

# Biconnectivity, Chain Decomposition and $st$ -Numbering Using $O(n)$ Bits

Sankardeep Chakraborty<sup>1</sup>, Venkatesh Raman<sup>2</sup>, and Srinivasa Rao Satti<sup>3</sup>

- 1 The Institute of Mathematical Sciences, Homi Bhabha National Institute (HBNI), Chennai, India  
sankardeep@imsc.res.in
- 2 The Institute of Mathematical Sciences, Homi Bhabha National Institute (HBNI), Chennai, India  
vraran@imsc.res.in
- 3 Seoul National University, Seoul, South Korea  
ssrao@cse.snu.ac.kr

---

## Abstract

Recent work by Elmasry et al. (STACS 2015) and Asano et al. (ISAAC 2014) reconsidered classical fundamental graph algorithms focusing on improving the space complexity. Elmasry et al. gave, among others, an implementation of depth first search (DFS) of a graph on  $n$  vertices and  $m$  edges, taking  $O(m \lg \lg n)$  time<sup>1</sup> using  $O(n)$  bits of space improving on the time bound of  $O(m \lg n)$  due to Asano et al. Subsequently Banerjee et al. (COCOON 2016) gave an  $O(m+n)$  time implementation using  $O(m+n)$  bits, for DFS and its classical applications (including testing for biconnectivity, and finding cut vertices and cut edges). Recently, Kammer et al. (MFCS 2016) gave an algorithm for testing biconnectivity using  $O(n + \min\{m, n \lg \lg n\})$  bits in linear time.

In this paper, we consider  $O(n)$  bits implementations of the classical applications of DFS. These include the problem of finding cut vertices, and biconnected components, chain decomposition and  $st$ -numbering. Classical algorithms for them typically use DFS and some  $\Omega(\lg n)$  bits of information at each node. Our  $O(n)$ -bit implementations for these problems take  $O(m \lg^c n \lg \lg n)$  time for some small constant  $c$  ( $c \leq 3$ ). Central to our implementation is a succinct representation of the DFS tree and a space efficient partitioning of the DFS tree into connected subtrees, which maybe of independent interest for space efficient graph algorithms.

**1998 ACM Subject Classification** F.1.1 Models of Computation, F.2.2 Nonnumerical Algorithms and Problems, G.2.2 Graph Theory

**Keywords and phrases** biconnectivity,  $st$ -number, chain decomposition, tree cover, space efficient algorithms, read-only memory

**Digital Object Identifier** 10.4230/LIPIcs.ISAAC.2016.22

## 1 Introduction

Motivated by the rapid growth of huge data sets (“big data”), space efficient algorithms are becoming increasingly important than ever before. The proliferation of handheld or embedded devices that are equipped with only a small amount of general-purpose memory provides another motivation for studying space efficient algorithms. In some of these devices, writing in the memory is a costly operation in terms of both speed and time than reading.

---

<sup>1</sup> We use  $\lg$  to denote logarithm to the base 2.



In such scenarios, algorithms that do not modify the input and use only a limited amount of work space are very much desired.

The standard model to study space efficient algorithms is the read-only memory model, and there is a rich history in computational complexity theory of such algorithms which use as little space as possible. In particular, L (also known as LSPACE or DLOGSPACE) is the complexity class containing decision problems that can be solved by a deterministic Turing machine using only logarithmic amount of work space for computation. There are several important algorithmic results [15, 12] for this class, the most celebrated being Reingold's method [30] for checking reachability between two vertices in an undirected graph. Barnes et al [7] gave a slightly sublinear space (using  $n/2^{\Theta(\sqrt{\lg n})}$  bits) algorithm for directed  $s$ - $t$  connectivity with polynomial running time. Space-efficient algorithms for classical selection and sorting problems [26, 27], and problems in computational geometry have also been studied [5, 6]. Recent work has focused on space requirement in special classes of graphs like planar and H-minor free graphs [9, 2].

For most of these algorithms using small space i.e., sublinear bits, their running time is often some polynomial of very high degree. Tompa [36] showed that for directed  $s$ - $t$  connectivity, if the number of bits available is  $o(n)$  then some natural algorithmic approaches to the problem require superpolynomial time. Thus it is sensible to focus (as in the case of some the recent papers like that of [16, 1, 4, 3, 24]) on designing algorithms that use  $O(n)$  bits of workspace. Our main objective here is to reduce the working space of the classical algorithms to  $O(n)$  bits with little or no penalty in running time. In these recent series of papers [16, 1, 4, 3, 24] space-efficient algorithms for only a few basic algorithmic graph problems are discussed: DFS, BFS, topological sort, strongly connected components, sparse spanning biconnected subgraph, among others. We add to this growing body of space-efficient algorithm design literature by providing such algorithms for a few more classical algorithmic graph problems, namely biconnectivity,  $st$ -numbering and chain decomposition.

## 1.1 Our results and organization of the paper

Our starting point is an  $O(m+n)$  time and  $O(n \lg(m/n))$  bits implementation for DFS and for finding a 'chain decomposition' using which we can find cut vertices, bridges, maximal biconnected components and ear decomposition (see Section 2 for definitions). This improves an earlier  $O(m+n)$  time and  $O(m+n)$  bits implementation [3] (see Theorem 4). The space used by these algorithms, for some ranges of  $m$  (say  $\Theta(n \lg \lg n)^c$  for some constant  $c$ ), is even better than that of the recent work by Kammer et al [24], that computes cut vertices using  $O(n + \min\{m, n \lg \lg n\})$  bits. This implementation appears in Section 3.

Chain decomposition is an important preprocessing routine for an algorithm to find cut vertices and biconnected components and also to test 3-connectivity [31] among others. In Section 5, we give an algorithm that takes  $O(m \lg^3 n \lg \lg n)$  time using  $O(n)$  bits, improving on previous implementations that took  $\Omega(n \lg n)$  bits or  $\Theta(m+n)$  bits.

