There are no pure relational width 2 constraint satisfaction problems

Víctor Dalmau

Departament de tecnologies de la informació i les comunicacions,
Universitat Pompeu Fabra,
Estació de França, Passeig de la circumval.lació 8, 08003, Barcelona (Spain)

Abstract

In this note, we show that every constraint satisfaction problem that has relational width 2 has also relational width 1. This is achieved by means of an obstruction-like characterization of relational width which we believe to be of independent interest.

Key words: Computational Complexity, Constraint Satisfaction Problems.

1. Introduction

Let **B** be a finite relational structure. In a *constraint satisfaction problem with template* **B**, CSP(**B**), we are given a relational finite structure **A** and the goal is to decide whether **A** is homomorphic to **B**. Motivated by the Feder-Vardi dichotomy conjecture [9] stating that for each **B**, CSP(**B**) is either solvable in polynomial time or NP-complete, there has been a good wealth of research aimed to distinguish those templates **B** that give rise to tractable (i.e., solvable in polynomial time) CSPs from those that do not. The length of the list of tractable cases known so far (see [5, 7] for recent surveys) contrasts sharply with the number of algorithmic principles which is very limited. Indeed, all known tractable cases are solvable either by the query language Datalog [9], via the "few subpowers" property [10], or by a combination (sometimes very non-trivial) of the two. Whereas the few subpowers property is well understood [10], the reach of Datalog Programs

Email addresses: `victor.dalmau@upf.edu` (Víctor Dalmau)

as a tool to solve CSPs has not yet been precisely delineated, despite considerable effort (see [6] for a survey on the topic). Datalog Programs have been parameterized in several ways (number of variables per rule, arity of the IDBs) giving rise to different notions of *width*. Among them, the *relational width*, introduced by Bulatov [4], has received considerable interest (see [4, 1, 2, 3, 13, 11]). An interesting feature of relational width is its independence on the arity of the relations of **B**, which makes it particularly appealing for the so-called algebraic approach to the CSP [5]. The class of problems with relational width 1 corresponds, in artificial intelligence terminology, to the class of those solvable by the arc-consistency algorithm [8]. Feder and Vardi [9] gave a complete characterization leading to a decision procedure for deciding if a structure **B** gives rise to a constraint satisfaction problem, CSP(**B**) of relational width 1. Little is known for higher levels of relational width. For $k = 2$ or $k \geq 4$ we do not possess examples of *pure* relational width k problems, i.e., structures **B** that have relational width k but not $k-1$. In this note we address and solve the case $k = 2$ showing that there are not pure relational width 2 problems. This is achieved by providing an obstruction-like characterization of relational width.

2. Preliminaires and Statement of the Main Result

Most of the terminology introduced in this section is fairly standard. A *vocabulary* is a finite set of relation symbols or predicates. In what follows, τ always denotes a vocabulary. Every relation symbol P in τ has an *arity* $r = \rho(P) \geq 0$ associated to it. We also say that P is an r-ary relation symbol.

A τ-structure $\mathbf{A}$ consists of a set A, called the *universe* of $\mathbf{A}$, and a relation $P^{\mathbf{A}} \subseteq A^r$ for every relation symbol $P \in \tau$ where r is the arity of P. For ease of notation, we shall say that $P(a_1, \ldots, a_r)$ *holds* in $\mathbf{A}$ to indicate that $(a_1, \ldots, a_r) \in P^{\mathbf{A}}$. All structures in this paper are assumed to be *finite*, i.e., structures with a finite universe. Throughout the paper we use the same boldface and slanted capital letters to denote a structure and its universe, respectively.

A *homomorphism* from a τ-structure $\mathbf{A}$ to a τ-structure $\mathbf{B}$ is a mapping $h : A \to B$ such that for every r-ary $P \in \tau$ and every $(a_1, \ldots, a_r) \in P^{\mathbf{A}}$, we have $(h(a_1), \ldots, h(a_r)) \in P^{\mathbf{B}}$. We say that $\mathbf{A}$ is homomorphic to $\mathbf{B}$ and denote this by $\mathbf{A} \to \mathbf{B}$ if there exists a homomorphism from $\mathbf{A}$ to $\mathbf{B}$.

If $\mathbf{A}$ is a τ-structure and $f : A \to B$ a mapping with domain the universe of A and image a finite set B, we define the homomorphic image of $\mathbf{A}$ by

f, $f(\mathbf{A})$, to be the τ-structure with domain $f(A)$, and such that for every $P \in \tau$ of arity, say r,

$$P^{f(\mathbf{A})} = \{(f(a_1), \ldots, f(a_r)) | (a_1, \ldots, a_r) \in P^{\mathbf{A}}\}$$

We define the union $\mathbf{A} \cup \mathbf{B}$ of τ-structures $\mathbf{A}$ and $\mathbf{B}$ to be the τ-structure with universe $A \cup B$ and such that $P^{\mathbf{A} \cup \mathbf{B}} = P^{\mathbf{A}} \cup P^{\mathbf{B}}$ for every $P \in \tau$. Notice that the union is not necessarily disjoint.

