SIAM J. COMPUT.
Vol. 38, No. 5, pp. 1782–1802
© 2009 Society for Industrial and Applied Mathematics

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

THE CSP DICHOTOMY HOLDS FOR DIGRAPHS WITH NO SOURCES AND NO SINKS (A POSITIVE ANSWER TO A CONJECTURE OF BANG-JENSEN AND HELL)*

LIBOR BARTO†, MARCIN KOZIK‡, AND TODD NIVEN†

Abstract. Bang-Jensen and Hell conjectured in 1990 (using the language of graph homomorphisms) a constraint satisfaction problem (CSP) dichotomy for digraphs with no sources or sinks. The conjecture states that the CSP for such a digraph is tractable if each component of its core is a cycle and is NP-complete otherwise. In this paper we prove this conjecture and, as a consequence, a conjecture of Bang-Jensen, Hell, and MacGillivray from 1995 classifying hereditarily hard digraphs. Further, we show that the CSP dichotomy for digraphs with no sources or sinks agrees with the algebraic characterization conjectured by Bulatov, Jeavons, and Krokhin in 2005.

Key words. constraint satisfaction problem, graph homomorphism, smooth digraphs

AMS subject classifications. 68R10, 08A70

DOI. 10.1137/070708093

1. Introduction.

The history of the constraint satisfaction problem (CSP) goes back more than thirty years and begins with the work of Montanari [Mon74] and Mackworth [Mac77]. Since that time many combinatorial problems in artificial intelligence and other areas of computer science have been formulated in the language of CSPs. The study of such problems, under this common framework, has applications in database theory [Var00], machine vision recognition [Mon74], temporal and spatial reasoning [SV98], truth maintenance [DD96], technical design [NL], scheduling [LALW98], natural language comprehension [All94], and programming language comprehension [Nad]. Numerous attempts to understand the structure of different CSPs has been undertaken, and a wide variety of tools ranging from statistical physics (e.g., [ANP05, KMRT$^+$07]) to universal algebra (e.g., [JCG97]) has been employed. Methods and results developed in seemingly disconnected branches of mathematics transformed the area. The conjecture proved in this paper resisted the approaches based in combinatorics and theoretical computer science for nearly twenty years. Only recent developments in the structural theory of finite algebras provided tools strong enough to solve this problem.

For the last ten years the study of CSPs has also been a driving force in theoretical computer science. The dichotomy conjecture of Feder and Vardi, published in [FV99], has origins going back to 1993. The conjecture states that a CSP, for any fixed language, is solvable in polynomial time or NP-complete. Therefore the class of CSPs would be a subclass of NP avoiding problems of intermediate difficulty. The logical

*Received by the editors November 14, 2007; accepted for publication (in revised form) August 1, 2008; published electronically January 9, 2009. A part of this article appeared, in a preliminary form, in the Proceedings of the 40th ACM Symposium on Theory of Computing, STOC'08. This work was partly supported by the Eduard Cech Center grant LC505.

http://www.siam.org/journals/sicomp/38-5/70809.html

†Eduard Čech Center, Charles University, 186 75 Praha, Czech Republic (barto@karlin.mff.cuni.cz, niven@karlin.mff.cuni.cz). The first author was supported by the Grant Agency of the Czech Republic under grant 201/06/0664 and by the project of Ministry of Education under grant MSM 0021620839.

‡Eduard Čech Center, Charles University, 186 75 Praha, Czech Republic, and Algorithmics Research Group, Jagiellonian University, 30-072 Kraków, Poland (marcin.kozik@tcs.uj.edu.pl).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

characterization of the class of CSPs (see [FV99] and [Kun]) provides arguments in support of the dichotomy; nevertheless the conjecture remains open.

One of the results of [FV99] shows that the CSP dichotomy conjecture is equivalent to the CSP dichotomy conjecture restricted to digraphs. Therefore the CSPs can be defined in terms of the (di)graph homomorphisms studied in graph theory for over forty years (cf. [Sab61, HP64, Lev73]). It adds a new dimension to a well-established problem and shows the importance of solving CSPs for digraphs. The classification of the complexity of the **H**-coloring problems for undirected graphs, discovered by Hell and Nešetřil [HN90], is an important step and provides a starting point towards proving, or refuting, the CSP dichotomy conjecture. There have since appeared many papers on the complexity of digraph coloring problems (see, e.g., [BJH90, BJHM95, Fed01, GWW92, HNZ96a, HNZ96b, HNZ96c, HZZ93, Mac91, Zhu95]), but as yet, no plausible conjecture on a graph theoretical classification has been proposed. Bang-Jensen and Hell [BJH90] did, however, conjecture a classification (implying the dichotomy) for the class of digraphs with no sources or sinks. Their conjecture significantly generalizes the result of Hell and Nešetřil.

In 1995, Bang-Jensen, Hell, and MacGillivray (in [BJHM95]) introduced the notion of hereditarily hard digraphs and conjectured their classification. Surprisingly, they were able to show that this conjecture and the one given in [BJH90] are equivalent. In this paper we prove the conjecture of Bang-Jensen and Hell and, as a consequence, the conjecture of Bang-Jensen, Hell, and MacGillivray.

Our paper relies on the interconnection between the CSP and algebra as first discovered by Jeavons, Cohen, and Gyssens in [JCG97] and refined by Bulatov, Jeavons, and Krokhin in [BJK05]. Using this connection, Bulatov, Jeavons, and Krokhin conjectured a full classification of the NP-complete CSPs. For a small taste of results in the direction of proving this classification, see [BIM$^+$06, Bul06, Dal05, Dal06, KV07]. A particularly interesting example, demonstrating the potency of the algebraic approach, is Bulatov's proof of the result of Hell and Nešetřil (see [Bul05]). A recent, purely algebraic result of Maróti and McKenzie [MM07] is one of the key ingredients in the proof of the conjecture of Bang-Jensen and Hell. This provides further evidence supporting the extremely strong bond between the CSP and universal algebra.

2. Preliminaries. We assume that the reader possesses a basic knowledge of universal algebra and graph theory. For an easy introduction to the notions of universal algebra that are not defined in this paper, we invite the reader to consult the monographs [BS81] and [MMT87]. Further information concerning the structural theory of finite algebras (called tame congruence theory) can be found in [HM88]. For an explanation of the basic terms in graph theory and graph homomorphisms, we recommend [HN04]. Finally, for an introduction to the connections between universal algebra and the CSP, we recommend [BJK05].

Throughout the paper we deviate from the standard definition of the CSP, with respect to a fixed language (found in, e.g., [BKJ00]), in favor of an equivalent definition from [FV99, LZ06]. The definitions of a *relational structure*, a *homomorphism*, or a *polymorphism* are presented further in this section in their full generality as well as in restriction to directed graphs.

A *directed graph* (or *digraph*) is a pair $\mathbf{G} = (V, E)$, where V is a set of vertices and $E \subseteq V \times V$ is a set of edges. More generally *a relational structure* $\mathcal{T} = (T, \mathcal{R})$ is an ordered pair, where T is a finite nonempty set and $\mathcal{R}$ is a finite set of finitary relations on T indexed by a set J. Let d_j denote the arity of the relation $R_j \in \mathcal{R}$. The indexed set of all the d_j constitutes the *signature* of $\mathcal{T}$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

A vertex of a digraph is called a *source* (resp., a *sink*) if it has no incoming (resp., outgoing) edges. An *oriented walk* is a sequence of vertices $(v_0, \dots, v_{n-1})$ such that $(v_i, v_{i+1}) \in E$ or $(v_{i+1}, v_i) \in E$ for any $i < n-1$ and the *length* of such a walk is $n-1$. A *walk* is an oriented walk such that $(v_i, v_{i+1}) \in E$ for any $i < n-1$. A *closed walk* is a walk such that $v_0 = v_{n-1}$. Given a digraph $\mathbf{G}$, we sometimes denote the set of vertices of $\mathbf{G}$ by $V(\mathbf{G})$ and similarly the edges of $\mathbf{G}$ by $E(\mathbf{G})$. A graph with n vertices is a *cycle* if its vertices can be ordered (i.e., $V = \{v_0, \dots, v_{n-1}\}$) in such a way that $E = \{(v_i, v_j) | \, j = i+1 \mod n\}$.

A *graph homomorphism* is a function between sets of vertices of two graphs mapping edges to edges. A graph is 3-colorable if and only if it maps homomorphically to the complete graph on three vertices (without loops). The notion of *colorability* is generalized using graph homomorphisms: a digraph, say $\mathbf{G}$, is $\mathbf{H}$-colorable if there exists a homomorphism mapping $\mathbf{G}$ to $\mathbf{H}$. For two relational structures of the same signature, say $\mathcal{T} = (T, \mathcal{R})$ and $\mathcal{U} = (U, \mathcal{S})$, a map $h : T \to U$ is a *homomorphism* if $h(T_j) \subseteq R_j$ for all $j \in J$ (where $h(T_j)$ is computed pointwise).

A *digraph polymorphism* is a homomorphism from a finite Cartesian power of a graph to the graph itself. Precisely, for a digraph $\mathbf{G} = (V, E)$ a function $h : V^n \to V$ is a *polymorphism* of $\mathbf{G}$ if, for any vertices $a_0, \dots, a_{n-1}, b_0, \dots, b_{n-1} \in V$,

$$\text{if } (a_i, b_i) \in E \text{ for all } i < n, \text{ then } (h(a_0, \dots, a_{n-1}), h(b_0, \dots, b_{n-1})) \in E.$$

The notion of a polymorphism is defined for relational structures as well. A *polymorphism* h of $\mathcal{T}$ is an operation $h : T^n \to T$ such that, for all relations $R \in \mathcal{R}$ of arity m, if

$$(a_{i,0}, a_{i,1}, \dots, a_{i,m-1}) \in R \qquad \text{for all } i < n,$$

then

$$(h(a_{0,0}, a_{1,0}, \dots, a_{n-1,0}), \dots, h(a_{0,m-1}, a_{1,m-1}, \dots, a_{n-1,m-1})) \in R.$$

A digraph $\mathbf{G} = (V, E)$ *retracts* to an induced subgraph $\mathbf{H} = (W, F)$ if there is an endomorphism $h : V \to V$ such that $h(V) = W$ and $h(a) = a$ for all $a \in W$. Such a map h is called a *retraction*. A *core* of a digraph is a minimal induced subgraph to which the digraph retracts. The definition of retraction and core clearly generalize to relational structures. It is a trivial fact that, for any digraph $\mathbf{H}$, and for a core of $\mathbf{H}$, say $\mathbf{H}'$, the set of $\mathbf{H}$-colorable digraphs coincides with the set of $\mathbf{H}'$-colorable digraphs.

