

Deterministic Soliton Graphs

Miklós Bartha
 Department of Computer Science
 Memorial University of Newfoundland
 St. John’s, NL, Canada
 E-mail: bartha@cs.mun.ca

Miklós Krész
 Department of Computer Science
 Juhász Gyula Teacher Training College
 University of Szeged, Hungary
 E-mail: kresz@jgytf.u-szeged.hu

Keywords: soliton automata, matching theory

Received: March 23, 2005

Soliton graphs are studied in the context of a reduction procedure that simplifies the structure of graphs without affecting the deterministic property of the corresponding automata. It is shown that an elementary soliton graph defines a deterministic automaton iff it reduces to a graph not containing even-length cycles. Based on this result, a general characterization is given for deterministic soliton graphs using chestnut graphs, generalized trees, and graphs having a unique perfect matching.

Povzetek: Članek obravnava grafe brez lihiv ciklov.

1 Introduction

One of the most ambitious goals of research¹ in modern bioelectronics is to develop a molecular computer. The introduction of the concept “soliton automaton” in [5] has been inspired by this research, with the intention to capture the phenomenon called soliton waves [4] through an appropriate graph model.

Soliton graphs and automata have been systematically studied by the authors on the grounds of matching theory in a number of papers. Perhaps the most significant contribution among these is [2], where soliton graphs have been decomposed into elementary components, and these components have been grouped into pairwise disjoint families based on how they can be reached by alternating paths starting from external vertices. This paper can also serve as a source of further references on soliton automata for the interested reader.

Since soliton automata are proposed as switching devices, deterministic automata are in the center of investigations. The results reported in this paper are aimed at providing a complete characterization of deterministic soliton automata. The two major aspects of this characterization are:

1. Describing elementary deterministic soliton graphs.
2. Recognizing that deterministic soliton graphs having an alternating cycle follow a simple hierarchical pattern called a chestnut.

¹Partially supported by Natural Science and Engineering Research Council of Canada, Discovery Grant #170493-03

An important tool in the study of both aspects is a reduction procedure, which might be of interest by itself. It allows elementary deterministic soliton graphs to be reduced to generalized trees, and it can also be used to reduce chestnut graphs to really straightforward ones, called baby chestnuts.

2 Soliton graphs and automata

By a graph $G = (V(G), E(G))$ we mean an undirected graph with multiple edges and loops allowed. A vertex $v \in V(G)$ is called *external* if its degree is one, and *internal* if the degree of v is at least two. An internal vertex is *base* if it is adjacent to an external one. External edges are those of $E(G)$ that are incident with at least one external vertex, and internal edges are those connecting two internal vertices. Graph G is called *open* if it has at least one external vertex.

A *walk* in a graph is an alternating sequence of vertices and edges, which starts and ends with a vertex, and in which each edge is incident with the vertex immediately preceding it and the vertex immediately following it. The *length* of a walk is the number of occurrences of edges in it. A *trail* is a walk in which all edges are distinct and a *path* is a walk in which all vertices are distinct. A cycle is a trail which can be decomposed into a path and an edge connecting the endpoints of the path. If $\alpha = ve_1 \dots e_n w$ is a walk from v to w and $\beta = wf_1 \dots f_k z$ is a walk from w to z , then the *concatenation* of α and β is the walk $\alpha \parallel \beta = ve_1 \dots e_n wf_1 \dots f_k z$ from v to z .

A *matching* M of graph G is a subset of $E(G)$ such that

no vertex of G occurs more than once as an endpoint of some edge in M . It is understood by this definition that loops cannot participate in any matching. The endpoints of the edges contained in M are said to be *covered* by M . A *perfect internal matching* is a matching that covers all of the internal vertices. An edge $e \in E(G)$ is *allowed (mandatory)* if e is contained in some (respectively, all) perfect internal matching(s) of G . *Forbidden edges* are those that are not allowed. We shall also use the term *constant edge* to identify an edge that is either forbidden or mandatory. An open graph having a perfect internal matching is called a *soliton graph*. A soliton graph G is *elementary* if its allowed edges form a connected subgraph covering all the external vertices. Observe that if G is elementary, then it cannot contain a mandatory edge, unless G is a mandatory edge by itself with a number of loops incident with one of its endpoints.

Let G be an elementary soliton graph, and define the relation \sim on $Int(G)$ as follows: $v_1 \sim v_2$ if an extra edge e connecting v_1 with v_2 becomes forbidden in $G + e$. It is known, cf. [6, 2], that \sim is an equivalence relation, which determines the so called *canonical partition* of (the internal vertices of) G . The reader is referred to [6] for more information on canonical equivalence, and on matching theory in general.

Let G be a graph and M be a matching of G . An edge $e \in E(G)$ is said to be M -positive (M -negative) if $e \in M$ (respectively, $e \notin M$). An M -alternating path (cycle) in G is a path (respectively, even-length cycle) stepping on M -positive and M -negative edges in an alternating fashion. An M -alternating loop is an odd-length cycle having the same alternating pattern of edges, except that exactly one vertex has two negative edges incident with it. Let us agree that, if the matching M is understood or irrelevant in a particular context, then it will not be explicitly indicated in these terms. An *external alternating path* is one that has an external endpoint. If both endpoints of the path are external, then it is called a *crossing*. An alternating path is *positive* if it is such at its internal endpoints, meaning that the edges incident with those endpoints are positive.