In Section 6, we give improved algorithms for finding cut vertices and biconnected components by giving a space efficient implementation of Tarjan's classical lowpoint algorithm. This takes  $O(m \lg n \lg \lg n)$  time.

Given a biconnected graph, and two distinguished vertices  $s$  and  $t$ ,  $st$ -numbering is a numbering of the vertices of the graph so that  $s$  gets the smallest number,  $t$  gets the largest and every other vertex is adjacent both to a lower-numbered and to a higher-numbered vertex. Finding an  $st$ -numbering is an important preprocessing routine for a planarity testing algorithm. In Section 7, we give an algorithm to determine an  $st$ -numbering of a

biconnected graph that takes  $O(m \lg^2 n \lg \lg n)$  time using  $O(n)$  bits. This improves the earlier implementations that take  $\Omega(n \lg n)$  bits.

**Techniques.** There are several approaches to find cut vertices and biconnected components. An algorithm due to Tarjan [34] is the standard ‘textbook’ algorithm, and another due to Schmidt [32] is based on chain decomposition of graphs. Both these approaches compute DFS and process the DFS tree in specific order maintaining some auxiliary information of the nodes. To implement these in  $O(n)$  bits, our main idea is to process the nodes of the DFS tree in batches of  $O(n/\lg n)$  nodes. Towards that, we use tree-cover algorithms (that are used in succinct representations of trees) that partition a tree into connected subtrees. This is described in detail in Section 4.

**Model of Computation.** Like all the recent research that focused on space-efficient graph algorithms [16, 1, 4, 3, 24], here also we assume that the input graph is given in a read-only memory (and so cannot be modified). If an algorithm must do some outputting, this is done on a separate write-only memory. When something is written to this memory, the information cannot be read or rewritten again. So the input is “read only” and the output is “write only”. In addition to the input and the output media, a limited random-access workspace is available. The data on this workspace is manipulated at word level as in the standard word RAM model, where the machine consists of words of size  $w = \Omega(\lg n)$  bits; and any logical, arithmetic, and bitwise operations involving a constant number of words take a constant amount of time. We count space in terms of the number of bits in the workspace used by the algorithms. Historically, this model is called the *register input model* and it was introduced by Frederickson [20] while studying some problems related to sorting and selection. We assume that the input graph  $G = (V, E)$  is represented using *adjacency array*, i.e., given a vertex  $v$  and an integer  $k$ , we can access the  $k$ th neighbor of vertex  $v$  in constant time. This representation was used in [16, 3, 24] recently to design various space efficient graph algorithms. We use  $n$  and  $m$  to denote the number of vertices and the number of edges respectively, in the input graph  $G$ . Throughout the paper, we assume that the input graph is a connected graph, and hence  $m \geq n - 1$ .

## 2 Preliminaries

**Rank-Select.** Given a bitvector  $B$  of length  $n$ , the rank and select operations are defined as follows:

- $\text{rank}_a(i, B) =$  number of occurrences of  $a \in \{0, 1\}$  in  $B[1, i]$ , for  $1 \leq i \leq n$ ;
- $\text{select}_a(i, B) =$  position in  $B$  of the  $i$ th occurrence of  $a \in \{0, 1\}$ .

The following theorem gives an efficient structure to support these operations.

► **Theorem 1** ([11]). *Given a bitstring  $B$  of length  $n$ , one can construct a  $o(n)$ -bit auxiliary structure to support rank and select operations in  $O(1)$  time. Also, such a structure can be constructed from the given bitstring in  $O(n)$  time.*

**Space-efficient DFS.** Elmasry et al. [16] showed the following tradeoff result for DFS,

► **Theorem 2** ([16]). *For every function  $t : \mathbb{N} \rightarrow \mathbb{N}$  such that  $t(n)$  can be computed within the resource bound of this theorem (e.g., in  $O(n)$  time using  $O(n)$  bits), the vertices of a graph  $G$  can be visited in depth first order in  $O((m + n)t(n))$  time with  $O(n + n \frac{\lg \lg n}{t(n)})$  bits.*

In particular, fixing  $t(n) = O(\lg \lg n)$ , we obtain a DFS implementation which runs in  $O(m \lg \lg n)$  time using  $O(n)$  bits. We build on top of this DFS algorithm to design all of our space-efficient algorithms.

**Graph theoretic terminology.** A cut vertex in an undirected graph is a vertex that when removed (with its incident edges) from a graph creates more components than previously in the graph. Similarly, a bridge is an edge that when removed (the vertices stay in place) from a graph creates more components than previously in the graph. A graph is biconnected if it is connected and contains at least 3 vertices, but no cut vertex. A graph is 2-edge-connected if it is connected and contains at least 2 vertices, but no bridge. Let  $G = (V, E)$  be a biconnected graph and  $s \neq t \in V$ . An ordering  $s = v_1, v_2, \dots, v_n = t$  of the vertices of  $G$  is called an  $st$ -ordering, if for all vertices  $v_j, 1 < j < n$ , there exist  $1 \leq i < j < k \leq n$  such that  $\{v_i, v_j\}, \{v_j, v_k\} \in E$ . It is well-known that  $G$  is biconnected if and only if, for every edge  $\{s, t\} \in E$ , it has an  $st$ -ordering.

**Chain decomposition and its application.** Schmidt [31] introduced a decomposition of the input graph that partitions the edge set of the graph into cycles and paths, called chains, and used this to design an algorithm to find cut vertices and biconnected components [32] and also to test 3-connectivity [31] among others. In this section we discuss the details of the decomposition algorithm and some of the applications for which we give space efficient implementations in the paper later.