An *algebra* is a tuple $\mathbf{A} = (A, f_0, \dots)$ consisting of a nonempty set A (called a *universe* of $\mathbf{A}$) and *operations* on A. An *operation* f_i is an n_i-ary function $f_i : A^{n_i} \to A$. With each operation we associate an *operation symbol* and, by an abuse of notation, denote it also by f_i. A set $B \subseteq A$ is a *subuniverse* of an algebra $\mathbf{A}$ if, for any number i, the operation f_i restricted to B^{n_i} has all the results in B. For a nonempty subuniverse B of an algebra $\mathbf{A}$ the algebra $\mathbf{B} = (B, f'_0, \dots)$ (where f'_i is a restriction of f_i to B^{n_i}) is a *subalgebra* of $\mathbf{A}$. A *power* of an algebra $\mathbf{A}$ has a universe A^k and the operations f''_i derived from the operations of $\mathbf{A}$ by computing coordinatewise. A subalgebra of a power of an algebra is often called a *subpower*. A *term function* of an algebra is any function that can be obtained by a composition using the operations of the algebra together with all the projections. A *term* is a formal way of denoting such a composition; i.e., a term function is attached to an algebra, but a term can be computed in a subalgebra or a power as well as in the original algebra. A set $C \subseteq A$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

generates a subuniverse B in an algebra $\mathbf{A}$ if B is the smallest subuniverse containing C—such a subuniverse always exists and can be obtained by applying all the term functions of the algebra $\mathbf{A}$ to all the choices of arguments coming from C.

In this paper all relational structures, digraphs, and algebras are assumed to be finite.

3. The main result. For a relational structure $\mathcal{T} = (T, R)$ we define the language $\mathrm{CSP}(\mathcal{T})$, of relational structures with the same signature as $\mathcal{T}$, to be

$$\mathrm{CSP}(\mathcal{T}) = \{\, \mathcal{U} \mid \text{there is a homomorphism from } \mathcal{U} \text{ to } \mathcal{T}\, \}.$$

Alternatively we can view $\mathrm{CSP}(\mathcal{T})$ as a decision problem:

INPUT: a relational structure $\mathcal{U}$ with the same signature as $\mathcal{T}$
QUESTION: does there exist a homomorphism from $\mathcal{U}$ to $\mathcal{T}$?

In either approach we are concerned with the computational complexity (of membership of the language, or of the decision problem, respectively) for a given relational structure. The CSP dichotomy conjecture proposed in [FV99] can be stated as follows.

THE CSP DICHOTOMY CONJECTURE. *For a relational structure $\mathcal{T}$ the problem $\mathrm{CSP}(\mathcal{T})$ is solvable in polynomial time or NP-complete.*

The (di)graph coloring problems can be viewed as special cases of the CSP. Although a digraph $\mathbf{H} = (W, F)$ is technically different from a relational structure, the set of $\mathbf{H}$-colorable digraphs is obviously polynomially equivalent to the CSP for an appropriate relational structure, and therefore we denote the class of all $\mathbf{H}$-colorable digraphs by $\mathrm{CSP}(\mathbf{H})$. Due to the reduction presented in [FV99], every CSP is polynomially equivalent to a digraph homomorphism problem. Thus we can restate the CSP dichotomy conjecture in the following way.

THE CSP DICHOTOMY CONJECTURE. *For a fixed digraph $\mathbf{H}$, deciding whether a given digraph is $\mathbf{H}$-colorable is either NP-complete or solvable in polynomial time.*

This brings us to the main problem of the paper, a conjecture nearly ten years older than the CSP dichotomy conjecture, and a special case of it. It deals with digraphs with no sources or sinks and was first formulated by Bang-Jensen and Hell in [BJH90].

THE CONJECTURE OF BANG-JENSEN AND HELL. *Let $\mathbf{H}$ be a digraph without sources or sinks. If each component of the core of $\mathbf{H}$ is a cycle, then $\mathrm{CSP}(\mathbf{H})$ is polynomially decidable. Otherwise $\mathrm{CSP}(\mathbf{H})$ is NP-complete.*

Note that the above conjecture is a substantial generalization of the $\mathbf{H}$-coloring result of Hell and Nešetřil [HN90].

The notion of hereditarily hard digraphs was introduced by Bang-Jensen, Hell, and MacGillivray in [BJHM95]. A digraph $\mathbf{H}$ is said to be *hereditarily hard* if the $\mathbf{H}'$-coloring problem is NP-complete for all loopless digraphs $\mathbf{H}'$ that contain $\mathbf{H}$ as a subgraph (not necessarily induced). The following conjecture was posed and shown to be equivalent to the Bang-Jensen and Hell conjecture in [BJHM95].

THE CONJECTURE OF BANG-JENSEN, HELL, AND MACGILLIVRAY. *Let $\mathbf{H}$ be a digraph. If the digraph $R(\mathbf{H})$ (which is obtained by iteratively removing the sources and sinks from $\mathbf{H}$ until none remain) does not admit a homomorphism to a cycle of length greater than one, then $\mathbf{H}$ is hereditarily hard. Otherwise there exists a loopless digraph $\mathbf{H}'$ containing $\mathbf{H}$ (as a not necessarily induced subgraph) such that $\mathbf{H}'$-coloring is solvable in a polynomial time.*

In this section we prove the Bang-Jensen and Hell conjecture and therefore the conjecture of Bang-Jensen, Hell, and MacGillivray. In this proof we assume Theorem 3.1, which will be proved in the subsequent sections of the paper. The reasoning

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

uses weak near unanimity operations[1] and Taylor operations (used only to connect Theorems 3.2 and 3.3, and therefore not defined here [HM88, Tay77, LZ06]).

THEOREM 3.1. *If a digraph without sources or sinks admits a weak near unanimity polymorphism, then it retracts to the disjoint union of cycles.*

It is easy to see that the colorability by digraphs retracting to a disjoint union of cycles is tractable (see, e.g., [BJH90]). It remains to prove the NP-completeness of the digraphs not retracting to such a union. Before we do so, we recall two fundamental results.

It follows from [HM88, Lemma 9.4 and Theorem 9.6] that a part of the result of Maróti and McKenzie [MM07, Theorem 1.1] can be stated as follows.

THEOREM 3.2 (see [MM07]). *A finite relational structure $\mathcal{T}$ admits a Taylor polymorphism if and only if it admits a weak near unanimity polymorphism.*

The following result was originally proved in [BKJ00] and [LZ03] and, as stated below, can be found in [LZ06, Theorem 2.3]. It relies on a connection between relational structures and varieties generated by algebras of their polymorphisms. A lack of a Taylor polymorphism in such an algebra implies an existence of a "trivial" algebra in a variety and NP-completeness of the associated CSP.

THEOREM 3.3 (see [LZ06]). *Let $\mathcal{T}$ be a relational structure which is a core. If $\mathcal{T}$ does not admit a Taylor polymorphism, then* $\mathrm{CSP}(\mathcal{T})$ *is NP-complete.*

If a digraph $\mathbf{H}$ without sources or sinks does not retract to a disjoint union of cycles, then its core $\mathbf{H}'$ also does not. Thus, by Theorem 3.1, it follows that $\mathbf{H}'$ does not admit a weak near unanimity polymorphism, and by Theorems 3.2 and 3.3 it follows that $\mathrm{CSP}(\mathbf{H}')$ is NP-complete, completing the proof of the conjecture of Bang-Jensen and Hell.

The conjecture (posed in [BKJ00]), classifying the CSPs from the algebraic point of view, can be stated as follows (see, e.g., [LZ06]).

THE ALGEBRAIC CSP DICHOTOMY CONJECTURE. *Let $\mathcal{T}$ be a relational structure that is a core. If $\mathcal{T}$ admits a Taylor polymorphism, then* $\mathrm{CSP}(\mathcal{T})$ *is polynomial time solvable. Otherwise* $\mathrm{CSP}(\mathcal{T})$ *is NP-complete.*

Note that the proof of the conjecture of Bang-Jensen and Hell immediately implies that the structure of the NP-complete digraph coloring problems agrees with the algebraic CSP dichotomy conjecture. The remainder of the paper is dedicated to the proof of the Theorem 3.1.

4. Notation.

In this section we introduce the notation required throughout the remainder of the paper.

4.1. Neighborhoods in graphs.

For a fixed digraph $\mathbf{G} = (V, E)$ we denote $(a, b) \in E$ by $a \to b$, and we use $a \xrightarrow{k} b$ to say that there is a directed walk from a to b of length precisely k. More generally we call a digraph $\mathbf{H}$ a *pattern* if $V(\mathbf{H}) = \{0, \dots, n-1\}$ and $(u, v) \in E$ if and only if $|u - v| = 1$ and $(v, u) \notin E$. We denote patterns by lowercase Greek letters and, for a pattern α, we write $a \xrightarrow{\alpha} b$ if there exists a homomorphism ϕ from α into $\mathbf{G}$ such that $\phi(0) = a$ and $\phi(n-1) = b$. In such a case we say that a and b can be connected via the pattern α. The oriented walk connecting vertices a and b and consisting of the images of elements of α under ϕ is a *realization* of the pattern. For any $W \subseteq V$ we define

$$W^{+n} = \{v \in V \mid (\exists w \in W)\ w \xrightarrow{n} v\}$$

[1] A weak near unanimity operation is a function such that, for any choice of arguments a, b, $w(b, a, \dots, a) = w(a, b, \dots, a) = \cdots = w(a, a, \dots, b)$ and $w(a, \dots, a) = a$. These operations are described in more detail in section 4.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

and similarly

$$W^{-n} = \{v \in V \mid (\exists w \in W)\ v \xrightarrow{n} w\}.$$

We define $W^0 = W$, and write a^{+n} (resp., a^{-n}, a^0) instead of $\{a\}^{+n}$ (resp., $\{a\}^{-n}, \{a\}^0$) for any $a \in V$. More generally, for a pattern α, we write

$$W^{\alpha} = \{v \in V \mid (\exists w \in W)\ w \xrightarrow{\alpha} v\}.$$

As before, we use a^{α} for $\{a\}^{\alpha}$. Sometimes, for ease of presentation, we write $a \xrightarrow{k,n} b$ to denote $a \xrightarrow{k} b$ and $a \xrightarrow{n} b$.

4.2. Digraph path powers. Let $\mathbf{G} = (V, E)$ be a digraph and α be a pattern. We define a path power of the digraph $\mathbf{G}$, which we denote by $\mathbf{G}^{\alpha}$, in the following way: the vertices of the power are the vertices of the digraph $\mathbf{G}$, and a pair $(c, d) \in V^2$ is an edge in $\mathbf{G}^{\alpha}$ if and only if $c \xrightarrow{\alpha} d$ in $\mathbf{G}$. Moreover, we set $\mathbf{G}^{+n} = \mathbf{G}^{\alpha}$ for the pattern α consisting of n arrows pointing forward. Note that if $f : V^m \to V$ is a polymorphism of $\mathbf{G}$, then it is also a polymorphism of any path power of this digraph. Path powers are special cases of primitive positive definitions (used in, e.g., [Bul05]) or indicator constructions introduced in [HN90] in order to deal with the colorability problem for undirected graphs.

