# CSP: dichotomy holds for digraphs with no sources and no sinks (a positive answer to the conjecture of Bang-Jensen and Hell)

Libor Barto, Marcin Kozik, Todd Niven

September 11, 2007

#### Abstract

We (hopefully) proved the result in the title.

# 1 Notation

By  $\overline{a}$  we understand a tuple  $(a, \ldots, a)$  and by a  $\overrightarrow{a}$  the tuple  $(a_0, \ldots, a_n)$ . By an abuse of the notation we sometimes identify a graph with the set of its edges. By a weak near unanimity operation we understand an operation  $w(x_0, \ldots, x_{h-1})$  such that

- $w(x, \ldots, x) \approx x$  and
- $w(y, x, \dots, x) \approx w(x, y, \dots, x) \approx \dots \approx w(x, x, \dots, y).$

A polymorphism of a relational structure is an operation compatible with all the relations.

# 2 Some graph notation

Let  $\mathbf{G} = (V, E)$  be a (directed) graph. Given a graph  $\mathbf{G}$  its set of edges is denoted by  $E(\mathbf{G})$  and the set of its vertices by  $V(\mathbf{G})$ . A vertex is called a source (resp. a sink), if it has no incoming (resp. outgoing) edge. A cycle is a closed path, while a circle is a cycle with no subcycles.

For a fixed graph  $\mathbf{G} = (V, E)$  we introduce the following notation. We denote  $(a, b) \in E$  by  $a \to b$ , and we use  $a \xrightarrow{k} b$  to say that there is a directed path from a to b of length exactly k. More generally for any oriented path p with endpoints c, d we write  $a \xrightarrow{p} b$  if there exists a homomorphism  $\phi$  from p into  $\mathbf{G}$  such that  $\phi(c) = a$  and  $\phi(d) = b$ . For any  $W \subseteq V$  we define

$$W^{+n} = \{ v \in V | \exists w \in W \text{ such that } w \xrightarrow{n} v \}$$

and similarly

$$W^{-n} = \{ v \in V | \exists w \in W \text{ such that } v \xrightarrow{n} w \}.$$

We write  $a^{+n}$  instead  $\{a\}^{+n}$  for any  $a \in V$ . More generally, for any oriented path p we write

$$W^p = \{ v \in V | \exists w \in W \text{ such that } w \xrightarrow{p} v \}.$$

As before we use  $a^p$  for  $\{a\}^p$ . Sometimes we use a notation  $a \xrightarrow{p,q} b$  to denote  $a \xrightarrow{p} b$  and  $a \xrightarrow{q} b$ .

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  $\overline{a^{+n}}$  contains all the tuples (of a fixed length) of the elements reachable by an *n*-path from a.

#### 3 Graph gadget powers

Let  $\mathbf{G} = (V, E)$  be a graph and p be an oriented path. We define a gadget graph power of the graph  $\mathbf{G}^p$  in the following way: the vertices of the power are the vertices of the graph  $\mathbf{G}$  and a pair  $(c, d) \in V^2$  is an edge in  $\mathbf{G}^p$  if and only if  $c \xrightarrow{p} d$  in  $\mathbf{G}$ . For a directed path p of length n we put  $\mathbf{G}^{+n} = \mathbf{G}^p$ .

Note that if  $f: V^m \to V$  is a polymorphism of **G** then it is also a polymorphism of any gadget power of this graph. Graph gadget powers can also be defined for general graphs instead of oriented paths, but we will not need this in our paper.

## 4 Basic facts about graphs

The first notion to be introduced is a notion of an algebraic length of a path. For any oriented path p we define the algebraic length al(p) to be

 $al(p) = #\{$ edges going forward in  $p\} - #\{$ edges going backward in  $p\}.$ 

For a graph  $\mathbf{G} = (V, E)$  we put

 $al(\mathbf{G}) = \min\{i > 0 | (\exists v \in V) (\exists p \text{ a path }) \text{ such that } v \xrightarrow{p} v \text{ and } al(p) = i\},\$ 

whenever the set on the right hand side is non-empty and  $\infty$  otherwise. We note that for graphs with no sources and no sinks (or with a directed cycle) the algebraic length of graph is always a natural number and that, for a connected graph **G**, the graph retracts to a directed circle if and only if there exists a directed cycle (equivalently circle) of length  $al(\mathbf{G})$  in **G**.

The following lemma describes a connection between algebraic lengths of graphs and their powers.

**Lemma 4.1.** Let p be a path of algebraic length k, let  $a \xrightarrow{p} b$  in **G** with no sources and no sinks. Then  $a \xrightarrow{q} b$  in  $\mathbf{G}^{+k}$ , where q is a path of algebraic length one.