Let G be a soliton graph, fixed for the rest of this section, and let M be a perfect internal matching of G . An M -alternating unit is either a crossing or an alternating cycle with respect to M . *Switching* on an alternating unit amounts to changing the sign of each edge along the unit. It is easy to see that the operation of switching on an M -alternating unit α creates a new perfect internal matching $S(M, \alpha)$ for G . Moreover, as it was proved in [1], every perfect internal matching M of G can be transformed into any other perfect internal matching M' by switching on a collection of pairwise disjoint alternating units. Consequently, an edge e of G is constant iff there is no alternating unit passing through e with respect to any perfect internal matching of G . A collection of pairwise disjoint M -alternating units will be called an M -alternating network, and the network transforming one perfect internal matching M into another M' will be denoted by $N(M, M')$. Clearly,

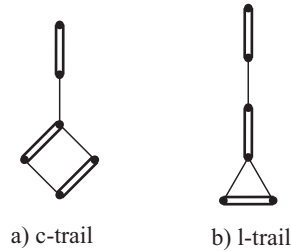


Figure 1: Soliton trails.

$N(M, M')$ is unique.

Now we generalize the alternating property to trails and walks. An *alternating trail* is a trail α stepping on positive and negative edges in such a way that α is either a path, or it returns to itself only in the last step, traversing a negative edge. The trail α is a *c-trail (l-trail)* if it does return to itself, closing up an even-length (respectively, odd-length) cycle. That is, $\alpha = \alpha_1 \parallel \alpha_2$, where α_1 is a path and α_2 is a cycle. These two components of α are called the *handle and circuit*, in notation, $h(\alpha)$ and $c(\alpha)$. The joint vertex on $h(\alpha)$ and $c(\alpha)$ is called the *center* of α . An *external alternating trail* is one starting out from an external vertex, and a *soliton trail* is a proper external alternating trail, that is, either a c-trail or an l-trail. See Fig. 1.

The collection of *external alternating walks* in G with respect to some perfect internal matching M , and the concept of switching on such walks are defined recursively as follows.

- (i) The walk $\alpha = v_0 e v_1$, where $e = (v_0, v_1)$ with v_0 being external, is an external M -alternating walk, and switching on α results in the set $S(M, \alpha) = M \Delta \{e\}$. (The operation Δ is symmetric difference of sets.)
- (ii) If $\alpha = v_0 e_1 \dots e_n v_n$ is an external M -alternating walk ending at an internal vertex v_n , and $e_{n+1} = (v_n, v_{n+1})$ is such that $e_{n+1} \in S(M, \alpha)$ iff $e_n \in S(M, \alpha)$, then $\alpha' = \alpha e_{n+1} v_{n+1}$ is an external M -alternating walk and

$$S(M, \alpha') = S(M, \alpha) \Delta \{e_{n+1}\}.$$

It is required, however, that $e_{n+1} \neq e_n$, unless $e_n \in S(M, \alpha)$ is a loop.

It is clear by the above definition that $S(M, \alpha)$ is a perfect internal matching iff the endpoint v_n of α is external, too. In this case we say that α is a *soliton walk*.

EXAMPLE Consider the graph G of Fig. 2, and let $M = \{e, h_1, h_2\}$. A possible soliton walk from u to v with respect to M is $\alpha = u e w g z_1 h_1 z_2 l_2 z_3 h_2 z_4 l_1 z_1 g w f v$. Switching on α then results in $S(M, \alpha) = \{f, l_1, l_2\}$.

Graph G gives rise to a *soliton automaton* \mathcal{A}_G , the states of which are the perfect internal matchings of G . The input alphabet for \mathcal{A}_G is the set of all (ordered) pairs of external vertices in G , and the state transition function δ is defined

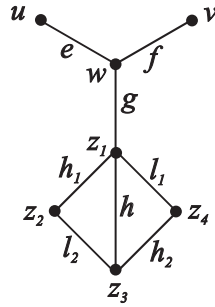


Figure 2: Example soliton graph G

by

$$\delta(M, (v, w)) = \{S(M, \alpha) \mid \alpha \text{ is a soliton walk, } v \text{ to } w\}.$$

Graph G is called *deterministic* if \mathcal{A}_G is such in the usual sense, that is, if for every state M and input (v, w) ,

$$|\delta(M, (v, w))| \leq 1.$$

EXAMPLE (Continued) Observe that the soliton automaton defined by the graph of Fig. 2 is non-deterministic, as $\alpha' = uewfv$ is also a soliton walk from u to v with respect to state M such that $S(M, \alpha) \neq S(M, \alpha')$.