4.3. Components. A *connected* digraph is a digraph such that there exists an oriented walk, consisting of at least one edge, between every choice of two vertices. A *strongly connected* digraph is a digraph such that, for every choice of two vertices, there is a walk connecting them. By a *component* (resp., *strong component*) of a digraph $\mathbf{G}$, we mean a maximal (under inclusion) induced subgraph that is connected (resp., strongly connected). Note that, according to this definition, a single vertex with the empty set of edges is not connected, and thus not every digraph decomposes into a union of components (or strong components). Given a digraph $\mathbf{G}$ with no sources or sinks, we say that a strong component $\mathbf{H}$ of $\mathbf{G}$ is a *top component* if $V(\mathbf{H})^{+1} = V(\mathbf{H})$. Similarly, we say that a strong component $\mathbf{H}$ of $\mathbf{G}$ is a *bottom component* if $V(\mathbf{H})^{-1} = V(\mathbf{H})$.

4.4. Algebraic length. The following definition is taken from [HNZ96b]. For a pattern α we define the *algebraic length* $al(\alpha)$ to be

$$al(\alpha) = |\{\text{edges going forward in } \alpha\}| - |\{\text{edges going backward in } \alpha\}|.$$

An *algebraic length of an oriented walk* is a shorthand expression for an algebraic length of a pattern which can be realized as such an oriented walk—the pattern is always clear from the context. For a digraph $\mathbf{G} = (V, E)$ we set

$$al(\mathbf{G}) = \min\{i > 0 \mid (\exists v \in V)\ (\exists \text{ a pattern } \alpha)\ v \xrightarrow{\alpha} v \text{ and } al(\alpha) = i\}$$

whenever the set on the right-hand side is nonempty and ∞ otherwise. In case of strongly connected digraphs in section 7 the algebraic length can be equivalently defined (cf. Corollary 5.7) as the greatest common divisor of the lengths of closed walks in a digraph. We note that for digraphs with no sources or sinks (or with a closed walk) the algebraic length of a nonempty digraph is always a natural number. It is folklore (cf. [HN04, Proposition 5.19]) that a connected digraph $\mathbf{G}$ retracts to a cycle if and only if it contains a closed walk of length $al(\mathbf{G})$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

4.5. Algebraic notation. By $\overline{a}$ we denote the tuple $(a, a, \ldots, a)$ (the arity will always be clear by the context), and by $\overrightarrow{a}$ we denote the tuple $(a_0, a_1, \ldots, a_n)$. Further, we extend the notation $\overline{a}$ to the sets in the following way. For a set W let $\overline{W}$ be an appropriate Cartesian power of W. Thus, for example, given a vertex a of a digraph $\mathbf{G}$, the set $\overline{a^{+n}}$ is the collection of all tuples whose coordinates are vertices reachable by a walk of length n from a.

An idempotent operation on a set A is an operation, say $f : A^n \to A$, such that $f(\overline{a}) = a$ for all $a \in A$. In accordance with [MM07], by a weak near unanimity operation we understand an idempotent operation $w(x_0, \ldots, x_{n-1})$ that satisfies

$$w(y, x, \ldots, x) = w(x, y, \ldots, x) = \cdots = w(x, x, \ldots, y),$$

for any choice of x and y in the underlying set. Moreover, for a term t of arity n, we define

$$t^{(i)}(x_0, x_1, \ldots, x_{n-1}) = t(x_{n-i}, x_{n-i+1}, \ldots, x_0, x_1, \ldots, x_{n-i-1}),$$

for each $0 \le i < n$, where addition on the indices is performed modulo n.

5. Preliminary results on digraphs. We start with a number of basic results describing the connection between digraphs and their path powers. The following lemma reveals the behavior of the algebraic lengths of oriented walks in powers of a digraph.

LEMMA 5.1. *Let $\mathbf{G}$ be a digraph without sources or sinks. Let α be a pattern of algebraic length k, and let $a \xrightarrow{\alpha} b$ in $\mathbf{G}$. Then $a \xrightarrow{\beta} b$ in $\mathbf{G}^{+k}$ for some pattern β of algebraic length one.*

Proof. For a fixed, large enough number j, consider all oriented walks in $\mathbf{G}$ of the form $a \xrightarrow{l_1} a_1 \xleftarrow{l_2} a_2 \xrightarrow{l_3} \cdots \leftrightarrows a_{l_j} = b$, where $l_1 - l_2 + \cdots \pm l_j = k$. We will show that at least one of these walks has all the l_i's divisible by k. Let us choose an oriented walk in which k divides all the l_i in a maximal initial segment of the i, and let l_{i_0} be the last element of this segment. If $i_0 + 1 < j$, then (assuming without loss of generality that i_0 is odd) the walk

$$a \xrightarrow{l_1} \cdots \xrightarrow{l_{i_0}} a_{i_0} \xleftarrow{l_{i_0+1}} a_{i_0+1} \xrightarrow{l_{i_0+2}} a_{i_0+2} \cdots$$

can be altered, using the fact that a_{i_0+1} (and possibly other vertices) is not a source, to obtain

$$a \xrightarrow{l_1} \cdots \xrightarrow{l_{i_0}} a_{i_0} \xleftarrow{l'_{i_0+1}} a'_{i_0+1} \xrightarrow{l'_{i_0+2}} a_{i_0+2} \cdots ,$$

where l'_{i_0+1} is greater than l_{i_0+1} and is divisible by k. This contradicts the choice of i_0.

If, on the other hand, $i_0 + 1 = j$, the number k divides $l_1 - l_2 + \cdots \pm l_{i_0}$ and, using the fact that $l_1 - l_2 + \cdots \pm l_{i_0} \mp l_{i_0+1} = k$, we infer that k divides l_{i_0+1}, again contradicting the choice of i_0. Thus $i_0 = j$ and we can find an oriented walk $a \xrightarrow{l_1} a_1 \xleftarrow{l_2} a_2 \xrightarrow{l_3} \cdots a_{l_j} = b$ with $l_1 - l_2 + \cdots \pm l_j = k$, where each l_i is divisible by k. This shows that a is connected to b via a pattern of algebraic length one in $\mathbf{G}^{+k}$. ∎

As a consequence we obtain the following fact.

COROLLARY 5.2. *Let $\mathbf{G}$ be a digraph, without sources or sinks, such that $al(\mathbf{G}) = 1$. Then $al(\mathbf{G}^{+k}) = 1$ for any natural number k.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Proof. Let $a \xrightarrow{\alpha} a$, where α is a pattern of algebraic length one. Then, by following a realization of α k-many times, we obtain $a \xrightarrow{\beta} a$ in $\mathbf{G}$ for a pattern β of algebraic length k. Now the statement follows from the previous lemma. ∎

Theorem 3.1 is proved in section 7 for strongly connected digraphs first, and therefore we need some preliminary results on such digraphs. The following very simple lemma is needed to prove some of the further corollaries in this section.

LEMMA 5.3. *Let c be a vertex in a strongly connected digraph. Then the greatest common divisor* (*GCD*) *of the lengths of the closed walks in this digraph is equal to the GCD of the lengths of the closed walks containing c.*

Proof. Suppose, for contradiction, that the GCD, say n', of the lengths of the closed walks containing c is bigger than the GCD of the lengths of the closed walks for the entire digraph. Then there exists a walk $d \xrightarrow{l} d$ of length l such that n' does not divide l. On the other hand, since the digraph is strongly connected, $c \xrightarrow{l'} d$ and $d \xrightarrow{l''} c$ for some numbers l', l''. The number n', by definition, divides $l' + l''$ and $l' + l + l''$ and thus divides l, a contradiction. ∎

Moreover, the following easy proposition holds.

PROPOSITION 5.4. *Let $\mathbf{G}$ be a connected digraph $\mathbf{G}$ and α be a pattern. If $a \xrightarrow{\alpha} a$ for a vertex a in $\mathbf{G}$, then the number $al(\mathbf{G})$ divides $al(\alpha)$.*

Proof. Let $\mathbf{G}$ be a connected digraph and, for some vertex a, $a \xrightarrow{\alpha} a$ via a pattern α. Let b be a vertex in $\mathbf{G}$ such that $b \xrightarrow{\beta} b$ for a pattern β satisfying $al(\beta) = al(\mathbf{G})$. Since $\mathbf{G}$ is connected there is a pattern γ such that $b \xrightarrow{\gamma} a$ and thus $b \xrightarrow{\gamma} a \xrightarrow{\alpha} a \xrightarrow{\gamma'} b$ with $al(\gamma') = -al(\gamma)$. Following appropriate walks, we can obtain an oriented walk, from b to b, of algebraic length $al(\alpha) - k \cdot al(\mathbf{G})$, for any number k. The minimality of $al(\mathbf{G})$ implies that $al(\mathbf{G})$ divides $al(\alpha)$. ∎

The following lemma is heavily used in the proof of Theorem 3.1 for strongly connected digraphs in section 7.

LEMMA 5.5. *If, for a strongly connected digraph $\mathbf{G} = (V, E)$, the GCD of the lengths of the closed walks in $\mathbf{G}$ is equal to one, then*

$$(\exists m)\ (\forall a, b \in V)\ (\forall n)\ \text{if } n \geq m,\ \text{then } a \xrightarrow{n} b.$$

Proof. Fix an arbitrary element $c \in V$. By Lemma 5.3 we find some closed walks containing c such that their lengths $k_1, \ldots, k_i$ satisfy $GCD(k_1, \ldots, k_i) = 1$. Thus c is contained in a closed walk of length l whenever l is a linear combination of $k_1, \ldots, k_i$ with nonnegative integer coefficients. It is easy to see that there is a natural number m' such that, for every $n' \geq m'$, n' can be expressed as such a linear combination; hence c is in a closed walk of length n' for each such n'. Now it suffices to set $m = m' + 2|V|$ since, for arbitrary vertices $a, b \in V$, there are walks of length at most $|V|$ from a to c and from c to b. ∎

The following easy corollary follows.

COROLLARY 5.6. *For a strongly connected digraph $\mathbf{G}$ with GCD of the lengths of the closed walks equal to one, and for any number n, the digraph $\mathbf{G}^{+n}$ is strongly connected.*

For strongly connected digraphs, the GCD of the lengths of the closed walks and the algebraic length of the digraph coincide.

COROLLARY 5.7. *For a strongly connected digraph, the GCD of the lengths of the closed walks is equal to the algebraic length of the digraph.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Proof. Let us fix a digraph $\mathbf{G} = (V, E)$ and denote by n the GCD of the lengths of the closed walks in $\mathbf{G}$. Since, by Proposition 5.4, the algebraic length of $\mathbf{G}$ divides the length of every closed walk in $\mathbf{G}$, $al(\mathbf{G})$ divides n.