*Proof.* Any path  $a \xrightarrow{l_1} a_1 \xleftarrow{l_2} a_2 \xrightarrow{l_3} \ldots a_{l_k} = b, l_1 - l_2 + \cdots = k$  can be made into a path, where all of the numbers  $l_i$  are divisible by k. This should be clear from the following example. Consider the path  $a \xrightarrow{1} \xleftarrow{6} \xrightarrow{8}$  of algebraic length 3. We can also write (using no sources and no sinks)  $a \xrightarrow{3} \xleftarrow{8} \xrightarrow{8}$  and then  $a \xrightarrow{3} \xleftarrow{9} \xrightarrow{9}$ . Now  $a \xrightarrow{l_1/k} a_1 \xleftarrow{l_2/k} \ldots b$  is a path of algebraic length one in  $\mathbf{G}^{+k}$ . We leave a proof of the following easy lemma to the reader.

**Lemma 4.2.** If, for strongly connected graph  $\mathbf{G} = (V, E)$ , the GCD of the all the lengths of the cycles in this graph is equal to one then there exists m such that for any  $a, b \in V$  and for any  $n \geq m$  there is a directed path of length n connecting a to b.

The following easy corollary follows.

**Corollary 4.3.** For a strongly connected graph  $\mathbf{G}$  with GCD of the cycles equal to one, and for any number n the graph  $\mathbf{G}^{+n}$  is strongly connected.

There is a direct connection between the greatest common diviser of the length of all the cycles and the algebraic length of the graph, for strongly connected graphs.

**Corollary 4.4.** For a strongly connected graph the GCD of lengths of all the cycles is equall to the algebraic length of the graph.

*Proof.* Let us fix a graph **G** and denote by *n* the greatest common divisor of the lengths of the cycles in **G**. In the graph  $\mathbf{G}^{+n}$  (by reasoning similar to the proof of Lemma 4.1) we can find a strongly connected component with the GCD of its cycles equall to one. Applying Lemma 4.2 to this graph we obtain  $a, b \in V(\mathbf{G}^{+n})$  such that  $a \xrightarrow{m,m+1} b$  in  $\mathbf{G}^{+n}$ . Thus  $a \xrightarrow{nm,nm+n} b$  in **G** and further  $a \xrightarrow{nm} b \xleftarrow{nm+n} a$  proving  $al(\mathbf{G}) \leq n$ .

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 the path 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.  $\Box$ 

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

### 5 A connection between graphs and algebra

Let  $\mathbf{G} = (V, E)$  be a graph admitting a weak near unanimity polymorphism  $w(x_0, \ldots, x_{h-1})$  (we denote by  $w_{\sigma}(x_0, \ldots, x_{h-1}) = w(x_{\sigma(0)}, \ldots, w_{\sigma(h-1)})$  for each permutation  $\sigma$ ).

For such a graph we define an algebra  $\mathbf{A} = (V, w(x_0, \dots, x_{h-1}))$ . In the remaining part of this section we investigate some trivial connection between the structure of a graph and the properties of the algebra connected to it. Note that using this notation we can say that E is a subuniverse of  $\mathbf{A}^2$ . Moreover for any subuniverse of  $\mathbf{A}$  say W we can define a graph  $\mathbf{H} = (W, E \cap W \times W)$  (or  $(W, E_{|W})$ ) which admits a weak near unanimity polymorphism and its algebra is exactly the subalgebra of  $\mathbf{A}$  with a universe equal to W.

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

**Lemma 5.1.** For any subuniverse W of an algebra A the sets  $W^+$  and  $W^-$  are subuniverses of A as well.

*Proof.* Take any elements  $a_0, \ldots, a_{h-1}$  from  $W^+$  let  $b_0, \ldots, b_{h-1} \in W$  such that  $b_i \to a_i$ . Then  $w(b_0, \ldots, b_{h-1}) \to w(a_0, \ldots, a_{h-1})$  showing  $w(a_0, \ldots, a_{h-1}) \in W^+$  which proves the claim.

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

**Corollary 5.2.** For any  $a \in V$  and for any path p and for any number n the sets  $a^{+n}$  and  $a^p$  are subuniverses of  $\mathbf{A}$ .

Subuniverses of **A** can also be obtained in a more complicated way.

**Lemma 5.3.** Every strongly connected component of  $\mathbf{G}$  with the GCD of its cycles equal to one is a subuniverse of  $\mathbf{A}$ .

*Proof.* Fix a vertex a in a strongly connected component of  $\mathbf{G}$  satisfying the assumptions of the lemma. Using Lemma 4.2 we find a number  $n_0$  such that there is path of length  $n_0$  from a to any element of the strongly connected component. Using the same lemma for  $\mathbf{G}$  with arrows reversed we get  $n_1$  such that there is a path of length  $n_1$  from each element of the component to a. Then  $a^{+n_0} \cap a^{-n_1}$  is the subuniverse in question.

