

# On Maltsev Digraphs

Catarina Carvalho<sup>1</sup>, László Egri<sup>2</sup>, Marcel Jackson<sup>3</sup>, and Todd Niven<sup>3\*</sup>

<sup>1</sup> School of Physics, Astronomy and Mathematics, University of Hertfordshire, UK,  
and Centro de Álgebra da Universidade de Lisboa, Portugal,  
[ccarvalho@cii.fc.ul.pt](mailto:ccarvalho@cii.fc.ul.pt)

<sup>2</sup> School of Computer Science, McGill University, Canada, [legri1@cs.mcgill.ca](mailto:legri1@cs.mcgill.ca)

<sup>3</sup> Department of Engineering and Mathematical Sciences, La Trobe University,  
Australia, [{m.g.jackson,t.niven}@latrobe.edu.au](mailto:{m.g.jackson,t.niven}@latrobe.edu.au)

**Abstract.** We study digraphs preserved by a Maltsev operation, Maltsev digraphs. We show that these digraphs retract either onto a directed path or to the disjoint union of directed cycles, showing that the constraint satisfaction problem for Maltsev digraphs is in logspace, L. (This was observed in [19] using an indirect argument.) We then generalize results in [19] to show that a Maltsev digraph is preserved not only by a majority operation, but by a class of other operations (e.g., minority, Pixley) and obtain a  $O(V_G^4)$ -time algorithm to recognize Maltsev digraphs. We also prove analogous results for digraphs preserved by conservative Maltsev operations which we use to establish that the list homomorphism problem for Maltsev digraphs is in L. We then give a polynomial time characterisation of Maltsev digraphs admitting a conservative 2-semilattice operation. Finally, we give a simple inductive construction of directed acyclic digraphs preserved by a Maltsev operation.

## 1 Introduction

The study of relational structures and, in particular, digraphs preserved by certain operations from universal algebra became extremely important during the last decade. The main driving force behind this is the algebraic constraint satisfaction problem (CSP) dichotomy conjecture, which states that a constraint satisfaction problem  $\text{CSP}(\mathbf{B})$  is tractable if the relational structure  $\mathbf{B}$  is preserved by a weak-near-unanimity (weak-NU) operation, and is NP-complete otherwise [5, 6, 20]. Generalizing the dichotomy theorem of Hell and Nešetřil [16], the conjecture has been established for digraphs with no sinks and no sources by showing that such digraphs are very structured, in fact, they retract onto a disjoint union of directed cycles [3]. Other results relating the complexity of CSPs on digraphs to the existence of operations that preserve the digraph can be found, for example, in [2, 1].

---

\* The first author was supported by grants SFRH/BPD/26216/2006 and ISFL-1-143 of CAUL financed by FCT and FEDER. The second author was supported by the National Sciences and Engineering Research Council of Canada (NSERC). The third and fourth authors were supported by ARC Discovery Project DP1094578.

Once the tractability of a CSP is established, one also wishes to know the fine-grained complexity of that CSP, i.e. is the CSP in some subclass of  $\mathbf{P}$ , such as  $\mathbf{L}$  or  $\mathbf{NL}$ ? To establish the membership of a CSP in a complexity class inside  $\mathbf{P}$  it is important to study the structure of relational structures (and digraphs) preserved by operations which are more “restrictive” than weak-NU operations, i.e. operations that imply the presence of a weak-NU operation. Two important results in this direction, which we will invoke, are that if a relational structure is preserved by a majority operation, then the corresponding CSP is in the complexity class  $\mathbf{NL}$  [9]; and if  $\text{CSP}(\mathbf{B})$  is definable in Datalog and  $\mathbf{B}$  is preserved by a Maltsev operation, then  $\text{CSP}(\mathbf{B})$  is in  $\mathbf{L}$  [10, 12].

We study the structure of digraphs preserved by a Maltsev operation, that we call *Maltsev digraphs*. We show that these digraphs retract either onto the disjoint union of directed cycles or to a directed path. This gives a direct proof that the corresponding CSP is in constant width symmetric Datalog and therefore in  $\mathbf{L}$ . (Membership of these CSPs in symmetric Datalog, without the constant width guarantee, was independently shown by Kazda [19], however, his proof is rather indirect.) We then generalise other results in [19] to show that a Maltsev digraph is preserved not only by a majority polymorphism, but also by a class of polymorphism obeying certain restrictions (e.g. minority, Pixley). We also extend the results to the conservative setting, i.e. we show that a conservative Maltsev digraph is preserved by a class of conservative polymorphisms.

A generalization of the rectangularity [4] property of digraphs is introduced. We call this rectangularity *total rectangularity*, and we establish that a digraph is preserved by a Maltsev operation iff it is totally rectangular. Similarly, we show that a digraph is preserved by a conservative Maltsev operation iff it is *universally rectangular*, a specific form of total rectangularity.

We apply our results to the list homomorphism problem,  $\text{LHOM}$ , for directed graphs. While the complexity of  $\text{LHOM}$  for undirected graphs is completely understood [13], for the directed version, the only result known is a  $\mathbf{P}$  vs.  $\mathbf{NP}$  dichotomy [17]. We show that  $\text{LHOM}$  for Maltsev digraphs is in  $\mathbf{L}$ .

It is shown that a digraph preserved by a Maltsev operation is also preserved by a conservative 2-semilattice operation iff the digraph satisfies a certain combinatorial property. We note that if a structure is preserved by a 2-semilattice operation, then  $\text{CSP}(\mathbf{B})$  is in Datalog and therefore in  $\mathbf{P}$  [18]. We also characterise Maltsev digraphs preserved by a conservative 2-semilattice operation and show that these digraphs can be recognised in  $\mathbf{NL}$ .