Conversely, let $a = a_0 \xrightarrow{l_0} b_0 \xleftarrow{k_0} a_1 \xrightarrow{l_1} \cdots \xleftarrow{k_{m-1}} a_m = a$ be a realization of a pattern of algebraic length $al(\mathbf{G})$. Let k_i' be such that $b_i \xleftarrow{k_i} a_{i+1} \xleftarrow{k_i'} b_i$ for all i. Note that n divides $k_i + k_i'$ and $\sum_{i<m} l_i + \sum_{i<m} k_i'$. Thus n divides $\sum_{i<m} l_i - \sum_{i<m} k_i = al(\mathbf{G})$, which shows that $n \leq al(\mathbf{G})$, and the lemma is proved. ∎

Finally, we remark that if α is a pattern of algebraic length one and $\mathbf{G}$ has no sources and no sinks, then $E(\mathbf{G}^\alpha) \supseteq E(\mathbf{G})$. In particular, if $al(\mathbf{G}) = 1$, then $al(\mathbf{G}^\alpha) = 1$.

6. A connection between graphs and algebra. In this section we present basic definitions and results concerning the connection between digraphs and algebras. Let $\mathbf{G} = (V, E)$ be a digraph admitting a weak near unanimity polymorphism $w(x_0, x_1, \dots, x_{h-1})$. We associate with $\mathbf{G}$ an algebra $\mathbf{A} = (V, w)$ and note that E is a subuniverse of $\mathbf{A}^2$. Note that for any subuniverse of $\mathbf{A}$, say W, we can define the digraph $\mathbf{G}_{|W} = (W, E \cap W \times W)$ (or $(W, E_{|W})$) which admits the weak near unanimity polymorphism $w|_{W^h}$, and the algebra $(W, w|_{W^h})$ is a subalgebra of $\mathbf{A}$. For the remainder of this section we assume that $\mathbf{G}$ and $\mathbf{A}$ are as above.

The first lemma describes the influence of the structure of the digraph on the subuniverses of the algebra.

LEMMA 6.1. *For any subuniverse W of $\mathbf{A}$ the sets W^{+1} and W^{-1} are subuniverses of $\mathbf{A}$.*

Proof. Take any elements $a_0, \dots, a_{h-1}$ from W^{+1} and choose $b_0, \dots, b_{h-1} \in W$ such that $b_i \to a_i$ for all i. Then $w(b_0, \dots, b_{h-1}) \to w(a_0, \dots, a_{h-1})$ showing that $w(a_0, \dots, a_{h-1}) \in W^{+1}$, and the claim is proved. The proof for W^{-1} is similar. ∎

Since the weak near unanimity operation is idempotent, all the one element subsets of V are subuniverses of $\mathbf{A}$. Using the previous lemma, the following result follows trivially.

COROLLARY 6.2. *For any $a \in V$, any pattern α, and any number n, the sets a^{+n}, a^{-n}, and a^α are subuniverses of $\mathbf{A}$.*

Subuniverses of $\mathbf{A}$ can also be obtained in another way.

LEMMA 6.3. *Let $\mathbf{H}$ be a strong component of $\mathbf{G}$. Assume that the GCD of the lengths of the cycles in $\mathbf{H}$ is equal to one. Then $V(\mathbf{H})$ is a subuniverse of $\mathbf{A}$.*

Proof. Using Lemma 5.5, we find a number m such that there is a walk $b \xrightarrow{m} c$ in $\mathbf{H}$ for all $b, c \in V(\mathbf{H})$. Fix a vertex $a \in V(\mathbf{H})$. There is a walk $a \xrightarrow{m} b$ for all $b \in V(\mathbf{H})$ and a walk $c \xrightarrow{m} a$ for all $c \in V(\mathbf{H})$. Thus, $V(\mathbf{H}) = a^{+m} \cap a^{-m}$ is a subuniverse. ∎

We present a second construction leading to a subuniverse of the algebra.

LEMMA 6.4. *If $\mathbf{H} = (W, F)$ is the largest induced subgraph of $\mathbf{G}$ without sources or sinks, then W is a subuniverse of $\mathbf{A}$.*

Proof. Clearly, the vertices of $\mathbf{H}$ can be described as those having arbitrarily long walks to and from them. Since $\mathbf{G}$ is finite, there exists a natural number k such that

$$W = \{w \mid (\exists v, v' \in V)\ v \xrightarrow{k} w \ \text{ and } \ w \xrightarrow{k} v'\}.$$

Thus $W = V^{+k} \cap V^{-k}$, and we are done, since both sets on the right-hand side are subuniverses. ∎

7. Strongly connected digraphs. In this section we present a proof Theorem 3.1 in the case of strongly connected digraphs.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

THEOREM 7.1. *If a strongly connected digraph of algebraic length k admits a weak near unanimity polymorphism, then it contains a closed walk of length k (and thus retracts to a cycle of length k).*

Using Corollary 5.7, the result can be restated in terms of the GCD of the lengths of closed walks in $\mathbf{G}$, and we will freely use this duality. Theorem 7.1 is a consequence of the following result.

THEOREM 7.2. *If a strongly connected digraph $\mathbf{G}$ of algebraic length one admits a weak near unanimity polymorphism, then it contains a loop.*

We present a proof of Theorem 7.1, assuming Theorem 7.2, and devote the remainder of this section to proving Theorem 7.2.

Proof of Theorem 7.1. Fix an arbitrary vertex c in a strongly connected digraph of algebraic length k. Using Lemma 5.3 and Corollary 5.7, we obtain closed walks containing c with the GCD of their lengths equal to k. Thus, in the path power $\mathbf{G}^{+k}$, the GCD of lengths of closed walks containing c is equal to one. Let $\mathbf{H}$ be the strong component of $\mathbf{G}^{+k}$ containing c. Using Lemma 6.3, we infer that $V(\mathbf{H})$ is a subuniverse of the algebra $(V(\mathbf{G}^{+k}), w)$, and thus $\mathbf{H}$ admits a weak near unanimity polymorphism. The algebraic length of $\mathbf{H}$ (again by Corollary 5.7) is one, and therefore by Theorem 7.2 it follows that there is a loop in $\mathbf{G}^{+k}$. This trivially implies a closed walk of length k in $\mathbf{G}$, and the theorem is proved using the folklore proposition from section 4.4. ∎

The remaining part of this section is devoted to the proof of Theorem 7.2. We start by choosing a digraph $\mathbf{G} = (V, E)$ to be a minimal (with respect to the number of vertices) counterexample to Theorem 7.2. We fix a weak near unanimity polymorphism $w(x_0, \ldots, x_{h-1})$ of this digraph and associate with it the algebra $\mathbf{A} = (V, w)$. The proof will proceed by a number of claims.

CLAIM 7.3. *The digraph $\mathbf{G}$ can be chosen to contain a closed walk of length* 2.

Proof. Using Lemma 5.5, we find a minimal k such that a closed walk of length 2^k is contained in $\mathbf{G}$. Consider the path power $\mathbf{G}^{+2^{k-1}}$. It contains a closed walk of length 2 and admits a weak near unanimity polymorphism. Moreover, since k was chosen to be minimal and $\mathbf{G}$ did not contain a loop, the path power $\mathbf{G}^{+2^{k-1}}$ does not contain a loop either. By Corollary 5.6 the path power is strongly connected, and by Corollary 5.2 it has algebraic length equal to one. Thus, the digraph $\mathbf{G}^{+2^{k-1}}$ is also a counterexample to Theorem 7.2 (with the same number of vertices as $\mathbf{G}$), and therefore we can use it as a substitute for $\mathbf{G}$. ∎

From this point on we assume that $\mathbf{G}$ contains a closed walk of length 2 (an undirected edge). The next claim allows us to choose and fix an undirected edge with special properties.

CLAIM 7.4. *There are vertices $a, b \in V$ forming an undirected edge in $\mathbf{G}$ and a binary term t of $\mathbf{A}$ such that $a = t(w(\overline{a}, b), w(\overline{b}, a))$.*

Proof. Let $M \subseteq V$ be a minimal (under inclusion) subuniverse of $\mathbf{A}$ containing an undirected edge, and let $a, b \in M$ be vertices in such an edge. Since vertices $w(\overline{a}, b)$, $w(\overline{b}, a) \in M$ form an undirected edge in $\mathbf{G}$, the set $\{w(\overline{a}, b), w(\overline{b}, a)\}$ generates, in the algebraic sense, the set M (by the minimality of M). Since every vertex in a subuniverse is a result of an application of some term function to the generators of the subuniverse, there exists a term t such that $t(w(\overline{a}, b), w(\overline{b}, a)) = a$. ∎

In the following claims we fix vertices a, b and a term $t(x, y)$ such that $a \to b \to a$ and $a = t(w(\overline{a}, b), w(\overline{b}, a))$ (provided by the previous claim). Note that, by the definition of the operation $w(x_0, \ldots, x_{h-1})$, for any numbers $i, j < h$, we obtain $a = t(w^{(i)}(\overline{a}, b), w^{(j)}(\overline{b}, a))$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Using Lemma 5.5, we find and fix a minimal number n such that $a^{+(n+1)} = V$. We put $W = a^{+n}$ and $F = (W \times W) \cap E$ so that $\mathbf{H} = (W, F)$ is an induced subgraph of the digraph $\mathbf{G}$. Using Corollary 6.2, we infer that W is a subuniverse of $\mathbf{A}$ and thus $\mathbf{H}$ admits a weak near unanimity polymorphism. In the following claims we will show that the algebraic length of some strong component of $\mathbf{H}$ is one, which will contradict the minimality of $\mathbf{G}$.

CLAIM 7.5. *For any vertex in W there exists a closed walk in $\mathbf{H}$ and a walk (also in $\mathbf{H}$) connecting the closed walk to this vertex.*

Proof. Let d_0 denote an arbitrary vertex of W. Since $a^{+(n+1)} = W^{+1} = V$ there is $d_1 \in W$ such that $d_1 \to d_0$. Similarly, there exists $d_2 \in W$ such that $d_2 \to d_1$. By repeating this procedure, we get both statements of the claim. ∎

The next claim will allow us to fix some more vertices necessary for further construction.

CLAIM 7.6. *There exist vertices $c, c' \in W$ and a number k such that*
1. $c' \to a$,
2. $c \xrightarrow{k} c$ *in* $\mathbf{H}$, *and*
3. $c \xrightarrow{k-n-1} c'$ *in* $\mathbf{H}$.

Proof. Since $W^{+1} = V$ there exists $c' \in W$ such that $c' \to a$. Let l be the length of a closed walk provided by Claim 7.5 for $c' \in W$. For a sufficiently large multiple k of l there is a walk in $\mathbf{H}$ of length $k - n - 1$ from some vertex of the closed walk to c'; we call this vertex c. This finishes the proof. ∎

From this point on we fix vertices c and c' in W and a number k to satisfy the conditions of the last claim. The following claims focus on uncovering the structure of the strong component containing c in $\mathbf{H}$.

CLAIM 7.7. *For any $m \leq n$ either $a^{+m} \subseteq a^{+n}$ or $a^{+m} \subseteq b^{+n}$.*

Proof. Since a is in a closed walk of length 2, we obviously have $a^{+n} \supseteq a^{+(n-2)} \supseteq a^{+(n-4)} \cdots$, which proves the claim for even m's. If, on the other hand, m is odd, we have $b^{+n} \supseteq a^{+(n-1)} \supseteq a^{+(n-3)} \cdots$, completing the proof. ∎

The next two claims are of major importance for the proof of Theorem 7.2. They are used to show that the algebraic length of the strong component of $\mathbf{H}$ containing c is one.

