A Combinatorial Approach to Collapsing Words

A. Cherubini[1], P. Gawrychowski[2], A. Kisielewicz[2], B. Piochi[3]

[1] Politecnico di Milano, Department of Mathematics, Milano, Italy,
`aleche@mate.polimi.it`
[2] University of Wrocław, Department of Mathematics and Computer Science,
Wrocław, Poland, `kisiel@math.uni.wroc.pl`
[3] Università di Firenze, Department of Mathematics, Firenze, Italy,
`piochi@math.unifi.it`

Abstract. Given a word w over a finite alphabet Σ and a finite deterministic automaton $\mathcal{A} = \langle Q, \Sigma, \delta \rangle$, the inequality $|\delta(Q, w)| \leq |Q| - n$ means that under the natural action of the word w the image of the state set Q is reduced by at least n states. The word w is n-collapsing if this inequality holds for any deterministic finite automaton that satisfies such an inequality for at least one word. In this paper we present a new approach to the topic of collapsing words, and announce a few results we have obtained using this new approach. In particular, we present a direct proof of the fact that the language of n-collapsing words is recursive.

1 Introduction

In this paper by an automaton $\mathcal{A} = \langle Q, \Sigma, \delta \rangle$ we mean a finite deterministic automaton with the state set Q, the input alphabet Σ, and the transition function $\delta : Q \times \Sigma \to Q$. The action of Σ on Q given by δ will be denoted simply by concatenation: $qa = \delta(q, a)$. This action extends naturally on the action of the words of Σ^* on Q. Given a word $w \in \Sigma^*$, we will be interested in the difference of the cardinalities $|Q| - |Qw|$, called the *deficiency* of the word w with respect to $\mathcal{A}$ and denoted $\mathrm{df}_{\mathcal{A}}(w)$.

Let $n \geq 1$, a word $w \in \Sigma^*$ is called *n-compressing* for $\mathcal{A}$, if $\mathrm{df}_{\mathcal{A}}(w) \geq n$. An automaton $\mathcal{A}$ is *n-compressible*, if there exists an n-compressing word for $\mathcal{A}$. A word $w \in \Sigma^*$ is *n-collapsing* (over Σ), if it is n-compressing for every n-compressible automaton with the input alphabet Σ.

It has been proved in [11] that n-collapsing words always exist, for any Σ and any $n \geq 1$. In [7] it is shown that, over a fixed alphabet Σ, each n-collapsing word is *n-full*, i.e., it contains any word of length n among its subwords. An n-compressible automaton $\mathcal{A}$ is called *proper* ([1]) if no word of length n is n-compressing for it. Thus to check whether a word $w \in \Sigma$ is n-collapsing it is enough to consider only proper n-compressible automata. For other results and connections with the Černý conjecture see [1, 3, 4, 9].

In [1] certain characterizations of 2-compressing words were given by associating to every word a family of finitely generated subgroups in some finitely generated free groups; it was proved that the property of being 2-collapsing is

connected with the subgroup indices in this context. A more geometric version of this idea has been developed in [2]. Unfortunately, these characterizations did not allow either to settle the natural complexity problem concerning collapsing words or to generalize to n-collapsing words (cf. remarks in [4]). In [1] the authors ask a few questions in hope to simplify the characterization.

In this paper we answer all these questions in negative. This is done by using another more combinatorial characterization. Our characterization made also possible to solve important complexity problems, and to tackle more general problems. It forms the base for a series of papers in the area (see e.g. [5, 6]).

We view an automaton $\mathcal{A} = \langle Q, \Sigma, \delta \rangle$ as a set of transformations labelled by letters of Σ rather than as a standard triple. By the *transformations of* $\mathcal{A}$ we mean those transformations of Q that are induced via δ by letters of Σ. Note that to define an automaton it is enough to assign just to any letter of Σ a transformation of Q.