Finally, an inductive construction of directed acyclic graphs preserved by a Maltsev operation is given. The main motivation behind this construction is that we suspect that extending this construction to  $n$ -permutable digraphs (2-permutable digraphs are precisely the Maltsev digraphs [15]) could make progress towards identifying all list homomorphism problems for digraphs in  $\mathbf{L}$ . We note that in [13], an inductive construction of “conservative”  $n$ -permutable digraphs is key to the identification of all graphs whose  $\text{LHOM}$  is in  $\mathbf{L}$ .

## 2 Preliminaries

### 2.1 Algebra

We describe the algebraic definitions for digraphs, however, note that these definitions are straightforward to generalize to relational structures. Let  $G = (V_G, E_G)$  and  $H = (V_H, E_H)$  be digraphs. A *homomorphism* from  $G$  to  $H$  is a map  $f$  from  $V_G$  to  $V_H$ , such that for every edge  $(u, v) \in E_G$  we have that  $(f(u), f(v))$  is an edge in  $H$ , i.e.  $(f(u), f(v)) \in E_H$ . A digraph  $G$  is called a *core* if every homomorphism from  $G$  to itself is an automorphism, i.e. a permutation on  $V_G$ . Let  $G'$  be a subgraph of  $G$ . We say that  $G$  *retracts onto*  $G'$  if there is a homomorphism  $h : G \rightarrow G'$  such that  $h$  is the identity map on  $G'$ . For a digraph  $H$ , we can then define  $\text{CSP}(H)$  as the class of all digraphs that admit a homomorphism to  $H$ .

An  $n$ -ary operation on a set  $A$  is a function  $f : A^n \rightarrow A$ . Given a digraph  $G$  and an  $n$ -ary operation  $f$  on  $V_G$ , we say that  $f$  *preserves*  $G$ , or that  $f$  is a *polymorphism* of  $G$ , if for any  $n$  edges  $(u_1, v_1), \dots, (u_n, v_n) \in E_G$  (not necessarily distinct), the pair  $(f(u_1, \dots, u_n), f(v_1, \dots, v_n)) \in E_G$ . For an  $n$ -ary operation  $f$ , we write  $f(x_1, \dots, x_n) \approx f(y_1, \dots, y_n)$  if  $f(x_1, \dots, x_n) = f(y_1, \dots, y_n)$  for all possible values of the  $x_i, y_i, i = 1, \dots, n$ . A ternary operation  $m$  is *Maltsev* if it satisfies  $m(x, x, y) \approx m(y, x, x) \approx y$ , *Pixley* if it satisfies  $m(x, x, y) \approx m(y, x, x) \approx m(y, x, y) \approx y$ , *majority* if it satisfies  $m(x, x, y) \approx m(x, y, x) \approx m(y, x, x) \approx x$ , and *minority* if it satisfies  $m(x, x, y) \approx m(x, y, x) \approx m(y, x, x) \approx y$ . A binary operation  $*$  is *2-semilattice* if it satisfies  $x * x \approx x, x * y \approx y * x \approx x * (x * y)$ .

### 2.2 Graph Theory

Since all graphs in the paper are directed, we use the terms graph and digraph interchangeably. For a natural number  $n$  we write  $[n] = \{1, 2, \dots, n\}$ . An *oriented path* is a sequence of, not necessarily distinct, vertices  $v_1, \dots, v_n$  such that for every  $i \in [n - 1]$ , either  $(v_i, v_{i+1})$  (a *forward edge*) or  $(v_{i+1}, v_i)$  (a *backward edge*) is an edge. We use the terms path and oriented path interchangeably. A *cycle* is an oriented path with starting point  $v_1$  and endpoint  $v_m$  such that either  $(v_m, v_1)$  or  $(v_1, v_m)$  is an edge. The *net length* of a path  $P$ ,  $\text{net}(P)$ , is the number of forward edges minus the number of backward edges in  $P$ . A *(reverse) dipath* is a sequence of, not necessarily distinct, vertices  $v_1, \dots, v_n$  such that for every  $i \in [n - 1]$ ,  $(v_i, v_{i+1})$  ( $(v_{i+1}, v_i)$ ) is an edge. A *directed cycle* is a dipath  $v_1, \dots, v_n$  such that  $(v_n, v_1)$  is also an edge. For a (reverse) dipath  $P$ , we let  $\text{len}(P)$  denote the number of edges in  $P$ . We use the term *simple* dipath or (directed) cycle to indicate that all vertices of the dipath or (directed) cycle are distinct.

A *component* of digraph  $G$  is a maximal subgraph,  $H$ , of  $G$  such that for every pair of vertices  $u, v \in V_H$ , there is an oriented path from  $u$  to  $v$ . A digraph with one component is said to be *connected*. A digraph is a *directed acyclic graph* (DAG) if it contains no directed cycles. A DAG  $G$  is *layered* if there exists  $q \in \mathbb{N}$  such that the vertices of  $G$  can be partitioned into  $q$  levels  $L_0, \dots, L_{q-1}$ , such that any edge of  $G$  goes from  $L_i$  to  $L_{i+1}$ , for some  $i = 0, \dots, q - 2$ .

Let  $G$  be a digraph, and  $x$  a vertex of  $G$ . We define  $x^{+1} = \{y \in V_G : (x, y) \in E_G\}$ , and  $x^{-1} = \{y \in V_G : (y, x) \in E_G\}$ . We call a vertex  $v$  a *source* if  $v^{-1} = \emptyset$ , and a *sink* if  $v^{+1} = \emptyset$ . If  $u$  and  $v$  are vertices of  $G$ ,  $u \xrightarrow{k} v$  denotes the existence of a dipath from  $u$  to  $v$  of length  $k$ ;  $u \rightarrow v$  denotes  $u \xrightarrow{1} v$ .

### 3 Retracts of Maltsev digraphs

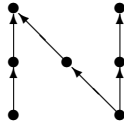
**Definition 1 (totally rectangular).** A digraph  $G$  is  $k$ -rectangular if the following implication holds for all vertices  $x, y, u, v$ :

$$x \xrightarrow{k} u \ \& \ y \xrightarrow{k} u \ \& \ y \xrightarrow{k} v \Rightarrow x \xrightarrow{k} v.$$

A digraph is rectangular if it is 1-rectangular, and totally rectangular if it is  $k$ -rectangular for every  $k \in \mathbb{N}$ .