CLAIM 7.8. *For any $m \leq n$ and for any $0 \leq i, j < h$ the following inclusion holds:*

$$t(w^{(i)}(\overline{a^{+n}}, a^{+m}), w^{(j)}(\overline{a^{+m}}, a^{+n})) \subseteq a^{+n}.$$

Proof. Note that $a = t(w^{(i)}(\overline{a}, b), w^{(j)}(\overline{b}, a))$ and therefore, for any choice of arguments of the term reachable by walks of length n from corresponding arguments of $t(w^{(i)}(\overline{a}, b), w^{(j)}(\overline{b}, a))$, the result is reachable by a walk of the same length from a, i.e.,

$$a^{+n} \supseteq t(w^{(i)}(\overline{a^{+n}}, b^{+n}), w^{(j)}(\overline{b^{+n}}, a^{+n})).$$

By the same token, using $a = t(w^{(i)}(\overline{a}, a), w^{(j)}(\overline{a}, a))$ provided by the idempotency of the terms, we obtain

$$a^{+n} \supseteq t(w^{(i)}(\overline{a^{+n}}, a^{+n}), w^{(j)}(\overline{a^{+n}}, a^{+n})).$$

Now the claim follows directly from Claim 7.7. ∎

The following technical claim will allow us to find walks in the strong component of $\mathbf{H}$ containing c.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

CLAIM 7.9. *The following implication holds in* $\mathbf{H}$ (*i.e., all the walks and vertices lie inside* $\mathbf{H}$). *For any numbers* $0 \le i, j < h$ *and all* $e, e', f \in W$ *and* $\overrightarrow{d}, \overrightarrow{d'}, \overrightarrow{g} \in \overline{W}$,

$$\text{if} \quad \begin{array}{c} e \\ \downarrow^{k} \\ e' \end{array} \quad \text{and} \quad \begin{array}{c} d_l \\ \downarrow^{k} \\ d'_l \end{array} \quad \text{for all } l, \quad \text{then} \quad \begin{array}{c} t(w^{(i)}(\overrightarrow{d}, c), w^{(j)}(\overline{c}, e)) \\ \downarrow^{k} \\ t(w^{(i)}(\overrightarrow{d'}, f), w^{(j)}(\overrightarrow{g}, e')). \end{array}$$

Proof. Note that, by looking at the tuples of vertices pointwise, we can find the following walks in $\mathbf{G}$:

$$\begin{array}{cccc} \overrightarrow{d} & c & \overline{c} & e \\ \Big| & \downarrow^{k-n-1} & \downarrow^{k-n-1} & \Big| \\ \Big| & c' & \overline{c'} & \Big| \\ \Big|^{k} & \downarrow & \downarrow & \Big|^{k} \\ \Big| & a & \overline{a} & \Big| \\ \downarrow & \downarrow^{n} & \downarrow^{n} & \downarrow \\ \overrightarrow{d'} & f & \overrightarrow{g} & e' \end{array}$$

where the walks from c to c' are provided by Claim 7.6 and lie entirely in $\mathbf{H}$. Applying the appropriate term to the consecutive vertices of the walks (rows in the diagram above), we obtain a walk of length k connecting $t(w^{(i)}(\overrightarrow{d}, c), w^{(j)}(\overline{c}, e))$ to $t(w^{(i)}(\overrightarrow{d'}, f), w^{(j)}(\overrightarrow{g}, e'))$. It remains to prove that all the vertices of this walk are in W. The first $k - n - 1$ vertices of the walks are in W, since W is a subuniverse and they are results of an application of a term to vertices of the subuniverse. For $m \ge 0$, the $(k-n+m)$th vertex of the walk is a member of $t(w^{(i)}(\overline{a^{+n}}, a^{+m}), w^{(j)}(\overline{a^{+m}}, a^{+n}))$ and thus in W by Claim 7.8. ∎

We now construct a closed walk in $\mathbf{H}$, that contains c, of length coprime to k.

CLAIM 7.10. *There exists a closed walk* $c \xrightarrow{(h+1)k-1} c$ *in digraph* $\mathbf{H}$.

Proof. In the proof of this claim we use only vertices and walks that lie inside $\mathbf{H}$. Fix $d \in W$ (provided by Claim 7.6) such that $c \to d \xrightarrow{k-1} c$ in $\mathbf{H}$. By repeatedly applying Claim 7.9 we obtain

$$\begin{array}{ccc} t(w(c,\ldots,c,c,c), w(c,c\ldots,c)) & & \\ \downarrow^{k} & & \\ t(w(c,\ldots,c,c,d), w(d,c,\ldots,c)) & = & t(w^{(1)}(c,\ldots,c,d,c), w^{(1)}(c,\ldots,c,d)) \\ & & \downarrow^{k} \\ t(w^{(2)}(c,\ldots,d,d,c), w^{(1)}(c,\ldots,c,d)) & = & t(w^{(1)}(c,\ldots,c,d,d), w^{(1)}(c,\ldots,c,d)) \\ \vdots & & \\ \downarrow & = & t(w^{(h-1)}(d,\ldots,d,d,c), w^{(1)}(c,\ldots,c,d)) \\ & & \downarrow^{k} \\ & & t(w^{(h-1)}(d,\ldots,d,d,d), w^{(1)}(d,\ldots,d,d)) \end{array}$$

and since the algebra is idempotent, the starting point of this walk is c and the ending point is d. Thus $c \xrightarrow{hk} d$ (for h the arity of the operation $w(x_0, \ldots, x_{h-1})$), which immediately gives us the claim. ∎

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

By Claims 7.6 and 7.10, the strong component of $\mathbf{H}$ containing c has GCD of the lengths of its closed walks equal to one, and thus, by Lemma 6.3, its vertex set forms a subuniverse of the algebra $\mathbf{A}$. As a digraph it admits a weak near unanimity polymorphism. By Corollary 5.7 it has algebraic length one, and (as an induced subgraph of $\mathbf{G}$) it has no loops. Since $\mathbf{H}$ was chosen to be strictly smaller than $\mathbf{G}$ we obtain a contradiction with the minimality of $\mathbf{G}$, and the proof of Theorem 7.2 is complete.

8. The general case. In this section we prove Theorem 3.1 in its full generality. Nevertheless the majority of this section is devoted to the proof of the following result.

THEOREM 8.1. *If a digraph with no sources or sinks has algebraic length one and admits a weak near unanimity polymorphism, then it contains a loop.*

Using the above result, we prove the core theorem of the paper, Theorem 3.1.

Proof of Theorem 3.1. Let $\mathbf{G}$ be a digraph with no sources or sinks which admits a weak near unanimity polymorphism. Let n be the algebraic length of some component of $\mathbf{G}$. The path power $\mathbf{G}^{+n}$ admits a weak near unanimity polymorphism, has no sources or sinks, and, by Lemma 5.1, has algebraic length equal to one. Thus, Theorem 8.1 applied to $\mathbf{G}^{+n}$ provides a loop in the path power and therefore a closed walk of length n in $\mathbf{G}$.

Let n be minimal, under divisibility, in the set of algebraic lengths of components of $\mathbf{G}$. Since the algebraic length of a component divides (by Proposition 5.4) the length of any closed walk in it, every closed walk of length n (for such a minimal n) forms a subgraph which is a cycle. Moreover, by the same reasoning, cycles obtained for two different minimal n's cannot belong to the same component. Thus each component of $\mathbf{G}$ maps homomorphically to an n-cycle (for any minimal n dividing the algebraic length of this component), and it is not difficult to see that these homomorphisms can be chosen so that their union is a retraction. This proves the theorem. ∎

Therefore the only missing piece of the proof to the conjecture of Bang-Jensen and Hell is Theorem 8.1. We prove this result by way of contradiction. Suppose that $\mathbf{G} = (V, E)$ is a minimal (with respect to the number of vertices) counterexample to Theorem 8.1, and let $\mathbf{A} = (V, w(x_0, \dots, x_{h-1}))$ be the algebra associated with $\mathbf{G}$, in the sense of section 6, for some weak near unanimity polymorphism $w(x_0, \dots, x_{h-1})$.

The first part of the proof is dedicated to finding a particular counterexample satisfying more restrictive conditions than $\mathbf{G}$. To do so we need to define a special family of digraphs called *tambourines*. The n-tambourine is the digraph $(\{d_0, \dots, d_{n-1}, u_0, \dots, u_{n-1}\}, F_n)$ such that

$$F_n = \bigcup_i \{(d_i, d_{i+1}), (d_i, u_i), (d_i, u_{i+1}), (u_i, u_{i+1})\},$$

where the addition on the indices is computed modulo n. The 12-tambourine can be found in Figure 1. We begin the proof of the theorem with the following claim.

CLAIM 8.2. *We can choose a digraph* $\mathbf{G}$ *and a number* n *such that*

1. *the* n*-tambourine maps homomorphically to* $\mathbf{G}$,
2. *every vertex of* $\mathbf{G}$ *is in a closed walk of length* n, *and*
3. $\mathbf{G}^{+(mn+1)} = \mathbf{G}$ *for any number* m.

To prove this claim, we begin with an easy subclaim and work towards replacing $\mathbf{G}$ with a particular path power of $\mathbf{G}$ which satisfies the additional conditions. Note that, for any pattern α, the path power G^{α} admits $w(x_0, \dots, x_{h-1})$ as a polymorphism and has no sources or sinks. If such a path power has algebraic length one and does not contain a loop, then it can be taken as a substitute for $\mathbf{G}$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

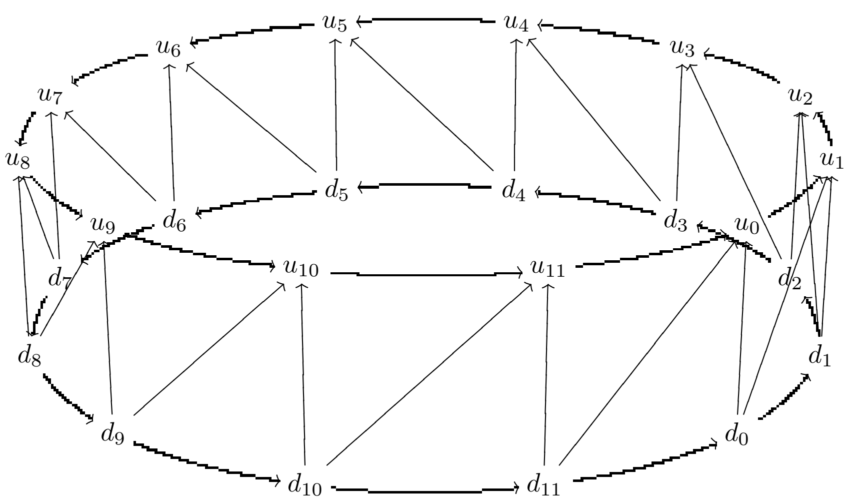


FIG. 1. *The 12-tambourine.*

SUBCLAIM 8.2.1. *The digraph* $\mathbf{G}$ *contains vertices* d *and* u *such that* $d \xrightarrow{|V|,|V|+1} u$.