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

**Lemma 5.4.** Let  $\mathbf{H} = (W, F)$  be a maximal subgraph of  $\mathbf{G}$  having no sources and no sinks. Then W is a subuniverse of  $\mathbf{A}$ .

*Proof.* Clearly, the vertices of  $\mathbf{H}$  can be described as those having arbitrarily long path from them and to them. Since our graphs are 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.

## 6 Strongly connected graphs

In this section we work with the notion of the greatest common divisor of the lengths of the cycles instead of the algebraic length, which seems more natural in case of the strongly connected graphs. We note that, for a connected graph  $\mathbf{G}$ , the graph retracts to a directed circle if and only if there exists a directed cycle (or equivalently circle) of length  $al(\mathbf{G})$  in  $\mathbf{G}$ . We prove the following theorem.

**Theorem 6.1.** If a strongly connected  $\mathbf{G}$  admits a weak near unanimity polymorphism and the GCD of all the lengths of the cycles in this graph is equal to one, then the graph contains a loop.

We start the proof by choosing a graph  $\mathbf{G} = (V, E)$  to be a minimal (with respect to the number of vertices) counterexample to Theorem 6.1.

Claim 6.2. We can assume that graph G has a 2-cycle.

*Proof.* By Lemma 4.2 we can find a minimal number k such that the graph **G** contains a  $2^k$ -cycle. Then the graph gadget power  $\mathbf{G}^{+2^{k-1}}$  has a 2-cycle and admits a weak near unanimity polymorphism. The graph  $\mathbf{G}^{+2^{k-1}}$  does not contain a loop since **G** didn't and k was chosen to be minimal. Moreover, by Corrolary 4.3, it is strongly connected. Finally if  $m_0, \ldots, m_l$  are the lengths of the cycles witnessing the divisibility condition, then by following the cycles multiple times we can find cycles (in **G**) of lengths  $2^{k-1}m_0, \ldots, 2^{k-1}m_l$  which witness the GCD condition for  $\mathbf{G}^{+2^{k-1}}$ .

Let  $a, b \in V$  be any vertices in the 2-cycle in **G**. The following claim requires no proof.

**Claim 6.3.** For any n the intersection  $a^{+n} \cap \{a, b\}$  is not empty.

Moreover the following fact can be easily proved.

**Claim 6.4.** For any  $m \leq n$  either  $a^{+m} \subseteq a^{+n}$  or  $a^{+m} \subseteq b^{+n}$ .

*Proof.* Since a is in a 2-cycle, we obviously have  $a^{+n} \supseteq a^{+n-2} \supseteq a^{+n-4} \dots$  If  $m \in \{n-1, n-3, \dots\}$  we have  $b^{+n} \supseteq a^{+n-1} \supseteq a^{+n-3} \dots$  This proves the claim.

The following claim is an easy consequence of the fact that  $w(x_0, \ldots, x_{h-1})$  is a polymporphism of **G**.

**Claim 6.5.** There are vertices  $a, b \in V$  in a 2-cycle and a binary term t such that  $a = t(w(\overline{a}, b), w(\overline{b}, a))$ .

*Proof.* Let  $M \subseteq V$  be a minimal (wrt. inclusion) subuniverse of **A** containing a 2-cycle. Let  $a, b \in M$  be vertices in a 2-cycle. Since vertices  $w(\overline{a}, b), w(\overline{b}, a) \in M$  form a 2-cycle, the set  $\{w(\overline{a}, b), w(\overline{b}, a)\}$  generates M (due to the minimality of M). Therefore there exists a term t such that  $t(w(\overline{a}, b), w(\overline{b}, a)) = a$ .  $\Box$ 

Let t be a term satisfying  $a = t(w(\overline{a}, b), w(\overline{b}, a))$ . Then, for any permutations  $\sigma, \sigma'$ , we obtain  $a = t(w_{\sigma}(\overline{a}, b), w_{\sigma'}(\overline{b}, a))$ .

**Claim 6.6.** For any n, any  $m \leq n$  and for any permutations  $\sigma, \sigma'$  the following inclusion holds

$$t(w_{\sigma}(\overline{a^{+n}}, a^{+m}), w_{\sigma'}(\overline{a^{+m}}, a^{+n})) \subseteq a^{+n}$$

*Proof.* Note that  $a = t(w_{\sigma}(\overline{a}, b), w_{\sigma'}(\overline{b}, a))$  implies

$$a^{+n} \supseteq t(w_{\sigma}(\overline{a^{+n}}, b^{+n}), w_{\sigma'}(\overline{b^{+n}}, a^{+n})).$$

Similarly  $a = t(w_{\sigma}(\overline{a}, a), w_{\sigma'}(\overline{a}, a))$  implies

$$a^{+n} \supseteq t(w_{\sigma}(\overline{a^{+n}}, a^{+n}), w_{\sigma'}(\overline{a^{+n}}, a^{+n})).$$

Now the claim follows directly from Claim 6.4.