It is not difficult to see that any proper 2-compressible automaton $\mathcal{A}$ has to have both permutation and non-permutation transformations, and the later correspond to letters a with $\mathrm{df}_{\mathcal{A}}(a) = 1$. Thus, for each such non-permutation transformation a there is a uniquely determined state $z \in Q$ which does not belong to the image Qa and two different states $x, y \in Q$ satisfying $xa = ya$; such a transformation will be referred to as a *transformation of type* $\{x, y\} \backslash z$ (read: x, y identified, z missing). Using this notions we classify proper 2 -compressible automata as follows

Proposition 1. *An automaton $\mathcal{A}$ is proper 2-compressible if and only if $\mathcal{A}$ satisfies one of the following conditions:*

(i) *there are x, y such that all non-permutation transformations are of the same type $\{x, y\} \backslash x$, and the group of permutations fixes neither the element x nor the set $\{x, y\}$;*

(ii) *there is x such that each non-permutation transformation is of type $\{x, z\} \backslash x$ for some z, at least two different types occur, and the group of permutations does not fix x;*

(iii) *there are x, y such that each non-permutation transformations is of type $\{x, y\} \backslash x$ or $\{x, y\} \backslash y$, both the types occur, and the group of permutations does not fix the set $\{x, y\}$.*

This classification corresponds closely to the one in [1], where the automata in cases (i) and (ii) are called MONO, and those satisfying (iii) are called STEREO. We shall call MONO1 and MONO2 the automata in the cases (i) and (ii), respectively.

We note that the use of a different language here is connected with a different view and leads to a different characterization theorem; one that allows a natural generalization. In particular, in general case we speak about transformations a of type $\{S_1, \ldots, S_k\} \backslash M$, where S_i are the sets of states having the same image under a, and M is the the set of non-images under a. Our idea is that this information is generally enough to approach problems on collapsing words.

2   New characterization of 2-collapsing words

We wish to show that for a word $w \in \Sigma^*$ being 2-collapsing over an alphabet Σ is equivalent to the nonexistence of nontrivial solutions to certain systems of conditions on permutations. Consider partitions of Σ into blocks, where blocks are intended to represent types of transformations and closely correspond to the role assignments introduced in [1]. A nontrivial partition of Σ with a distinguished block P, $\emptyset \subseteq P \subseteq \Sigma$, will be called a *DB-partition* and will be denote by (P, Υ), where $\Upsilon = \{B_2, \ldots, B_h\}$ is the induced partition of $\Sigma \setminus P$ ($h \geq 2$). Let w be a 2-full word over Σ. To each subword of w of the form $\alpha v \beta$, where v is a nonempty word whose all letters belong to P (i.e. $v \in P^+$), while $\alpha \notin P$ and $\beta \in B_j$, we assign a permutation condition of the form

$$1v \in \{1, j\},$$

where the letters of P are treated as permutation variables. Thus, the condition means that the image of 1 under the product v of permutations belongs to the set $\{1, j\}$. The resulting set of permutation conditions (containing all conditions corresponding to subwords of w with the properties described above) will be denoted $\Gamma_w(P, \Upsilon)$ and referred to as the system of permutation conditions determined by a word w and a DB-partition (P, Υ). Note that different orderings of blocks in $\{B_2, \ldots, B_h\}$ lead to systems which are "equivalent" in the sense that both of them have or have not non-trivial solutions; so we don't care the orderings of blocks.

We say that this system *has a solution* if there exists an assignment of permutations on a finite set $\{1, 2, \ldots, N\}$ to letters in P such that all the conditions in $\Gamma_w(P, \Upsilon)$ are satisfied. A trivial solution is one with all permutations fixing 1. Also, in the special case when Υ consists of a unique block B_2 (and in consequence, all j's on the right hand side of the conditions are equal 2), a solution with all permutations fixing the set $\{1, 2\}$ is considered trivial. The remaining solutions are *nontrivial*.

A DB-partition $(P, \{B_1, B_2\})$ of Σ (into exactly 3 blocks, with a distinguished block P) will be called a *3DB-partition*. For such a partition, we define an additional system of permutation conditions as follows. To each subword of w of the form $\alpha v \beta$, with $\alpha \in B_i, \beta \in B_j$, $i, j \in \{1, 2\}$, and $v \in P^+$, we assign a permutation condition of the form

$$iv \in \{1, 2\}$$

(the image of i under v belongs to $\{1, 2\}$). The resulting set of permutation conditions will be denoted by $\Gamma'_w(P, \{B_1, B_2\})$. For such a system, a solution in permutations is *nontrivial* if the image of the set $\{1, 2\}$ does not remain fixed under all the permutations.

Theorem 1. *A word $w \in \Sigma^*$ is 2-collapsing if and only if it is 2-full and the following conditions holds:*

(i) $\Gamma_w(P,\Upsilon)$ has no nontrivial solution for any DB-partition (P,Υ) of Σ;

(ii) $\Gamma'_w(P,\{B_1,B_2\})$ has no nontrivial solution for any 3DB-partition $(P,\{B_1,B_2\})$ of Σ.