Proof. Let α be the pattern

$$\underbrace{\rightarrow \ \cdots \rightarrow}_{|V|+1} \underbrace{\leftarrow \ \cdots \leftarrow}_{|V|}.$$

Using the fact that $al(\alpha) = 1$ and that $\mathbf{G}$ has no sources or sinks, it follows that $E(\mathbf{G}) \subseteq E(\mathbf{G}^\alpha)$. Moreover, let a, b be vertices in $\mathbf{G}$ such that b is contained in a closed walk and $a \xrightarrow{k} b$ for some k. Then $a \xrightarrow{k'} b$ for some $k' \leq |V|$, and choosing b' (from the closed walk containing b) such that $b' \xrightarrow{k'+1} b$, we obtain

$$b' \xrightarrow{k'+1} b \xrightarrow{(|V|+1)-(k'+1)} c \xleftarrow{(|V|+1)-(k'+1)} b \xleftarrow{k'} a \quad \text{ for some } c.$$

Thus $b' \xrightarrow{\alpha} a$, and this implies that every component of $\mathbf{G}$ becomes a strong component of $\mathbf{G}^\alpha$.

Let $\mathbf{H} = (W, F)$ be a component of $\mathbf{G}$ with a closed walk realizing a pattern of algebraic length one. Then, for an appropriate F', containing F, the digraph $\mathbf{H}' = (W, F')$ is a strong component of $\mathbf{G}^\alpha$. The digraph $\mathbf{H}'$ contains $\mathbf{H}$ as a subgraph, and therefore its algebraic length is one. The path power $\mathbf{G}^\alpha$ admits $w(x_0, \ldots, x_{h-1})$ as a polymorphism, and thus, by Lemma 6.3, the digraph $\mathbf{H}'$ admits an appropriate restriction of $w(x_0, \ldots, x_{h-1})$. Theorem 7.2 provides a loop in $\mathbf{H}'$, which in turn implies the existence of vertices $d, u \in W$ such that $d \xrightarrow{|V|,|V|+1} u$ in $\mathbf{G}$. ∎

Proof of Claim 8.2. We fix $n = |V|!$ and argue that, for some k, the path power $\mathbf{G}_k = \mathbf{G}^{+(kn+1)}$ satisfies the assertions of the claim and therefore can be taken as a substitute for $\mathbf{G}$. Note that, for any number k, the digraph $\mathbf{G}_k$ admits $w(x_0, \ldots, x_{h-1})$ as a polymorphism, has no sources or sinks, and, by Corollary 5.2, has algebraic length one.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

We first prove that, for all k, the digraph $\mathbf{G}_k$ does not contain a loop. If $\mathbf{G}_k$ does contain a loop, then there exists a closed walk of length $kn+1$ in some strong component of $\mathbf{G}$. In the same strong component in $\mathbf{G}$ there exists a closed walk of length smaller than n and thus coprime to $kn+1$; therefore the GCD of the lengths of closed walks in this strong component is one, and, using Corollary 5.7, Lemma 6.3, and Theorem 7.2, we obtain a loop in this strong component and therefore also in $\mathbf{G}$, a contradiction. Thus, to prove the claim, it remains to verify the additional required properties.

We now show that, for the fixed number n, the n-tambourine maps homomorphically to $\mathbf{G}_k$ for $k \geq 4$. Let d, u be vertices of $\mathbf{G}$ provided by Subclaim 8.2.1. Since $\mathbf{G}$ has no sources or sinks, we can find vertices d', u', each contained in a closed walk, such that d' is connected by a walk to d and u is connected by a walk to u'. By following the closed walks containing d' and u' multiple times, we get d'_0, u'_0, each contained in a closed walk, such that $d'_0 \xrightarrow{3n,3n+1} u'_0$. Moreover, again following the closed walks multiple times, we obtain

$$d'_0 \xrightarrow{n} d'_0 \xrightarrow{3n,3n+1} u'_0 \xrightarrow{n} u'_0.$$

Let d'_i denote the ith vertex of the closed walk $d'_0 \xrightarrow{n} d'_0$ and, similarly, u'_i the ith vertex of the closed walk $u'_0 \xrightarrow{n} u'_0$. Then, for any number $k \geq 4$ and any $i < n$, we have $d'_i \xrightarrow{kn+1} u'_i$ and $d'_i \xrightarrow{kn+1} u'_{(i+1) \bmod n}$. On the other hand, $d'_i \xrightarrow{kn+1} d'_{(i+1) \bmod n}$ and $u'_i \xrightarrow{kn+1} u'_{(i+1) \bmod n}$. Thus, for any $k \geq 4$, the map $d_i \mapsto d'_i$, $u_i \mapsto u'_i$ is a homomorphism from the n-tambourine in the path power $\mathbf{G}_k$.

To prove the second assertion of the claim we need to show that if $k \geq 4$, then any vertex of $\mathbf{G}_k$ is in a closed walk of length n. We fix such a number k and let $W \subset V$ be the subuniverse of $\mathbf{A}$ generated by $\{d'_0, \dots, d'_{n-1}, u'_0, \dots, u'_{n-1}\}$. Let $\mathbf{G}'_k$ be the subgraph induced by $\mathbf{G}_k$ on W. The digraph $\mathbf{G}'_k$ obviously admits a restriction of $w(x_0, \dots, x_{h-1})$ and (since the n-tambourine maps homomorphically to it) has algebraic length one. Choose an arbitrary $a \in W$. Then, by the definition of W, we have a term $t(x_0, \dots, x_{n-1}, y_0, \dots, y_{n-1})$ such that $a = t(d'_0, \dots, d'_{n-1}, u'_0, \dots, u'_{n-1})$. Therefore,

$$\left.\begin{array}{c} t(d'_0, \dots, d'_{n-2}, d'_{n-1}, u'_0, \dots, u'_{n-2}, u'_{n-1}) \\ \downarrow \\ t(d'_1, \dots, d'_{n-1}, d'_0, u'_1, \dots, u'_{n-1}, u'_0) \\ \vdots \\ t(d'_{n-1}, \dots, d'_{n-3}, d'_{n-2}, u'_{n-1}, \dots, u'_{n-3}, u'_{n-2}) \\ \downarrow \\ t(d'_0, \dots, d'_{n-2}, d'_{n-1}, u'_0, \dots, u'_{n-2}, u'_{n-1}) \end{array}\right\} n,$$

and thus a is in a closed walk of length n. This proves that $\mathbf{G}'_k$ has no sources and no sinks, and since it cannot be a counterexample smaller than $\mathbf{G}$, we infer that $W = V$. Therefore the second assertion holds for all the digraphs $\mathbf{G}_k$ with $k \geq 4$.

In the digraph $\mathbf{G}_4$ every vertex is in a closed walk of length n, and therefore $E({\mathbf{G}_4}^{+(nm+1)}) \subseteq E({\mathbf{G}_4}^{+(n(m+1)+1)})$ for any number m. Thus, there is a number l such that for any $m \geq l$ we have ${\mathbf{G}_4}^{+(nm+1)} = {\mathbf{G}_4}^{+(nl+1)}$. Take

$$\mathbf{G}' = {\mathbf{G}_4}^{+(nl+1)} = \mathbf{G}^{+(4n+1)(nl+1)} = \mathbf{G}_{(4nl+l+4)n+1}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

and note that, according to the previous paragraphs of this proof, such a digraph satisfies all but the last assertion of the claim. Let m be arbitrary. Then $(\mathbf{G}')^{+(mn+1)} = \mathbf{G}_4{}^{+((mnl+l+m)n+1)} = \mathbf{G}_4{}^{+(nl+1)} = \mathbf{G}'$, and thus $\mathbf{G}'$ can be taken to substitute for $\mathbf{G}$ and the claim is proved. ∎

From this point on we substitute $\mathbf{G}$ with a digraph provided by the previous claim and fix it together with the number n. For ease of notation we denote the number modulo n using brackets (e.g., $[n+1] = 1$). We already know that the n-tambourine maps homomorphically to $\mathbf{G}$, but we must choose such a homomorphism carefully.

CLAIM 8.3. *The n-tambourine can be mapped homomorphically to $\mathbf{G}$ in such a way that, for some term $t(x_0, \ldots, x_{n-1})$ of algebra $\mathbf{A}$,*

$$d'_i = t^{(i)}(w(\overline{d'_0}, d'_1), w(\overline{d'_1}, d'_2), \ldots, w(\overline{d'_{n-1}}, d'_0)) \quad \text{for all } i < n,$$

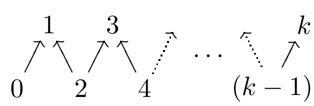
where d'_i is the image of d_i.

Proof. Let $d_i \mapsto d'_i, u_i \mapsto u'_i$ be a homomorphism from the n-tambourine to $\mathbf{G}$. Then, for any i, we have

$$\begin{tikzcd}
w(\overline{u'_i}, u'_{[i+1]}) \arrow[r] & w(\overline{u'_{[i+1]}}, u'_{[i+2]}) \arrow[r] & \cdots \\
 & & \cdots \\
w(\overline{d'_i}, d'_{[i+1]}) \arrow[r] \arrow[uu] \arrow[uur] & w(\overline{d'_{[i+1]}}, d'_{[i+2]}) \arrow[r] \arrow[uu] & \cdots
\end{tikzcd}$$

and thus $d_i \mapsto w(\overline{d'_i}, d'_{[i+1]}), u_i \mapsto w(\overline{u'_i}, u'_{[i+1]})$ is also a homomorphism from the n-tambourine to $\mathbf{G}$. By repeating this procedure, we obtain an infinite sequence of homomorphisms from the n-tambourine to $\mathbf{G}$, and thus some homomorphism has to appear twice in this sequence. This homomorphism satisfies the claim, since the term $t(x_0, \ldots, x_{n-1})$ can be easily obtained as a composition of the polymorphism $w(x_0, \ldots, x_{h-1})$ used in the construction of the sequence. ∎

In the remaining part of the proof we fix vertices $d'_0, \ldots, d'_{n-1}, u'_0, \ldots, u'_{n-1}$ provided by the previous claim and a term $t(x_0, \ldots, x_{n-1})$ associated with them. Let φ_k be the pattern $0 \xrightarrow{\varphi_k} k$



with exactly k edges. (The last edge of the pattern is pointing forward for odd k, as in the above picture, and backward for even k.)

