



**HAL**  
open science

## Fast Self-Stabilizing Minimum Spanning Tree Construction

Lélia Blin, Shlomi Dolev, Maria Potop-Butucaru, Stephane Rovedakis

► **To cite this version:**

Lélia Blin, Shlomi Dolev, Maria Potop-Butucaru, Stephane Rovedakis. Fast Self-Stabilizing Minimum Spanning Tree Construction. 2010. hal-00492398v1

**HAL Id: hal-00492398**

**<https://hal.science/hal-00492398v1>**

Preprint submitted on 15 Jun 2010 (v1), last revised 21 Jul 2010 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Fast Self-Stabilizing Minimum Spanning Tree Construction

Using Compact Nearest Common Ancestor Labeling Scheme

Lélia Blin      Shlomi Dolev      Maria Gradinariu Potop-Butucaru  
Stephane Rovedakis

June 15, 2010

## Abstract

We present a novel self-stabilizing algorithm for minimum spanning tree (MST) construction. The space complexity of our solution is  $O(\log^2 n)$  bits and it converges in  $O(n^2)$  rounds. Thus, this algorithm improves the convergence time of all previously known self-stabilizing asynchronous MST algorithms by a multiplicative factor  $\Theta(n)$ , to the price of increasing the best known space complexity by a factor  $O(\log n)$ . The main ingredient used in our algorithm is the design, for the first time in self-stabilizing settings, of a labeling scheme for computing the nearest common ancestor with only  $O(\log^2 n)$  bits.

# 1 Introduction

Since its introduction in a centralized context [13, 12], the minimum spanning tree (or MST) problem gained a benchmark status in distributed computing thanks to the seminal work of Gallager, Humblet and Spira [7].

The emergence of large scale and dynamic systems, often subject to transient faults, revives the study of scalable and self-stabilizing algorithms. A scalable algorithm does not rely on any global parameter of the system (e.g. upper bound on the number of nodes or the diameter). Self-stabilization introduced first by Dijkstra in [5] and later publicized by several books Dolev[6] deals with the ability of a system to recover from catastrophic situation (i.e. the global state may be arbitrarily far from a legal state) without external (e.g. human) intervention in finite time.

Although there already exist self-stabilizing solutions for the MST construction, none of them considered the extension of the Gallager, Humblet and Spira algorithm (GHS) to self-stabilizing settings. Interestingly, this algorithm unifies the best properties for designing large scale MSTs: it is fast and totally decentralized and it does not rely on any global parameter of the system. Our work proposes an extension of this algorithm to self-stabilizing settings. Our extension uses only logarithmic memory and preserves all the good characteristics of the original solution in terms of convergence time and scalability.

Gupta and Srimani presented in [11] the first self-stabilizing algorithm for the MST problem. The MST construction is based on the computation of all shortest paths (for a certain cost function) between all pairs of nodes. While executing the algorithm, every node stores the cost of all paths from it to all the other nodes. To implement this algorithm, the authors assume that every node knows the number  $n$  of nodes in the network, and that the identifiers of the nodes are in  $\{1, \dots, n\}$ . Every node  $u$  stores the weight of the edge  $e_{u,v}$  placed in the MST for each node  $v \neq u$ . Therefore the algorithm requires  $\Omega(\sum_{v \neq u} \log w(e_{u,v}))$  bits of memory at node  $u$ . Since all the weights are distinct integers, the memory requirement at each node is  $\Omega(n \log n)$  bits. The main drawback of this solution is its lack of scalability since each node has to know and maintain information for all the nodes in the system. A different approach for the message-passing model, was proposed by Higham and Liang [9]. The algorithm performs roughly as follows: every edge checks whether it eventually belongs to the MST or not. To this end, every non tree-edge  $e$  floods the network to find a potential cycle, and when  $e$  receives its own message back along a cycle, it uses the information collected by this message (i.e., the maximum edge weight of the traversed cycle) to decide whether  $e$  could potentially be in the MST or not. If the edge  $e$  has not received its message back after the time-out interval, it decides to become tree edge. The memory used by each node is  $O(\log n)$  bits, but the information exchanged between neighboring nodes is of size  $O(n \log n)$  bits, thus only slightly improving that of [11]. This solution also assume that each node has access to a global parameter of the system: the diameter. Its computation is expensive in large scale systems and becomes even harder in dynamic settings. Note also that the time complexity announced by the authors,  $O(n)$  stays only in the particular synchronous settings considered by the authors. In asynchronous setting the complexity is  $O(n^2)$  rounds.