The concept of relational width was introduced initially by Bulatov in [4]. The presentation given here follows [6].

For any mapping f and $I \subseteq \text{dom}(f)$ we denote by f_I the restriction of f to I. For every f, g partial mappings from A to B, we write $f \subseteq g$ to indicate that $\text{dom}(f) \subseteq \text{dom}(g)$ and that $g_{\text{dom}(f)} = f$. We also say that g is an *extension of* f or alternatively that f is a *restriction* of g.

Definition 1. *Let* $\mathbf{A}$,$\mathbf{B}$ *be* τ*-structures and let* $k \geq 1$. *A* k*-minimal family for* $(\mathbf{A}, \mathbf{B})$ *is a nonemtpy set* H *of partial mappings of arity at most* k *from* A *to* B *such that for every* $h \in H$*:*

(i) *for every tuple* $P(a_1, \ldots, a_m)$ *in* $\mathbf{A}$ *there exists some tuple* $P(b_1, \ldots, b_m)$ *in* $\mathbf{B}$ *such that* $h(a_i) = b_i$ *for every* $a_i \in \text{dom}(h)$ *and such that for every subset* I *of* $\{a_1, \ldots, a_m\}$ *with* $|I| \leq k$, *there exists a mapping* h' *in* H *such that* $h'(a_i) = b_i$ *for every* $a_i \in I$.

(ii) $h' \in H$ *for every* $h' \subseteq h$.

(iii) *if* $\text{dom}(h) < k$ *then for every* $a \in A$, *there exists some* $h' \in H$ *with* $a \in \text{dom}(h')$ *and* $h \subseteq h'$

Observation 1. *Let* $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ *be* τ*-structures and let* $k \geq 1$. *If* $\mathbf{A}$ *is homomorphic to* $\mathbf{B}$ *and there is a* k*-minimal family* H *for* $(\mathbf{B}, \mathbf{C})$ *then there is a* k*-minimal family for* $(\mathbf{A}, \mathbf{C})$.

Proof. Let f be the homomorphism from $\mathbf{A}$ to $\mathbf{B}$ and define J to be the set containing for every $I \subseteq A$ of size at most k, and every mapping h in H of domain $f(I)$, the mapping $h \circ f_I$. It is easy to verify that J is a k-minimal family. ■

There exists a very simple procedure, called the *k-minimal test* [4], that decides, given two relational structures $\mathbf{A}$ and $\mathbf{B}$, whether there exists a k-minimal family for $(\mathbf{A}, \mathbf{B})$ (and actually finds one). The k-minimal test starts by placing in the hypothetical k-minimal family H all partial mappings from A to B of domain size at most k. Then in an iterative fashion it removes from H all mappings that do not satisfy any of conditions (1-3) of k-minimal family until the process stabilizes. Since the number of partial mappings from A to B with domain size k is bounded by $|A||B|^k$ the k-minimal test runs in polynomial time. We say that $(\mathbf{A}, \mathbf{B})$ passes the k-minimal test if the resulting H is nonempty and that fails otherwise. A structure $\mathbf{B}$ has *relational width* k if $\mathbf{A} \to \mathbf{B}$ for every structure $\mathbf{A}$ such that $(\mathbf{A}, \mathbf{B})$ passes the k-minimal test.

The main result of this paper is the following

Theorem 1. *Every structure with relational width* 2 *also has relational width* 1.

3. Proof of Theorem 1

The proof has two ingredients: The first one is an obstruction-like characterization of relational width (Theorem 4). The second ingredient is the Sparse Incomparability Lemma [12].

Let $m \geq 1$. A *cycle of length m in a τ-structure* $\mathbf{A}$ is a collection of m different tuples $P_0(a_1^0, \ldots, a_{r_0}^0), \ldots, P_{m-1}(a_1^{m-1}, \ldots, a_{r_{m-1}}^{m-1})$ that hold in $\mathbf{A}$ such that the cardinality of the set $\{a_j^i \mid 0 \leq i \leq m-1, 1 \leq j \leq r_i\}$ is less than $1 + \sum_{0 \leq i \leq m-1}(r_i - 1)$.

A *loop* is a cycle of length 1. The *girth* of a τ-structure is the length of its shortest cycle.

Theorem 2. *(Sparse Incomparability Lemma) Let k, l be positive integers and let $\mathbf{A}$ be a structure. Then there exists a structure $\mathbf{G}$ with the following properties:*

1. $\mathbf{G}$ *is homomorphic to* $\mathbf{A}$
2. *For every structure* $\mathbf{B}$ *with at most k elements,* $\mathbf{A}$ *is homomorphic to* $\mathbf{B}$ *iff* $\mathbf{G}$ *is homomorphic to* $\mathbf{B}$
3. $\mathbf{G}$ *has girth* $\geq l$.

The following definition introduces the new notion of k-reltree. It will be shown in Theorem 4 that k-reltrees are precisely the obstructions corresponding to the k-minimal test.