CLAIM 8.4. *The neighborhood $(d'_0)^{\varphi_n}$ contains all vertices of $\mathbf{G}$.*

Proof. Note that, in the n-tambourine, we have

$$(d_0)^{\varphi_n} = \{d_0, \ldots, d_{n-1}, u_0, \ldots, u_{n-1}\},$$

and thus in the digraph $\mathbf{G}$ we have

$$(d'_0)^{\varphi_n} \supseteq \{d'_0, \ldots, d'_{n-1}, u'_0, \ldots, u'_{n-1}\}.$$

Let $\mathbf{G}'$ denote the subgraph of $\mathbf{G}$ induced on the set $(d'_0)^{\varphi_n}$. Then, by Corollary 6.2, $\mathbf{G}'$ admits a restriction of $w(x_0, \ldots, x_{h-1})$ as a polymorphism and has algebraic length one. Further restricting the digraph $\mathbf{G}'$, denote the largest induced subgraph of $\mathbf{G}'$ without sources or sinks by $\mathbf{G}''$. By Lemma 6.4 $\mathbf{G}''$ admits a weak near unanimity

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

polymorphism. Moreover, the vertices $\{d'_0, \dots, d'_{n-1}, u'_0, \dots, u'_{n-1}\}$ are among the vertices of $\mathbf{G}''$. Thus $\mathbf{G}''$ is a counterexample to Theorem 8.1 and therefore has to be equal to $\mathbf{G}$. This proves the claim. ∎

We choose (and fix) k to be a minimal number such that $(d'_0)^{\varphi_{k+1}} = V$. Define $W_i = (d'_i)^{\varphi_k}$, for each $i < n$. We set

$$W = \bigcap_{i<n} W_i,$$

and since W is an intersection of subuniverses of $\mathbf{A}$, by Corollary 6.2, it is itself a subuniverse of $\mathbf{A}$. We denote by $\mathbf{H}$ the subgraph of $\mathbf{G}$ induced by W and prove that $\mathbf{H}$ is a counterexample to Theorem 8.1, contradicting the minimality of $\mathbf{G}$.

The most involved part of the proof deals with constructing a closed realization of a pattern with the algebraic length one in $\mathbf{H}$. Two following claims introduce tools for "projecting" certain walks from $\mathbf{G}$ to $\mathbf{H}$.

Claim 8.5. *There exists a term $s(x_0, \dots, x_{p-1})$ of algebra $\mathbf{A}$ such that for every coordinate $q < p$ there exists i such that*

$$s^{(q)}(W_l, W, \dots, W) \subseteq W_{[i-l]} \cap W_{[i-l+1]} \quad \textit{for any } l < n.$$

Proof. Let $p = hn$ and let $s(x_0, \dots, x_{p-1})$ be defined by

$$t\left(w(x_0, \dots, x_{h-1}), w(x_h, \dots, x_{2h-1}), \dots, w(x_{(n-1)h}, \dots, x_{hn-1})\right).$$

For all $q < p$, let i be maximal such that $q = ih + q''$ for some nonnegative q''. Then, for all $l < n$

$$\begin{aligned}
s^{(q)}(W_l, \overline{W}) &\subseteq t^{(i)}\left(w^{(q'')}(W_l, \overline{W}), w(\overline{W}), \dots, w(\overline{W})\right)\\
&\subseteq t^{(i)}\left(w^{(q'')}(\overline{W_l}, W_{[l+1]}), w(\overline{W_{[l+1]}}, W_{[l+2]}), \dots, w(\overline{W_{[l+n-1]}}, W_l)\right)\\
&= t^{([i-l])}\left(w(\overline{W_0}, W_1), \dots, w^{(q'')}(\overline{W_l}, W_{[l+1]}), \dots, w(\overline{W_{n-1}}, W_0)\right)\\
&\subseteq W_{[i-l]},
\end{aligned}$$

where the last inclusion follows from Claim 8.3 and the fact that

$$\begin{aligned}
d'_{[i-l]} &= t^{([i-l])}(w(\overline{d'_0}, d'_1), \dots, w(\overline{d'_l}, d'_{[l+1]}), \dots, w(\overline{d'_{n-1}}, d'_0))\\
&= t^{([i-l])}(w(\overline{d'_0}, d'_1), \dots, w^{(q'')}(\overline{d'_l}, d'_{[l+1]}), \dots, w(\overline{d'_{n-1}}, d'_0)).
\end{aligned}$$

Similar reasoning shows that

$$\begin{aligned}
s^{(q)}(W_l, \overline{W}) &\subseteq t^{(i)}\left(w^{(q'')}(W_l, \overline{W}), w(\overline{W}), \dots, w(\overline{W})\right)\\
&\subseteq t^{(i)}\left(w^{(q'')}(W_l, \overline{W_{[l-1]}}), w(W_{[l+1]}, \overline{W_l}), \dots, w(W_{[l+n-1]}, \overline{W_{[l+n-2]}})\right)\\
&= t^{[i-l+1]}\left(w(W_1, \overline{W_0}), \dots, w^{(q'')}(W_l, \overline{W_{[l-1]}}), \dots, w(W_0, \overline{W_{n-1}})\right)\\
&\subseteq W_{[i-l+1]},
\end{aligned}$$

and the proof is finished. ∎

Further, using the term constructed in the last claim, we can construct a term satisfying stronger conditions.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

CLAIM 8.6. *There exists a term $r(x_0,\dots,x_{m-1})$ of algebra* $\mathbf{A}$ *such that for every coordinate $q<m$*

$$r^{(q)}\left(\bigcup_{l<n} W_l, W,\dots,W\right) \subseteq W.$$

Proof. Let $s(x_0,\dots,x_{p-1})$ be the p-ary term provided by the previous claim. Note that the term

$$s_2(x_0,x_1,\dots,x_{p^2-1}) = s(s(x_0,\dots,x_{p-1}),\dots,s(x_{p^2-p},\dots,x_{p^2-1}))$$

has the property that for every coordinate $q<p^2-1$ there exists an i such that

$$s_2^{(q)}(W_l,\overline{W}) \subseteq W_{[i-l]}\cap W_{[i-l+1]}\cap W_{[i-l+2]}.$$

To prove a more general statement we recursively define a sequence of terms

- $s_1(x_0,\dots,x_{p-1}) = s(x_0,\dots,x_{p-1})$ and
- $s_{j+1}(x_0,\dots,x_{p^j-1}) = s(s_j(x_0,\dots,x_{p^{j-1}-1}),\dots,s_j(x_{(p-1)p^{j-1}},\dots,x_{p^j-1}))$.

We claim that for any j, any $q<p^j$, and any $l<n$ there is an i such that

$$s_j^{(q)}\big(W_l,W,\dots,W\big) \subseteq W_{[i-l]}\cap\dots\cap W_{[i-l+j]}.$$

We prove this fact by induction on j. The first step of the induction holds via Claim 8.5. Assume that the fact holds for j; then for any l (setting q' to be the result of integer division of q by p^{j-1}, and q'' the remainder of this division) there exist i and i' such that

$$\begin{aligned}
s_{j+1}^{(q)}(W_l,\overline{W}) &\subseteq s^{(q')}\big(s_j^{(q'')}(W_l,\overline{W}), s_j(\overline{W}),\dots,s_j(\overline{W})\big)\\
&\subseteq s^{(q')}\big(W_{[i-l]}\cap\dots\cap W_{[i-l+j]},\overline{W}\big)\\
&\subseteq W_{[i'+i-l]}\cap\dots\cap W_{[i'+i-l+(j+1)]},
\end{aligned}$$

where the second inclusion follows from the induction step and the last one from Claim 8.5. Setting $r(x_0,\dots,x_{m-1})$ equal to $s_{n-1}(x_0,\dots,x_{p^n-1})$ proves the claim. ∎

From this point on we fix a term $r(x_0,\dots,x_{m-1})$ (of arity m) provided by the previous claim. To prove additional properties of the set W (e.g., the fact that it is not empty) we require the following easy claim.

CLAIM 8.7. *Let α be a pattern, and let $a_0\to a_1$ and $b_0\to b_1$ be edges that belong to closed walks. If $a_0\xrightarrow{\alpha} b_0$, then $a_1\xrightarrow{\alpha} b_1$.*

Proof. We prove the claim by induction with respect to the number of edges in α. Let the vertices a_0,a_1,b_0,b_1 be as in the statement of the claim. Assume that $a_0\to b_0$. If i is the length of the closed walk containing the edge $a_0\to a_1$, then, following this walk almost n times, $a_1\xrightarrow{in-1} a_0\to b_0\to b_1$ and, by Claim 8.2, $a_1\to b_1$. The same reasoning can be applied to the case of $a_0\leftarrow b_0$, and the first step of the induction is proved.

For a pattern α consisting of more than one edge we can assume, without loss of generality, that the last edge is going forward. Then $a_0\xrightarrow{\alpha'} a_0'\to b_0$ for some vertex a_0' (where α' is the pattern obtained by removing the last edge of α). By Claim 8.2, it follows that a_0' is in a closed walk of length n, and therefore $a_0'\to a_1'\xrightarrow{n-1} a_0'$ for some a_1'. By the induction hypothesis, $a_1\xrightarrow{\alpha'} a_1'$ and, by the first step of the induction, $a_1'\to b_1$, which proves the claim. ∎

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

We recall the definition of the top and bottom components of the graph from subsection 4.3 and prove some basic properties of W.

Claim 8.8. *The digraph* $\mathbf{H}$ *has no sources and no sinks and*

1. *if* k *is even, then every bottom component is contained in* W*, and*
2. *if* k *is odd, then every top component is contained in* W*.*

Proof. First we show that, for any vertices a, b such that $a \xrightarrow{i} b \xrightarrow{j} a$ in $\mathbf{G}$ for some i, j,

$$\text{if } a \in W_l, \quad \text{then } b \in W_{[l+i]}.$$

To see this note that if $d'_l \xrightarrow{\varphi_k} a$ and $a \to b \xrightarrow{j} a$, then, using Claim 8.7 and the edge $d'_l \to d'_{[l+1]}$, we infer that $d'_{[l+1]} \xrightarrow{\varphi_k} b$. The same procedure repeated i-many times provides the result for arbitrary i.

Let $a \in W$ be arbitrary and b be such that $a \xrightarrow{i} b \xrightarrow{j} a$ for some numbers i, j. Since $a \in W$ it follows, using the note above, that $b \in \bigcap_{l<n} W_{[l+i]} = W$, and this implies that W is a union of strong components. Since, by Claim 8.2, every vertex in $\mathbf{G}$ belongs to a closed walk of length n, the digraph $\mathbf{H}$ has no sources or sinks.