The following fact is crucial for the proof of the main theorem of this section.

**Claim 6.7.** Let  $c \in a^{+n}$  and  $c \xrightarrow{l-1} c^{l-1} \to a$  such that all of the elements of the path, except possibly a, are in  $a^{+n}$ . Then for any  $\overrightarrow{d}, e, \overrightarrow{d'}, e', f, \overrightarrow{g} \in a^{+n}$  if

$$\overrightarrow{d} \xrightarrow{l+n} \overrightarrow{d'} and e \xrightarrow{l+n} e$$

then

$$t(w(\overrightarrow{d},c),w(\overline{c},e)) \xrightarrow{l+n} t(w(\overrightarrow{d'},f),w(\overrightarrow{g},e'))$$

where all the paths are in  $a^{+n}$ . Moreover the same claim holds for w's replaced with  $w_{\sigma}$  and  $w_{\sigma'}$  respectively for any permutations  $\sigma, \sigma'$ .

*Proof.* Note that  $d_i \xrightarrow{n+l} d'_i, e \xrightarrow{n+l} e', c \xrightarrow{l-1} c^{l-1}$  in  $a^{+n}$ , and denote the k-th element (starting from 0) of the first path  $d^k_i$  and the rest analogously. Thus

$$t(w(\overrightarrow{d},c),w(\overline{c},e)) \xrightarrow{l-1} t(w(\overrightarrow{d^{l-1}},c^{l-1}),w(\overline{c^{l-1}},e^{l-1})) \text{ in } a^{+n}$$

Moreover

$$t(w(\overrightarrow{d^{l-1}}, c^{l-1}), w(\overrightarrow{c^{l-1}}, e^{l-1})) \to t(w(\overrightarrow{d^{l}}, a), w(\overline{a}, e^{l}))$$

and  $t(w(\overrightarrow{d^{l}}, a), w(\overline{a}, e^{l})) \in a^{+n}$  by Claim 6.6 used with m = 0. Denote by  $a = g_i^0 \to g_i^1 \cdots \to g_i^n = g_i$  (and similarly for f) paths of length n from a to  $\overrightarrow{g}$  and f. Note that

$$t(w(\overrightarrow{d^{l+m}}, f^m), w(\overrightarrow{g^m}, e^{l+m})) \to t(w(\overrightarrow{d^{l+m+1}}, f^{m+1}), w(\overrightarrow{g^{m+1}}, e^{l+m+1}))$$

for any  $0 \le m \le n-1$ , and that both elements belong to  $a^{+n}$  (by Claim 6.6). This proves the claim.

Let n be minimal such that  $a^{+(n+1)} = V$ . The existence of such a number follows from Lemma 4.2.

**Claim 6.8.** There exists a cycle in  $a^{+n}$  and a path (in  $a^{+n}$ ) connecting the cycle to either a or b.

*Proof.* Either a or b is in  $a^{+n}$ . Assume the first possibility  $a \in a^{+n}$ . Since  $a^{+(n+1)} = V$  there is  $a_1$  such that  $a \xrightarrow{n} a_1 \to a$ . Similarly, there exists  $a_2$  such that  $a \xrightarrow{n} a_2 \to a_1$ . By repeating this procedure, we get both statements of the claim. The case  $b \in a^{+n}$  is similar.

**Claim 6.9.** There exists an element  $c \in a^{+n}$  and a number k such that:

- 1.  $c \xrightarrow{k} c$  in  $a^{+n}$ ,
- 2.  $c \xrightarrow{k-n} a$  with all the elements of the path, except possibly a, in  $a^{+n}$ .

*Proof.* Let l be the length of a cycle inside  $a^{+n}$ . Take a sufficiently big multiplicity k of l so that there is a path of length k - n from some element in the cycle to a (with all the elements of the path, except possibly a, are in  $a^{+n}$ ) and call this element c.

**Claim 6.10.** The element c from Claim 6.9 is in two cycles of coprime lengths in  $(a^{+n}, E_{|a^{+n}})$ .

*Proof.* Let d be such that  $c \to d \xrightarrow{k-1} c$ . Claim 6.7 applied to the elements  $t(w(\overline{c}, c), w(\overline{c}, c))$  and  $t(w(\overline{c}, d), w(\overline{c}, c))$  shows that these element can be connected by a path of length k in  $a^{+n}$ . Apply the same claim (choosing an appropriate permutation) to  $t(w(\overline{c}, c, d), w(\overline{c}, c))$  and  $t(w(\overline{c}, d, d), w(\overline{c}, c))$ . Repeat this procedure to connect, in  $a^{+n}$ , c to  $t(w(c, c, \overline{d}), w(\overline{c}, c))$  by a path of length equal to the some multiplicity of k. Again apply Claim 6.7 to the pair of elements  $t(w(c, c, \overline{d}), w(c, \overline{c}, c))$  and  $t(w(c, d, \overline{d}), w(d, \overline{c}, c))$  and finally to  $t(w(c, d, \overline{d}), w(d, \overline{c}, c))$  and  $t(w(d, d, \overline{d}), w(d, \overline{d}, d))$  to prove the claim.