Definition 2. *Let* $\mathbf{T}$ *be a relational structure and let* I *be a subset of nodes of* $\mathbf{T}$ *with* $|I| \leq k$. *The pair* $(\mathbf{T}, I)$ *is called a* k*-reltree (from relational tree) if*

(1) $\mathbf{T}$ *contains only one tuple with no repeated elements, or*

(2) there is a finite collection $(\mathbf{T}_j, I_j), j \in J$ *of* k*-reltrees and distinct* $e_1, \ldots, e_n \in T$, $n \geq 0$ *such that for all* $j \in J$, $T_j \cap \{e_1, \ldots, e_n\} \subseteq I_j$ *and for all* $i, j \in J$, $T_i \cap T_j \subseteq \{e_1, \ldots, e_n\}$, *and*

  (a) $\mathbf{T}$ *is the union of the tuple* $P(e_1, \ldots, e_n)$ *(for some* n*-ary* $P \in \tau$*) and* $\bigcup_{j \in J} \mathbf{T}_j$*,and* $I \subseteq \{e_1, \ldots, e_n\}$ *or*

  (b) $\mathbf{T} = \bigcup_{j \in J} \mathbf{T}_j$ *and* $I = \{e_1, \ldots, e_n\}$, *or*

(3) there is a k*-reltree* $(\mathbf{T}, I')$ *with* $I \subseteq I'$.

Finally, a structure $\mathbf{T}$ *is a* k*-reltree if* $(\mathbf{T}, \emptyset)$ *is a* k*-reltree.*

Generally, a relational structure $\mathbf{A}$ is called a *tree* if it is cycle-free. In our terminology, trees are precisely 1-reltrees.

Theorem 3. *Let* $\mathbf{A}, \mathbf{B}$ *be structures and let* $k \geq 1$. *The following are equivalent:*

(a) $(\mathbf{A}, \mathbf{B})$ *passes the* k*-minimal test*

(b) there is a k*-minimal family for* $(\mathbf{A}, \mathbf{B})$

Furthermore if $\mathbf{A}$ *is loop-free then (a) and (b) are also equivalent to the following statement:*

(c) every k*-reltree homomorphic to* $\mathbf{A}$ *is homomorphic to* $\mathbf{B}$

Proof.

$[(a) \Leftrightarrow (b)]$. This is precisely the proof of the correctness of the k-minimal test, which is straightforward.
$[(b) \Rightarrow (c)]$ Let H be a k-minimal family for $(\mathbf{A}, \mathbf{B})$. We shall prove that if $(\mathbf{T}, I)$ is a k-reltree, f is a homomorphism from $\mathbf{T}$ to $\mathbf{A}$ and h is a mapping in H with $\text{dom}(h) = f(I)$ then there exists a homomorphism g from $\mathbf{T}$ to $\mathbf{B}$ such that $g_I = (h \circ f_I)$. The proof is by structural induction on $(\mathbf{T}, I)$.

(1) $\mathbf{T}$ is simply a tuple $P(e_1,\ldots,e_n)$ and I is any subset of $\{e_1,\ldots,e_n\}$ with $|I|\leq k$. Let $P(a_1,\ldots,a_n)$ be the image of $P(e_1,\ldots,e_n)$ according to f. Let $P(b_1,\ldots,b_n)$ be the tuple in $\mathbf{B}$ guaranteed to exist because h satisfies condition (i) of a k-minimal family. The mapping $g:\{e_1,\ldots,e_n\}\to B$, $g(e_i)=b_i$, $1\leq i\leq n$, satisfies the required conditions.

(2a) Let $P(a_1,\ldots,a_n)$ be the image of $P(e_1,\ldots,e_n)$ according to f. Let $P(b_1,\ldots,b_n)$ be the tuple in $\mathbf{B}$ that is guaranteed to exist because h satisfies condition (i) of a k-minimal family. Set $g(e_i)=b_i$ for $1\leq i\leq n$. In order to define g over the rest of T do the following:

For $j\in J$, consider the the mapping $h'_j : f(I_j)\cap\{a_1,\ldots,a_n\}\to B$ defined by $h'_j(a_i)=b_i$, $a_i\in \text{dom}(h'_j)$. Condition (ii) of a k-minimal family guarantees that $h'_j\in H$. Furthermore, by condition (iii) of a k-minimal family, H contains an extension h_j of h'_j with domain $f(I_j)$. By induction hypothesis there exists a homomorphism g_j from $\mathbf{T}_j$ to $\mathbf{B}$ such that $g_j(e)=h_j(f(e))$ for every $e\in I_j$. Define $g(e)=g_j(e)$ for every $j\in J$ and every $e\in T_j$. Mapping g satisfies the required conditions.

(2b) $(\mathbf{T},I)$ is obtained by rule $(2b)$. Define $g(e)=h(f(e))$ for all $e\in I$ and extend g over the rest of T as in the previous case.

(3) $(\mathbf{T},I)$ is obtained by rule (3) from $(\mathbf{T},I')$ with $I\subseteq I'$. By property (iii) of H there exists h' defined over $f(I')$ that extends h. The mapping g guaranteed to exist for $(\mathbf{T},I')$, f and h' satisfies the required conditions.