Let k be even and let a be a member of a bottom component. Since every vertex of the graph, by Claim 8.2, belongs to a closed walk, there exists b in the bottom component containing a such that $a \to b$. Since $(d'_0)^{\varphi_{k+1}} = V$, we have $d'_0 \xrightarrow{\varphi_{k-1}} c \leftarrow a' \to b$ for some a' and c. The vertex a is in a bottom component, and therefore a' must be a member of the same bottom component. This implies that $a' \to b \xrightarrow{i} a'$, for some i, and following the closed walk containing b almost n times, $a \to b \xrightarrow{n(i+1)-1} a' \to c$. Thus, by Claim 8.2, we have $a \to c$ and $a \in W_0$. Therefore every bottom component is contained in W_0. To see that every a from a bottom component is contained in an arbitrary W_l we find a b satisfying $a \xrightarrow{l} b \xrightarrow{i} a$ for some i and apply the note from the beginning of the proof of the claim. The claim is proved for even k's, and the same reasoning provides a proof for odd k and top components. ∎

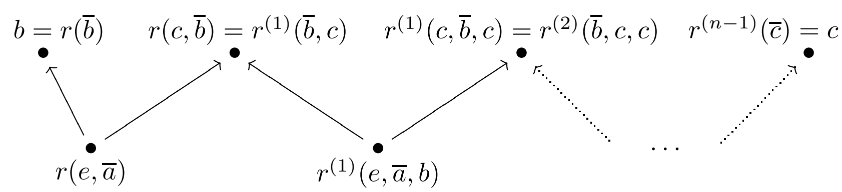
Now we are ready to prove the final claim of this section.

Claim 8.9. *The algebraic length of* $\mathbf{H}$ *is one.*

Proof. In the case where k is odd, we want to find $a, b, c \in W$ and $e \in W_0$ such that

$$\begin{tikzcd} a \arrow[r] & b \arrow[r] & c \\ & e \arrow[u] \arrow[ur] & \end{tikzcd}$$

To find such vertices we set $e = d'_1$ and find, using Claim 8.8, $b \in W$ from a top component such that $u'_{[2]} \xrightarrow{in-1} b$ for some i. There exist a and c in the same component (and thus in W by Claim 8.8) such that $a \to b \to c$. Since $d'_1 \xrightarrow{1,2} u'_{[2]}$, we have $e \xrightarrow{in+1} b$ and $e \xrightarrow{in+1} c$, and therefore, by Claim 8.2, the vertices a, b, c, and e satisfy the required properties. Then, using the term $r(x_0, \ldots, x_{m-1})$, we produce the following oriented walk:



Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

By Claim 8.6, all the vertices of this walk belong to W. Thus we have constructed an oriented walk in $\mathbf{H}$ realizing a pattern of algebraic length zero connecting b to c. Since $b \to c$ we immediately obtain that the algebraic length of $\mathbf{H}$ is one.

In the case where k is even, we similarly find $a, b, c \in W$ and $e \in W_0$ (using u_1' for e) such that

$$\begin{tikzcd} & & e & \\ a & b \arrow[l] \arrow[ur] & & c \arrow[ll] \arrow[ul] \end{tikzcd}$$

The construction of a closed oriented walk realizing a pattern of algebraic length one is the same as it is for odd k, with the exception that the direction of the edges is reversed. $\square$

Thus $\mathbf{H}$ is a digraph without sources or sinks (by Claim 8.8), admitting a weak near unanimity polymorphism and, by the last claim, having algebraic length equal to one. Since, by the definition of W, the number of vertices in $\mathbf{H}$ is strictly smaller than the number of vertices in $\mathbf{G}$, we obtain a contradiction with the minimality of $\mathbf{G}$, and Theorem 8.1 is proved.

REFERENCES

[All94] J. Allen, *Natural Language Understanding*, 2nd ed., Benjamin Cummings, San Francisco, CA, 1994.

[ANP05] D. Achlioptas, A. Naor, and Y. Peres, *Rigorous location of phase transitions in hard optimization problems*, Nature, 435 (2005), pp. 759–764.

[BIM+06] J. Berman, P. Idziak, P. Marković, R. McKenzie, M. Valeriote, and R. Willard, *Varieties with few subalgebras of powers*, Trans. Amer. Math. Soc., to appear.

[BJH90] J. Bang-Jensen and P. Hell, *The effect of two cycles on the complexity of colourings by directed graphs*, Discrete Appl. Math., 26 (1990), pp. 1–23.

[BJHM95] J. Bang-Jensen, P. Hell, and G. MacGillivray, *Hereditarily hard H-colouring problems*, Discrete Math., 138 (1995), pp. 75–92.

[BJK05] A. Bulatov, P. Jeavons, and A. Krokhin, *Classifying the complexity of constraints using finite algebras*, SIAM J. Comput., 34 (2005), pp. 720–742.

[BKJ00] A. A. Bulatov, A. A. Krokhin, and P. Jeavons, *Constraint satisfaction problems and finite algebras*, in Automata, Languages and Programming (Geneva, 2000), Lecture Notes in Comput. Sci. 1853, Springer, Berlin, 2000, pp. 272–282.

[BS81] S. Burris and H. P. Sankappanavar, *A Course in Universal Algebra*, Graduate Texts in Math. 78, Springer-Verlag, New York, 1981.

[Bul05] A. A. Bulatov, *H-coloring dichotomy revisited*, Theoret. Comput. Sci., 349 (2005), pp. 31–39.

[Bul06] A. A. Bulatov, *A dichotomy theorem for constraint satisfaction problems on a* 3*-element set*, J. ACM, 53 (2006), pp. 66–120.

[Dal05] V. Dalmau, *A new tractable class of constraint satisfaction problems*, Ann. Math. Artif. Intell., 44 (2005), pp. 61–85.

[Dal06] V. Dalmau, *Generalized majority-minority operations are tractable*, Log. Methods Comput. Sci., 2 (2006), pp. 1–15.

[DD96] R. Dechter and A. Dechter, *Structure-driven algorithms for truth maintenance*, Artificial Intelligence, 82 (1996), pp. 1–20.

[Fed01] T. Feder, *Classification of homomorphisms to oriented cycles and of k-partite satisfiability*, SIAM J. Discrete Math., 14 (2001), pp. 471–480.

[FV99] T. Feder and M. Y. Vardi, *The computational structure of monotone monadic SNP and constraint satisfaction: A study through Datalog and group theory*, SIAM J. Comput., 28 (1998), pp. 57–104.

[GWW92] W. Gutjahr, E. Welzl, and G. Woeginger, *Polynomial graph-colorings*, Discrete Appl. Math., 35 (1992), pp. 29–45.

[HM88] D. Hobby and R. McKenzie, *The Structure of Finite Algebras*, Contemp. Math. 76, American Mathematical Society, Providence, RI, 1988.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 12/27/12 to 129.173.72.87. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

[HN90] P. Hell and J. Nešetřil, *On the complexity of H-coloring*, J. Combin. Theory Ser. B, 48 (1990), pp. 92–110.

[HN04] P. Hell and J. Nešetřil, *Graphs and Homomorphisms*, Oxford Lecture Series in Math. Appl. 28, Oxford University Press, Oxford, 2004.

[HNZ96a] P. Hell, J. Nešetřil, and X. Zhu, *Complexity of tree homomorphisms*, Discrete Appl. Math., 70 (1996), pp. 23–36.

[HNZ96b] P. Hell, J. Nešetřil, and X. Zhu, *Duality and polynomial testing of tree homomorphisms*, Trans. Amer. Math. Soc., 348 (1996), pp. 1281–1297.

[HNZ96c] P. Hell, J. Nešetřil, and X. Zhu, *Duality of graph homomorphisms*, in Combinatorics, Paul Erdős is Eighty, Vol. 2 (Keszthely, 1993), Bolyai Soc. Math. Stud. 2, János Bolyai Mathematical Society, Budapest, 1996, pp. 271–282.

[HP64] Z. Hedrlín and A. Pultr, *Relations (graphs) with given finitely generated semigroups*, Monatsh. Math., 68 (1964), pp. 213–217.

[HZZ93] P. Hell, H. S. Zhou, and X. Zhu, *Homomorphisms to oriented cycles*, Combinatorica, 13 (1993), pp. 421–433.

[JCG97] P. Jeavons, D. Cohen, and M. Gyssens, *Closure properties of constraints*, J. ACM, 44 (1997), pp. 527–548.

[KMRT+07] F. Krząkała, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborová, *Gibbs states and the set of solutions of random constraint satisfaction problems*, Proc. Natl. Acad. Sci. USA, 104 (2007), pp. 10318–10323.

[Kun] G. Kun, *Constraints, MMSNP and Expander Relational Structures*, manuscript, 2007; http://arxiv.org/abs/0706.1701.

[KV07] E. Kiss and M. Valeriote, *On tractability and congruence distributivity*, Log. Methods Comput. Sci., 3 (2007), 2:6 (electronic).

[LALW98] D. Lesaint, N. Azarmi, R. Laithwaite, and P. Walker, *Engineering dynamic scheduler for Work Manager*, BT Technology J., 16 (1998), pp. 16–29.

[Lev73] L. A. Levin, *Universal enumeration problems*, Problemy Peredači Informacii, 9 (1973), pp. 115–116.

[LZ03] B. Larose and L. Zádori, *The complexity of the extendability problem for finite posets*, SIAM J. Discrete Math., 17 (2003), pp. 114–121.

[LZ06] B. Larose and L. Zádori, *Taylor terms, constraint satisfaction and the complexity of polynomial equations over finite algebras*, Internat. J. Algebra Comput., 16 (2006), pp. 563–581.

[Mac77] A. Mackworth, *Consistency in networks of relations*, Artificial Intelligence, 8 (1977), pp. 99–118.

[Mac91] G. MacGillivray, *On the complexity of colouring by vertex-transitive and arc-transitive digraphs*, SIAM J. Discrete Math., 4 (1991), pp. 397–408.

[MM07] M. Maróti and R. McKenzie, *Existence theorems for weakly symmetric operations*, Algebra Universalis, to appear.

[MMT87] R. N. McKenzie, G. F. McNulty, and W. F. Taylor, *Algebras, Lattices, Varieties, Vol.* I, The Wadsworth & Brooks/Cole Mathematics Series, Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA, 1987.

[Mon74] U. Montanari, *Networks of constraints: Fundamental properties and applications to picture processing*, Information Sci., 7 (1974), pp. 95–132.

[Nad] B. A. Nadel, *Constraint satisfaction in Prolog: Complexity and theory-based heuristics*, Inform. Sci., 83 (1995), pp. 113–131.

[NL] B. A. Nadel and J. Lin, *Automobile transmission design as a constraint satisfaction problem: Modeling the kinematic level*, Artificial Intelligence for Engineering Design, Analysis and Manufacturing (AI EDAM), 53 (1991), pp. 137–171.

[Sab61] G. Sabidussi, *Graph derivatives*, Math. Z., 76 (1961), pp. 385–401.

[SV98] E. Schwalb and L. Vila, *Temporal constraints: A survey*, Constraints, 3 (1998), pp. 129–149.

[Tay77] W. Taylor, *Varieties obeying homotopy laws*, Canad. J. Math., 29 (1977), pp. 498–527.

[Var00] M. Vardi, *Constraint satisfaction and database theory: A tutorial*, in Proceedings of 19th ACM Symposium on Principles of Database Systems (PODS'00), ACM, New York, 2000, pp. 76–85.

[Zhu95] X. Zhu, *A polynomial algorithm for homomorphisms to oriented cycles*, J. Algorithms, 19 (1995), pp. 333–345.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.