It is not hard to verify that a Maltsev digraph must be totally rectangular, but in Section 4 (see Corollary 1) we show that the two properties are equivalent.

*Example 1.* The digraph in Fig. 1 is rectangular but not 2-rectangular. While the digraph in Fig. 4 is totally rectangular.



**Fig. 1.** A rectangular digraph that is not 2-rectangular.

We now state the main result of this section.

**Theorem 1.** Let  $G$  be a totally rectangular digraph. If  $G$  is acyclic then  $G$  retracts onto a simple dipath. Otherwise  $G$  retracts onto the disjoint union of simple directed cycles.

The proof of Theorem 1 is a direct consequence of Lemma 4 below. We begin with some definitions and simple observations.

**Lemma 1.** Let  $G$  be a digraph. Then  $G$  is layered iff for every pair of vertices  $u, v$  in  $G$ , and any pair of oriented paths  $P$  and  $Q$  from  $u$  to  $v$ , it holds that  $net(P) = net(Q)$ .

**Definition 2.** Let  $G$  be a digraph that contains a directed cycle. Let  $C$  be a shortest directed cycle in  $G$  and assume it has length  $m$ . We say that  $G$  is inconsistent if there exist two vertices  $u, v$  in  $G$  such that, there are two different oriented paths of net lengths  $\ell_1$  and  $\ell_2$  from  $u$  to  $v$  such that  $\ell_1 \not\equiv \ell_2 \pmod{m}$ . Otherwise we say that  $G$  is consistent.

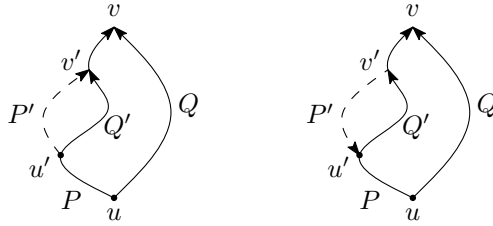
**Proposition 1.** Let  $G$  be a digraph that contains a directed cycle. Let  $C$  be a shortest directed cycle in  $G$ . Then  $G$  retracts onto  $C$  iff  $G$  is consistent.

**Lemma 2.** *Let  $G$  be a totally rectangular digraph and  $u, v$  be vertices in  $G$ . Let  $P$  and  $Q$  be two dipaths in  $G$  from  $u$  to  $v$ , such that  $\text{len}(P) > \text{len}(Q)$ . Set  $k = \text{len}(P)$ ,  $\ell = \text{len}(Q)$ , and  $d = k - \ell$ . Then one of the following two cases occurs:*

1. *If  $2\ell > k$ , then  $G$  contains vertices  $u', v'$  and dipaths  $P', Q'$  from  $u'$  to  $v'$  with the following property:  $\text{len}(P') = \ell$ ,  $\text{len}(Q') = 2\ell - k$ , and  $\text{len}(P') - \text{len}(Q') = d$ ;*
2. *If  $2\ell \leq k$ , then  $G$  contains a directed cycle of length  $d$ .*

*Proof.* See Fig. 2. In the first case, let  $u'$  be the vertex of  $P$  such that the subpath  $P_{u'v}$  of  $P$  from  $u'$  to  $v$  has length  $\ell$ . Let  $v'$  be the vertex of  $P$  such that the subpath  $P_{uv'}$  of  $P$  from  $u$  to  $v'$  has length  $\ell$ . Applying the  $\ell$ -rectangularity of  $G$  to  $P_{u'v}$ ,  $Q$ , and  $P_{uv'}$ , we obtain the desired dipath  $P'$  with  $\text{len}(P') = \ell$ . The other required dipath  $Q'$  is the subpath of  $P$  from  $u'$  to  $v'$ . Since  $k = 2\ell - \text{len}(Q')$  we have that  $\text{len}(Q') = 2\ell - k$ , and  $\text{len}(P') - \text{len}(Q') = \ell - (2\ell - k) = k - \ell = d$ .

In the second case, the two paths  $P'$  and  $Q'$  form a cycle of length  $\ell + (k - 2\ell) = d$  because  $2\ell \leq k$ .



**Fig. 2.** Case 1 (left) and Case 2 (right) of Lemma 2.

**Lemma 3.** *Let  $G$  be a totally rectangular digraph and  $u, v \in V_G$ . Let  $P_1$  and  $P_2$  be oriented paths in  $G$  from  $u$  to  $v$ . Assume that  $\text{net}(P_1) > \text{net}(P_2)$ , and set  $d = \text{net}(P_1) - \text{net}(P_2)$ . Then there are vertices  $s, t \in V_G$  and dipaths  $Q_1$  and  $Q_2$  in  $G$  from  $s$  to  $t$ , such that  $\text{len}(Q_1) - \text{len}(Q_2) = d$ .*

**Lemma 4.** *Let  $G$  be a connected totally rectangular digraph. If  $G$  is a DAG then  $G$  retracts onto a simple dipath. Otherwise  $G$  retracts onto a simple directed cycle.*

*Proof.* Assume first that  $G$  is a DAG. We claim that  $G$  must be layered. Assume, for a contradiction, that  $G$  is not layered. By Lemma 1, there exist  $u, v \in V_G$  and oriented paths  $P$  and  $Q$  from  $u$  to  $v$ , such that  $\text{net}(P) \neq \text{net}(Q)$ . Using Lemma 3, we can assume that  $P$  and  $Q$  are dipaths of different length. Now we repeatedly apply Case 1 of Lemma 2 as long as it is possible, and then applying Case 2 yields a cycle, a contradiction. So  $G$  is layered.