Let α be a soliton c -trail with respect to M . It is easy to see that the walk

$$s(\alpha) = h(\alpha) \parallel c(\alpha) \parallel h(\alpha)^R$$

is a soliton walk, and the effect of switching on $s(\alpha)$ is the same as switching on the cycle $c(\alpha)$ alone. (For any walk β , β^R denotes the reverse of β .) If α is a soliton l -trail, then $s(\alpha)$ is defined as the soliton walk

$$s(\alpha) = h(\alpha) \parallel c(\alpha) \parallel c(\alpha) \parallel h(\alpha)^R.$$

Clearly, this walk induces a self-transition of \mathcal{A}_G , that is, no state change is observed. In the sequel, all perfect internal matchings of G will simply be called *states* for obvious reasons.

Recall from [5] that an edge e of G is *impervious* if there is no soliton walk passing through e in any state of G . Edge e is *viable* if it is not impervious. See Fig. 2, edge h for an example of an impervious edge.

We are going to give a simpler characterization of impervious edges in terms of alternating paths, rather than walks. To this end, we need the following lemma.

Lemma 2.1. *If α is an external M -alternating walk from v to u , then there exists an M -alternating network Γ and an external M -alternating trail β from v to u such that*

1. Γ consists of cycles only, and it is disjoint from β ;
2. $S(M, \alpha) = S(S(M, \Gamma), \beta)$.

Proof. Easy induction on the length of α , left to the reader. \square

An internal vertex $v \in V(G)$ is called *accessible* with

respect to state M if there exists a positive external M -alternating path leading to v . It is easy to see, cf. [2], that vertex v is accessible with respect to some state M iff v is accessible with respect to all states of G .

Proposition 2.2. *For every edge $e \in E(G)$, e is impervious iff both endpoints of e are inaccessible.*

Proof. If either endpoint of e is accessible, then e is clearly viable. Assume therefore that both endpoints u_1 and u_2 of e are inaccessible, and let α be an arbitrary external M -alternating walk from $v \in Ext(G)$ to either u_1 or u_2 , say u_1 . By Lemma 2.1, there exists a suitable external M -alternating trail β from v to u_1 . Each internal edge lying on β has an accessible endpoint, so that e is not among them. Moreover, the edge of $h(\beta)$ incident with u_1 must be positive with respect to $S(M, \alpha)$, otherwise $h(\beta)$ would be a positive external alternating path with respect to state $S(M, \Gamma)$. (Recall that $h(\beta)$ is the handle of β , and take $h(\beta) = \beta$ if β is just a path.) But then e must be negative with respect to $S(M, \alpha)$, or else $h(\beta)eu_2$ would be a positive external alternating path leading to u_2 (with respect to $S(M, \Gamma)$, or even M , since Γ is disjoint from β). We conclude that the walk α cannot continue on e , because it must take the two positive edges incident with u_1 before and after hitting that vertex. Thus, every time an external alternating walk reaches either endpoint of e , it will miss e as a possible continuation. In other words, e is impervious. \square

3 Elementary decomposition of soliton graphs

Again, let us fix a soliton graph G for the entire section. In general, the subgraph of G determined by its allowed edges has several connected components, which are called the *elementary components* of G . Notice that an elementary component can be as small as a single external vertex of G . Such elementary components are called *degenerate*, and they are the only exception from the general rule that each elementary component is an elementary graph. A *mandatory elementary component* is a single mandatory edge $e \in E(G)$, which might have a loop around one or both of its endpoints.

The structure of elementary components in a soliton graph G has been analysed in [2]. To summarize the main results of this analysis, we first need to review some of the key concepts introduced in that paper. Elementary components are classified as *external* or *internal*, depending on whether or not they contain an external vertex. An elementary component of G is *viable* if it does not contain impervious allowed edges. A viable internal elementary component C is *one-way* if all external alternating paths (with respect to any state M) enter C in vertices belonging to the same canonical class of C . This unique class, as well as the vertices belonging to this class, are called *principal*. Furthermore, every external elementary component is considered a priori one-way (with no principal canonical class,

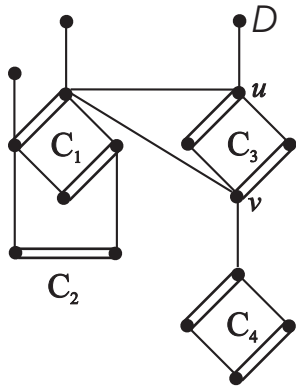


Figure 3: Elementary components in a soliton graph

of course). A viable elementary component is *two-way* if it is not one-way. An *impervious* elementary component is one that is not viable.

EXAMPLE The graph of Fig. 3 has five elementary components, among which C_1 and D are external, while C_2 , C_3 and C_4 are internal. Component C_3 is one-way with the canonical class $\{u, v\}$ being principal, while C_2 is two-way and C_4 is impervious.