The algorithm first performs a depth first search on  $G$ . Let  $r$  be the root of the DFS tree  $T$ . DFS assigns an index to every vertex  $v$ , namely, the time vertex  $v$  is discovered for the first time during DFS – call it the depth-first-index of  $v$  ( $DFI(v)$ ). Imagine that the back edges are directed away from  $r$  and the tree edges are directed towards  $r$ . The algorithm decomposes the graph into a set of paths and cycles called chains as follows. First we mark all the vertices as unvisited. Then we visit every vertex starting at  $r$  in the increasing order of DFI, and do the following. For every back edge  $e$  that originates at  $v$ , we traverse a directed cycle or a path. This begins with  $v$  and the back edge  $e$  and proceeds along the tree towards the root and stops at the first visited vertex or the root. During this step, we mark every encountered vertex as visited. This forms the first chain. Then we proceed with the next back edge at  $v$ , if any, or move towards the next vertex in the increasing DFI order and continue the process. Let  $D$  be the collection of all such cycles and paths. Notice that the cardinality of this set is exactly the same as the number of back edges in the DFS tree as each back edge contributes to a cycle or a path. Also, as initially every vertex is unvisited, the first chain would be a cycle as it would end in the starting vertex. Schmidt proved the following theorem.

► **Theorem 3** ([32]). *Let  $D$  be a chain decomposition of a connected graph  $G(V, E)$ . Then  $G$  is 2-edge-connected if and only if the chains in  $D$  partition  $E$ . Also,  $G$  is 2-vertex-connected if and only if  $\delta(G) \geq 2$  (where  $\delta(G)$  denotes the minimum degree of  $G$ ) and  $D_1$  is the only cycle in the set  $D$  where  $D_1$  is the first chain in the decomposition. An edge  $e$  in  $G$  is bridge if and only if  $e$  is not contained in any chain in  $D$ . A vertex  $v$  in  $G$  is a cut vertex if and only if  $v$  is the first vertex of a cycle in  $D \setminus D_1$ .*

Banerjee et al. [3] gave a space-efficient implementation of Theorem 3. En route they also provided an improved implementation for DFS (over Theorem 2) in sparse graphs ( $m = O(n)$  edges). In particular, they proved the following,

► **Theorem 4** ([3]). *A DFS traversal of an undirected or directed graph  $G$  can be performed in  $O(m + n)$  time using  $O(m + n)$  bits. In the same amount of time and space, given a connected undirected graph  $G$ , we can perform a chain decomposition of  $G$ , and using that we can determine whether  $G$  is 2-vertex (and/or edge) connected. If not, in the same amount of time and space, we can compute all the bridges, cut vertices, and output 2-vertex (and edge) connected components.*

Kammer et al. [24] recently improved the space bound for finding cut vertices, still using linear time to  $O(n + \min\{m, n \lg \lg n\})$  bits.

### 3 DFS and applications using $O(n \lg(m/n))$ bits

One can easily implement the tests in Theorem 3 in  $O(m)$  time using  $O(m)$  words, by storing the DFIs and the entire chain decomposition,  $D$ . It is not too hard to improve the space to  $O(n)$  words, still maintaining the  $O(m)$  running time. Theorem 4 shows how to perform the tests using  $O(m + n)$  bits and  $O(m)$  time. The central idea there is to maintain the DFS tree using  $O(m + n)$  bits using an unary encoding of the degree sequence of the graph. We first show how the space for the DFS tree representation can be improved to  $O(n \lg m/n)$  bits.

► **Lemma 5.** *Given the adjacency array representation of an undirected graph  $G$  on  $n$  vertices with  $m$  edges, using  $O(m)$  time, one can construct an auxiliary structure of size  $O(n \lg(m/n))$  bits that can store a “pointer” into an arbitrary position within the adjacency array of each vertex. Also, updating any of these pointers (within the adjacency array) takes  $O(1)$  time.*

**Proof.** We first scan the adjacency array of each vertex and construct a bitvector  $B$  as follows: starting with an empty bitvector  $B$ , for  $1 \leq i \leq n$ , if  $d_i$  is the length of the adjacency array of vertex  $v_i$  (i.e., its degree), then we append the string  $0^{\lceil \lg d_i \rceil - 1} 1$  to  $B$ . The length of  $B$  is  $\sum_{i=1}^n \lceil \lg d_i \rceil$ , which is bounded by  $O(n \lg(m/n))$ . We construct auxiliary structures to support *select* queries on  $B$  in constant time, using Theorem 1. We now construct another bitvector  $P$  of the same size as  $B$ , which stores pointers into the adjacency arrays of each vertex. The pointer into the adjacency array of vertex  $v_i$  is stored using the  $\lceil \lg d_i \rceil$  bits in  $P$  from position  $\text{select}(i - 1, B) + 1$  to position  $\text{select}(i, B)$ , where  $\text{select}(0, B)$  is defined to be 0. Now, using *select* operations on  $B$  and using constant time word-level read/write operations, one can access and/or modify these pointers in constant time. ◀

► **Lemma 6.** *Given a graph  $G$  with  $n$  vertices and  $m$  edges, in the adjacency array representation in the read-only memory model, the representation of a DFS tree can be stored using  $O(n \lg(m/n))$  additional bits, which can be constructed on the fly during the DFS algorithm.*

**Proof.** We use the representation of Lemma 5 to store *parent* pointers into the adjacency array of each vertex. In particular, whenever the DFS outputs an edge  $(u, v)$ , where  $u$  is the parent of  $v$ , we scan the adjacency array of  $v$  to find  $u$  and store a pointer to that position (within the adjacency array of  $v$ ). The additional time for scanning the adjacency arrays adds upto  $O(m)$  which would be subsumed by the running time of the DFS algorithm. ◀

We call the representation of the DFS tree of Lemma 6 as the *parent pointer representation*. Now given Lemma 5 and 6, we can simulate the DFS algorithm of [3] (Theorem 4) to obtain an  $O(n \lg(m/n))$  bits and  $O(m + n)$  time (see [3] for details) DFS implementation. The proof of Theorem 4 then uses another  $O(m + n)$  bits to construct the chain decomposition of  $G$  and perform the tests as mentioned in Theorem 3, and we show here how even the space for the construction of a chain decomposition and performing the tests can be improved. We summarize our results in the following theorem below:

► **Theorem 7** ( $\spadesuit^2$ ). *A DFS traversal of an undirected or directed graph  $G$  can be performed in  $O(m + n)$  time using  $O(n \lg(m/n))$  bits of space. In the same amount of time and space, given a connected undirected graph  $G$ , we can perform a chain decomposition of  $G$ , and using that we can determine whether  $G$  is 2-vertex (and/or edge) connected. If not, in the same amount of time and space, we can compute all the bridges, cut vertices, and output 2-vertex (and edge) connected components.*

The above result for DFS improves the tradeoff result of Theorem 2 for relatively sparse graphs. Specifically, to achieve  $O(m + n)$  time for DFS, the algorithm of Theorem 2 uses  $O(n \lg \lg n)$  bits. This is  $\Omega(n \lg(m/n))$  for all values of  $m$  where  $m = O(n \lg n)$ . Hence, for sparse graphs we obtain a better tradeoff. Also, it improves the space bound of Theorem 4, from  $O(m + n)$  to  $O(n \lg(m/n))$ , while maintaining the same linear running time. In addition, it improves the algorithm for finding the cut vertices by Kammer et al. [24] from  $O(n + \min\{m, n \lg \lg n\})$  to  $O(n \lg(m/n))$ .

#### 4 Tree Cover and Space Efficient Construction

Central to all of our algorithms is a decomposition of the DFS tree. For this we use the well-known tree covering technique which was first proposed by Geary et al. [21] in the context of succinct representation of rooted ordered trees. The high level idea is to decompose the tree into subtrees called *minitrees*, and further decompose the mini-trees into yet smaller subtrees called *microtrees*. The microtrees are tiny enough to be stored in a compact table. The root of a minitree can be shared by several other minitrees. To represent the tree, we only have to represent the connections and links between the subtrees. Later He et al. [23] extended this approach to produce a representation which supports several additional operations. Farzan and Munro [18] modified the tree covering algorithm of [21] so that each minitree has at most one node, other than the root of the minitree, that is connected to the root of another minitree. This simplifies the representation of the tree, and guarantees that in each minitree, there exists at most one non-root node which is connected to (the root of) another minitree. The tree decomposition method of Farzan and Munro [18] is summarized in the following theorem:

► **Theorem 8** ([18]). *A rooted ordered tree with  $n$  nodes can be decomposed into  $\Theta(n/L)$  minitrees of size at most  $2L$  which are pairwise disjoint aside from the minitree roots. Furthermore, aside from edges stemming from the minitree root, there is at most one edge leaving a node of a minitree to its child in another minitree. The decomposition can be performed in linear time.*

In our algorithms, we apply Theorem 8 with  $L = n/\lg n$ . For this parameter  $L$ , since the number of minitrees is only  $O(\lg n)$ , we can represent the structure of the minitrees within the original tree (i.e., how the minitrees are connected with each other) using  $O(\lg^2 n)$  bits. The decomposition algorithm of [18] ensures that each minitree has at most one ‘child’ minitree (other than the minitrees that share its root) in this structure. (We use this property crucially in our algorithms.) We refer to this as the *minitree-structure*.

Explicitly storing all the minitrees requires  $\omega(n)$  bits. One way to represent them efficiently is to store them using any linear-bit encoding of a tree. But this representation doesn’t allow us to efficiently compute the preorder numbers of the nodes, for example. Instead, we encode

<sup>2</sup> For proofs of results marked with ( $\spadesuit$ ), please refer to the full version [10].

the entire tree structure using a linear-bit encoding, and store pointers into this encoding to represent the minitrees, as described below. We first encode the tree using the *balanced parenthesis* (BP) representation [28], summarized in the following theorem.

► **Theorem 9** ([28]). *Given a rooted ordered tree  $T$  on  $n$  nodes, it can be represented as a sequence of balanced parentheses of length  $2n$ . Using an additional  $o(n)$  bits, we can support subtree size and navigational queries on  $T$  based on preorder and postorder.*

We now represent each minitree by storing pointers to the set of all *chunks* in the BP representation that together constitute the minitree. Farzan et al. [19, Lemma 2] show that each minitree is split into a constant number of consecutive chunks in the BP sequence. Hence, one can store a representation of the minitrees by storing a bitvector of length  $n$  that marks the starting positions of these chunks in the BP sequence, together with an  $O(\lg^2 n)$ -bit structure that stores, for each minitree, pointers to all the chunks in BP sequence indicating the starting positions of the chunks corresponding to the minitrees. The bit vector has  $O(\lg n)$  1's since there are  $O(\lg n)$  minitrees, and each minitree is split into a constant number of chunks. We refer to the representation obtained using this tree covering (TC) approach as the TC representation of the tree.

The following lemma shows that we can construct the TC representation of the DFS tree of a given graph, using  $O(n)$  additional bits.

► **Lemma 10** (♠). *Given a graph  $G$  on  $n$  vertices and  $m$  edges, if there is an algorithm that takes  $t(n, m)$  time and  $s(n, m)$  bits to perform DFS on  $G$ , then one can create the TC representation of the DFS tree in  $t(n, m) + O(n)$  time, using  $s(n, m) + O(n)$  bits.*

We use the following lemma in the description of our algorithms.