In [2] we proposed a self-stabilizing loop-free algorithm for the MST problem. Contrary to previous self-stabilizing MST protocols, this algorithm does not make any assumption on the network size (including upper bounds) or the unicity of the edge weights. The proposed solution improves on the memory space usage since each participant needs only  $O(\log n)$  bits while preserving the same time complexity as the algorithm in [9].

Clearly, in the self-stabilizing implementation of the MST algorithms there is a trade-off between the memory complexity and their time complexity (see Table 1). The challenge we

	a priori knowledge	space complexity	convergence time
[11]	network size and the nodes in the network	$O(n \log n)$	$O(n^2)$
[9]	upper bound on diameter	$O(\log n)$ +messages of size $O(n \log n)$	$O(n^3)$
[2]	none	$O(\log n)$	$O(n^3)$
This paper	none	$O(\log^2 n)$	$O(n^2)$

Table 1: Distributed Self-Stabilizing algorithms for the MST problem

address in this paper is to design fast and scalable self-stabilizing MST with little memory. Our approach brings together two worlds: the time efficient MST constructions and the memory compact informative labeling schemes. Therefore, we extend the GHS algorithm to self-stabilizing settings and keep compact its memory space by using a self-stabilizing extension of the nearest common ancestor labeling scheme of [1]. Note that labeling schemes have already been used in order to infer a broad set of information such as vertex adjacency, distance, tree routing or tree ancestry [4], however none of these schemes have been studied in self-stabilizing settings.

Our contribution is therefore twofold. We propose for the first time in self-stabilizing settings a  $O(\log^2 n)$  bits scheme for computing the nearest common ancestor. Furthermore, based on this scheme, we describe a new silent self-stabilizing algorithm for the MST problem. Our algorithm does not make any assumption on the network size (including upper bounds) or the existence of an a priori known root. Moreover, our solution is the best space/time compromise over the existing self-stabilizing MST solutions. The convergence time is  $O(n^2)$  asynchronous rounds and the memory space per node is  $O(\log^2 n)$  bits. Interestingly, our work is the first to prove the effectiveness of an informative labeling scheme in self-stabilizing settings and therefore opens a wide research path in this direction.

## 2 Model and notations

We consider an undirected weighted connected network  $G = (V, E, w)$  where  $V$  is the set of nodes,  $E$  is the set of edges and  $w : E \rightarrow \mathbb{R}^+$  is a positive cost function. Nodes represent processors and edges represent bidirectional communication links. Additionally, we consider that  $G = (V, E, w)$  is a network in which the weight of the communication links may change value.

The processors asynchronously execute their programs consisting of a set of variables and a finite set of rules. The variables are part of the shared register which is used to communicate with the neighbors. A processor can read and write its own registers and can read the shared registers of its neighbors. Each processor executes a program consisting of a sequence of guarded rules. Each rule contains a guard (Boolean expression over the variables of a node and its neighborhood) and an action (update of the node variables only). Any rule whose guard is true is said to be enabled. A node with one or more enabled rules is said to be privileged and may make a move executing the action corresponding to the chosen enabled rule.

A local state of a node is the value of the local variables of the node and the state of its program counter. A configuration of the system  $G = (V, E)$  is the cross product of the local states of all nodes in the system. The transition from a configuration to the next one is produced by the execution of an action at a node. A computation of the system is defined as a weakly fair, maximal sequence of configurations,  $e = (c_0, c_1, \dots, c_i, \dots)$ , where each configuration  $c_{i+1}$

follows from  $c_i$  by the execution of a single action of at least one node. During an execution step, one or more processors execute an action and a processor may take at most one action. Weak fairness of the sequence means that if any action in  $G$  is continuously enabled along the sequence, it is eventually chosen for execution. Maximality means that the sequence is either infinite, or it is finite and no action of  $G$  is enabled in the final global state.

In the sequel we consider the system can start in any configuration. That is, the local state of a node can be corrupted. Note that we don't make any assumption on the bound of corrupted nodes. In the worst case all the nodes in the system may start in a corrupted configuration. In order to tackle these faults we use self-stabilization techniques.