$[(c)\Rightarrow(a)]$ We shall show that for every mapping h removed from H by the k-minimal test there exists a k-reltree $(\mathbf{T},I)$, some homomorphism f from $\mathbf{T}$ to $\mathbf{A}$, with f_I one-to-one, $f(I)=\text{dom}(h)$, and such that for every homomorphism $g:\mathbf{T}\to\mathbf{B}$, $g_I\neq(h\circ f_I)$. We shall prove it by induction on the elimination order of h.

If h is removed in the first iteration, then necessarily condition (i) of k-minimal family is falsified by h. Set $\mathbf{T}$ to be the structure containing only the tuple $P(a_1,\ldots,a_n)$ given by the condition, define f to be the identity mapping, and let $I=dom(h)$.

Assume now that h is removed in some subsequent iteration. We do a case by case analysis depending on which condition of k-minimal family is falsified by h.

(i) Let $P(a_1, \ldots, a_n)$ be the tuple that forces h to be eliminated and let h_j, $j \in J$ be the set of mappings with domain entirely contained in $\{a_1, \ldots, a_n\}$ that have been previously removed from H. For each $j \in J$, let $(\mathbf{T}_j, I_j)$ and f_j be the k-reltree and mapping respectively for h_j. By renaming adequately the nodes of $\mathbf{T}_j$ we can assume that f_j restricted to I_j is the identity and that all the other variables are new, i.e., $I_j = T_j \cap \{a_1, \ldots, a_n\}$. We can also assume that apart from the elements in $\{a_1, \ldots, a_n\}$ any two of these structures do not share any other element, i.e., for every $i \neq j \in J$, $T_i \cap T_j \subseteq \{a_1, \ldots, a_n\}$. We are now in a position to define $(\mathbf{T}, I)$ and f. $(\mathbf{T}, I)$ is constructed by rule $(2b)$ from $(\mathbf{T}_j, I_j)$, $j \in J$, the tuple $P(a_1, \ldots, a_m)$, and $I = \text{dom}(h)$. $f(x)$ is defined to be the identity if $x \in \{a_1, \ldots, a_n\}$ and $f_j(x)$ if $x \in T_j$, otherwise. It is easy to verify that $(\mathbf{T}, I)$ and f satisfy the required conditions.

(ii) There exists some $h \subseteq h'$ such that h' was previously removed from H. Let $(\mathbf{T}', I')$ and f' be guaranteed by the hypothesis condition. In this case we only need to set $\mathbf{T} = \mathbf{T}'$, $I = \text{dom}(h)$, and $f = f'$.

(iii) In this case, h is eliminated because $|\text{dom}(h)| = n < k$ and there exists some a such that H does not contain any extension of h defined over a. Hence, every possible extension $h_j : \text{dom}(h) \cup \{a\}$, $j \in J$ of h has been previously removed from H. For every $j \in J$, there exists suitable $(\mathbf{T}_j, I_j)$, and f_j. Let $\text{dom}(h) = \{a_1, \ldots, a_n\}$ and rename the variables of the structures $\mathbf{T}_j$, $j \in J$ so that for every $j \in J$, $T_j \cap \{a_1, \ldots, a_n\} \subseteq I_j$, f_j is the identity on $T_j \cap \{a_1, \ldots, a_n\}$, and for all $i \neq j \in J$, $T_i \cap T_j \subseteq \{a_1, \ldots, a_n\}$. We set $\mathbf{T}$ to be $\bigcup_{j \in J} \mathbf{T}_j$, $I = \{a_1, \ldots, a_n\}$, and set $f(x)$ to be the identity if $x \in \{a_1, \ldots, a_n\}$ and $f_j(x)$ where $x \in T_j$, otherwise. $(\mathbf{T}, I)$ and f satisfy the required conditions.

Finally we prove the contrapositive of the implication. If the k-minimal test fails then the mapping h with empty domain is removed. This implies that condition (c) is false.

■

In order to prove our main theorem we will use an obstruction-like characterization of relational width.

Definition 3. *Let* $\mathbf{B}$ *be a* τ*-structure. A set* $\mathcal{O}$ *of* τ*-structures is an obstruction set of* $\mathbf{B}$ *if for every* τ*-structure* $\mathbf{A}$

$$\mathbf{A} \to \mathbf{B} \textit{ iff } \forall\, \mathbf{O} \in \mathcal{O}, \mathbf{O} \not\to \mathbf{A}$$

Observe that as direct application of Theorem 3 it can be shown that a structure has relational width k iff it has an obstruction set consisting of k-reltrees. Although this would be enough in order to prove our main theorem we believe that it is interesting to introduce here another class of relational structures, which we call k-greltrees (from *generalized* reltrees). The notion of k-greltree is a proper generalization of that of k-reltree but as we will show in Theorem 4 both concepts are equivalent when it comes to define obstructions. The reason why we believe the notion of k-greltree might be appealing is because it is defined in terms of tree-decompositions as several other related notions such as treewidth.