Thus restricting to graph  $(a^{+n}, E_{|a^{+n}})$  and further to a strongly connected component of c we obtain a graph strictly smaller than **G** (by the choice of n + 1) admiting a weak near unanimity polymorphism, having GCD of cycle equal to one and without loops – this contradicts the minimality of **G**. The proof of Theorem 6.1 is concluded.

**Corollary 6.11.** If a strongly connected **G** admits a weak near unanimity polymorphism and the GCD of all the lengths of the cycles in this graph is equal to k, then the graph contains a cycle (and circle) of length k (and thus retracts onto it).

*Proof.* The graph gadget power  $\mathbf{G}^{+k}$  is strongly connected and GCD of length of cycles equal to one. According to the Theorem 6.1 it contains a loop. This simply means that there is a cycle of length k in  $\mathbf{G}$ .

It is easy to notice (using Corrolary 4.4) that a strongly connected graph **G** retracts to a circle if and only if **G** contains a  $al(\mathbf{G})$ -circle. Whence we have:

**Corollary 6.12.** Let **G** be a strongly connected graph.  $CSP(\mathbf{G})$  is tractable, if **G** retracts to a circle. Otherwise it is NP-complete.

The above corollary is a special case of Corollary 7.16.

# 7 No sources and no sinks

The main result of this section is the following theorem.

**Theorem 7.1.** Let **G** be a graph with no sources and no sinks. If  $al(\mathbf{G}) = 1$  and **G** admits a weak near unanimity polymorphism then it contains a loop.

The following corollary proves very usefull.

**Corollary 7.2.** Every strongly connected component of a counterexample to the Theorem 7.1 has a GCD of lengths of cycles different than one.

*Proof.* If there is a strongly connected component in **G** with the GCD of lengths of the cycles equal to one, then, by Lemma 5.3 it is a subalgebra of an algebra associated with **G** and, by Theorem 6.1, it contains a loop – a contradiction.  $\Box$ 

Let  $\mathbf{G} = (V, E)$  be a minimal counterexample to the Theorem 7.1. Let  $v_l$  denote the oriented path  $a \xrightarrow{l} \frac{l-1}{l} b$ .

**Claim 7.3.** We can assume that there exist  $d, u \in V$  such that  $d \xrightarrow{1,2} u$ .

*Proof.* First we show that there exists  $d \in V$  such that  $d \xrightarrow{k,k+1} u$  for some  $d \in V$  and a natural number k. We suppose otherwise and will prove, for sufficiently large k, that the graph  $\mathbf{G}^{v_k}$  is strongly connected. Then  $al(\mathbf{G}^{v_k}) = 1$  (since this gadget power preserves the edges of the original graph) and such a graph contradicts Corollary 7.2. To prove strong connectivity let d, u be elements of strongly connected components of  $\mathbf{G}$  each containing a cycle and such that  $d \xrightarrow{l} u$ . We can choose an element d' in the strongly connected component of d such that  $d' \xrightarrow{l-1} d$ . Then  $u \to d'$  in  $\mathbf{G}^{v_l}$  and thus d and u are in the same strongly connected component of the power. Choosing sufficiently large l allows us to obtain a strongly connected graph.

Now let k be minimal with the property that there exist d, u such that  $d \xrightarrow{k,k+1} u$ , namely  $d \to c \xrightarrow{k} u$  and  $d \to b \xrightarrow{k-1} u$ . According to the minimality of k, the gadget power  $\mathbf{G}^{v_{k-1}}$  does not contain a loop and thus we can choose  $\mathbf{G}^{v_{k-1}}$  to be a new  $\mathbf{G}$  (the algebraic length remains one, see the note above again). In this power  $d \xrightarrow{1,2} b$ , that is  $d \to b$  and  $d \to c \to b$ , where the last arrow follows from  $c \xrightarrow{k} u \xrightarrow{k-1} b$ .

A pair  $(D = \{d_0, \ldots, d_{n-1}\}, U = \{u_0, \ldots, u_{n-1}\})$  is called an *n*-tambourine, if  $d_i \to d_{i+1}, u_i \to u_{i+1}, d_i \to u_i, d_i \to u_{i+1}$  for all  $i \in \{0, \ldots, n-1\}$ , where the subscripts are computed modulo n.

#### Claim 7.4. We can assume that **G** contains an n-tambourine (D, U).

*Proof.* Compute the GCD of length of the cycles for every strongly component. Multiply these numbers and call the product k. For any natural number l,  $al(\mathbf{G}^{+(kl+1)}) = 1$  (see Lemma 4.1),  $\mathbf{G}^{+(kl+1)}$  contains no loop otherwise we contradict Corollary 7.2 in some strongly connected component.

