mathbf{X}', v) + (\mathbf{Y}, v)]$ with $\mathbf{B}' \in \mathcal{O}$. We do the transformation step by step. In each step we have some $\mathbf{B}_i \in \mathcal{O}$ and a homomorphism $h_i : \mathbf{B}_i \to [(\mathbf{X}, v)_+ + (\mathbf{X}', v) + (\mathbf{Y}, v)]$ using "less and less" the part of $[(\mathbf{X}, v)_+ + (\mathbf{X}', v) + (\mathbf{Y}, v)]$ coming from $(\mathbf{X}, v)_+$.

Consider the substructure $\mathbf{A}_0$ of $\mathbf{B}$ induced by the vertices h maps to $(\mathbf{X}, v)_-$. Clearly, $(\mathbf{X}, v)$ is compatible with $\mathcal{C}' = \mathbf{A}_0/\approx$. If h restricted to $\mathbf{A}_0$ is a non-retraction $\mathbf{A}_0 \to (\mathbf{X}, v)_-$, then $(\mathbf{X}, v)$ is strongly compatible with $\mathcal{C}'$. So $(\mathbf{X}', v)$ must also be compatible with $\mathcal{C}'$ and, in the latter case, also strongly compatible with $\mathcal{C}'$. Let $\mathbf{A}_0' \in \mathcal{C}'$ and the homomorphism $g_0' : \mathbf{A}_0' \to (\mathbf{X}', v)_-$ show this, that is g_0' is a non-retraction if $(\mathbf{X}, v)$ is strongly compatible with $\mathcal{C}'$. Let $\mathbf{B}_1 = \mathbf{A}_0' + \mathbf{B}^*$, where $\mathbf{B}^*$ is the substructure of $\mathbf{B}$ outside $\mathbf{A}_0$. From $\mathbf{B} = \mathbf{A}_0 + \mathbf{B}^* \in \mathcal{O}$ and $\mathbf{A} \approx_\mathcal{O} \mathbf{A}'$ we have $\mathbf{B}_1 \in \mathcal{O}$. We define $h_1 : \mathbf{B}_1 \to [(\mathbf{X}, v)_+ + (\mathbf{X}', v) + (\mathbf{Y}, v)]$ by making its restriction to $\mathbf{A}_0'$ be g_0' (this maps to $(\mathbf{X}', v)_-$) and making its restriction to $\mathbf{B}^*$ be the same as the corresponding restriction of h (mapping to $[(\mathbf{X}, v)_+ + (\mathbf{Y}, v)]$).

Next we give the recursive step. Assume $\mathbf{B}_i \in \mathcal{O}$ and $h_i : \mathbf{B}_i \to [(\mathbf{X}, v)_+ + (\mathbf{X}', v) + (\mathbf{Y}, v)]$ is given. We define the *old-parts* of $\mathbf{B}_i$ to be the maximal connected substructures of $\mathbf{B}_i$ that h_i maps into $[(\mathbf{X}, v)_+]$ (considered as a substructure of $[(\mathbf{X}, v)_+ + (\mathbf{X}', v) + (\mathbf{Y}, v)]$). We exclude single vertex substructures mapped to v and do not consider these old-parts. We measure progress by the decreasing number of old-parts, that is, we make sure that $\mathbf{B}_{i+1}$ has fewer old-parts than $\mathbf{B}_i$. This ensures that the procedure terminates with no old-parts left. If $\mathbf{B}_i$ has no old-parts, then h_i maps the entire structure $\mathbf{B}_i$ to $[(\mathbf{X}', v) + (\mathbf{Y}, v)]$. In this case we set $\mathbf{B}' = \mathbf{B}_i$ and $h' = h_i$.