Let C be an elementary component of G , and M be a state. An M -alternating C -ear is a negative M -alternating path or loop having its two endpoints, but no other vertices, in C . The endpoints of the ear will necessarily be in the same canonical class of C . We say that elementary component C' is *two-way accessible* from component C with respect to any (or all) state(s) M , in notation $C\rho C'$, if C' is covered by an M -alternating C -ear. It is required, though, that if C is one-way and internal, then the endpoints of this ear *not be* in the principal canonical class of C . As it was shown in [2], the two-way accessible relationship is matching invariant. A *family* of elementary components in G is a block of the partition induced by the smallest equivalence relation containing ρ . A family \mathcal{F} is called external if it contains an external elementary component, otherwise \mathcal{F} is internal. A *degenerate family* is one that consists of a single degenerate external elementary component. Family \mathcal{F} is *viable* if every elementary component in \mathcal{F} is such, and \mathcal{F} is *impervious* if it is not viable. As it turns out easily, the elementary components of an impervious family are all impervious. Soliton graph G is *viable* if all of its families are such.

EXAMPLE (CONTINUED) Our example graph in Fig. 3 has four families: $\mathcal{F}_1 = \{C_1, C_2\}$, $\mathcal{F}_2 = \{D\}$, $\mathcal{F}_3 = \{C_3\}$, $\mathcal{F}_4 = \{C_4\}$. Family \mathcal{F}_1 is external, \mathcal{F}_2 is degenerate, and \mathcal{F}_3 is internal. These families are all viable, whereas family \mathcal{F}_4 is impervious.

The first group of results obtained in [2] on the structure of elementary components of G can now be stated as follows.

Theorem 3.1. *Each viable family of G contains a unique*

one-way elementary component, called the root of the family. Each vertex in every member of the family, except for the principal vertices of the root, is accessible. The principal vertices themselves are inaccessible, but all other vertices are only accessible through them.

A family \mathcal{F} is called *near-external* if each forbidden viable edge incident with any principal vertex of its root is external. For two distinct viable families \mathcal{F}_1 and \mathcal{F}_2 , \mathcal{F}_2 is said to follow \mathcal{F}_1 , in notation $\mathcal{F}_1 \mapsto \mathcal{F}_2$, if there exists an edge in G connecting any non-principal vertex in \mathcal{F}_1 with a principal vertex of the root of \mathcal{F}_2 . The reflexive and transitive closure of \mapsto is denoted by \mapsto^* . The second group of results in [2] characterizes the edge connections between members inside one viable family, and those between two different families.

Theorem 3.2. *The following four statements hold for the families of G .*

1. *An edge e inside a viable family \mathcal{F} is impervious iff both endpoints of e are in the principal canonical class of the root. Every forbidden edge e connecting two different elementary components in \mathcal{F} is part of an ear to some member $C \in \mathcal{F}$.*
2. *For every edge e connecting a viable family \mathcal{F}_1 to any other family (viable or not) \mathcal{F}_2 , at least one endpoint of e is principal in \mathcal{F}_1 or \mathcal{F}_2 . If the endpoint of e in \mathcal{F}_1 is not principal, then \mathcal{F}_2 is viable and it follows \mathcal{F}_1 .*
3. *The relation \mapsto^* is a partial order between viable families, by which the external families are minimal elements.*
4. *If \mathcal{F} and \mathcal{G} are families such that $\mathcal{F} \mapsto^* \mathcal{G}$, then each non-principal vertex u of \mathcal{G} is accessible from \mathcal{F} , meaning that for every state M there exists a positive M -alternating path to u either from a suitable external vertex of \mathcal{F} , if \mathcal{F} is external, or from an arbitrary principal vertex of \mathcal{F} , if \mathcal{F} is internal. The path α runs entirely in the families between \mathcal{F} and \mathcal{G} according to \mapsto^* .*

For convenience, the inverse of the partial order \mapsto^* will be referred to as \leq_G . Theorems 3.1 and 3.2 are fundamental regarding the structural decomposition of soliton graphs, and they will be applied liberally throughout the forthcoming sections (especially in Section 4). There will be no explicit reference made, however, to these theorems whenever they apply.

4 Chestnuts

Chestnuts have been introduced in [5] as a group of deterministic soliton graphs having a very simple and easily recognizable structure.

Definition 1. A connected graph G is called a *chestnut* if it has a representation in the form $G = \gamma + \alpha_1 + \dots + \alpha_k$ with $k \geq 1$, where γ is a cycle of even length and each α_i ($i \in [k]$) is a tree subject to the following conditions:

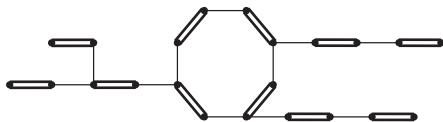


Figure 4: A chestnut.

- (i) $V(\alpha_i) \cap V(\alpha_j) = \emptyset$ for $1 \leq i \neq j \leq k$;
- (ii) $V(\alpha_i) \cap V(\gamma)$ consists of a unique vertex – denoted by v_i – for each $i \in [k]$;
- (iii) v_i and v_j are at even distance on γ for any distinct $i, j \in [k]$;
- (iv) any vertex $w_i \in V(\alpha_i)$ with $d(w_i) > 2$ is at even distance from v_i in α_i for each $i \in [k]$.

See Fig. 4 for an example chestnut.