Assume that  $G$  has levels  $L_0, \dots, L_{q-1}$ . Fix vertices  $s \in L_0$  and  $t \in L_{q-1}$ , and let  $O$  be any oriented path from  $s$  to  $t$  (such a path exists because  $G$  is

connected). Applying the total rectangularity of  $G$  to appropriate subpaths of  $O$ , it is easy to see that there exists a dipath  $D$  of length  $q - 1$  from  $s$  to  $t$  in  $G$ . Clearly,  $G$  retracts onto  $D$ .

Suppose that  $G$  contains a directed cycle. By Proposition 1, it is enough to show that  $G$  is consistent. Assume this is not the case. Let  $C$  be a shortest directed cycle in  $G$ , and assume it has length  $m$ . Because  $G$  is inconsistent, we can find vertices  $u, v \in V_G$  and oriented paths  $P_1$  and  $P_2$  from  $u$  to  $v$ , such that  $\text{net}(P_1) \not\equiv \text{net}(P_2) \pmod{m}$ . Set  $\ell_1 = \text{net}(P_1)$  and  $\ell_2 = \text{net}(P_2)$ . Assume w.l.o.g. that  $\ell_1 > \ell_2$ , and that  $u$  is a vertex of  $C$ . Note that if  $u$  is not a vertex of  $C$ , then we fix a vertex  $c$  of  $C$  and find any oriented path  $S$  from  $c$  to  $u$ . Then attaching  $S$  to  $P_1$  and  $P_2$  at vertex  $u$  gives us the desired oriented paths. Furthermore, we can assume that  $\ell_1 - \ell_2 = d < m$ , because if not, we can add  $C$ -loops from  $u$  to  $u$  to  $P_2$  to increase its length by a multiple of  $m$ , until  $\ell_1 - \ell_2 < m$ . Using Lemma 3 we obtain directed paths  $Q_1$  and  $Q_2$  such that  $\text{len}(Q_1) - \text{len}(Q_2) = d$ , and then, by applying Lemma 2, we obtain a cycle of length  $d$  in  $G$ , a contradiction.

By Lemma 4, each connected component of  $G$  retracts either onto a simple dipath or to a simple directed cycle. The trivial observation that a dipath homomorphically maps to a cycle completes the proof Theorem 1.

## 4 Characterisations, Polymorphisms and Algorithms

### 4.1 Rectangular Characterisations and Other Polymorphisms

In this section we generalise a technique of Kazda [19] to characterise digraphs that admit Maltsev and conservative Maltsev polymorphisms as those which are totally rectangular and universally rectangular respectively and to provide polynomial time algorithms for recognising the relevant properties. Furthermore, we show that Maltsev digraphs also admit many other polymorphisms, and under certain conditions, they also admit conservative 2-semilattice polymorphisms.

**Definition 3 (conservatively  $k$ -rectangular, universally rectangular).**

We say that a graph is conservatively  $k$ -rectangular if it satisfies the following sentence:

$$\left. \begin{array}{l} x \rightarrow x_1 \rightarrow \cdots \rightarrow x_{k-1} \rightarrow u \\ y \rightarrow y_1 \rightarrow \cdots \rightarrow y_{k-1} \rightarrow u \\ y \rightarrow z_1 \rightarrow \cdots \rightarrow z_{k-1} \rightarrow v \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \text{There is a path } x \rightarrow w_1 \rightarrow \cdots \rightarrow \\ w_{k-1} \rightarrow v \text{ with } w_i \in \{x_i, y_i, z_i\} \\ \text{for each } i. \end{array} \right. \quad (1)$$

A graph that is conservatively  $k$ -rectangular for all  $k \geq 1$  will be called universally rectangular.

*Example 2.* The digraph in Fig. 4 is conservatively rectangular but not conservatively 2-rectangular. While the digraph in Fig. 3 is universally rectangular.

**Definition 4.** Let  $G$  be a digraph. Define the binary relations  $R^-$  on  $V_G$  by  $x R^- y$  if  $x^{-1} \cap y^{-1} \neq \emptyset$ . The dual relation  $R^+$  is defined by  $x R^+ y$  if  $x^{+1} \cap y^{+1} \neq \emptyset$ .

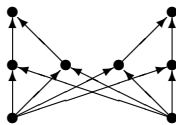


Fig. 3. A universally rectangular digraph

The relation  $R^+$  is an equivalence relation on the set  $\{x \in V_G : x^+ \neq \emptyset\}$ , the set of vertices of  $G$  that are not sinks. The relation  $R^-$  is an equivalence relation on the set  $\{x \in V_G : x^- \neq \emptyset\}$ , the set of vertices of  $G$  that are not sources. So it makes sense to consider the respective factor graphs, this was observed in [19]. We use the notation  $G/R^+$  to denote the graph on the  $R^+$ -classes of  $G$ . Given  $R^+$ -classes  $A, B$ , we write  $A \rightarrow B$  if there is some  $a \in A$  and  $b \in B$  with  $a \rightarrow b$ . Similarly,  $G/R^-$  denotes the same construction, but using the relation  $R^-$ . Note that  $G/R^+$  is not strictly an actual graph quotient of  $G$ , only a quotient of an induced subgraph of  $G$ . Nevertheless, we sometimes refer to it as “the quotient of  $G$  by  $R^+$ ”. The proof of the following lemma is routine.

**Lemma 5.** *Let  $G$  be a rectangular digraph and  $k > 1$ .*

1.  *$G$  is  $\ell$ -rectangular for all  $\ell = 1, \dots, k$  if and only if  $G/R^+$  is  $\ell$ -rectangular for all  $\ell = 1, \dots, k - 1$ .*
2. *If  $G$  is conservatively  $\ell$ -rectangular for all  $\ell = 1, \dots, k$  then  $G/R^+$  is conservatively  $\ell$ -rectangular for all  $\ell = 1, \dots, k - 1$ .*

For a totally rectangular graph  $G$ , define  $G_0 = G$  and  $G_{i+1} = G_i/R^+$ ,  $i \geq 1$ . From Lemma 5 it follows that  $G_i$  is defined for all positive integers  $i$ , and eventually  $G_i$  will either be empty or a disjoint union of directed cycles (the only situations that  $R^+$  can be trivial). We define  $G_\infty = G_k$ , where  $k$  is such that  $G_k = G_{k+1}$  (i.e.  $G_\infty$  is either empty or a disjoint union of directed cycles).