The reader may observe an explicit similarity with the characterization in [1, Theorem 3.3]. Yet while, indeed, there is a correspondence in the general structure, our approach is almost converse: rather then looking into an algebraic structure behind, we reduce the problem to the simplest conditions on permutations.

Proof. Our general idea is to compute the deficiency of a word w proceeding letter by letter and observing if and how the deficiency increase. To give an idea of how this technique works, we sketch the proof of (i).

To prove the ,,only if" part, recall that if w is 2-collapsing, then it is 2-full. By way of contradiction assume that the system $\Gamma_w(P,\Upsilon)$ has a nontrivial solution for some DB-partition (P,Υ), and that this solution consists of permutations on a set $Q = \{1,2,\ldots,n\}$. Let $\mathcal{A}$ be an automaton over Σ with the set Q of states, where the letters in P act as the permutations in the solution and the letters in each of blocks $B_i \in \Upsilon$ act as (arbitrary) transformations of type $\{1,i\}\backslash 1$. Since the solution is nontrivial, the group of permutations does not fix 1, and if $\Upsilon = \{B_2\}$ then the group of permutations does not fix the set $\{1,2\}$. Thus, by Proposition 1, in each case $\mathcal{A}$ is a proper 2-compressible MONO automaton.

Now, if w has no subword of the form $\alpha v\beta$, with $v \in P^+$, $\alpha \in B_i$ and $\beta \in B_j$, then $w = v\alpha_{i_1}\ldots\alpha_{i_m}u$ with $v,u \in P^*$ and $\alpha_{i_j} \in \Upsilon$, hence $Qw = Q - \{1\}$ and $\mathrm{df}_{\mathcal{A}}(w) = 1$, a contradiction.

Now, let $\alpha v\beta$ be a subword of w with $v \in P^+$, $\alpha \in B_i$ and $\beta \in B_j$, and assume that it is the first subword of this type in w. It follows that $w = s\alpha v\beta t$, where $s,t \in \Sigma^*$, with $\mathrm{df}_{\mathcal{A}}(s\alpha) = 1$, and 1 is missing in the image $Qs\alpha$. Since v is nonempty, the permutation condition $1v \in \{1,j\}$ is in $\Gamma_w(P,\Upsilon)$. It means that 1 is moved into 1 or j, and consequently 1 or j is missing in the image $Qs\alpha v$. Since β identifies 1 and j, $\mathrm{df}_{\mathcal{A}}(s\alpha v\beta) = 1$, and again 1 is missing in the image $Qs\alpha v\beta$. Proceed letter by letter to get that also in this case $\mathrm{df}_{\mathcal{A}}(w) = 1$: contradiction.

To prove the ,,if" part, assume that w is 2-full but not 2-collapsing, i.e. there exists a proper 2-compressible automaton $\mathcal{A}$ over Σ, with the set of states $Q = \{1,2,\ldots,n\}$, for which w is not 2-compressing. If $\mathcal{A}$ is of type MONO, then consider the DB-partition of Σ, where P represent permutations of $\mathcal{A}$, and $B_2,\ldots,B_h$ represent transformations of types $\{1,2\}\backslash 1$, $\ldots$, $\{1,h\}\backslash 1$, respectively (we assume without loss of generality that $x = 1$ is the distinguished state). The fact that w is not 2-compressing for $\mathcal{A}$ means that, computing deficiency letter by letter, after encountering in w the first letter of any B_i, the deficiency decrease by one, but it does not decrease on further letters. The only segments where the deficiency may decrease are those of the form $\alpha v\beta$, with $\alpha \in B_i, \beta \in B_j$ and $v \in P^+$ when $1v \notin \{1,j\}$. Since the deficiency does not decrease on these segments then the permutations satisfy the corresponding conditions $1v \in \{1,j\}$, as required. The solution they form is nontrivial because of respective conditions (i) or (ii) in Proposition 1.

From our approach, one can easily see that classifying automata as MONO and STEREO has no natural generalization, and is only a very special feature of proper 2-compressible automata. Yet, to demonstrate that the problems in [1] have negative answers we need the following corollary, which can be obtained by more detailed proof as above.