Assume now that there is still at least one old-part of $\mathbf{B}_i$. As $\mathrm{Inc}(\mathbf{B}_i)$ is a forest with the old-parts being pairwise disjoint subtrees we can choose an old-part $\mathbf{A}_i$ that does not separate two further old-parts in this graph. Let us fix such an old-part $\mathbf{A}_i$ and let $g_i : \mathbf{A}_i \to [(\mathbf{X}, v)_+]$ denote the restriction of h_i to $\mathbf{A}_i$. There is a unique way to express $\mathbf{B}_i$ as an $(\mathbf{E}_i, \mathbf{E}_{i,0})$-extension of $\mathbf{A}_i$: we set $\mathbf{E}_{i,0}$ to be the forest consisting of the components of $\mathbf{B}_i$ other than the component $\mathbf{A}_i^*$ containing $\mathbf{A}_i$, while for $w \in g_i^0$ we set $\mathbf{E}_i(w)$ to be the substructure of $\mathbf{A}_i^*$ that is separated from $\mathbf{A}_i$ by w in the tree $\mathrm{Inc}(\mathbf{A}^*)$. We make $\mathbf{E}_i(w)$ include w as its root. Note that for $w \in V(\mathbf{A}_i) \setminus g_i^0$ the image $g_i(w)$ of w is a not v, so such a w cannot separate the old-part $\mathbf{A}_i$ from any points in $\mathbf{A}_i^*$ as otherwise $\mathbf{A}_i$ would not be a maximal connected substructure mapped to $[(\mathbf{X}, v)_+]$.

By our choice of $\mathbf{A}_i$, if there are further old-parts in $\mathbf{A}_i^*$ they must all be contained in a single substructure $[\mathbf{E}_i(w)]$. In this case we denote the corresponding vertex $w \in g_i^0$ as w_i. In case $\mathbf{A}_i$ is the only old-part of $\mathbf{B}_i$ no vertex w_i is specified.

Let T_i be the signature of the $(\mathbf{E}_i, \mathbf{E}_{i,0})$-extension $\mathbf{B}_i$ of $\mathbf{A}_i$. Now $\mathbf{A}_i \in \mathbb{T}$, $g_i : \mathbf{A} \to [(\mathbf{X}, v)_+]$ and $\mathbf{B}_i \in \mathcal{O}$ shows that $(\mathbf{X}, v)$ is compatible with T_i. By our assumption $(\mathbf{X}', v)$ must also be compatible with the signature T_i. Let $\mathbf{A}_i' \in \mathbb{T}$, $g_i' : \mathbf{A}' \to [(\mathbf{X}', v)_+]$ and its $(\mathbf{E}_i', \mathbf{E}_{i,0}')$-extension $\mathbf{B}_i' \in \mathcal{O}$ show this, that is, assume this extension has signature T_i.

It is tempting at this point to define $\mathbf{B}_{i+1} = \mathbf{B}_i'$ but then we would need to find a homomorphism $h_{i+1} : \mathbf{B}_i' \to [(\mathbf{X}, v)_+ + (\mathbf{X}', v) + (\mathbf{Y}, v)]$,

which may not exist as some of the forests $[\mathbf{E}'_i(w)]$ or $\mathbf{E}'_{i,0}$ may not have a homomorphism to $[(\mathbf{X}, v)_+ + (\mathbf{X}', v) + (\mathbf{Y}, v)]$. We need to replace the rooted trees $\mathbf{E}'_i(w)$ with a $\sim_{\mathcal{O}}$-equivalent rooted tree $\mathbf{E}(w')$ first to make the homomorphism possible.

Let $T_i = (L_{i,1}, L_{i,2}, \mathcal{C}'_i)$. We define $f_i : g'^0_i \to g^0_i$ as follows. For each $j = 1, 2$ and $w \in g'^j_i$ we choose $f_i(w) \in g^j_i$ with $\mathbf{E}_i(f_i(w)) \sim_{\mathcal{O}} \mathbf{E}'_i(w)$. The number of possible choices for $w' = f_i(w)$ is $|\{w' \in g^j_i \mid \mathbf{E}_i(w') \sim_{\mathcal{O}} \mathbf{E}'_i(w)\}| \geq L_{i,j}(\mathbf{E}'_i(w)/\sim_{\mathcal{O}}) \geq 1$, so such a choice is always available. If there are old-parts in $\mathbf{A}^*_i$ beyond $\mathbf{A}_i$, we pick $f_i(w) = w_i$ for at most a single $w \in g'^0_i$. This is also possible because if we have more than one w for which $f_i(w) = w_i$ is possible at all, then $L_{i,j}(\mathbf{E}_i(w_i)/\sim_{\mathcal{O}}) = 2$ for the corresponding value of $j = 1$ or 2, so we have the freedom not to choose $f_i(w) = w_i$ for any w.