Definition 4. *Let* $\mathbf{A}$ *be a* τ*-structure. A tree-decomposition of* $\mathbf{A}$ *is a pair* (T, φ) *where* T *is a tree and* $\varphi : V(T) \to \mathcal{P}(A)$ *is a mapping that assigns to every node of* T *a set of elements of* A*, satisfying the following conditions:*

1. *nodes containing any given element of* A *form a subtree,*
2. *for any tuple in any relation of* $\mathbf{A}$*, there is a node in* T *containing all elements from that tuple.*

Note: for ease of notation we say that a node $v \in V(T)$ contains an element $a \in A$ if $a \in \varphi(v)$.

Definition 5. *A* τ*-structure* $\mathbf{A}$ *is a* k*-generalized relational tree (or* k*-greltree) if there exists a tree-decomposition* (T, φ) *of* $\mathbf{A}$ *such that:*

(i) two different nodes of T *share at most* k *elements*

(ii) for every node t *of* T *there exists a tuple of* $\mathbf{A}$ *that contains every element of* t *or* t *has size at most* k*.*

Observe also that if all predicates in τ have arity at most k then a τ-structure is a k-greltree iff its Gaifman graph has treewidth at most $k-1$.

Lemma 1. *Let* $\mathbf{A}$ *be a structure and let* $k \geq 1$*. If* $\mathbf{A}$ *is a* k*-reltree then it is also a* k*-geltree.*

Proof. It is easily shown by structural induction that if $(\mathbf{A}, I)$ is a k-reltree then there is a tree-decomposition (T, φ) of $\mathbf{A}$ and a node $v \in T$ such that $\varphi(v) = I$.

■

The converse is not true as in particular a k-reltree, for any k, cannot have loops. But this is not the only reason: consider for example a structure $\mathbf{A}$ with $k \geq 2$ nodes $a_1, \ldots, a_k$ and with only one k-ary relation with tuples $(a_1, \ldots, a_k)$ and $(a_2, \ldots, a_k, a_1)$. Structure $\mathbf{A}$ is not a $(k-1)$-reltree but it is certainly a 1-geltree as shown by the tree decomposition containing one single node v with $\varphi(v) = \{a_1, \ldots, a_k\}$. However one can show that if a structure has an obstruction set consisting of k-greltrees then it also has one containing only k-reltrees.

Theorem 4. *Let* $\mathbf{B}$ *be a structure and* $k \geq 1$*. The following are equivalent:*

(a) $\mathbf{B}$ *has relational width* k

(b) $\mathbf{B}$ *has an obstruction set consisting of* k*-reltrees*

(c) $\mathbf{B}$ *has an obstruction set consisting of* k*-greltrees.*

Proof. The equivalence between (a) and (b) is a direct consequence of Theorem 3 although an small adjustment needs to be done as, in Theorem 3, structure $\mathbf{A}$ is assumed to be loop-free.

$[(a) \Rightarrow (b)]$ We need to show that if $\mathbf{B}$ is a structure with relational width k and $\mathbf{A}$ is a structure (not necessarily loop-free) homomorphic to $\mathbf{B}$ then $\mathbf{A}$ admits an homomomorphism from some k-reltree not homomorphic to $\mathbf{B}$. By the Sparse Incomparability Lemma, if $\mathbf{A}$ is not homomorphic to $\mathbf{B}$ there exists some loop-free structure $\mathbf{G}$ that is homomorphic to $\mathbf{A}$ and not homomorphic to $\mathbf{B}$. Theorem 3 shows that there exists a k-reltree $\mathbf{C}$ homomorphic to $\mathbf{G}$ (and hence to $\mathbf{A}$) and not homomorphic to $\mathbf{B}$.

$[(b) \Rightarrow (a)]$ Let $\mathbf{B}$ be a structure satisfying condition (b) and let $\mathbf{A}$ be a structure not homomorphic to $\mathbf{B}$. Again by the Sparse Incomparability Lemma, there exists some loop-free structure $\mathbf{G}$ that is homomorphic to $\mathbf{A}$ and not to $\mathbf{B}$. By Theorem 3, there is no k-minimal strategy for $(\mathbf{G}, \mathbf{B})$, which by Observation 1 implies that there is no k-minimal strategy for $(\mathbf{A}, \mathbf{B})$.