**Definition 1 (self-stabilization)** *Let  $\mathcal{L}_A$  be a non-empty legitimacy predicate<sup>1</sup> of an algorithm  $A$  with respect to a specification predicate  $Spec$  such that every configuration satisfying  $\mathcal{L}_A$  satisfies  $Spec$ . Algorithm  $A$  is self-stabilizing with respect to  $Spec$  iff the following two conditions hold:*

- (i) *Every computation of  $A$  starting from a configuration satisfying  $\mathcal{L}_A$  preserves  $\mathcal{L}_A$  (closure).*
- (ii) *Every computation of  $A$  starting from an arbitrary configuration contains a configuration that satisfies  $\mathcal{L}_A$  (convergence).*

### 3 Overview of our solution

We propose to extend the Gallager, Humblet and Spira (GHS) algorithm, [7], to self-stabilizing settings via a compact informative labeling scheme. Thus, the resulting solution presents several advantages appealing for large scale systems: it is compact since it uses only logarithmic memory in the size of the network, it scales well since it does not rely on any global parameter of the system, it is fast — its time complexity is the better known in self-stabilizing settings. Additionally, it self-recovers from any transient fault.

The central notion in the GHS approach is the *fragment*. A fragment is a partial spanning tree of the graph, i.e, a fragment is a tree which spans a subset of nodes. Note that a fragment can be limited to a single node. An outgoing edge of a fragment  $F$  is an edge with a unique endpoint in  $F$ . The minimum-weight outgoing edge of a fragment  $F$  is denoted in the following as  $ME_F$ . In the GHS construction, initially each node is a fragment. For each fragment  $F$ , the GHS algorithm in [7] identifies the  $ME_F$  and merges the two fragments endpoints of  $ME_F$ . Note that, with this scheme, more than two fragments may be merged concurrently. The merging process is recursively repeated until a single fragment remains. The result is a MST. The above approach is often called “blue rule” for MST construction.

This scheme is particularly appealing when transient faults yield to a forest of fragments (sub-trees of a MST). The direct application of the blue rule makes the system recover. However, when more severe faults hit the system the process variables may be corruption at the point where the configuration of the network may be a set of fragments that are not sub-trees of some MST. That is, it may be a spanning tree but not of minimum weight, or it can contain cycles. In this case only the blue rule application is not sufficient to recover. To overcome this difficulty, we combine the blue rule with another method, referred in the literature as the “red rule”. The red rule removes the heaviest edge from every cycle. The resulting configuration is a MST. We use the red rule as follows: given a spanning tree  $T$  of  $G$ , every edge  $e$  of  $G$  that is not in  $T$  is added to  $T$ , thus creating a (unique) cycle in  $T \cup \{e\}$ . This cycle is called a *fundamental cycle*, denoted by  $C_e$ . If  $e$  is not the edge of maximum weight in  $C_e$ , then, according to the red rule, there exists an edge  $f \neq e$  in  $C_e$  with  $w(f) > w(e)$ . The edge of maximum weight can be removed since it is not part of any MST.

---

<sup>1</sup>A legitimacy predicate is defined over the configurations of a system and is an indicator of its correct behavior.

Our MST construction combines both the blue and red rules. The blue rule application needs that each node identify its own fragment. The red rule requires that nodes identify the corresponding fundamental cycle for every non-tree-edge. In both case we use a self-stabilizing labeling scheme, NCA-L, which provides every node with a distinct informative label such that the nearest common ancestor of two nodes can be identified based only on the labels of these nodes (see Section 3.1). Thus, the advantage of this labeling is twofold. First the labeling helps nodes to identify their fragments. Second, given any non-tree edge  $e = (u, v)$ , the path in the tree going from  $u$  to the nearest common ancestor of  $u$  and  $v$ , then from there to  $v$ , and finally back to  $u$  by traversing  $e$ , is the fundamental cycle  $C_e$ .

To summarize, our algorithm will use the blue rule to construct a spanning tree, and the red rule to recover from invalid configurations. In both case, it uses our algorithm NCA-L to identify both fragments and fundamental cycles. Note that, in [3] a distributed algorithm using the blue and red rules to construct a MST in a dynamic network is proposed, however this algorithm is not self-stabilizing.