Corollary 1. *For a fixed finite alphabet Σ:*

(i) *A word w is 2-compressing for each 2-compressible MONO1 automaton if and only if the system $\Gamma_w(P, \Upsilon)$ has no nontrivial solution for any DB-partition (P, Υ) with $|\Upsilon| = 1$.*

(ii) *A word w is 2-compressing for each 2-compressible MONO2 automaton if and only if the system $\Gamma_w(P, \Upsilon)$ has no nontrivial solution for any DB-partition (P, Υ) with $|\Upsilon| > 1$.*

(iii) *A word w is 2-compressing for each 2-compressible STEREO automaton if and only if the system $\Gamma_w(P, \Upsilon)$ has no nontrivial solution for any 3DB-partition (P, Υ).*

3 Complexity of recognizing 2-collapsing words

If $\Sigma = \{\alpha, \beta\}$, then we have only two DB-partitions, i.e. only two corresponding system of permutation conditions, each in only one variable. Let

$$E_\alpha(w) = \{k \geq 1 \,:\, \beta\alpha^k\beta \text{ is a subword of } w\};$$
$$E_\beta(w) = \{k \geq 1 \,:\, \alpha\beta^k\alpha \text{ is a subword of } w\}.$$

Using Theorem 1 one almost immediately gets the following result, which closely resembles Proposition 3 of [10]:

Lemma 1. *A word $w \in \{\alpha, \beta\}^*$ is 2-collapsing if and only if it is 2-full and for all integers $n \geq 3$, $0 < r < n$ non of the sets $E_\alpha(w)$ and $E_\beta(w)$ modulo n is included in $\{0, r\}$.*

In order to find an algorithm to check whether a word w is 2-collapsing, we obviously may assume that $n \leq N$, where N is the minimum of the second smallest elements in $E_\alpha(w)$ and in $E_\beta(w)$. Since $N < |w|$, this yields the following corollary (suggested in [4]):

Corollary 2. *For a 2-element alphabet Σ, checking whether a word $w \in \Sigma^*$ is 2-collapsing may be done in polynomial time with respect to $|w|$.*

Moreover, having such a simple insight into the problem, a further question arises naturally: can the problem be solved still in polynomial time when the words are given in the compressed form $\alpha^{k_1}\beta^{k_2}\ldots\beta^{k_n}$ (where $\Sigma = \{\alpha, \beta\}$, k_1 and k_n are allowed to be 0, and the numbers are given in the decimal or binary encodings)? Our new approach makes possible to obtain the following

Theorem 2. *For a 2-element alphabet Σ, checking whether a word $w \in \Sigma^*$ is 2-collapsing may be done in polynomial time with respect to the size of w in the compressed form.*

The proof is a little bit tricky. The idea is that although the number N of cases to be checked may be too large with respect to the size of w in the compressed form, one observes that it is enough to check only the prime divisors of $n < N$ and $n = 4$. The number of such divisors is polynomial with respect to the size of w in the compressed form, but the problem still remains how to overcome the fact that factorization into primes is not known to be in P. Here one applies a certain factorization of the product $k_1 \cdot \ldots \cdot k_n$ into larger divisors, which is enough for our purposes.

At this point we would like also to mention two further results which have been obtained in smaller teams, and which use Theorem 1, as a starting point.

Theorem 3 ([5]). *The problem of recognizing 2-collapsing words over a fixed alphabet Σ with more than 2 letters is co-NP-complete.*

It seems that the same methods may be used for n-collapsing words, but the number of independent systems of permutation conditions is much larger.

A word w is n-synchronizing if it is n-collapsing for every n-compressible automaton with $n + 1$ states (i.e. brings all the states to a single state; c.f. [1]). Since there are only finitely many n-state automata over a fixed alphabet Σ, one can observe that the problem of recognizing n-synchronizing words can be solved in polynomial time. Yet, we have:

Theorem 4 ([6]). *The general problem of recognizing 2-synchronizing words, where the input consists of an alphabet and a word, is co-NP-complete.*

4   Decidability

For some time it was not even clear that the problem of recognizing of n-collapsing words is decidable. In fact, it is the main result announced in [4], where a large sketch of the proof of this fact is given. The full proof in [8] consists of several lemmas and occupies more than 10 pages. We prove this fact in this section.