The next lemma is obtained by applying the Maltsev property to the columns of the premise of (1).

**Lemma 6.** *Let  $G$  be a digraph.*

1. *If  $G$  has a Maltsev polymorphism, then  $G$  is totally rectangular.*
2. *If  $G$  has a conservative Maltsev polymorphism, then  $G$  is universally rectangular.*

**Lemma 7.** *Let  $a, b$  be vertices in a totally rectangular digraph  $G$  satisfying conservative 2-rectangularity and assume that neither  $a$  nor  $b$  is a source or sink. If  $a/R^+ \cap b/R^-$  is nonempty then either  $b \in a/R^+$  or  $a \in b/R^-$ .*

*Proof.* Let  $c \in a/R^+ \cap b/R^-$ . There are vertices  $e, f, g, h$  such that  $\{a, c\} \subseteq e^{-1}$ ,  $b \in f^{-1}$ ,  $a \in g^+$  and  $\{c, b\} \in h^+$ . However  $G$  is conservatively 2-rectangular so that there is either an edge from at least one of  $a, c$  to  $f$  or there is an edge from  $g$  to at least one of  $\{b, c\}$ . Then 1-rectangularity shows that either there is an edge from  $a$  to  $f$  or from  $g$  to  $b$ .

**Theorem 2.** *Consider a property  $C$  of digraphs defined by the existence of polymorphisms  $t_1, t_2, \dots, t_k$  (not necessarily distinct) satisfying a single equational sequence*

$$t_1(x_{1,1}, x_{1,2}, \dots, x_{1,n_1}) \approx \dots \approx t_k(x_{k,1}, x_{k,2}, \dots, x_{k,n_k}) \approx x,$$

where  $\{x_{1,1}, \dots, x_{1,n_1}\} = \dots = \{x_{k,1}, \dots, x_{k,n_k}\}$  and  $x \in \{x_{1,1}, \dots, x_{1,n_1}\}$ . The following statements are true provided that the equation  $x \approx y$  does not follow from  $C$ .

1. *Let  $G$  be a totally rectangular digraph. Then  $G$  has property  $C$  if and only if  $G_\infty$  has property  $C$ .*
2. *Let  $G$  be universally rectangular. Then  $G$  has property  $C$  with each  $t_i$  conservative if and only if  $G_\infty$  has property  $C$  with each of the  $t_i$  conservative.*

*The same conclusions can be made without the requirement that  $\approx x$  be included in the equational sequence, and if the polymorphisms are required to be idempotent.*

*Proof.* Our proof is very similar to the main proof in [19]; we use Lemma 5 rather than the assumption of the Maltsev property directly. We focus only on the conservative case (not considered in [19]), as the non-conservative case is obtained by following this proof and missing some steps. We give only a sketch here.

It is easy to see that if  $G$  has conservative property  $C$  then so does  $G/R^+$  by defining  $t_i(x_1/R^+, \dots, x_n/R^+) = t_i(x_1, \dots, x_n)/R^+$ . Thus, it suffices to show that if  $G/R^+$  satisfies some conservative property  $C$  then so does  $G$ . This uses only total rectangularity (to ensure that successive quotients are well defined).

The reverse direction is shown by backward induction over successive quotients by  $R^+$ : essentially, provided  $G/R^+$  has polymorphisms  $t_i$  witnessing property  $C$ , then so does  $G$ . This part of the argument is mostly identical to that given in [19] (in the case of majority polymorphisms), except Lemma 7 is invoked to guarantee a conservative choice of the polymorphisms.

It is useful to note that instead of explicit use of universal rectangularity, this argument used only the fact that on each successive quotient by  $R^+$ , both rectangularity and the conclusion of Lemma 7 hold.

Some instances of polymorphisms satisfying the conditions of Theorem 2 are majority, Maltsev and Pixley. In these cases  $G_\infty$  always has the desired polymorphism, giving the following corollary.

**Corollary 1.** *Let  $G$  be a digraph.*

1.  *$G$  admits a (conservative) Maltsev polymorphism iff it admits a (conservative) Pixley operation iff it is totally (universally) rectangular.*
2. *If  $G$  is totally (universally) rectangular then  $G$  admits a (conservative) minority polymorphism and a (conservative) majority polymorphism.*



*Remark 1.* The first part of Corollary 1 strengthens the result given in Lemma 4 of [11], for the case of digraphs. The relational clone  $\langle \mathbf{B} \rangle$  of a structure  $\mathbf{B}$  is the set of all relations that can be expressed with primitive positive first-order formulas (i.e. only existential quantification, conjunction, and equality is allowed) from  $\mathbf{B}$ . When we restrict [11, Lemma 4] to digraphs, it can be stated as follows: A digraph  $G$  is preserved by a Maltsev operation iff every binary relation in  $\langle G \rangle$  is rectangular. It is easy to see and well-known that every binary relation in  $\langle G \rangle$  can be expressed as  $B_G(S, a, b) = \{(h(a), h(b)) \mid h \text{ is a homomorphism from } S \text{ to } G\}$  for a structure  $S$  with two distinguished vertices  $a$  and  $b$ . Then Corollary 1 implies that for a digraph  $G$  to be preserved by a Maltsev operation it is enough to require that only those binary relations in  $\langle G \rangle$  that can be expressed as  $B_G(S, a, b)$ , where  $S$  is a directed path with initial vertex  $a$  and terminal vertex  $b$ , are rectangular.