Pick vertices d, u from Claim 7.3, vertices  $d_0$  and  $u_0$  and a number n such that

- $u_0$  is contained a cycle and  $d_0$  is a contained in a cycle (and the lengths of the cycles divide n),
- $d_0 \xrightarrow{kn} d \to u \xrightarrow{kn-1} u_0.$

We denote by  $d_0, d_1, \ldots, d_{n-1}$  and  $u_0, u_1, \ldots, u_{n-1}$  the elements of *n*-cycles containing  $d_0$  and  $u_0$ . Now,  $d_i \xrightarrow{3kn+1} d_{i+1}$  and  $u_i \xrightarrow{3kn+1} u_{i+1}$ . Moreover

$$d_i \xrightarrow{kn-i} d_0 \xrightarrow{kn} d \xrightarrow{1,2} u \xrightarrow{kn-1} u_0 \xrightarrow{i} u_i \to u_{i+1}$$

which implies  $d_i \xrightarrow{3kn+1} u_i$  and  $d_i \xrightarrow{3kn+1} u_{i+1}$ . Now  $(D = \{d_0, \ldots, d_{n-1}\}, U = \{u_0, \ldots, u_{n-1}\})$  is an *n*-tambourine in  $\mathbf{G}^{+(3kn+1)}$ . Since, by the first paragraph, the graph  $\mathbf{G}^{+(3kn+1)}$  contains no loop, it can be taken to replace  $\mathbf{G}$ .

Now fix n and assume that (D, U) is an n-tambourine in **G**.

Claim 7.5. Every element is in an n-cycle.

*Proof.* Since the graph induced by  $D \cup U$  has neither sources nor sinks and has algebraic length one,  $D \cup U$  generates the whole algebra **A** according to Lemma 5.4. Therefore every vertex is of the form  $t(d_0, \ldots, d_{n-1}, u_0, \ldots, u_{n-1})$  for some term t. Then vertices  $t(d_i, \ldots, d_{i+n-1}, u_i, \ldots, u_{i+n-1})$ , for  $0 \le i < n$  form an n-cycle.

Claim 7.6. We can assume  $\mathbf{G}^{+(mn+1)} = \mathbf{G}$  for any natural number m.

*Proof.* The graph  $\mathbf{G}^{+(mn+1)}$  contains no loop – otherwise we have a vertex in both *n*-cycle (Claim 7.5) and (mn + 1)-cycle – contradiction to Corollary 7.2.

From Claim 7.5 it follows that  $E(\mathbf{G}^{+(mn+1)}) \supseteq E(\mathbf{G})$  for any m, hence if we replace  $\mathbf{G}$  by  $\mathbf{G}^{+(mn+1)}$  sufficiently many times, we obtain the sought after  $\mathbf{G}$ .

An application of the claim above produces  $\mathbf{G}^{+(n+1)} = \mathbf{G}$  and we can easily deduce the following claim.

**Claim 7.7.** Let  $(a_0, a_1)$  and  $(b_0, b_1)$  be edges in n-cycles. If  $a_0 \rightarrow b_0$ , then  $a_1 \rightarrow b_1$ .

*Proof.*  $a_1 \xrightarrow{n-1} a_0 \to b_0 \to b_1$ , thus  $a_1 \xrightarrow{n+1} b_1$  and finally, by Claim 7.6,  $a_1 \to b_1$ .

Repeated application of Claim 7.7 gives us:

**Claim 7.8.** Let p be an oriented path and  $(a_0, a_1)$  and  $(b_0, b_1)$  be edges in ncycles. If  $a_0 \xrightarrow{p} b_0$ , then  $a_1 \xrightarrow{p} b_1$ .

**Claim 7.9.** There is an n-tambourine  $(D = \{d_0, \ldots, d_{n-1}\}, U)$  and a term  $t(x_0, \ldots, x_{n-1})$  such that

 $t(w(\overline{d_i}, d_{i+1}), w(\overline{d_{i+1}}, d_{i+2}), \dots, w(\overline{d_{i+n-1}}d_{i+n})) = d_i$ 

for every  $i \in \{0, ..., n-1\}$ .

*Proof.* It is easy to see that if (D, U) is an *n*-tambourine, then the pair (D', U'), where

$$D' = \{w(\overline{d_0}, d_1), w(\overline{d_1}, d_2), \dots, w(\overline{d_{n-1}}d_0)\}$$
  
$$U' = \{w(\overline{u_0}, u_1), w(\overline{u_1}, u_2), \dots, w(\overline{u_{n-1}}u_0)\}$$

is an *n*-tambourine as well. By continuing this process we eventually get into an *n*-tambourine we already constructed. Call this *n*-tambourine (D, U), thus we have  $(D, U) = (D^{''...'}, U^{''...'})$  and the statement follows.