Theorem 5. *Let w be a word over a fixed alphabet Σ. If w is not n-collapsing, then there exists an n-compressible automaton $\mathcal{A}$ with number of states $|Q| < 3(n-1)|w|$ such that $\mathrm{df}_{\mathcal{A}}(w) < n$.*

Proof. Suppose that w is not n-collapsing, and let $\mathcal{A}' = \langle Q', \Sigma, \delta' \rangle$ be an n-compressible automaton such that $\mathrm{df}_{\mathcal{A}'}(w) < n$. By n-compressibility it follows that there is also a word w' such that $\mathrm{df}_{\mathcal{A}'}(w') \geq n$, and we may assume that $w' = wu$ extends w. Our aim is to construct an automaton $\mathcal{A} = \langle Q, \Sigma, \delta \rangle$ having the same properties, with $Q \subseteq Q'$, small enough.

Let Σ consist of letters α_i and β_i, chosen so that α_i's represent non-permutation transformations of $\mathcal{A}'$, while β_i represent permutations. We may assume that

$$w = \alpha_{i_0} \Gamma_1 \alpha_{i_1} \ldots \Gamma_d \alpha_{i_d},$$

where each $\Gamma_j = \beta_{t_1} \ldots \beta_{t_i}$ is a product of β_i's, possibly empty (permutations at the beginning and at the end may be ignored). We assume that α_i is of type $\{S_{i,1}, \ldots, S_{i,m_i}\}/M_i$, that is, M_i is the set of non-images under α_i, and $S_{i,1}, \ldots, S_{i,m_i}$ are the sets having the same image under α_i. Note that we may assume that the total number of elements in $S_{i,1}, \ldots, S_{i,m_i}$ is less than $2(n-1)$, and $|M_i| \leq n-1$ (otherwise $\mathrm{df}_{\mathcal{A}'}(\alpha_i) \geq n$ on the letter α_i alone). Also $m_i \leq n-1$, so the number of images $S_{i,j}\alpha_i$ is less than n. Consequently, by rough estimation, the total number of elements in all the subsets M_i and $S_{i,j}$ and in all the images of $S_{i,j}$ does not exceed $4(n-1)s$, where $s = |\Sigma|$.

Now consider the information we need to determine the exact value of the deficiency $\mathrm{df}_{\mathcal{A}'}(w)$. Consider the partial deficiency sets $X_j = Q' \setminus Q'\alpha_{i_0} \ldots \alpha_{i_j}$ for all initial segments of w terminating with some α_i; note that $X_0 = M_{i_0}$, and $X_d = Q' \setminus Q'w$. In general

$$M_{i_j} \subseteq X_j \subseteq M_{i_j} \cup X_{j-1} \Gamma_j \alpha_{i_j}.$$

For $x \in X_{j-1}$, it may happen that $x\Gamma_j \alpha_{i_j} \notin X_j$ if and only if there is a suitable $y \in Q'$ with $y\alpha_{i_j} = x\alpha_{i_j}$ and this is determined by the information on the types of α_i's.

Thus, we need only one more partial information in order to compute all X_j's, namely the following one. For each $j \geq 0$, if $\Gamma_j = \beta_{t_1} \ldots \beta_{t_j}$, we need to know the values of β_{t_1} on the states $x \in X_{j-1}$, the values of β_{t_2} on the states $x \in X_{j-1}\beta_{t_1}$, and so on; and finally the values of α_{i_j} on the states $x \in X_{j-1}\Gamma_j$. Since $|X_j| < n$ for all j, this yields totally no more than $(n-1)(|w|-1)$ new states involved (recall that $X_0 = M_{i_0}$). As a result, a partial information on δ' involving less than $(n-1)(4s+|w|-1)$ states determines the required properties of the word w. Indeed, it should be clear that we obtain an automaton $\mathcal{A}$ with no more than $5(n-1)|w|$ satisfying $\mathrm{df}_{\mathcal{A}}(w) < n$, while closing in a natural way cycles in permutations and closing the cycles in non-permutations in such way that they preserve the sets M_i's.

To complete the proof of decidability it is enough to observe that essentially the same argument can be applied to the word wu and to recall the fact that the length of u may be bounded from above by a function of n. The latter follows, for example, from the proof of the basic result [11] concerning the very existence of n-collapsing words (c.f. [4, Section 1]). It follows that there exists an n-compressible automaton $\mathcal{A} = \langle Q, \Sigma, \delta \rangle$, with $|Q|$ bounded by a function of n and $|w|$, witnessing that w is not n-collapsing.