The above corollary yields an algorithm for verifying if a graph has a Maltsev (or Pixley) polymorphism. Indeed, the rectangularity of a digraph is equivalent to the following property of its adjacency matrix: when two rows (or two columns) share a common 1 they are identical. On an  $n$ -vertex digraph this property may be verified in  $O(n^3)$  steps. A digraph has a Maltsev polymorphism if and only if each (of at most  $n$ ) successive quotient by  $R^+$  is rectangular, with the process stopping once there are no  $R^+$ -classes of size more than 1 (which happens after at most  $n$  quotients). Overall this takes  $O(n^4)$  steps (quadratic in terms of the size of the adjacency matrix).

Universal rectangularity (equivalently, the existence of a conservative Maltsev polymorphism) can also be verified in polynomial time by verifying total rectangularity and conservative 2-rectangularity at each successive quotient by  $R^+$ . In fact, the proof of Theorem 2 is sufficiently constructive to construct the desired polymorphisms (when they exist): simply work backwards from their definition of  $G_\infty$ .

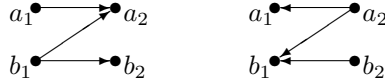
## 4.2 Conservative 2-Semilattice Polymorphisms

A disjoint union of directed cycles admits a conservative commutative binary (ccb) polymorphism (which coincide with conservative 2-semilattice operations) if and only if it contains no even cycles. This provides a case where Theorem 2 characterises a proper subclass of conservative Maltsev digraphs. In this section we classify Maltsev digraphs admitting a ccb polymorphism. A corollary of the result will be a sort of converse to Theorem 2: a Maltsev digraph with a ccb polymorphism is necessarily conservative Maltsev (Proposition 2 below).

Consider any digraph  $G$  and let  $*$  be any conservative commutative binary operation on  $V_G$ . The operation  $*$  has an easy interpretation as a colouring of the nondiagonal elements of the cartesian square  $V_G^2 \setminus \{(v, v) \mid v \in V_G\}$ : the pair  $(a, b)$  is coloured  $L$  if  $a * b = a$  and  $R$  if  $a * b = b$ ; commutativity is equivalent to  $(a, b)$  having different colour to  $(b, a)$ . (The exclusion of the diagonal elements is only for convenience.) We now examine the consequences of  $*$  being a ccb polymorphism.

For any digraph  $G$  we define a structure—the *ccb graph* of  $G$ —on the non-diagonal elements of the cartesian square  $V_G^2 \setminus \{(v, v) \mid v \in V_G\}$ . The ccb graph is a graph with two kinds of edges: “orienting” edges, which are directed, and “straight” edges which are considered as having no direction.

- A (directed) orienting edge is placed from  $(a_1, b_1)$  to  $(a_2, b_2)$  if there are parallel edges connecting each of the following pairs in  $G$ :  $a_1$  and  $a_2$ ;  $b_1$  and  $b_2$ ; and  $b_1$  and  $a_2$  but not  $a_1$  and  $b_2$ . The following diagram depicts two situations that an orienting edge pointing from  $(a_1, b_1)$  to  $(a_2, b_2)$  can arise:



- An (undirected) straight edge is placed from  $(a_1, b_1)$  to  $(a_2, b_2)$  if there are parallel edges connecting the following pairs:  $a_1$  and  $a_2$ ;  $b_1$  and  $b_2$ ; but not  $a_1$  and  $b_2$  or  $a_2$  and  $b_1$ .

A directed path in the ccb graph is a path of orienting edges and straight edges, in which each orienting edge is traversed in a forward direction. We now adopt the notation  $\rightsquigarrow$  to denote directed paths: note that  $(a, b) \rightsquigarrow (c, d)$  if and only if  $(d, c) \rightsquigarrow (b, a)$ . An important observation is: if  $(a, b)$  is not connected to  $(c, d)$  in the ccb graph of  $G$  then the compatibility of the ccb operation with the edges of  $G$  does not fail at the pairs  $(a, b)$  and  $(c, d)$ . Thus *the digraph  $G$  admits a ccb polymorphism iff the ccb graph of  $G$  can be coloured by  $L$  and  $R$  such that  $L$  and  $R$  are preserved across straight edges,  $L$  is preserved forward across orienting edges and  $R$  is preserved backward across orienting edges*. Expressed in terms of  $\rightsquigarrow$  this becomes the following rules.

- (L) If  $(u, v)$  is coloured  $L$  and  $(u, v) \rightsquigarrow (x, y)$  then:
  - (1)  $(x, y)$  is coloured  $L$ .
  - (2)  $(y, x)$  is coloured  $R$ .
- (R) If  $(u, v)$  is coloured  $R$  and  $(x, y) \rightsquigarrow (u, v)$  then:
  - (1)  $(x, y)$  is coloured  $R$ .
  - (2)  $(y, x)$  is coloured  $L$ .

**Theorem 3.** *A digraph  $G$  admits a ccb polymorphism  $*$  if and only if for every distinct  $a, b \in G$ , the ccb-graph of  $G$  does not contain a directed path both from  $(a, b)$  to  $(b, a)$  and from  $(b, a)$  to  $(a, b)$ . When a ccb polymorphism exists, it can be constructed in a polynomial number of steps.*

*Proof.* The forward implication is simply the statement that the colouring of the ccb-graph must colour each pair  $(a, b)$  oppositely to its reverse  $(b, a)$ , and the colouring rules (L)(1)–(R)(2) must be obeyed.

Now we show the converse: assume that for every  $a, b \in G$ , the ccb-graph of  $G$  does not contain a directed path from  $(a, b)$  to  $(b, a)$  and from  $(b, a)$  to  $(a, b)$ . We construct (in a polynomial number of steps) a successful colouring by  $L$  and  $R$ , whence a ccb polymorphism.

**First phase.** Begin by finding any pairs  $(a, b)$  from which there is a directed path to the reverse pair  $(b, a)$  (obviously, such a path must contain an orienting edge). In every case, colour  $(a, b)$  by  $R$  and  $(b, a)$  by  $L$ . By assumption, no pair is coloured two different colours simultaneously.