► **Lemma 11** (♠). *Let  $G$  be a graph, and  $T$  be its DFS tree. If there is an algorithm that takes  $t(n, m)$  time and  $s(n, m)$  bits to perform DFS on  $G$ , then, using  $s(n, m) + O(n)$  bits, one can reconstruct any minitree given by its ranges in the BP sequence of the TC representation of  $T$ , along with the labels of the corresponding nodes in the graph in  $O(t(n, m))$  time.*

## 5 Chain decomposition using $O(n)$ bits

Theorem 7 gives a chain decomposition algorithm that runs in  $O(m + n)$  time, using  $O(n \lg(m/n))$  bits. In this section we describe how one can implement Schmidt's chain decomposition algorithm described in Section 2 using only  $O(n)$  bits using our partition of the DFS tree of Section 4. The main idea of our implementation is to process all the back edges out of each minitree, in the *preorder* of the minitrees. Also, when processing back edges out of a minitree  $\tau$ , we process all the back edges that go from  $\tau$  to the other minitrees in their *postorder*, processing all the edges from  $\tau$  to a minitree  $\tau_1$  before processing any other back edges going out of  $\tau$  to a different minitree. This requires us to go through all the edges out of each minitree at most  $O(\lg n)$  (number of minitrees) times (although it is subsumed by the other parts of the computation, and doesn't affect the overall running time). Thus the order in which we process the back edges is different from the order in which we process them in Schmidt's algorithm, but we argue that this does not affect the correctness of the algorithm. In particular, we observe that Schmidt's algorithm correctly produces a chain decomposition

- even if we change the order in which we process vertices to any other order (instead of preorder), as long as we process a vertex  $v$  only after all its ancestors are also processed – for example, in level order.



This also implies that as long as we process the back edges coming to a vertex  $v$  (from any of its descendants) only after we process all the back edges going to any of its ancestors from any of  $v$ 's descendants, we can produce a chain decomposition correctly. To process (i.e., to output the chain containing) a back edge  $(u, v)$  between a pair of minitrees  $\tau_1$  and  $\tau_2$ , where  $u$  belongs to  $\tau_1$ ,  $v$  belongs to  $\tau_2$ , and  $\tau_1$  is an ancestor of  $\tau_2$  in the minitree-structure, we first output the edge  $(u, v)$ , and then traverse the path from  $v$  to the root of  $\tau_2$ , outputting all the traversed edges. We then start another DFS to produce the minitree  $\tau_p$  containing the parent  $p$  of the root of  $\tau_2$ , and output the path from  $p$  to the root of  $\tau_p$ , and continue the process until we reach a vertex that has already been output as part of any chain (including the current chain). We maintain a bitvector of length  $n$  to keep track of the vertices that have been output as part of any chain, to perform this efficiently. A crucial observation that we use in bounding the runtime is that once we produce a minitree  $\tau_p$  for a particular pair  $(\tau_1, \tau_2)$  of minitrees, we don't need to produce it again, as the root of  $\tau_2$  will be marked after the first time we output it as part of a chain. Also, once we generate the two minitrees  $\tau_1$  and  $\tau_2$ , we go through all the vertices of  $\tau_1$  in preorder, and process all the edges that go between  $\tau_1$  and  $\tau_2$ . For a particular minitree  $\tau_1$ , once we process the back edges between  $\tau_1$  and all its descendant minitrees (i.e., descendants of the node corresponding to  $\tau_1$  in the minitree-structure), we finally process all the back edges that go within the minitree  $\tau_1$ .

The time taken for the initial part, where we construct the DFS tree, decompose it into minitrees, and construct the auxiliary structures, is  $O(m \lg \lg n)$ , using Theorem 2 with  $t(n) = \lg \lg n$ . The running time of the algorithm is dominated by the cost of processing the back edges. For each pair of minitrees, we may, in the worst-case, need to generate  $O(\lg n)$  minitrees. Since there are  $O(\lg^2 n)$  pairs of minitrees, and since generating each minitree requires  $O(m \lg \lg n)$  time (using the DFS algorithm), the total running time is  $O(m \lg^3 n \lg \lg n)$ . Thus, we obtain the following.

► **Theorem 12.** *Given an undirected graph  $G$  on  $n$  vertices and  $m$  edges, we can output a chain decomposition of  $G$  in  $O(m \lg^3 n \lg \lg n)$  time using  $O(n)$  bits of space.*

## 6 Finding cut vertices and biconnected components using $O(n)$ bits

A naïve algorithm to test for biconnectivity of a graph  $G = (V, E)$  is to check if  $(V \setminus \{v\}, E)$  is connected, for each  $v \in V$ . Using the  $O(n)$  bits and  $O(m + n)$  time BFS algorithm [3] for checking connectivity, this gives a simple  $O(n)$  bits algorithm running in time  $O(mn)$ . Another approach is to use Theorem 12, as in the proof of Theorem 7, to test biconnectivity and output cut vertices in  $O(m \lg^3 n \lg \lg n)$  time using  $O(n)$ .

Here we show that using  $O(n)$  bits we can design an even faster algorithm running in  $O(m \lg n \lg \lg n)$  time. If  $G$  is not biconnected, then we also show how to find out all the cut-vertices and biconnected components within the same time and space bounds. We implement the classical low-point algorithm of Tarjan [34]. Recall that, the algorithm computes for every vertex  $v$ , a value  $\text{lowpoint}[v]$  which is defined as

$$\text{lowpoint}[v] = \min\{DFI(v) \cup \{\text{lowpoint}[s] \mid s \text{ is a child of } v\} \cup \{DFI(w) \mid (v, w) \text{ is a back-edge}\}$$

Tarjan proved that if vertex  $v$  is not the root, then  $v$  is a cut vertex if and only if  $v$  has a child  $s$  such that  $\text{lowpoint}[s] \geq v$ . (The root of a DFS tree is a cut vertex if and only if the root has more than one child.) Since the lowpoint values requires  $\Omega(n \lg n)$  bits in the worst case, this poses the challenge of efficiently testing the condition for biconnectivity when only  $O(n)$  bits. To deal with this, as in the case of the chain decomposition algorithm, we compute lowpoint values in  $O(\lg n)$  batches using our tree covering algorithm. Cut vertices



encountered in the process, if at all, are stored in a separate bitmap. We show that each batch can be processed in  $O(m \lg \lg n)$  time using DFS, resulting in an overall runtime of  $O(m \lg n \lg \lg n)$ .