Our first observation regarding chestnuts is that they are bipartite graphs. Let us call a vertex of a chestnut G *outer* if its distance from any of the v_i 's is even, and *inner* if this distance is odd. Then the inner and outer vertices indeed define a bipartition of G . Moreover, the degree of each inner vertex is at most 2. It is easy to come up with a perfect internal matching for G : just mark the cycle γ in an alternating fashion, then the continuation is uniquely determined by the structure of the trees α_i . Thus, G has exactly two states. It is also easy to see that the inner internal vertices are accessible, while the outer ones are inaccessible. Thus, the cycle γ forms an internal elementary component with its outer vertices constituting the principal canonical class of this component. Moreover, γ forms a stand-alone internal family in G . The rest of G 's families are all single mandatory edges along the trees α_i , or they are degenerate ones consisting of a single inner external vertex. Their rank in the partial order \leq_G is consistent with their position in the respective trees α_i , following a decreasing order from the leafs to the root. The family $\{\gamma\}$ is the minimum element of \leq_G , and G has no impervious edges.

By the description above, every chestnut G is a deterministic soliton graph. Moreover, G is strongly deterministic [5] in the sense that, for each pair (v_1, v_2) of external vertices, there exists at most one soliton walk from v_1 to v_2 in each state of G . We are going to show that for every connected soliton graph G having no impervious edges, but possessing a non-mandatory internal elementary component, G is deterministic iff G is a chestnut.

Lemma 4.1. *Let α be an external M -alternating path of a soliton graph G leading to a principal vertex v in some internal family \mathcal{F} . Then α can be extended to a soliton trail, the center of which lies in a family $\mathcal{H} \leq_G \mathcal{F}$.*

Proof. Every possible continuation β of α as an M -alternating path can only leave the family \mathcal{F} by entering another family $\mathcal{G} <_G \mathcal{F}$. It is therefore inevitable that, when β finally returns to itself, this will happen in a family $\mathcal{H} \leq_G \mathcal{F}$. The path β must eventually return to itself, since G is finite. \square

Lemma 4.2. *Let β be a soliton c -trail of a deterministic soliton graph G starting from $v \in Ext(G)$ with respect to state M . Then, starting from v , there exists no soliton trail with respect to M that is different from β .*

Proof. Assume, by contradiction, that an unwanted $\delta \neq \beta$ exists. Clearly, $c(\delta) = c(\beta)$, otherwise the soliton walks $s(\beta)$ and $s(\delta)$ would define two different transitions of \mathcal{A}_G in state M on input (v, v) . Therefore we have $h(\beta) \neq h(\delta)$. Starting from v , let z be the first vertex on both $h(\beta)$ and $h(\delta)$ where these paths split into two different directions (or just use a pair of parallel edges to reach the same vertex). Thus, $\beta = \gamma \parallel \beta'$ and $\delta = \gamma \parallel \delta'$, where γ is a suitable path from v to z , and β', δ' are M -alternating c -trails starting out from z on different edges. Clearly, the last edge of γ incident with z is positive, and the first edges of β' and δ' incident with z are negative. Therefore the walk

$$\chi = \gamma \parallel \beta' \parallel h(\beta')^R \parallel \delta' \parallel h(\delta')^R \parallel \gamma^R$$

is a soliton walk from v to v , which defines a self-transition of \mathcal{A}_G . This transition, however, is different from the one defined by the walks $s(\beta)$ and $s(\gamma)$; a contradiction. \square

Theorem 4.3. *Let G be a deterministic soliton graph with no impervious edges, and assume that G has a non-mandatory elementary component C lying in an internal family \mathcal{F} . Then C consists of a single even-length cycle, $\mathcal{F} = \{C\}$ and \mathcal{F} is a minimal element with respect to the partial order \leq_G .*

Proof. Let M be an arbitrary state of G , and α be a negative external M -alternating path from some $v \in Ext(G)$ to a principal vertex z of the root R of \mathcal{F} . Furthermore, let γ be an M -alternating cycle in C . Since C is accessible through z , we can fix a soliton c -trail β with respect to M such that β starts out from v and $c(\beta) = \gamma$. We can assume, without loss of generality, that $R = C$. For, we need to rule out the only possible scenario that is incompatible with this assumption, namely when R is a single mandatory edge $e = (z, w)$. Since in this case there are two-way members of the family \mathcal{F} (e.g. C), there exists an M -alternating R -ear (loop) ϵ around w . The loop ϵ gives rise to a soliton l -trail $\delta = \alpha\epsilon\epsilon$, which, by Lemma 4.2, cannot co-exist with β .

Now let us assume that $R = C$ has an allowed edge different from the ones along γ . Clearly, C then has another M -alternating cycle $\gamma' \neq \gamma$. As above, it is possible to extend α to a soliton c -trail β' with respect to M such that $c(\beta') = \gamma'$. Again, this contradicts Lemma 4.2, knowing that $\beta \neq \beta'$.

We have seen so far that $R = C$, and C is spanned by γ . Moreover, the principal and non-principal vertices determine the two classes of C 's canonical partition. No two principal vertices can be connected in C by an edge, since such an edge would be impervious in G . Suppose that there exists an edge e connecting two non-principal vertices u_1 and u_2 of γ . The edge e divides γ into two M -alternating loops χ_1 and χ_2 originating from u_1 and u_2 , respectively.