**Variables used by NCA-L and MST modules** For any node  $v \in V(G)$ , we denote by  $N(v)$  the set of all neighbors of  $v$  in  $G$ . Algorithm NCA-L and MST maintains the set  $N(v)$  at every node  $v$ . We use the following notations:

- $p_v$ : the parent of  $v$  in the current spanning tree;
- $\ell_v$ : the label of  $v$  composed by a list of pair of variables, each pair is one identifier and one integer;
- $size_v$ : a pair of variables, the first one is an integer the number of nodes in a sub-tree of  $v$  and the second one is the identifier of the children of  $v$  which has the maximum number of nodes;
- $mwe_v$ : the minimum weighted edge composed by a pair of variables, the first one is an integer, the weight of the edge and the second one is a label.

### 3.1 Self-stabilizing Nearest Common Ancestor Labeling

Our labeling scheme, called in the following NCA-L, uses the notions of *heavy* and *light* edges introduced in [8]. In a tree, a *heavy* edge is an edge between a node  $u$  and one of its children with a maximum number of nodes in its sub-tree. The other edges between  $u$  and its other children are tagged as *light* edges. The idea of the scheme is as follows. A tree is recursively divided into disjoint paths: the heavy and the light paths following the edges these paths use.

Each node  $v$  stores a tuple  $size_v$  of variables. The first variable is the local estimation of the number of nodes in the sub-tree rooted at  $v$ . The second variable is the identifier of a child of  $v$  with maximum number of nodes. That is, it indicates the heavy edge. The computation of  $size_v$  is processed from the leaves to the root (see Predicate SizeC( $v$ ) in Figure 3). A leaf has no child, therefore the first variable of  $size_v$  of a leaf is tagged by  $\perp$  (see Predicate Leaf( $v$ ) in Figure 3 and Rule  $R_\ell$ ).

A label  $\ell_v$  of a node  $v$  is a list of pair of variables, each pair is one identifier and one integer. The first element of the list contains the identifier of the tree (or fragment) and the distance at this root via a heavy path (a path including only heavy edges). For a root  $v$  the label  $\ell_v$  is the following pair  $(Id_v, 0)$ , respectively the identifier of  $v$  and the distance to itself, i.e., zero (see Rule  $R_\odot$ ). When a node  $u$  is tagged by its parent like heavy (i.e.,  $size_{p_v}[1] = Id_u$ ), then the node  $u$  takes the same label with its parent but it increases by one the distance of the last pair of the parent label. Examples of these cases are given in Figure 2. When a node  $u$  is tagged

- Child  $\equiv \{u \in N(v) : p_u = \text{Id}_v\}$
- SizeC( $v$ )  $\equiv \text{Leaf}(v) \vee (\text{size}_v = (1 + \sum_{u \in \text{Child}(v)} \text{size}_u, \arg \max\{\text{size}_u : u \in \text{Child}(v)\}))$
- Leaf( $v$ )  $\equiv (\nexists u \in N(v), p_u = \text{Id}_v) \wedge \text{size}_v = (1, \perp)$
- Label( $v$ )  $\equiv \text{Label}_R(v) \vee \text{Label}_{Nd}(v)$
- Label $_R$ ( $v$ )  $\equiv (p_v = \emptyset \wedge \ell_v = (\text{Id}_v, 0))$
- Label $_{Nd}$ ( $v$ )  $\equiv p_v \in N(v) \wedge (\text{Heavy}(v) \vee \text{Light}(v))$
- Heavy( $v$ )  $\equiv \text{size}_{p_v}[1] = \text{Id}_v \wedge \text{size}_v[0] < \text{size}_{p_v}[0] \wedge \text{last}(\ell_{p_v})[1] + 1 = \text{last}(\ell_v)[1]$
- Light( $v$ )  $\equiv \text{size}_{p_v}[1] \neq \text{Id}_v \wedge \text{size}_v[0] \leq \text{size}_{p_v}[0]/2 \wedge \ell_v = \ell_{p_v} \cdot (\text{Id}_v, 0)$
- $nca(u, v) \equiv \begin{cases} \ell.(a_0, a_1) & \text{s.t. } \ell_u \cap \ell_v = \ell, \ell_u = \ell.(a_0, a_1).\ell'_u \quad \text{if } (a_0 = b_0 \vee \ell \neq \emptyset) \wedge \ell_u \prec \ell_v \\ & \text{and } \ell_v = \ell.(b_0, b_1).\ell'_v \\ \ell.(b_0, b_1) & \text{s.t. } \ell_u \cap \ell_v = \ell, \ell_u = \ell.(a_0, a_1).\ell'_u \quad \text{if } (a_0 = b_0 \vee \ell \neq \emptyset) \wedge \ell_v \prec \ell_u \\ & \text{and } \ell_v = \ell.(b_0, b_1).\ell'_v \\ \emptyset & \text{otherwise} \end{cases}$
- Cycle( $v$ )  $\equiv \ell_v \subset \ell_{p_v} \vee \ell_v \prec \ell_{p_v}$
- MinEnabled( $v$ )  $\equiv \text{Enabled}(v) \wedge (\forall u \in N(v), \text{Enabled}(u) \wedge \text{Id}_v < \text{Id}_u)$