### 6.1 Computing lowpoint, cut vertices and biconnected components

We first obtain a TC representation of the DFS tree using the decomposition algorithm of Theorem 8 with  $L = n/\lg n$ . We then *process* each minitree, in the postorder of the minitrees in the minitree structure. To process a minitree, we compute the lowpoint values of each of the nodes in the minitree (except possibly the root). in overall  $O(m)$  time. During the processing of any minitree, if we determine that a vertex is a cut vertex, we store this information by marking the corresponding node in a separate bit vector. Each minitree can be reconstructed in  $O(m \lg \lg n)$  time using Lemma 11. The lowpoint value of a node is a function of the lowpoints of all its children. However the root of a minitree may have children in other minitrees. Hence for the root of the minitree, we store the partial lowpoint value (till that point) which will be used to update its value when all its subtrees have computed their lowpoint values (possibly in other minitrees). Computing the lowpoint values in each of the minitrees is done in a two step process. In the first step, we compute the DFI number of the deepest back edge node of each node in the minitree. Here the deepest back edge node of a node  $v$  is defined the smallest *DFI* value among the vertices  $w$  such that  $(v, w)$  is a back edge. Banerjee et al. show in [3] how one can compute the deepest back edge from any node while discussing a space-efficient implementation for computing a sparse spanning biconnected subgraph of a given biconnected graph. The corresponding algorithm makes two passes of DFS and hence takes  $O(m \lg \lg n)$  time using  $O(n)$  bits. We use that subroutine here to compute the deepest back edges. As there are only  $\Theta(n/\lg n)$  nodes, we have space to store these values. In the second step, we do another DFS starting at the root of this minitree and compute the lowpoint values as we will do in a normal DFS (as deepest back edge values have been stored).

To compute the effect of the roots of the minitrees on the lowpoint computation, we store various  $\Theta(\lg n)$  bit information with each of the  $\Theta(\lg n)$  minitree roots including their partial/full lowpoint values, the rank of its first/last child in its subtree. After we process one minitree, we generate the next minitree, in postorder, and process it in a similar fashion and continue until we exhaust all the minitrees.

As we can mark all the cut vertices (if any) in a bitvector of length  $n$ , reporting them and computing 2-connected components is a routine task. Clearly we have taken  $O(n)$  space and the total running time is  $O(m \lg \lg n \lg n)$  as we run the DFS algorithm  $O(\lg n)$  times. We formalize our theorem below.

► **Theorem 13.** *Given an undirected graph  $G$ , in  $O(m \lg n \lg \lg n)$  time and  $O(n)$  bits of space we can determine whether  $G$  is 2-vertex connected. If not, in the same amount of time and space, we can compute all the cut vertices of the graph and also output all the 2-vertex connected components.*

## 7 st-numbering

The *st*-ordering of vertices of an undirected graph is a fundamental tool for many graph algorithms, e.g. in planarity testing, graph drawing. The first linear-time algorithm for *st*-ordering the vertices of a biconnected graph is due to Even and Tarjan [17], and is further simplified by Ebert [14], Tarjan [35] and Brandes [8]. All these algorithms, however, preprocess

the graph using depth-first search, essentially to compute lowpoints which in turn determine an (implicit) open ear decomposition. A second traversal is required to compute the actual  $st$ -ordering. All of these algorithms take  $O(n \lg n)$  bits of space. We give an  $O(n)$  bits implementation of Tarjan's [35] algorithm. We first describe the two pass classical algorithm of Tarjan without worrying about the space requirement. Let us denote by  $p(v)$  the parent of the vertex  $v$  in the DFS tree.  $DFI(v)$  and  $lowpoint(v)$  have the usual meaning as defined previously. The first pass is a depth first search during which for every vertex  $v$ ,  $p(v)$ ,  $DFI(v)$  and  $lowpoint(v)$  are computed and stored. The second pass constructs a list  $L$ , which is initialized with  $[s, t]$ , such that if the vertices are numbered in the order in which they occur in  $L$ , then we obtain an  $st$ -ordering. In addition, we also have a sign array of  $n$  bits, initialized with  $sign[s] = -$ . The second pass is a preorder traversal of the spanning tree starting from the root  $s$  of the DFS tree. It is easy to see that the procedure runs in linear time using  $O(n \lg n)$  bits of space. To make it space efficient, we use ideas similar to our biconnectivity algorithm. At a high level, we generate the lowpoint values of the first  $n/\lg n$  vertices in depth first order and process them. Due to space restriction, we cannot store the list  $L$  as in Tarjan's algorithm; instead we use the BP sequence of the DFS tree and augment it with some extra information to 'encode' the final  $st$ -ordering, as described below.

Similar to our algorithms in the last two sections, this algorithm also runs in  $O(\lg n)$  phases and in each phase it processes  $n/\lg n$  vertices. In the first phase, to obtain the lowpoint values of the first  $n/\lg n$  vertices in depth first order, we run as in our biconnectivity algorithm a procedure to store explicitly for these vertices their lowpoint values in an array. Also during the execution of the biconnectivity algorithm, the BP sequence is generated and stored in the BP array. We create two more arrays, of size  $n$  bits, that have one to one correspondence with the BP array. First array is for storing the sign for every vertex as in Tarjan's algorithm, and call it Sign. To simulate the effect of the list  $L$ , we create the second array, called  $P$ , where we store the relative position, i.e., "before" or "after", of every vertex with respect to its parent. Namely, if  $u$  is the parent of  $v$ , and  $v$  comes before (after, respectively)  $u$  in the list  $L$  in Algorithm 3, then we store  $P[v] = b$  ( $P[v] = a$ , respectively). One crucial observation is that, even though the list  $L$  is dynamic, the relative position of the vertex  $v$  does not change with respect to the position of  $u$ , and is determined at the time of insertion of  $v$  into the list  $L$  (new vertices may be added between  $u$  and  $v$  later). In what follows, we show how to use the BP sequence, and the array  $P$  to emulate the effect of list  $L$  and produce the  $st$ -ordering.