Let  $f_z$  denote the oriented path  $a = a_0 \to a_1 \leftarrow a_2 \to a_3 \leftarrow \ldots a_z = b$ . Let z be minimal such that  $d_0^{f_{z+1}} = V$ . Such a z exists since, for big enough z', we have  $D \cup U \subseteq d_0^{f_{z'}}$  which implies  $d_0^{f_{z'}} = V$ . Put

$$P = \bigcap_{i=0}^{n-1} d_i^{f_z}.$$

The set P is a proper subset of V. It will serve us as the smaller counterexample for the theorem.

The following claim is crucial.

Claim 7.10. There is a term q, satisfying

$$q(P, P, \dots, P, \bigcup_{i=0}^{n-1} d_i^{f_z}, P, \dots, P) \subseteq P,$$

where the union is at arbitrary coordinate.

 $\mathit{Proof.}$  Recall that h denotes the arity of our weak near unanimity operation. Let

$$s(x_0, x_1, \dots, x_{hn-1}) = = t(w(x_0, x_1, x_{h-1}), w(x_h, x_{h+1}, \dots, x_{2h-1}), \dots, w(x_{(n-1)h}, \dots, x_{hn-1}))$$

For a natural number j we define the j-th power of  $s(x_0, \ldots, x_{hn-1})$  (to be a  $(hn)^j$ -th ary term) in a recursive way:

- $s^1(x_0, \ldots, x_{hn-1}) = s(x_0, \ldots, x_{hn-1})$  and
- for bigger k

$$s^{k}(x_{0}, \dots, x_{(hn)^{k}-1}) =$$
  
=  $s^{k-1}(s(x_{0}, \dots, x_{hn-1}), \dots, s(x_{hn((hn)^{k-1}-1)}, \dots, x_{(hn)^{k}-1})).$ 

We will proof by induction on k that for every coordinate j in  $s^k$  and every  $i \in \{0, \ldots, n-1\}$  there is a number  $l \in \{0, \ldots, n-1\}$  such that

$$s^k(P,\ldots,P,d_l^{f_z}\cup d_{l+1}^{f_z}\cup\ldots d_{l+k}^{f_z},P,\ldots,P)\subseteq d_i^{f_z},$$

where the union is on the *j*-th coordinate. Then we can put  $q = s^n$  to satisfy the claim.

In the first step (k = 1), we prove more: For every coordinate j, there exists l such that for all  $i, s(P, \ldots, P, d_{i+l}^{f_z} \cup d_{i+l+1}^{f_z}, P \ldots, P) \subseteq d_i^{f_z}$ . Claim 7.9 tells us that

$$t(w_{\sigma}(\overline{d_i}, d_{i+1}), w_{\sigma}(\overline{d_{i+1}}, d_{i+2}), \dots, w_{\sigma}(\overline{d_{i+n-1}}d_{i+n})) = d_i,$$

therefore

$$t(w_{\sigma}(\overline{d_i^{f_z}}, d_{i+1}^{f_z}), w_{\sigma}(\overline{d_{i+1}^{f_z}}, d_{i+2}^{f_z}), \dots, w_{\sigma}(\overline{d_{i+n-1}^{f_z}} d_{i+n}^{f_z})) \subseteq d_i^{f_z},$$

thus l can be taken as  $j \div h$  (integer division of j by h), since we can have  $d_{i+l}^{f_z}$  or  $d_{i+l+1}^{f_z}$  at the l-th coordinate (by choosing appropriate  $\sigma$ ) and the set at any coordinate is a subset of P.

To prove the induction step, let  $i \in \{0, \ldots, n-1\}$  and j be a coordinate of  $s^{k+1}$ 

$$s^{k+1}(x_0,\ldots,x_j,\ldots)=s^k(\ldots,\ldots,s(\ldots,x_j,\ldots),\ldots),$$

where the inner s which contains  $x_j$  is at coordinate j' in  $s^k$  and  $x_j$  is at coordinate j'' in the inner s. More precisely  $j' = j \div (hn)^k$  and  $j'' = j \mod (hn)^k$ . From the induction step, we have l' such that

$$s^k(P,\ldots,P,d_{l'}^{f_z}\cup d_{l'+1}^{f_z}\cup\ldots d_{l'+k}^{f_z},P,\ldots,P)\subseteq d_i^{f_z}.$$

From the first step, we get l'' such that for all i'

$$s(P,\ldots,P,d_{i'+l''}^{f_z}\cup d_{i'+l''+1},P\ldots,P)\subseteq d_{i'}^{f_z}.$$

Now we can put l = l' + l'', since

$$s^{k}(P,\ldots,P,s(P,\ldots,d_{l'+l''}^{f_{z}}\cup\cdots\cup d_{l'+l''+k+1}^{f_{z}},P,\ldots,P),P\ldots,P) \subseteq$$
$$\subseteq s^{k}(P,\ldots,P,d_{l'}^{f_{z}}\cup\cdots\cup d_{l'+k}^{f_{z}},P,\ldots,P) \subseteq d_{i}^{f_{z}}.$$

**Claim 7.11.** Every element of P is in an n-cycle inside P. In particular P has neither sources nor sinks.

Proof. It follows from Claim 7.8.