**Second phase.** Routine arguments show that after phase 1, no uncoloured pair is forced to be coloured by rules (L)(1)–(R)(2). Complete the colouring by iterating the following process until all pairs are coloured: take any uncoloured pair,  $(a, b)$  say, and colour it arbitrarily; colour any other pairs for which rules (L)(1)–(R)(2) apply starting from  $(a, b)$ .

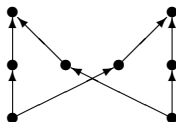
Routine arguments show that, at each iteration, no pair is coloured twice by different colours, hence the desired colouring (and ccb-polymorphism) is eventually achieved.

Note that the second property in Theorem 3 can be verified using Reachability in the ccb-graph of  $G$ , so is solvable in nondeterministic logarithmic space. Hence deciding if a digraph has a ccb polymorphism is in NL too.

**Proposition 2.** *The following are equivalent for a Maltsev graph  $G$ :*

1.  $G$  has a conservative commutative idempotent binary polymorphism;
2.  $G$  has a conservative commutative idempotent binary polymorphism and a conservative Maltsev polymorphism;
3.  $G$  has a conservative Maltsev polymorphism and  $G_\infty$  is empty or has no even length cycles.

*Proof.* (2) $\Rightarrow$ (3) $\Rightarrow$ (1) follow from Theorem 2, because a disjoint union of directed cycles has a ccb iff it has no even length cycles. For (1) $\Rightarrow$ (2) we use Theorem 3 to prove a version of Lemma 7 with conservative 2-rectangularity replaced by ccb. The routine proof, which we omit, uses the digraph in Fig. 4.



**Fig. 4.** A Maltsev digraph with no conservative Maltsev polymorphism

Note that the digraph in Fig. 4 is easily seen to admit a conservative majority polymorphism, a Maltsev polymorphism, but no conservative Maltsev polymorphism. So conservative majority cannot replace ccb in Proposition 2.

### 4.3 A simple inductive construction of Maltsev DAGs

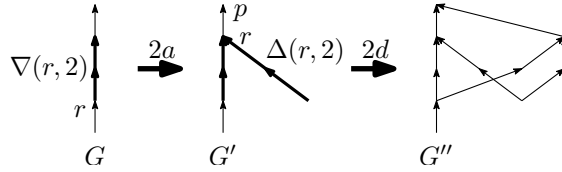
In this section we provide a simple inductive characterisation of totally rectangular DAGs. We note that in [19, Corollary 16] Kazda gives an inductive construction of Maltsev digraphs, however, this construction is not fully satisfying in the sense that it is non-deterministic, i.e. it does not specify how to obtain

the desired preimages, and it is not clear if it can be made deterministic. The construction described below consists of repeated applications of two straightforward steps (and their reverse versions) which clearly specify how to obtain a new Maltsev digraph from an already constructed one by a certain copying process. We need the following definitions.

**Definition 5 ((Reverse) arborescence).** *An (reverse) arborescence is a directed tree with root  $r$  such that every edge points away from (towards)  $r$ .*

**Definition 6 ( $\nabla(r, h)$  and  $\Delta(r, h)$ ).** *Let  $G$  be a digraph,  $r \in V_G$ , and  $h \in \mathbb{N}$ .  $\nabla(r, h)$  ( $\Delta(r, h)$ ) is defined to be the subgraph of  $G$  whose vertices and edges are the vertices and edges of all (reverse) sub-dipaths of  $G$  which have initial vertex  $r$  and length  $h$ . A vertex  $v \in \nabla(r, h) - r$  ( $\Delta(r, h) - r$ ) is called an endpoint of  $\nabla(r, h)$  ( $\Delta(r, h)$ ) if there is a (reverse) dipath of length  $h$  from  $r$  to  $v$ . Otherwise  $v$  is called an inner vertex of  $\nabla(r, h)$  ( $\Delta(r, h)$ ).*

**Definition 7 (isolated  $\nabla(r, h)$ ).** *Let  $G$  be a digraph. Consider  $\nabla(r, h)$  ( $\Delta(r, h)$ ) for some  $r \in V_G$  and  $h \in \mathbb{N}$ . We say that  $\nabla(r, h)$  ( $\Delta(r, h)$ ) is isolated in  $G$  if for every inner vertex  $v$  of  $\nabla(r, h)$ , both the in-neighbourhood and the out-neighbourhood of  $v$  belongs to  $\nabla(r, h)$  ( $\Delta(r, h)$ ).*



**Fig. 5.** Construction of a totally rectangular DAG.

We are ready to define the construction formally in Fig. 6. This construction can be used, for example, to define a minority operation for a totally rectangular DAG. Also, a more restricted version of the construction can be defined, but it becomes slightly more technical.

*Example 3.* Consider the totally rectangular DAG  $G''$  in Fig. 5. To construct it using the method in Fig. 6, we start with the dipath  $G$  and first apply Step 2a to  $G$  to obtain  $G'$ . Next we apply Step 2d to  $G'$  obtain  $G''$ . The thick edges indicate the subgraphs  $\nabla(r, 2)$  and  $\Delta(r, 2)$ , which are the subgraphs to be copied and attached appropriately.

**Theorem 4.** *The class of totally rectangular DAGs is the set of digraphs  $\mathcal{M}$  defined in Fig. 6.*

1.  $\mathcal{C}$  contains dipaths of all possible lengths  $n \in \mathbb{N}_0$ ;
2.  $\mathcal{C}$  is closed under applying the following operations:
  - (a) Given a digraph  $G$ , let  $r \in V_G$  and  $h \in \mathbb{N}$  such that  $\nabla(r, h)$  is an arborescence. Let  $\nabla'$  be a copy of  $\nabla(r, h)$ . Join  $\nabla'$  to  $G$  by identifying the corresponding endpoints of  $\nabla'$  and  $\nabla(r, h)$ . Let the resulting graph be  $G'$ ;
  - (b) Given a digraph  $G$ , let  $r \in V_G$  be such that  $r$  has exactly one incoming edge  $(p, r)$ , and  $h \in \mathbb{N}$  such that  $\nabla(r, h)$  is an isolated arborescence. Let  $\nabla'$  be a copy of  $\nabla(r, h)$  with root  $r'$ . Join  $\nabla'$  to  $G$  by identifying the corresponding endpoints of  $\nabla'$  and  $\nabla(r, h)$ , and adding the edge  $(p, r')$ . Let the resulting graph be  $G'$ ;
  - (c) The reverse version of Step 2a (defined in the natural way);
  - (d) The reverse version of Step 2b (defined in the natural way).
3.  $\mathcal{M}$  is the set of digraphs that can be obtained by taking disjoint unions of digraphs in  $\mathcal{C}$ .