We first describe how to reconstruct the list  $L$  using the BP sequence and the  $P$  array. The main observation we use in the reconstruction  $L$  is that a node  $v$  appears in  $L$  after all the nodes in its child subtrees whose roots are marked with  $b$  in  $P$ , and also before all the nodes in its child subtrees whose roots are marked with  $a$  in  $P$ . Also, all the nodes in a subtree appear "together" (consecutively) in the list  $L$ . Thus by looking at the  $P[v]$  values of all the children of a node  $u$ , and computing their subtree sizes, we can determine the position in  $L$  of  $u$  among all the nodes in its subtree. With this approach, we can reconstruct the list  $L$ , and hence output the  $st$ -numbers of all the nodes in linear time, if  $L$  can be stored in memory - which requires  $O(n \lg n)$  bits. Now to perform this step with  $O(n)$  bits, we repeat the whole process of reconstruction  $\lg n$  times, where in the  $i$ -th iteration, we reproduce sublist  $L[(i-1)n/\lg n + 1, \dots, in/\lg n]$  - by ignoring any node that falls outside this range - and reporting all these nodes with  $st$ -numbers in the range  $[(i-1)n/\lg n + 1, \dots, in/\lg n]$ . As each of these reconstruction takes  $O(m \lg n \lg \lg n)$  time, we obtain the following.

► **Theorem 14.** *Given an undirected biconnected graph  $G$  on  $n$  vertices and  $m$  edges, and two distinct vertices  $s$  and  $t$ , we can output an  $st$ -numbering of all the vertices of  $G$  in  $O(m \lg^2 n \lg \lg n)$  time, using  $O(n)$  bits of space.*

## 7.1 An application of $st$ -numbering

Given vertices  $a_1, \dots, a_k$  of a graph  $G$  and natural numbers  $c_1, \dots, c_k$  with  $c_1 + \dots + c_k = n$ , we want to find a partition of  $V$  into sets  $V_1, \dots, V_k$  with  $a_i \in V_i$  and  $|V_i| = c_i$  for every  $i$  such that every set  $V_i$  induces a connected graph in  $G$ . This problem is called the  $k$ -partitioning problem. The problem is NP-hard even when  $k = 2$ ,  $G$  is bipartite and the condition  $a_i \in V_i$  is relaxed [13]. But, Györi [22] and Lovász [25] proved that such a partition always exists if the input graph is  $k$ -connected. Thus, let  $G$  be  $k$ -connected. In particular, if  $k = 2$ , the  $k$ -partitioning problem can be solved in the following manner [33, 29]: Let  $v_1 := a_1$  and  $v_n := a_2$ , compute an  $v_1 v_n$ -numbering  $v_1, v_2, \dots, v_n$  and note that for any vertex  $v_i$  (in particular for  $i = c_1$ ) the graphs induced by  $v_1, \dots, v_i$  and by  $v_i, \dots, v_n$  are connected. Thus applying Theorem 14, we obtain the following:

► **Theorem 15.** *Given an undirected biconnected graph  $G$ , two distinct vertices  $a_1, a_2$ , and two natural numbers  $c_1, c_2$  such that  $c_1 + c_2 = n$ , we can obtain a partition  $(V_1, V_2)$  of the vertex set  $V$  of  $G$  in  $O(m \lg^2 n \lg \lg n)$  time, using  $O(n)$  bits of space, such that  $a_1 \in V_1$  and  $a_2 \in V_2$ ,  $|V_1| = c_1$ ,  $|V_2| = c_2$ , and both  $V_1$  and  $V_2$  induce connected subgraph on  $G$ .*

## 8 Conclusions and Open Problems

We have given  $O(m \lg^c n)$ -time,  $O(n)$ -bit algorithms for a number of important applications of DFS. Obtaining linear time algorithms for them while maintaining  $O(n)$  bits of space usage is an interesting open problem. Note that while there are  $O(n)$ -bit,  $O(m + n)$ -time algorithms for BFS, obtaining such an implementation for DFS is open. Another open problem is whether our  $O(n)$ -bit  $st$ -numbering algorithm can be used to design a  $O(m \lg^c n)$  time planarity test using  $O(n)$  bits of extra space.

---

### References

- 1 T. Asano, T. Izumi, M. Kiyomi, M. Konagaya, H. Ono, Y. Otachi, P. Schweitzer, J. Tarui, and R. Uehara. Depth-first search using  $O(n)$  bits. In *25th ISAAC*, pages 553–564, 2014. doi:10.1007/978-3-319-13075-0\_44.
- 2 T. Asano, D.G. Kirkpatrick, K. Nakagawa, and O. Watanabe.  $\tilde{O}(\sqrt{n})$ -space and polynomial-time algorithm for planar directed graph reachability. In *39th MFCS LNCS 8634*, pages 45–56, 2014. doi:10.1007/978-3-662-44465-8\_5.
- 3 N. Banerjee, S. Chakraborty, and V. Raman. Improved space efficient algorithms for BFS, DFS and applications. In *22nd COCOON*, volume 9797, pages 119–130. Springer, 2016.
- 4 N. Banerjee, S. Chakraborty, V. Raman, S. Roy, and S. Saurabh. Time-space tradeoffs for dynamic programming in trees and bounded treewidth graphs. In *21st COCOON*, volume 9198 of *LNCS*, pages 349–360. Springer, 2015.
- 5 L. Barba, M. Korman, S. Langerman, K. Sadakane, and R.I. Silveira. Space-time tradeoffs for stack-based algorithms. *Algorithmica*, 72(4):1097–1129, 2015. doi:10.1007/s00453-014-9893-5.
- 6 L. Barba, M. Korman, S. Langerman, and R.I. Silveira. Computing a visibility polygon using few variables. *Comput. Geom.*, 47(9):918–926, 2014. doi:10.1016/j.comgeo.2014.04.001.
- 7 G. Barnes, J. Buss, W. Ruzzo, and B. Schieber. A sublinear space, polynomial time algorithm for directed  $s$ - $t$  connectivity. *SIAM J. Comput.*, 27(5):1273–1282, 1998. doi:10.1137/S0097539793283151.