Figure 1: Predicates used by the algorithm NCA-L for the labeling procedure.

by its parent as light (i.e.,  $\text{size}_{p_v}[1] = w$  where  $w \neq u$ ), then the node  $u$  takes the label of its parent to which it adds a pair composed by its identifier and a zero distance 0 (see Figure 2).

This labeling scheme is used in second part of this article in MST algorithm to find the minimum weighted edges, but it also used to detect and destroy cycle since the initial configuration may not be a spanning tree. To this end we define an order  $\prec$  on the labels of nodes. Let  $a$  and  $b$  be two nodes and  $\ell_a$  and  $\ell_b$  be their respective labels such that  $\ell_a = \ell.(a_0, a_1).\ell'_a$  and  $\ell_b = \ell.(b_0, b_1).\ell'_b$  with  $\ell_a \cap \ell_b = \ell$ .  $\ell_a \prec \ell_b$  if (1)  $a_0 = \emptyset$  and  $b_0 \neq \emptyset$ , or (2)  $a_0 < b_0$  or (3)  $a_0 = b_0$  and  $a_1 < b_1$ .

A node  $u$  can detect the presence of a cycle by only comparing its label with the label of its parent. That is, if its label is contained in the label of its parent, or it is inferior to the one of its parent then  $u$  is part of a cycle (see Predicate Cycle( $v$ ) in Figure 3). In this case node  $u$  becomes root in order to break the cycle (see bellow Rule  $R_{\odot}$ ).

Algorithm NCA-L is composed of two rules. Rule  $R_{\odot}$  creates a root or breaks cycles while rule  $R_{\ell}$  produces of proper labeling. Note that the last predicates in rules  $R_{\odot}$  and  $R_{\ell}$  (the part in gray) are used only for insuring the exclusivity of rules execution when the labeling scheme works together with the MST scheme.

A node  $v$  with an incoherent parent (which is not one of its neighbors) or present in a cycle executes  $R_{\odot}$ . Following the execution of this rule node  $v$  becomes root, it sets its parent to void and its label to  $(\text{Id}_v, 0)$ .

Rule  $R_{\ell}$  helps a node  $v$  to compute the number of nodes in its sub-tree ( $\text{size}_v$ ) and also provides it with a coherent label.

$R_{\odot}$ : [ **Root creation** ]

**If**  $p_v \notin N(v) \vee (p_v = \emptyset \wedge \ell_v \neq (\text{Id}_v, 0)) \vee (\text{Cycle}(v) \wedge \neg \text{NeedReorientation}(v))$   
**Then**  $p_v := \emptyset$ ;  $\ell_v = (\text{Id}_v, 0)$ ;

$R_{\ell}$ : [ **Label correction** ]