**Fig. 6.** Inductive construction of the set  $\mathcal{M}$  of totally rectangular DAGs.

## 5 Some Applications to the Constraint Satisfaction Problem

The logic programming language Datalog is one of the main tools to solve CSPs in P. The fragments of Datalog called linear and symmetric Datalog are conjectured to contain all CSPs in NL and L, respectively, see [7, 8, 14]. A minor technicality is that it is actually the complement of a CSP that can be defined in Datalog and its fragments, not the actual CSP.

By Theorem 1, the core of a Maltsev digraph is either a directed path or a disjoint union of cycles, and for such digraphs the following is not difficult to show.

**Corollary 2.** *Let  $H$  be a Maltsev digraph. Then the complement of  $\text{CSP}(H)$  can be defined in symmetric Datalog of constant width and therefore  $\text{CSP}(H)$  is in L.*

The *list homomorphism problem* for a digraph  $H$ ,  $\text{LHOM}(H)$ , is the following decision problem. Given an input digraph  $G$  and for each vertex  $v \in V_G$  a list  $L_v \subseteq V_H$ , determine if there is a homomorphism  $h$  from  $G$  to  $H$  such that for each  $v \in V_G$ ,  $h(v) \in L_v$ . This problem is exactly  $\text{CSP}(H_u)$  where  $H_u$  is the structure obtained by expanding the digraph  $H$  with unary relations  $U$ , where  $U$  runs through all non-empty subsets of  $V_H$ . Using Corollary 1, the following corollary is easy to deduce.

**Corollary 3.** *The complement of  $\text{LHOM}(H)$  for a conservative Maltsev digraph  $H$  can be defined in symmetric Datalog, and therefore  $\text{LHOM}(H)$  is in L.*

## References

1. L. Barto and J. Bulín. CSP dichotomy for special polyads. *Submitted*, 2011.
2. L. Barto, M. Kozik, M. Maróti, and T. Niven. CSP dichotomy for special triads. *Proceedings of the AMS*, 137:2921–2934, 2009.
3. L. Barto, M. Kozik, and T. Niven. Graphs, polymorphisms and the complexity of homomorphism problems. In *Proceedings of the 40th annual ACM symposium on Theory of computing*, STOC '08, pages 789–796, 2008.
4. A. Bulatov and V. Dalmau. A simple algorithm for Mal'tsev constraints. *SIAM Journal on Computing*, 36(1):16–27, 2006.
5. A. Bulatov, P. Jeavons, and A. Krokhin. Constraint satisfaction problems and finite algebras. In *Proceedings of the 27th International Colloquium on Automata, Languages and Programming*, ICALP '00, pages 272–282, 2000.
6. A. Bulatov, P. Jeavons, and A. Krokhin. Classifying the complexity of constraints using finite algebras. *SIAM Journal on Computing*, 34(3):720–742, 2005.
7. A. Bulatov, A. Krokhin, and B. Larose. Dualities for constraint satisfaction problems. In *LNCS Surveys on Complexity of Constraints*, volume 5250, pages 93–124, 2008.
8. V. Dalmau. Linear Datalog and bounded path duality of relational structures. *Logical Methods in Computer Science*, 1:1–32, 2005.
9. V. Dalmau and A. Krokhin. Majority constraints have bounded pathwidth duality. *European Journal of Combinatorics*, 29(4):821–837, 2008.
10. V. Dalmau and B. Larose. Maltsev + Datalog  $\Rightarrow$  Symmetric Datalog. In *Proceedings of the 23rd IEEE Symposium on Logic in Computer Science*, LICS '08, pages 297–306, 2008.
11. M. E. Dyer and D. Richerby. On the complexity of  $\#\text{CSP}$ . In *Proceedings of the 42nd ACM Symposium on Theory of Computing*, STOC '10, pages 725–734, 2010.
12. L. Egri. On CSPs below P and the  $\text{NL} \neq \text{P}$  conjecture. *Submitted*, 2011.
13. L. Egri, A. A. Krokhin, B. Larose, and P. Tesson. The complexity of the list homomorphism problem for graphs. In *Proceedings of the 27th International Symposium on Theoretical Aspects of Computer Science*, STACS '10, pages 335–346, 2010.
14. L. Egri, B. Larose, and P. Tesson. Symmetric Datalog and constraint satisfaction problems in logspace. In *Proceedings of the 22nd Annual IEEE Symposium on Logic in Computer Science*, LICS '07, pages 193–202, 2007.
15. J. Hagemann and A. Mitschke. On  $n$ -permutable congruences. *Algebra Universalis*, 3:8–12, 1973.
16. P. Hell and J. Nešetřil. On the complexity of  $H$ -coloring. *Journal of Combinatorial Theory, Series B*, 48:92–110, 1990.
17. P. Hell and A. Rafiey. The dichotomy of list homomorphisms for digraphs. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms*, SODA '11, pages 1703–1713, 2011.
18. P. Jeavons, D. Cohen, and M. Gyssens. Closure properties of constraints. *J. ACM*, 44(4):527–548, 1997.
19. A. Kazda. Maltsev digraphs have a majority polymorphism. *European Journal of Combinatorics*, 32:390–397, 2011.
20. M. Maróti and R. McKenzie. Existence theorems for weakly symmetric operations. *Algebra Universalis*, 59(3-4):463–489, 2008.