- 8 U. Brandes. Eager  $st$ -ordering. In *10th ESA*, pages 247–256, 2002. doi:10.1007/3-540-45749-6\_25.
- 9 D. Chakraborty, A. Pavan, R. Tewari, N. V. Vinodchandran, and L. Yang. New time-space upperbounds for directed reachability in high-genus and  $h$ -minor-free graphs. In *FSTTCS*, pages 585–595, 2014. doi:10.4230/LIPIcs.FSTTCS.2014.585.
- 10 S. Chakraborty, V. Raman, and S.R. Satti. Biconnectivity, chain decomposition and  $st$ -numbering using  $O(n)$  bits. *CoRR*, abs/1606.08645, 2016. URL: <http://arxiv.org/abs/1606.08645>.
- 11 D.R. Clark. *Compact Pat Trees*. PhD thesis. University of Waterloo, Canada, 1996.
- 12 S. Datta, N. Limaye, P. Nimbhorkar, T. Thierauf, and F. Wagner. Planar graph isomorphism is in log-space. In *24th CCC*, pages 203–214, 2009. doi:10.1109/CCC.2009.16.
- 13 M.E. Dyer and A.M. Frieze. On the complexity of partitioning graphs into connected subgraphs. *Discrete Applied Mathematics*, 10(2):139–153, 1985.
- 14 J. Ebert.  $st$ -ordering the vertices of biconnected graphs. *Computing*, 30(1):19–33, 1983. doi:10.1007/BF02253293.
- 15 M. Elberfeld, A. Jakoby, and T. Tantau. Logspace versions of the theorems of bodlaender and courcelle. In *51th FOCS*, pages 143–152, 2010. doi:10.1109/FOCS.2010.21.
- 16 A. Elmasry, T. Hagerup, and F. Kammer. Space-efficient basic graph algorithms. In *32nd STACS*, pages 288–301, 2015. doi:10.4230/LIPIcs.STACS.2015.288.
- 17 S. Even and R.E. Tarjan. Computing an  $st$ -numbering. *Theor. Comput. Sci.*, 2(3):339–344, 1976. doi:10.1016/0304-3975(76)90086-4.
- 18 A. Farzan and J. Ian Munro. Succinct representation of dynamic trees. *Theor. Comput. Sci.*, 412(24):2668–2678, 2011. doi:10.1016/j.tcs.2010.10.030.
- 19 A. Farzan, R. Raman, and S.S. Rao. Universal succinct representations of trees? In *36th ICALP*, pages 451–462, 2009. doi:10.1007/978-3-642-02927-1\_38.
- 20 G.N. Frederickson. Upper bounds for time-space trade-offs in sorting and selection. *J. Comput. Syst. Sci.*, 34(1):19–26, 1987. doi:10.1016/0022-0000(87)90002-X.
- 21 R.F. Geary, R. Raman, and V. Raman. Succinct ordinal trees with level-ancestor queries. *ACM Trans. Algorithms*, 2(4):510–534, 2006. doi:10.1145/1198513.1198516.
- 22 E. Györi. Partition conditions and vertex-connectivity of graphs. *Combinatorica*, 1(3):263–273, 1981.
- 23 M. He, J.I. Munro, and S.R. Satti. Succinct ordinal trees based on tree covering. *ACM Trans. Algorithms*, 8(4):42, 2012. doi:10.1145/2344422.2344432.
- 24 F. Kammer, D. Kratsch, and M. Laudahn. Space-efficient biconnected components and recognition of outerplanar graphs. In *41st MFCS*, 2016.
- 25 L. Lovasz. A homology theory for spanning tress of a graph. *Acta Mathematica Hungarica*, 30(3-4):241–251, 1977.
- 26 J. I. Munro and M. Paterson. Selection and sorting with limited storage. *Theor. Comput. Sci.*, 12:315–323, 1980. doi:10.1016/0304-3975(80)90061-4.
- 27 J. I. Munro and V. Raman. Selection from read-only memory and sorting with minimum data movement. *Theor. Comput. Sci.*, 165(2):311–323, 1996. doi:10.1016/0304-3975(95)00225-1.
- 28 J. I. Munro and V. Raman. Succinct representation of balanced parentheses and static trees. *SIAM J. Comput.*, 31(3):762–776, 2001. doi:10.1137/S0097539799364092.
- 29 S. Nakano, M.S. Rahman, and T. Nishizeki. A linear-time algorithm for four-partitioning four-connected planar graphs. *Inf. Process. Lett.*, 62(6):315–322, 1997. doi:10.1016/S0020-0190(97)00083-5.
- 30 O. Reingold. Undirected connectivity in log-space. *J. ACM*, 55(4), 2008. doi:10.1145/1391289.1391291.

- 31 J. M. Schmidt. *Structure and constructions of 3-connected graphs*. PhD thesis, Free University of Berlin, 2010. URL: [http://www.diss.fu-berlin.de/diss/receive/FUDISS\\_thesis\\_000000022941](http://www.diss.fu-berlin.de/diss/receive/FUDISS_thesis_000000022941).
- 32 J. M. Schmidt. A simple test on 2-vertex- and 2-edge-connectivity. *Inf. Process. Lett.*, 113(7):241–244, 2013. doi:10.1016/j.ipl.2013.01.016.
- 33 J. M. Schmidt. The mondshhein sequence. In *41st ICALP*, pages 967–978, 2014. doi:10.1007/978-3-662-43948-7\_80.
- 34 R. E. Tarjan. Depth-first search and linear graph algorithms. *SIAM J. Comput.*, 1(2):146–160, 1972. doi:10.1137/0201010.
- 35 R. E. Tarjan. Two streamlined depth-first search algorithms. *Fund. Inf.*, 9:85–94, 1986.
- 36 M. Tompa. Two familiar transitive closure algorithms which admit no polynomial time, sublinear space implementations. *SIAM J. Comput.*, 11(1):130–137, 1982. doi:10.1137/0211010.