To obtain our bound for $|Q|$, first note that we may not need the whole information on the sets $S_{i,1}, \ldots, S_{i,m_i}$. It is enough to know, for each α_{i_j}, single pairs x, y from some of these sets satisfying suitable equalities $y\alpha_{i_j} = x\alpha_{i_j}$; and we do not even need to know the exact value of the image. Thus, considering α_{i_j}, we need to know M_{i_j} together with one of the following: either the image

$x\alpha_{i_j}$ for $x \in X_{j-1}\Gamma_j$ or an element y with $y\alpha_{i_j} = x\alpha_{i_j}$. Then, one can see that a partial definition of δ involving a subset $S \subseteq Q'$ with $|S| \leq (n-1)(s+|w|-1)$ is enough to guarantee that $\mathrm{df}_\mathcal{A}(w) < n$ for any $\mathcal{A} = \langle Q, \Sigma, \delta \rangle$ having such a subset of states and with δ extending the partial definition.

Now, from the fact that $\mathrm{df}_{\mathcal{A}'}(wu) \geq n$, we infer that there are $x, y \in Q'w$ such that $xu = yu$. For $u = \gamma_1 \ldots \gamma_t$, denote $x_0 = x$, and $x_i = x_{i-1}\gamma_i$, and similarly, $y_0 = y$, and $y_i = y_{i-1}\gamma_i$, for all $1 \leq i \leq t$. Note that $x_t = xu = yu = y_t$.

If all $x_i, y_i \in S$, then we may choose $Q = S$, and to finish the proof as before (this part requires some more detailed analysis that it is possible to complete definitions of non-transformations without adding new states while keeping all the necessary properties). The resulting automaton $\mathcal{A} = \langle Q, \Sigma, \delta \rangle$ satisfies $\mathrm{df}_\mathcal{A}(w) = k < n$, and $\mathrm{df}_\mathcal{A}(w) = k+1$. If $k < n-1$, we simply add new states $q_1', \ldots, q_{n-1-k}'$, all of them transformed into q_1' by all the transformations: we obtain so an automaton with the required properties and with less than $|S| + (n-1) = (n-1)(s+|w|)$ states.

If all x_i's are in S, while some of y_i's are not, assume u be the shortest possible with the desired property, to argue that there are at most $(n-1)|w|$ states y_i not in S. Indeed let a be the smallest index for which $y_a\gamma_a \notin S$, and b the largest one with $y_b\gamma_b^{-1} \notin S$. Consider all the pairs $x_i\gamma_i = x_{i+1}$ with $a \leq i < b$. We may assume that there is no i, j such that $x_i = x_j$ and $\gamma_i = \gamma_j$; otherwise we could build u', shorter than u, with the same properties. But in the first part there are no more than $(n-1)|w|$ such pairs with determined images, whence the claim. Adjoin all x_i's and y_i's to S to obtain, as before, the automaon we are looking for, whose set Q of states fulfills $|Q| < (n-1)(s+2|w|)$.

At last, if both some x_i and some y_i are not in S, we may argue that, again, the total number of such states not in S does not exceed $(n-1)|w|$; otherwise we can modify u reducing ourselves to the former case (this part is left to the reader).

Theorem 5 obviously shows that, for each $n > 1$, the language of n-collapsing words over Σ is recursive (it is always enough to check a finite number of automata). The bound in our theorem is even slightly better than $3(n-1)|w|+n+1$ in [8, Theorem 1] (and [4, Theorem 1]). Remark that considering separately some extremal cases (like all the transformations are the same) there is a room to still improve the bound in the theorem.

5 Collapsing words on MONO and STEREO automata

To complete the paper we provide the ideas of solutions to the open problems from [1] which we mentioned above.

For a 3-element alphabet Σ it is not difficult to find a word, which is 2-compressing for each 2-compressible STEREO automaton, but fails to be 2-collapsing; the same question for MONO was stated as an open problem (Question 5.1 in [1]), intending to consider possible simplifications of the characterization given there. We give a negative answer to this and other two questions, producing suitable counterexamples.

Proposition 2. *Let* $\Sigma = \{\alpha, \beta, \gamma\}$ *There is a word* $w \in \Sigma^+$ *which is* 2-*compressing for any* MONO *automaton, with input alphabet* Σ, *and which fails to be* 2-*collapsing.*

In the proof below we use notation $\alpha = (i_1 i_2 ...)...$ to denote a permutation α which has a cycle sending i_1 into i_2 in its decomposition into a product of disjoint cycles.