$[(b) \Rightarrow (c)]$ follows from Lemma 1 so it only remains to show that $[(c) \Rightarrow (b)]$. Let $\mathbf{B}$ be a structure satisfying condition (c) and let $\mathcal{O}$ be a obstruction set of $\mathbf{B}$ consisting of k-geltrees. It is only necessary to show that every $\mathbf{A}$ not homomorphic to $\mathbf{B}$ admits an homomomorphism from some k-reltree $\mathbf{C}$ in $\mathcal{O}$. Again by the Sparse Incomparability Lemma, if $\mathbf{A}$ is not homomorphic to $\mathbf{B}$ there exists some structure $\mathbf{G}$ with girth at least 3 that is homomorphic to $\mathbf{A}$ and not homomorphic to $\mathbf{B}$. Consequently there exists some $\mathbf{C}$ in $\mathcal{O}$ that is homomormorphic to $\mathbf{G}$ (and hence to $\mathbf{A}$). We shall show that $\mathbf{C}$ is, indeed, a k-reltree. Let (T, φ) be a tree-decomposition of $\mathbf{C}$ satisfying the conditions of Definition 5. Observe that in a tree-decomposition T we can replace any edge (v, v') in T by two edges (v, u), (u, v') where u is a new node with $\varphi(u') = \varphi(v) \cap \varphi(v')$ obtaining again a tree-decomposition that satisfies the conditions of Definition 5. Hence we can assume wlog. that for every edge (v, v') of T $\varphi(v) \subseteq \varphi(v')$ or $\varphi(v') \subseteq \varphi(v)$. Furthermore, condition (ii) of a k-greltree gurantees that there is no edge in T between two nodes of size larger than k. We also assume by adding a node if necessary that T contains at least one node of size at most k. We shall prove by induction on the number of nodes of T that if v is a node in T with size at most k, then $(\mathbf{C}, \varphi(v))$ is a k-reltree. For the base case of the induction assume that T consists of a single node of size at most k. Hence $\mathbf{C}$ has necessarily at most k nodes and the result follows from the observation that by repeated application of rules (1),(2b) and (3) of a k-reltree it is possible to generate all structures with at most k nodes (indeed, it is easy to see that by iterative application of these rules one could generate any structure with a tree-decomposition consisting only of nodes of size at most k). For the inductive case, assume first that all neighbours v_j, $j \in J$, of v, have size at most k. Let T_j, $j \in J$, be each one of the connected components of T after removing node v, and let $\mathbf{C}_j$, $j \in J$, be the substructure of $\mathbf{C}$ induced by $\bigcup_{u \in T_j} \varphi(u)$. By the inductive hypothesis $(\mathbf{C}_j, \varphi(v_j))$, $j \in J$ is a k-reltree. Also, if $\mathbf{C}'$ is the substructure of $\mathbf{C}$ induced by $\varphi(v)$, then by induction hypothesis $(\mathbf{C}', \varphi(v))$ is a k-reltree. Finally $\mathbf{C}$ is obtained from $(\mathbf{C}_j, \varphi(v_j))$, $j \in J$, and $(\mathbf{C}', \varphi(v))$ by using rule (2b).

If v has an edge to a node v' of size larger than k then let v_j, $j \in J$ the set of neighbours (including v) of v', let T_j $j \in J$ be each one of the connected components of T after removing v', and let $\mathbf{C}_j$, $j \in J$ be the substructure of $\mathbf{C}$ induced by $\bigcup_{u \in T_j} \varphi(u)$. Since for every $j \in J$, v_j has size at most k, $(\mathbf{C}_j, \varphi(v_j))$ is a k-reltree. Now let us turn our attention to v'. Since the size of $\varphi(v')$ is larger than k there is a tuple $t = P(e_1, \ldots, e_n)$ of $\mathbf{C}$ containing all nodes in $\varphi(v')$. Also, condition (2) of tree-decomposition guarantees that all elements of t are contained in a node v^* of T. This node should be precisely

v' since otherwise the intersection $\varphi(v') \cap \varphi(v^*)$ which is $\varphi(v')$ would be larger than k. Hence $\varphi(v')$ is precisely $\{e_1, \ldots, e_n\}$. Let t_j, $j \in J'$ be the class of all tuples of $\mathbf{C}$ different than t and entirely contained in $\{e_1, \ldots, e_n\}$. We can infer that for every $j \in J'$, t_j has arity 1 because otherwise the image of tuples t and t' in $\mathbf{G}$ would be a cycle of lenght at most 2 which is impossible (Note here that it is crucial that all elements of t_j are contained in t as otherwise both tuples could have the same image in $\mathbf{G}$). For every $j \in J'$, the pair $(\mathbf{D}_j, \{e_{i_j}\})$ where $\mathbf{D}_j$ is the structure containing only tuple t_j and e_{i_j} its only element is a k-reltree. Finally, $(\mathbf{C}, \varphi(v))$ is obtained by applying rule $(2a)$ with $t = P(e_1, \ldots, e_n)$ and k-reltrees $(\mathbf{C}_j, v_j)$ $j \in J$ and $(\mathbf{D}_j, \{e_{i_j}\})$, $j \in J'$ (Here we are using also the fact that $\varphi(v)$ is necessarily contained in $\varphi(v')$). ∎

Lemma 2. *Every 2-greltree with girth at least 3 has no cycles.*

Proof. This is done by contradiction. Let

$$P_1(a_1^1, \ldots, a_{r_1}^1), \ldots, P_{m-1}(a_1^{m-1}, \ldots, a_{r_{m-1}}^{m-1})$$

be a cycle in $\mathbf{A}$ and let us assume that m is minimal. Hence $r_i \geq 2$ for $i = 1, \ldots, m-1$. Furthermore, by the minimality of m we can assume that there exists different elements $a_0, \ldots, a_{m-1} \in A$ such that for every $0 \leq i \neq j \leq m-1$, the ith and the jth tuple share only element a_i if $i + 1 = j \pmod{m}$ and none otherwise.