We set $\mathbf{E}''_i = \mathbf{E}_i \circ f_i$, that is for $w \in g'^0_i$ we have $\mathbf{E}''_i(w) = \mathbf{E}_i(f_i(w))$. We set $\mathbf{B}_{i+1}$ to be the $(\mathbf{E}''_i, \mathbf{E}_{i,0})$-extension of $\mathbf{A}'_i$. (Note here that $f_i(w) = f_i(w')$ for distinct root vertices of g'_i is possible as long as $f_i(w) \neq w_i$. In this case we have $\mathbf{E}''_i(w) = \mathbf{E}''_i(w') = \mathbf{E}_i(f_i(w))$, and the two substructures $[\mathbf{E}''_i(w)]$ and $[\mathbf{E}''_i(w')]$ are isomorphic, but naturally they are disjoint substructures. This shows the limits of our notation, but hopefully leads to no confusion.) We define $h_{i+1} : \mathbf{B}_{i+1} \to [(\mathbf{X}, v)_+ + (\mathbf{X}', v) + (\mathbf{Y}, v)]$ through its restrictions. The restriction to $\mathbf{A}'_i$ is g'_i, the restriction to $\mathbf{E}_{i,0}$ is the restriction of h_i to $\mathbf{E}_{i,0}$. Finally, for $w \in g'^0_i$ the restriction of h_{i+1} to $[\mathbf{E}''_i(w)]$ is the restriction of h_i to $[\mathbf{E}_i(f_i(w))]$. These restrictions uniquely define the homomorphism h_{i+1} as the given substructures cover $\mathbf{B}_{i+1}$, only the root vertices in g'^0_i are covered more than once and these points are mapped to v in all the given restrictions.

For the recursive definition to work we need to show that $\mathbf{B}_{i+1}$ is in $\mathcal{O}$ and it has fewer old-parts than $\mathbf{B}_i$.

We start with showing that $\mathbf{B}_{i+1} \in \mathcal{O}$. Note that $\mathbf{E}_{i,0} \approx_{\mathcal{O}} \mathbf{E}'_{i,0}$ since both of them are contained in the $\approx_{\mathcal{O}}$ equivalence class $\mathbf{C}'_i$ (part of the signature T_i). As the $(\mathbf{E}'_i, \mathbf{E}'_{i,0})$-extension of $\mathbf{A}'_i$ is $\mathbf{B}'_i \in \mathcal{O}$ the $(\mathbf{E}'_i, \mathbf{E}_{i,0})$-extension of $\mathbf{A}'_i$ must also be in $\mathcal{O}$. The structure $\mathbf{B}_{i+1}$ differs from this last structure by being obtained as an extension using the function $\mathbf{E}''_i$ instead of $\mathbf{E}'_i$. As we always have $\mathbf{E}'_i(w) \sim_{\mathcal{O}} \mathbf{E}''_i(w)$, a similar argument applies: a single such change does not alter membership in $\mathcal{O}$. Doing these changes one by one we obtain eventually that $\mathbf{B}_{i+1} \in \mathcal{O}$ as claimed.

Consider now the old-parts of $\mathbf{B}_{i+1}$. Some may be found in $\mathbf{E}_{i,0}$, but these are also old-parts of $\mathbf{B}_i$. Others may be found in some $[\mathbf{E}''_i(w)]$ but only if $[\mathbf{E}_i(f_i(w))]$ contains an old-part of $\mathbf{B}_i$. By our assumption this holds only for at most a single $w \in g'^0_i$ with $f_i(w) = w_i$, so we have no more old-parts in $\mathbf{B}_{i+1}$ than in $\mathbf{B}_i$. In fact, we have fewer as $\mathbf{A}_i$ was an old-part in $\mathbf{B}_i$ and it got replaced by $\mathbf{A}'_i$ that is mapped to $(\mathbf{X}', v)$.

We have defined the homomorphism $h' : \mathbf{B}' \to [(\mathbf{X}', v) + (\mathbf{Y}, v)]$. Before finishing the proof of Proposition 6.5 by showing that this is a non-retraction

we make an easy observation. We call two distinct points in the universe of a σ-structure *neighbors* if they appear in a common block, that is, if they are in distance two in the incidence graph. For any step i in the procedure above and any $w \in g_i'^0$ the homomorphism h_{i+1} maps all neighbors of w in $[\mathbf{E}_i''(w)] \subseteq \mathbf{B}_{i+1}$ to non-root points of $(\mathbf{Y}, v)$. This is shown by an easy induction together with the statement that no point in $\mathbf{B}_i$ has both a neighbor that h_i maps to a non-root point of $(\mathbf{X}, v)_+$ and another neighbor that h_i maps to a non-root point of $(\mathbf{X}', v)_+$.