Now we need to distinguish two possibilities – whether z is odd or even. Let us finish the proof assuming z is odd, the second possibility being similar.

**Claim 7.12.** Let  $u_0 \xrightarrow{i} u'$  for some *i*. Then there exists a set U' containing u' such that (D, U') is an n-tambourine.

*Proof.* It is a straightforward application of Claim 7.6.  $\Box$ 

**Claim 7.13.** There exists a set U' such that (D, U') is an n-tambourine and  $U' \subseteq P$ .

Proof. Let u' be arbitrary vertex such that  $u_0 \xrightarrow{i} u'$  for some i and for any u'' s.t.  $u' \to u'', u''$  is in the same strongly connected component as u' (In other words, u' is in a maximal strongly connected component above  $u_0$ .) Due to Claim 7.12, there is an n-tambourine  $(D, U' = \{u'_0, \ldots, u'_{n-1}\})$ , where U' contains u'. Let  $i \in \{0, \ldots, n-1\}$ . Since  $d_0^{f_{z+1}} = V$ , we have  $d_0 \xrightarrow{f_{z+1}} d_{i-1}$ , hence  $d_0 \xrightarrow{f_{z-1}} a \to b \leftarrow u'_{i-1}$ . Because b is in the same strongly connected component as  $u'_{i-1}$ , there is a path  $b \xrightarrow{k} u'_{i-1}$ . Now  $a \to b \xrightarrow{k} u'_{i-1} \xrightarrow{(k+1)(n-1)} u'_{i-1} \to u'_i$ , i.e.  $a \xrightarrow{(k+1)n+1} u'_i$  and thus  $a \to u'_i$  due to Claim 7.6. Therefore  $d_0 \xrightarrow{f_{z-1}} a \to u'_i$ , i.e.  $d_0 \xrightarrow{f_z} u'_i$ . We have proved that  $d_0^{f_z} \supseteq U'$ . The rest follows from Claim 7.8.

Claim 7.14.  $al(P, E_{|P}) = 1.$ 

*Proof.* The path

$$\begin{aligned} u_1' &= q(u_1', u_1', \dots, u_1') \leftarrow q(d_1, u_0', \dots, u_0') \to q(u_2', u_1', u_1', \dots, u_1') \leftarrow \\ q(u_1', d_1, u_0', \dots, u_0') \to q(u_2', u_2', u_1', \dots, u_1') \leftarrow \dots \to q(u_2', \dots, u_2') = u_2' \leftarrow u_1' \end{aligned}$$

has algebraic length 1 and lies inside P according to the previous claim and Claim 7.10.  $\hfill \Box$ 

*P* is a proper subuniverse of **A**, has no sources and no sinks and  $al(P, E_{|P}) = 1 - a$  contradiction with minimality of *V*. This finishes the proof of Theorem 7.1.

As an easy collorary, we have

.

**Corollary 7.15.** Let **G** be a graph with no sources and no sinks. If  $a \xrightarrow{p} a$ , where p is a path of algebraic length k, and **G** has a weak near unanimity polymorphism, then it contains a k-cycle.

*Proof.* The gadget power  $\mathbf{G}^{+k}$  has no sources and no sinks and has algebraic length equal to one (Lemma 4.1). Hence there is a loop in  $\mathbf{G}^{+k}$ , thus a k-cycle in  $\mathbf{G}$ .

And finally:

**Corollary 7.16.** Let  $\mathbf{G}$  be a graph with no sources and no sinks. If  $\mathbf{G}$  retracts to a disjoint union of circles, then  $CSP(\mathbf{G})$  is tractable. Otherwise it is NP-complete.

*Proof.* Assume that **G** has a weak near unanimity polymorphism. Let M be the set of algebraic length of all components of **G** and N be the set of minimal elements of M with respect to divisibility. According to the previous corollary, we have cycle  $C_n$  of length n for every  $n \in N$ . By minimality it follows that  $C_n$  is a circle for each  $n \in N$  and that  $C_n$ ,  $C_{n'}$  are disjoint for  $n, n' \in N$ ,  $n \neq n'$ . Now **G** retracts onto the disjoint union  $\bigcup_{n \in N} C_n$  and clearly  $CSP(\mathbf{G})$  is tractable.

If **G** has no weak near unanimity polymorphism and is a core, then it is NP-complete.  $\hfill \Box$