Let (T, φ) be a suitable tree-decomposition of $\mathbf{A}$ that certifies that $\mathbf{A}$ is a 2-greltree. By the definition of tree-decomposition, for every $0 \leq i \leq m-1$, T contains a node, let us call it n_i, that contains $\{a_1^i, \ldots, a_{r_i}^i\}$. Since $r_i \geq 2$ then , by definition 5, n_i should be precisely $\{a_1^i, \ldots, a_{r_i}^i\}$, as we cannot have two different tuples containing $\{a_1^i, \ldots, a_{r_i}^i\}$ as this would be a cycle of length 2. Consider the following walk in T: Start in n_0 and follow the unique path from n_0 to n_1, then continue following the unique path from n_1 to n_2, and proceed in the same way until by crossing the path from n_{m-1} to n_0 the walk returns to n_0. Let us start by showing that after reaching node n_1 for the first time, the walk must reverse direction. Indeed, let $i \geq 1$ such that n_1 is crossed back later when following the path from n_i to $n_{i+1} \pmod{m}$. By the definition of tree-decomposition every node in the path from n_i to n_{i+1} contains a_i and hence a_i belongs to n_1. But this is only possible if $i = 1$ and hence the walk must reverse direction.

The walk then proceeds by following the path from n_1 to n_2. Every node in this segment contains a_1 and hence by the same type of reasoning it cannot cross n_0. Hence there is some node u at which this path stops going towards n_0 and branches off in a different direction. Necessarily $\{a_0, a_1\} \subseteq u$ as u participates both in the path going from n_0 to n_1 and the path going from n_1 to n_2. Later on during the walk, u must be necessarily crossed back, say, when walking the path from node n_i to $n_{i+1} \pmod m$ for some $i \geq 2$. Hence u contains a_i as well. Since u has cardinality at least 3 there exists a tuple in $\mathbf{A}$ containing $\{a_0, a_1, a_i\}$. This tuple jointly with tuple $P_1(a_1^1, \ldots, a_{r_1}^1)$ constitutes a cycle of length 2, which is impossible.

■

Proof. (of Theorem 1)

Let $\mathbf{B}$ be a τ-structure with relational width 2. We shall show that if $\mathbf{A}$ is a structure not homomorphic to $\mathbf{B}$ then $(\mathbf{A}, \mathbf{B})$ fails the 1-minimal test. By the Sparse Incomparability Lemma, if $\mathbf{A}$ is not homomorphic to $\mathbf{B}$ there exists some structure $\mathbf{G}$ with girth at least 3 that is homomorphic to $\mathbf{A}$ and not homomorphic to $\mathbf{B}$. By Theorem 4 there exists some 2-greltree $\mathbf{C}$ that is homomorphic to $\mathbf{G}$ but not to $\mathbf{B}$. Pick such $\mathbf{C}$ with minimum number of tuples. We shall see that the girth of $\mathbf{C}$ is at least 3, and hence, by Lemma 2, $\mathbf{C}$ is a tree. By composition of homomorphisms $\mathbf{C}$ is homormorphic to $\mathbf{A}$ but not to $\mathbf{B}$. Therefore by Theorem 3, $(\mathbf{A}, \mathbf{B})$ fails the 1-minimal test.

It only remains to check that if $\mathbf{C}$ is a 2-greltree with minimum number of tuples homorphic to $\mathbf{G}$ but not to $\mathbf{B}$ then $\mathbf{C}$ does not have cycles of length at most 2. Clearly, if $\mathbf{C}$ has a cycle of length 1 then its image in $\mathbf{G}$ is, as well, a cycle of lenght 1 which is impossible. The same reasonning does not always apply to cycles of length 2. Indeed, if $P_0(a_1^0, \ldots, a_{r_0}^0)$, $P_1(a_1^1, \ldots, a_{r_1}^1)$ is a cycle of $\mathbf{C}$ and h is a homomorphism from $\mathbf{C}$ to $\mathbf{G}$ then it is possible that the image $P_0(h(a_1^0), \ldots, h(a_{r_0}^0))$ $P_1((a_1^1), \ldots, (a_{r_1}^1))$ is not a cycle of $\mathbf{G}$ *if the two tuples of the image are the same.* Hence we can assume that the two predicates are the same and for ease of notation we write $P = P_0 = P_1$ and $r = r_0 = r_1$.

Define the mapping $f : C \to C$ with $f(a_i^1) = a_i^0$ for all $i = 1, \ldots, r$ and f acting as the identity in all other cases. Clearly, the image of f, $f(\mathbf{C})$, is homomorphic to $\mathbf{G}$, because $h(a_i^0) = h(a_i^1)$ for all $i = 1, \ldots, r$, and not homomorphic to $\mathbf{B}$. We shall show that $f(\mathbf{C})$ is a 2-greltree contradicting the minimality of $\mathbf{C}$.