Now we turn to the proof of h' being a non-retraction. By Proposition 6.1 and since $h : \mathbf{B} \to [(\mathbf{X}, v) + (\mathbf{Y}, v)]$ is a non-retraction, $[(\mathbf{X}, v) + (\mathbf{Y}, v)]$ has a component $\mathbf{Z}$ with no restriction of h being an isomorphism to $\mathbf{Z}$.

If $\mathbf{Z} \subseteq [(\mathbf{Y}, v)]$, then the part of h_1 that maps to $\mathbf{Z}$ is copied from h, later the part of h_{i+1} that maps to $\mathbf{Z}$ is also copied from h_i, so as no restriction h is an isomorphism to $\mathbf{Z}$ the same can be said about the last function h'. Here $\mathbf{Z}$ is a component of $[(\mathbf{X}', v) + (\mathbf{Y}, v)]$ either because $\mathbf{Z} \subseteq (\mathbf{Y}, v)_-$ or because $\mathbf{Z} = [(\mathbf{Y}, v)_+]$ and $(\mathbf{X}, v)_+ = \mathbf{T}_0$, in the latter case forcing $(\mathbf{X}', v) = \mathbf{T}_0$. Again by Proposition 6.1 no restriction is an isomorphism to the component $\mathbf{Z}$ means that h' is a non-retraction as claimed.

If $\mathbf{Z} \subseteq (\mathbf{X}, v)_-$, then h restricted to $\mathbf{A}_0$ is a non-retraction to $(\mathbf{X}, v)_-$ as it has no restriction that is an isomorphism to $\mathbf{Z}$, so by the choice of $g_0' : \mathbf{A}_0' \to (\mathbf{X}', v)_-$ it is also a non-retraction and we have a component $\mathbf{Z}'$ of $(\mathbf{X}', v)_-$ such that no restriction of g_0' is an isomorphism to $\mathbf{Z}'$. This means that no restriction of h_1 is an isomorphism to $\mathbf{Z}'$ and, since the part mapping to $(\mathbf{X}', v)_-$ is copied from h_i to h_{i+1}, we conclude that no restriction of h' is an isomorphism to $\mathbf{Z}'$ and therefore h' is a non-retraction.

Finally we consider the $\mathbf{Z} = [(\mathbf{X}, v)_+ + (\mathbf{Y}, v)_+]$ case with $(\mathbf{X}, v)_+$ (and thus also $(\mathbf{X}', v)$) not being $\mathbf{T}_0$. Assume for a contradiction that h' is a retraction. By Proposition 6.1 we have a restriction of h' that is an isomorphism to $\mathbf{Z}' = [(\mathbf{X}', v)_+ + (\mathbf{Y}, v)_+]$. No restriction of h_1 is an isomorphism to either $\mathbf{Z}$ or $\mathbf{Z}'$, the former because the relevant part of h_1 is copied from h, the latter because the non-root vertices of $(\mathbf{X}', v)_+$ do not even appear in the image of h_1. Thus we must have an index i with no restriction of h_i being an isomorphism to either $\mathbf{Z}$ or $\mathbf{Z}'$ but such that the restriction of h_{i+1} to a substructure $\mathbf{D}$ of $\mathbf{B}_{i+1}$ is suddenly an isomorphism to one of $\mathbf{Z}$ or $\mathbf{Z}'$. We cannot have $\mathbf{D} \subseteq \mathbf{E}_{i,0}$ or $\mathbf{D} \subseteq [\mathbf{E}_i''(w)]$ for some $w \in g_i'^0$ as then the restriction of h_i to the corresponding substructure in $\mathbf{B}_i$ would also be an isomorphism to $\mathbf{Z}$ or $\mathbf{Z}'$. The substructure $\mathbf{D}$ is connected and it contains a single point $w \in V(\mathbf{D})$ mapped to the root v, so we must have $w \in g_i'^0$ and $\mathbf{D}$ is the union of its substructures $\mathbf{D}_0$ contained in $\mathbf{A}_i'$ and $\mathbf{D}_1$ contained in $[\mathbf{E}_i''(w)]$. Here $\mathbf{D}_0$ is not trivial and mapped to $[(\mathbf{X}', v)_+]$, so h_{i+1} must map $\mathbf{D}$ isomorphically to $\mathbf{Z}'$. We know that the neighbors of w in $\mathbf{E}_i''(w)$ are mapped by h_{i+1} to non-root vertices of $(\mathbf{Y}, v)$. As no point in $V(\mathbf{D})$ other than w is mapped to the root we must have that h_{i+1} maps $\mathbf{D}_1$ to $[(\mathbf{Y}, v)]$. To make the restriction of h_{i+1} to $\mathbf{D}$ an isomorphism to $\mathbf{Z}'$ we must have