Proof. Let A be a finite automaton with the set of states $Q = \{1, 2, 3\}$ and input alphabet Σ. Let α be of type $\{1, 2\}\backslash 1$, β of type $\{1, 2\}\backslash 2$, and $\gamma = (123)$ a permutation. By Proposition 1 A is a proper 2-compressible STEREO automata and it is easy to verify that all the words in $X \cup \{\alpha, \beta\}^+$, with $X = \{\alpha\gamma x, \alpha\gamma^3 x, \beta\gamma^2 x, \beta\gamma^3 x | \ x \in \{\alpha, \beta\}\}$, are not 2-compressing for A.

Now consider the 2-full word

$$w = \beta\alpha\gamma\beta^2\gamma^2\beta\alpha\gamma\alpha\beta\alpha\gamma\beta\alpha\beta^2\gamma^2\alpha\beta\gamma^3\beta\alpha^2\gamma\beta\gamma^2\alpha.$$

All of its subwords of the form $x\gamma^+y$ with $x, y \in \{\alpha, \beta\}$ are in X (whence it does not compress A and is not 2-collapsing), but it is 2-compressing for each MONO automaton with input alphabet $\{\alpha, \beta, \gamma\}$ by Corollary 1. In fact we shall prove that for any DB-partition (P, Υ) of $\{\alpha, \beta, \gamma\}$ the system $\Gamma_w(P, \Upsilon)$ has no nontrivial solution.

Assume $|P| = 1$ and denote by π the element of P. If $\Upsilon = B_2$, then the solutions of the system $\Gamma_w(P, \Upsilon)$ are all trivial. In fact for all choice of π, the conditions $1\pi \in \{1, 2\}$, $1\pi^2 \in \{1, 2\}$ occur in $\Gamma_w(P, \Upsilon)$: the solutions of the former condition either fix 1 or have the form $(12...)...$, and $\pi = (12...)...$ is a solution of the latter only if the cycle $(12...)$ reduces to (12). If $\Upsilon = \{B_2, B_3\}$, then the solutions of the system $\Gamma_w(P, \Upsilon)$ are still trivial, since for each choice of π the conditions $1\pi \in \{1, 2\}$, $1\pi \in \{1, 3\}$ occur in the system $\Gamma_w(P, \Upsilon)$, hence all the solutions must fix 1.

So assume that $|P| = 2$. Let $\Upsilon = \{\alpha\}$. The system $\Gamma_w(P, \Upsilon)$ is

(1) $1\gamma\beta^2\gamma^2\beta \in \{1, 2\}$,   (2) $1\gamma \in \{1, 2\}$,   (3) $1\beta \in \{1, 2\}$,
(4) $1\gamma\beta \in \{1, 2\}$,   (5) $1\beta^2\gamma^2 \in \{1, 2\}$,   (6) $1\beta\gamma^3\beta \in \{1, 2\}$,
(7) $1\gamma\beta\gamma^2 \in \{1, 2\}$,

The solutions of the condition (3) either fix 1 or are of the form $\beta = (12...)...$; similarly the solutions of the condition (ii) either fix 1 or are of the form $\gamma = (12...)...$. Suppose that it is not the case that both of them fix 1. Assume $1\gamma = 1$ and let $\beta = (12x...)...$: hence by condition (5) $x\gamma^2 = 1\beta^2\gamma^2 \in \{1, 2\}$. This yields $\beta = (12)...$; in fact if $x\gamma^2 = 1$ then $x = 1$, else by the first condition we get $x\gamma^2\beta = 2\beta \in \{1, 2\}$. Whence by conditions (6) and (7) $2\gamma^2 = 2 = 2\gamma^3$, thus γ fixes 2 too and the solutions are trivial. So let $\gamma = (12...)...$ Condition (4) yields $2\beta \in \{1, 2\}$, whence either β fixes 1 and 2 or it is of the form $\beta = (12)...$ In both cases by condition (5) it follows that $\gamma = (12)...$ and the solutions are always trivial.

The cases when $\Upsilon = \{\beta\}$ or $\Upsilon = \{\gamma\}$ are similar and still there exist only trivial solutions.