Let (T, φ) be a suitable tree-decomposition of $\mathbf{C}$ and let u_0 be a node of T containing $\{a_1^0, \ldots, a_r^0\}$. It is not difficult to see that indeed $\varphi(u_0)$ is precisely $\{a_1^0, \ldots, a_r^0\}$ since otherwise $\varphi(u_0)$ would have size at least 3 and hence necessarily all nodes in it would be contained in a tuple t. This would be impossible because the images according to h of t and $P(a_1^0, \ldots, a_r^0)$ would constitute a cycle of size 2 in $\mathbf{G}$. By the same reasonning there is a node u_1 in T such that $\varphi(u_1) = \{a_1^1, \ldots, a_r^1\}$. By condition (i) of 2-greeltree, tuples $P_0(a_1^0, \ldots, a_{r_0}^0)$, $P_1(a_1^1, \ldots, a_{r_1}^1)$ share exactly two elements and for ease of notation we shall assume that the common elements are precisely the first two and write $a_1 = a_1^0 = a_1^1$ and $a_2 = a_2^0 = a_2^1$,

The set of nodes of T can be partitioned in two sets of nodes V_0 and V_1 such that:

- V_0 and V_1 are connected in T,
- $\bigcup_{v \in V_0} \varphi(v) \cap \bigcup_{v \in V_1} \varphi(v) = \{a_0, a_1\}$, and
- $u_i \in V_i$ for $i = 0, 1$.

The partition can be obtained in the following way: define V_0 to be the set of all elements reachable from u_0 without crossing u_1 and V_1 to be the rest of nodes. It is clear that V_0 and V_1 satisfy all the requiered conditions.

For $i = 0, 1$, let $\mathbf{C}_i$ be the substructure of $\mathbf{C}$ induced by $\bigcup_{v \in V_i} \varphi(v)$. Then $f(\mathbf{C}) = f(\mathbf{C}_0) \cup f(\mathbf{C}_1) = \mathbf{C}_0 \cup f(\mathbf{C}_1)$. $\mathbf{C}_0$ is clearly a 2-reltree and indeed a tree-decompostion (T_0, φ_0) of $\mathbf{C}_0$ can be obtained by restricting (T, φ) to the nodes in V_0. Since f is a bijection over C_1, then a suitable tree-decomposition (T_1, φ_1) of $\mathbf{C}_1$ can be obtained by setting T_1 to be the restriction of T over V_1 and $\varphi_1(v) = \{f(a) \mid a \in \varphi(v)\}$, $v \in V_1$. Define T' to be the tree obtained by making the union of T_0 and T_1 and gluing toghether u_0 and u_1. Define $\varphi' : V(T') \to f(C)$ to be $\varphi_0(v)$ if $v \in T_0$ and $\varphi_1(v)$ if $v \in T_1$. The pair (T', φ') is a suitable tree-decomposition of $f(\mathbf{C})$.

■

Acknowledgements

We are grateful to Miklos Maroti for pointing out a mistake in a previous version of this paper and to Hubie Chen for several comments on a previous version of this paper.

References

[1] A. Bulatov. Tractable Conservative Constraint Satisfaction Problems. In *Proceedings of the 18th Annual IEEE Simposium on Logic in Computer Science, (LICS'03)*, pages 321–330, 2003.

[2] A. Bulatov. A Graph of a Relational Structure and Constraint Satisfaction Problems. In *Proceedings of the 19th IEEE Annual Symposium on Logic in Computer Science, (LICS'04)*, pages 448–457, 2004.

[3] A. Bulatov. A Dichotomy Theorem for Constraints on a Three-element Set. *Journal of the ACM*, 53(1):66–120, 2006.

[4] A. Bulatov. Combinatorial problems raised from 2-semilattices. *Journal of Algebra*, 298(2):321–339, 2006.

[5] A. Bulatov, A. Krokhin, and P. Jeavons. Classifying the Complexity of Constraints Using Finite Algebras. *SIAM J. on computing*, 34(3):720–742, 2005.

[6] A. Bulatov, A. Krokhin, and B. Larose. Dualities for Constraint Satisfaction Problems. Submitted for publication.

[7] A. Bulatov and M. Valeriote. Results on the Algebraic Approach to the CSP. In *Proceedings of the Dagstuhl Seminar on the Complexity of Constraints*. To appear.

[8] V. Dalmau and J. Pearson. Set Functions and Width 1. In *5th International Conference on Principles and Practice of Constraint Programming, (CP'99)*, volume 1713 of *Lecture Notes in Computer Science*, pages 159–173, Berlin/New York, 1999. Springer-Verlag.

[9] T. Feder and M.Y. Vardi. The Computational Structure of Monotone Monadic SNP and Contraint Satisfaction: A Study through Datalog and Group Theory. *SIAM J. Computing*, 28(1):57–104, 1998.

[10] P. Idziak, P. Markovic, R. MacKenzie, and M. Valeriote. Tractability and learnability arising from algebras with few subpowers. In *22nd IEEE Symposium on Logic in Computer Science (LICS'07)*, pages 213–224, 2007.

[11] E. Kiss and M. Valeriote. On Tractability and Congruence Distributivity. *Logical Methods in Computer Science*, 3(2), 2006.

[12] J. Nešetřil and V. Rödl. Chormatically Optimal Rigid Graphs. *J. Comb. Theory, Series B*, 46:122–141, 1989.

[13] M. Valeriote. A Subalgebra Intersection Property for Congruence Distributive Lattices. *Canadian Journal of Mathematics*. To appear.