Also notice that vertex z on γ lies outside one of χ_1 and χ_2 . Consequently, α can again be extended to a soliton l-trail δ such that $c(\delta) = \chi_1$ or $c(\delta) = \chi_2$. (Remember that z and $u_{1(2)}$ are in different canonical classes.) A contradiction with Lemma 4.2 is reached, showing that $C = \gamma$.

It remains to be seen that $\mathcal{F} = \{C\}$, and \mathcal{F} is minimal with respect to \leq_G . Any M -alternating C -ear ϵ originating from two different non-principal vertices u_1 and u_2 is equivalent to an edge connecting u_1 directly with u_2 , and so need not be considered separately. On the other hand, if the ear ϵ was an M -alternating loop, then α could again be trivially extended to a soliton l-trail δ with $c(\delta) = \epsilon$, which is impossible. Now assume that there exists a family $\mathcal{G} <_G \mathcal{F}$, and continue α to obtain a negative external M -alternating path α' leading to a principal vertex of the root of \mathcal{G} . By Lemma 4.1, α' can further be extended to a soliton trail δ having its center in a family $\mathcal{H} \leq \mathcal{G}$. The trail δ is therefore different from β , which contradicts Lemma 4.2. The proof is now complete. \square

Theorem 4.4. *Let G be a connected deterministic soliton graph having no impervious edges. If G has a non-mandatory internal elementary component, then G is a chestnut.*

Proof. Induction on the number of non-degenerate elementary components of G . Assume that a non-mandatory internal elementary component C exists in G . If C is the only non-degenerate elementary component in G , then C is an internal family by itself, and the statement of the theorem follows directly from Theorem 4.3. Now let G have more than one non-degenerate elementary component, and suppose that the statement is true for all appropriate soliton graphs having fewer non-degenerate elementary components than G . Let \mathcal{F} denote the family of C . By Theorem 4.3, if \mathcal{F} is internal, then $\mathcal{F} = \{C\}$ and \mathcal{F} is minimal with respect to \leq_G . Let \mathcal{G} be a non-degenerate family such that $\mathcal{F} \leq_G \mathcal{G}$, and \mathcal{G} is either external or near-external. Clearly, if \mathcal{F} is external, then $\mathcal{G} = \mathcal{F}$. Otherwise \mathcal{G} can be found by stepping upwards in the partial order \leq_G starting from family \mathcal{F} , which is a minimal element by Theorem 4.3. Notice that this search must reject \mathcal{F} itself, as \mathcal{F} being near-external would imply that its sole member C is the only non-degenerate elementary component in G . Thus, in this case, $\mathcal{G} \neq \mathcal{F}$.

Let R be the root of \mathcal{G} . We are going to prove that

1. R is mandatory,
2. $\mathcal{G} = \{R\}$, and
3. there is exactly one forbidden edge incident with R 's unique non-principal (or non-external) vertex.

Fix a state M for G , and choose an M -alternating cycle γ in C . Since C is accessible from \mathcal{G} , γ can be extended to a soliton c-trail β from $v \in \text{Ext}(G)$, where v is either in R , or it is adjacent to a principal vertex in R .

Proof of Statement 1. Assume, by contradiction, that R is non-mandatory, and distinguish the following two cases.

Case 1: R is external. Let u be the vertex in R where the path $h(\beta)$ finally leaves this elementary component. By

[1], there exists a crossing δ in R from v to another external vertex z via u with respect to some state \bar{M}_R of R . Modify M , so that its restriction to R is replaced by \bar{M}_R , and let M' denote the resulting state. Clearly, the straight crossing δ between v and z in either direction is a possible transition of \mathcal{A}_G in state M' . On the other hand, this crossing can also make a detour to include γ through the appropriate section of β that starts at u . Notice, however, that this detour is only available from one direction, depending on whether the M' -positive edge incident with u on δ points toward v or z . Nevertheless, the co-existence of these two different transitions violates the deterministic property.

Case 2: R is internal. Trivially, there exists a soliton c-trail δ with respect to M starting from v such that $c(\delta)$ runs entirely in R ; a contradiction with Lemma 4.2.

Proof of Statement 2. As in the proof of Theorem 4.3, the existence of an R -ear as an M -alternating loop would immediately contradict Lemma 4.2.

Proof of Statement 3. Assume that there is more than one forbidden edge going out from the non-principal (or non-external) vertex of \mathcal{G} to different internal families of G . By Lemma 4.1, each of these edges can be made part of a suitable soliton trail in G with respect to M , starting from vertex v . Since β is also such a trail, a contradiction with Lemma 4.2 is inevitable.