Proposition 2 answers (in negative) to Question 5.1 in [1]. The following Proposition gives a negative answer to Question 5.2.(i) in [1]:

Proposition 3. *Let $\Sigma = \{\alpha, \beta, \gamma\}$. There is a word $w \in \Sigma^+$ which is 2-compressing for any* MONO *automaton with one non-permutation transformation, but not for each* MONO *automaton with two non-permutation transformations.*

Proof. Let A be a finite automaton with set of states $Q = \{1, 2, 3\}$ and input alphabet Σ. Let α be of type $\{1, 2\}\backslash 1$, β of type $\{1, 3\}\backslash 1$, and $\gamma = (123)$ a permutation. A is a proper 2-compressible MONO automaton with 2 non permutation transformations. Consider the set $X = \{x\gamma^2\beta, x\gamma^3\beta, x\gamma\alpha, x\gamma^3\alpha \mid x \in \{\alpha, \beta\}\}$, it is easy to verify that all words in $X \cup \{\alpha, \beta\}^*$ have deficiency 1 with respect to A.

Consider the 2-full word

$$w = \gamma\alpha\beta^2\alpha\gamma^2\beta\gamma\alpha^2\beta\gamma^3\beta\alpha\gamma\alpha\beta\alpha^2\gamma^2\beta\alpha\beta\gamma^2\beta^2\alpha\gamma.$$

All of its subwords of the form $x\gamma^+y$ with $x, y \in \{\alpha, \beta\}$ are in X (whence it does not compress A and is not 2-collapsing), but is 2-compressible with respect to each MONO automaton with input alphabet $\{\alpha, \beta, \gamma\}$ by Corollary 1. In fact for any DB-partition (P, Υ) of Σ the system $\Gamma_w(P, \Upsilon)$ has no nontrivial solution; the proof is analogous to that of Proposition 2.

It is easy to deduce the following

Corollary 3. *Let $\Sigma = \{\alpha, \beta, \gamma\}$. There is a word $w \in \Sigma^+$ which is 2-compressing for each* MONO1 *automaton, and which fails to be 2-compressing for any* MONO2 *automaton with input alphabet Σ.*

Also Question 5.2.(ii) in [1], regarding STEREO automata, has a negative answer. The proof of the following proposition is similar to the former ones:

Proposition 4. *Let $\Sigma = \{\alpha, \beta, \gamma\, \delta\}$. There is a word $w \in \Sigma^+$ which is 2-compressing for each* STEREO *automaton with input alphabet Σ and with two non-permutation transformations, but fails to be 2-compressing for some* STEREO *automaton with input alphabet Σ with three non-permutation transformations.*

References

1. D. S. Ananichev, A. Cherubini, and M. V. Volkov, *Image reducing words and subgroups of free groups*, Theor. Comput. Sci. 307, no.1 (2003), 7792.
2. D. S. Ananichev, A. Cherubini, and M. V. Volkov, *An inverse automata algorithm for recognizing 2-collapsing words*, in: M. Ito, M. Toyama (eds.), Developments in Language Theory, Lect. Notes Comp. Sci. 2450, Springer, Berlin 2003, 270282.
3. D. S. Ananichev and I. V. Petrov, *Quest for short synchronizing words and short collapsing words*, WORDS. Proc. 4th Int. Conf., Univ. of Turku, Turku, 2003, 411418.
4. D. S. Ananichev, I. V. Petrov, and M. V. Volkov, *Collapsing words: A Progress Report*, in: C. De Felice and A. Restivo (eds.), Developments in Language Theory, Lect. Notes Comp. Sci. 3572, Springer, Berlin 2005, 11-21.

5. A. Cherubini and A. Kisielewicz, *Recognizing collapsing words is co-NP-complete*, Proceedings of DCFS 2006, to appear.
6. P. Gawrychowski and A. Kisielewicz, *Recognizing 2-synchronizing words*, preprint 2006.
7. S. W. Margolis, J.-E. Pin, and M. V. Volkov, *Words guaranteeing minimum image*, Internat. J. Foundations Comp. Sci. 15 (2004) 259276.
8. I. V. Petrov, *An algorithm for recognizing n-collapsing words*, preprint 2005.
9. J.-E. Pin, *On two combinatorial problems arising from automata theory*, Ann. Discrete Math. 17 (1983) 535548.
10. E. V. Pribavkina, *On some properties of the language of 2-collapsing words*, A, in: C. De Felice and A. Restivo (eds.), Developments in Language Theory, Lect. Notes Comp. Sci. 3572, Springer, Berlin 2005, 374-384.
11. N. Sauer and M. G. Stone, *Composing functions to reduce image size*, Ars Combinatoria 1 (1991) 171176.