that $\mathbf{D}_0$ is mapped isomorphically to $[(\mathbf{X}, v)_+]$ and $\mathbf{D}_1$ is mapped isomorphically to $[(\mathbf{Y}, v)_+]$. The restriction to $\mathbf{D}_0$ makes w a danger point of g'_i, i.e., $w \in g'^1_i$. Thus, we also have $f_i(w) \in g^1_i$ is a danger point of g_i. Let $\mathbf{D}' \subseteq \mathbf{A}_i$ be the substructure showing this, that is $\mathbf{D}'$ contains $f_i(w)$ and g_i maps $\mathbf{D}'$ isomorphically to $[(\mathbf{X}, v)_+]$. Now h_i maps the union of $\mathbf{D}'$ with the substructure corresponding to $\mathbf{D}_1 \subseteq [\mathbf{E}''_i(w)]$ in $[\mathbf{E}_i(f_i(w))]$ isomorphically to $\mathbf{Z}$. This contradicts our assumptions and proves that $h' : \mathbf{B}' \to [(\mathbf{X}', v) + (\mathbf{Y}, v)]$ is a non-retraction.

Proving that h' is a non-retraction finishes the proof of Proposition 6.5 and thus also completes the proof of Proposition 6.4, Theorem 5.4 and Theorem 5.3. □

[1] A. Bulatov - A. Krokhin - B. Larose: Dualities for Constraint Satisfaction Problems, *Complexity of Constraints (LNCS 5250)* (2008), 93–124.

[2] C. Carvalho - V. Dalmau - A. Krokhin: Caterpillar Duality for Constraint Satisfaction Problems, in Proceedings of the 23rd IEEE Symposium on Logic in Computer Science (LICS '08), 2008, 307–316.

[3] D. Duffus - P.L. Erdős - J. Nešetril - L. Soukup: Antichains in the homomorphism order of graphs, *Comment Math. Univ. Carol.* **48** (4) (2007), 571–583.

[4] P.L. Erdős - L. Soukup: How to split antichains in infinite posets, *Combinatorica* **27** (2) (2007), 147–161.

[5] P.L. Erdős - L. Soukup: No finite-infinite antichain duality in the homomorphism poset of directed graphs, *Order* **27** (3) (2010), 317–325.

[6] P.L. Erdős - C. Tardif - G. Tardos: On infinite-finite duality pairs of directed graphs, submitted (2012).

[7] P.L. Erdős - C. Tardif - G. Tardos: Caterpillar dualities and regular languages, submitted (2012).

[8] T. Feder - M. Y. Vardi: The Computational structure of monotone monadic SNP and constraint satisfaction: a study through datalog and group theory. *SIAM Journal of Computing* **28**, (1998), 57–104.

[9] J. Foniok - J. Nešetřil - C. Tardif: Generalized dualities and maximal finite antichains in the homomorphism order of relational structures, *Europ. J. Comb* **29** (2008), 881–899.

[10] B. Larose - C. Loten - C. Tardif: A Characterisation of first order definable constraint satisfaction problems, *Log. Methods Comput. Sci.* **3** (2007), no. 4, 4:6, 22 pp.

[11] L. Libkin: Logics for Unranked Trees: An Overview. Logical Methods in Computer Science 2(3): (2006).

[12] Nešetřil, J., Tardif, C.: Duality theorems for finite structures (characterising gaps and good characterisations), *J. Comb. Theory (B)* **80** (2000), 80–97.

[13] Nešetřil, J., Tardif, C.: Short answers to exponentially long questions: extremal aspects of homomorphism duality. *SIAM J. Disc. Math.* **19** (4) (2005), 914–920.