Now we are ready to synthesize statements 1, 2, 3, and finish the proof of Theorem 4.4. It has turned out that the case $\mathcal{F} = \mathcal{G}$ is not possible. Detach the mandatory family \mathcal{G} from G , keeping the unique forbidden edge specified in statement 3 as an external edge in the remainder graph G' . Notice that, if R is internal, then its principal vertex can only be adjacent to external vertices, or else G would have impervious edges. Observe that G' is also deterministic, connected, has no impervious edges, and still has the non-mandatory internal elementary component C in it. Apply the induction hypothesis to establish G' as a chestnut. Finally, conclude that G is also a chestnut by sticking back the mandatory family \mathcal{G} onto G' . The proof is complete. \square

5 Reducing soliton graphs

A *redex* r in graph G consists of two adjacent edges $e = (u, z)$ and $f = (z, v)$ such that $u \neq v$ are both internal and the degree of z is 2. The vertex z is called the *center* of r , while u and v (e and f) are the two *focal* vertices (respectively, *focal* edges) of r .

Let r be a redex in G . *Contracting* r in G means creating a new graph G_r from G by deleting the center of r and merging the two focal vertices of r into one vertex s . Now suppose that G is a soliton graph. For a state M of G , let M_r denote the restriction of M to edges in G_r . Clearly, M_r is a state of G_r . Notice that the state M can be reconstructed from M_r in a unique way. In other words, the connection $M \mapsto M_r$ is a one-to-one correspondence between the states of G and those of G_r .

For any walk α in G , let $\text{trace}_r(\alpha)$ denote the restriction

of α to edges in G_r . It is easy to see that if α is a soliton walk in G with respect to M , then so is $trace_r(\alpha)$ in G_r with respect to M_r . Moreover, the soliton walk α can again be uniquely recovered from $trace_r(\alpha)$. Consequently, the connection $\alpha \mapsto trace_r(\alpha)$ is also a one-to-one correspondence between soliton walks in G and soliton walks in G_r . Furthermore, $M' = S(M, \alpha)$ holds in G iff $(M')_r = S(M_r, trace_r(\alpha))$ holds in G_r . We thus have proved the following statement.

Proposition 5.1. *The soliton automata \mathcal{A}_G and \mathcal{A}_{G_r} are isomorphic.*

Notice, furthermore, that if an alternating unit goes through both focal vertices of r , then it must do so along the center of r . As a consequence we have:

Proposition 5.2. *The function $trace_r$ establishes a one-to-one correspondence between the alternating units of G and those of G_r .*

It follows from the previous two propositions that every edge e of G_r is allowed in G_r iff e is allowed in G . As to the two focal edges of r , they can either be allowed or not in G , even when G_r is elementary. This issue is addressed by Proposition 5.3 below.

Proposition 5.3. *Let r be a redex in soliton graph G , and assume that G_r is elementary. Then G is elementary iff both focal edges of r are allowed in G , or, equivalently, each focal vertex of r has at least one allowed edge of G_r incident with it.*

Proof. It is sufficient to note that either focal edge of r is forbidden in G iff the other focal edge is mandatory. Moreover, an arbitrary internal edge e of G is mandatory iff all edges adjacent to e are forbidden. \square

Another natural simplifying operation on graphs is the removal of a loop from around a vertex v if, after the removal, v still remains internal. Such loops will be called *secondary*. Let G_v denote the graph obtained from G by removing a secondary loop e at vertex v . Clearly, if G is a soliton graph, then so is G_v , and the states of G_v are exactly the same as those of G . The automata \mathcal{A}_G and \mathcal{A}_{G_v} , however, need not be isomorphic. This follows from the fact that any external alternating walk reaching v on a positive edge can turn back in G after having made the loop e twice, while this may not be possible for the same walk without the presence of e . Nevertheless, it is still true that for every elementary soliton graph G , G is deterministic iff G_v is such.

There are loops, however, the removal of which preserves isomorphism of soliton automata. These loops are exactly the ones around the inaccessible vertices of G . Each such loop is impervious, so that its removal does not affect the automaton behavior of G .

6 General characterization of deterministic soliton graphs

Graph G is said to be *reduced* if it is free from redexes and secondary loops. A *generalized tree* is a connected graph not containing any even-length cycles. By this definition, the odd-length cycles possibly present in a generalized tree must be pairwise edge-disjoint, which explains the terminology.

The proof of the following statement is left to the reader as a simple exercise.

Lemma 6.1. *Let α be an alternating cycle with respect to state M of an elementary soliton graph G . Then G has an M -alternating cycle α' and a crossing β that intersects α' .*

Proposition 6.2. *If G is nondeterministic, then G has an alternating cycle with respect to some state M . Conversely, if G is elementary and has an alternating cycle with respect to some state M , then G is nondeterministic.*

Proof. Assume that G is nondeterministic. Then there exists a state M and soliton walks α, β connecting the same pair of external vertices in such a way that $S(M, \alpha) \neq S(M, \beta)$. Consider the network $N(S(M, \alpha), S(M, \beta))$. This network is not empty, and it consists of alternating cycles only.

Now let G be elementary, having an M -alternating cycle α . By Lemma 6.1 we can assume that G also has a crossing β with respect to the same state M that intersects α . Consider the network $N(S(M, \beta), S(M, \alpha))$. This network will contain a crossing β' different from β , yet connecting the same two external vertices v_1, v_2 . Thus, for the state $M' = S(M, \beta)$, \mathcal{A}_G has two different transitions on input (v_1, v_2) resulting in states $S(M', \beta')$ and $S(M', \beta) = M$, respectively. \square

The key step to our results in this section is Theorem 6.3 below. The proof of this theorem is rather complex, therefore we do not present it here. The interested reader is referred to [3] for a complete proof.

Theorem 6.3. *Let G be a reduced elementary soliton graph. If G contains an even-length cycle, then it also has an alternating cycle with respect to some state of G .*

For an arbitrary graph G , contract all redexes and remove all secondary loops in an iterative manner to obtain a reduced graph $r(G)$. Observe that this reduction procedure has the so called Church-Rosser property, that is, if G admits two different one-step reductions to graphs G_1 and G_2 , then either G_1 is isomorphic to G_2 , or G_1 and G_2 can further be reduced to a common graph $G_{1,2}$. In this context, one reduction step means contracting a redex or removing a single secondary loop. As an immediate consequence of the Church-Rosser property, the graph $r(G)$ above is unique up to graph isomorphism. In a similar fashion, the process of contracting all redexes and removing all impervious secondary loops is called *i-reduction*, and

the graph obtained from G after i -reduction is denoted by $r_i(G)$.

Theorem 6.4. *For any graph G , if $r(G)$ is a generalized tree, then G is a deterministic soliton graph. Conversely, if G is a deterministic elementary soliton graph, then $r(G)$ is a generalized tree.*

Proof. Clearly, G is a soliton graph iff $r(G)$ is such. By Proposition 5.2, if $r(G)$ is a generalized tree, then G does not contain alternating cycles with respect to any of its states. Proposition 6.2 then implies that G is deterministic. Conversely, if G is a deterministic elementary soliton graph, then so is $r(G)$, containing no alternating cycles with respect to any of its states. (See again Propositions 5.2 and 6.2.) Thus, by Theorem 6.3, $r(G)$ is a generalized tree. \square

Corollary 6.5. *An elementary soliton graph is deterministic iff it reduces to a generalized tree.*

Definition 2. A baby chestnut is a chestnut $\gamma + \alpha_1 + \dots + \alpha_k$ such that γ is a pair of parallel edges and each tree α_i ($1 \leq i \leq k$) consists of one edge or two adjacent edges.

Theorem 6.6. *Let G be a viable connected soliton graph possessing a non-mandatory internal elementary component. Then G is deterministic iff $r_i(G)$ is a baby chestnut.*

Proof. ‘Only if’ By Theorem 4.4, G is a chestnut augmented by some impervious edges connecting the outer internal vertices with each other. Since each internal inner vertex, different from the base ones, is the center of a redex, we can eliminate all of these vertices using reduction, except of course the last inner vertex in γ , which will no longer identify a redex. After removing the secondary impervious loops generated during redex elimination, $r_i(G)$ becomes a baby chestnut.

‘If’ Blowing up γ by inverse redex elimination, or stretching the trees α_i in this manner preserves the property of being a chestnut, and any impervious loops added during this procedure may only stretch into impervious edges. Thus, the graph resulting from a baby chestnut after any number of blow-ups and stretches is still a chestnut with some additional impervious edges. \square

Now we are ready to state the main result of this paper.

Theorem 6.7. *Let G be a connected viable soliton graph. Then G is deterministic iff it satisfies one of the following two conditions.*

1. G i -reduces to a baby chestnut.
2. Each external component of G reduces to a generalized tree, and the subgraph of G determined by its internal components has a unique perfect matching.

Proof. Immediate by Theorems 4.4, 6.6, and Corollary 6.5. \square

7 Conclusion

We have presented a detailed analysis of deterministic soliton graphs. First we proved that every connected deterministic soliton graph having no impervious edges, but possessing a non-mandatory internal elementary component is a chestnut. Then we introduced a simple reduction procedure on graphs, and showed that an elementary soliton graph is deterministic iff it reduces to a generalized tree. Using i -reduction, we could provide a yet simpler description of chestnut graphs, and gave a characterization of deterministic soliton graphs in general.

References

- [1] M. Bartha, E. Gombás, On graphs with perfect internal matchings, *Acta Cybernetica* **12** (1995), 111–124.
- [2] M. Bartha, M. Krész, Structuring the elementary components of graphs having a perfect internal matching, *Theoretical Computer Science* **299** (2003), 179–210.
- [3] M. Bartha, M. Krész, Deterministic soliton automata defined by elementary graphs, in *Kalmár Workshop on Logic and Computer Science*, Technical Report, University of Szeged, 2003, pp. 69–79.
- [4] F. L. Carter, A. Schultz, D. Duckworth, Soliton switching and its implications for molecular electronics, in *Molecular Electronic Devices II* (F. L. Carter Ed.), Marcel Dekker Inc., New York, 1987, pp. 149–182.
- [5] J. Dassow, H. Jürgensen, Soliton automata, *J. Comput. System Sci.* **40** (1990), 154–181.
- [6] L. Lovász, M. D. Plummer, *Matching Theory*, North Holland, Amsterdam, 1986.