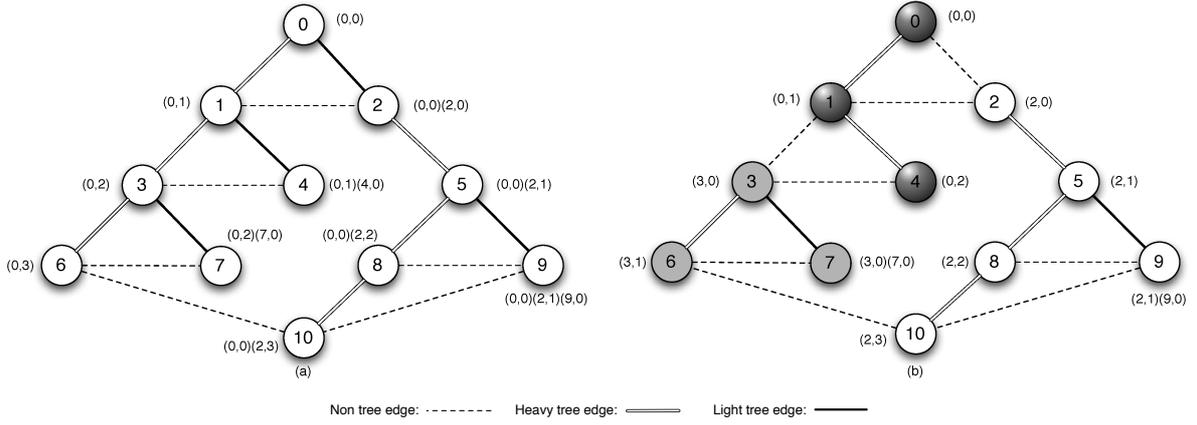


Figure 2: Labeling scheme

**If**  $\neg \text{Cycle}(v) \wedge (\neg \text{SizeC}(v) \vee \neg \text{Label}(v)) \wedge \text{MinEnabled}(v) \wedge \neg \text{TreeMerg}(v)$   
**Then If**  $\text{Leaf}(v)$  **then**  $\text{size}_v = (1, \perp)$   
**Else**  $\text{size}_v := (1 + \sum_{u \in \text{Child}(v)} \text{size}_u, \arg \max\{\text{size}_u : u \in \text{Child}(v)\})$ ;  
**If**  $\text{size}_{p_v}[1] = \text{ld}_v$  **then**  $\ell_v := \ell_{p_v}$ ;  $\text{last}(\ell_v)[1] := \text{last}(\ell_v)[1] + 1$ ;  
**Else**  $\ell_v = \ell_{p_v} \cdot (\text{ld}_v, 0)$

### 3.2 Self-stabilizing MST

In this section we describe our self-stabilizing MST algorithm. The algorithm executes two phases: the MST correction and the MST fragments merging. Recall that our algorithm use the blue rule to construct a spanning tree and the red rule to recover from invalid configurations. In both cases, it uses the nearest-common ancestor labeling scheme to identify fragments and fundamental cycles. We assume in the following that the *merging* operations have a higher priority than the *recovering* operations. That is, the system recovers from an invalid configuration if and only if no merging operation is possible. In the worst case, after a failure hit the system, a merging phase will be followed by a recovery phase and finally by a second merging phase.

#### The minimum weighted edge and MST Correction

Note that the scope of our labeling scheme is twofold. First, it allows a node to identify the neighbors that share the same fragment and consequently to select the outgoing edges of a fragment. Second, the labeling scheme may be used to identify cycles and recover the tree. To this end, the algorithm uses the nearest common ancestor predicate  $nca$  depicted in Figure 3. For two nodes  $u$  and  $v$  with  $e = (u, v)$  and  $e$  not tree or fragment edge ( $p_u \neq v$  and  $p_v \neq u$ ), if the nearest common ancestor does not exist then  $u$  and  $v$  are not in the same fragment ( $nca(u, v) = \emptyset$ ). Otherwise  $u$  and  $v$  are in a same fragment  $F$  and the addition of  $e$  to  $F$  generates a cycle. This fundamental cycle  $C_e$  is the following: let  $\text{path}(u, v)$  be the set of edges on the unique path between  $u$  and  $v$  in  $F$ , with  $u \in F$  and  $v \in F$ .  $C_e = \text{path}(u, nca(u, v)) \cup \text{path}(nca(u, v), v) \cup e$ . Consider the example depicted on Figure 2(b). The labels of nodes 10 and 6 are respectively  $\ell_{10} = (2, 3)$  and  $\ell_6 = (3, 1)$ . In this case  $nca(10, 6) = \emptyset$ , the edge  $(10, 6)$  is an outgoing edge. If  $(10, 6)$  is of minimum weight then  $(10, 6)$  can be used for a fusion between the fragment rooted

- MergeChild( $v$ )  $\equiv \min\{mwe_u : u \in \text{Child}(v) \wedge mwe_u[1] = \emptyset\}$
- MergeAdj( $v$ )  $\equiv (\min\{w(u, v) : u \in N(v) \setminus \text{Child}(v) \setminus \{p_v\} \wedge nca(u, v) = \emptyset\}, \emptyset)$
- MergeEdge( $v$ )  $\equiv \min\{\text{MergeChild}(v), \text{MergeAdj}(v)\}$
  
- FarLcaChild( $v$ )  $\equiv \arg \min_{\prec} \{mwe_u[1] : u \in \text{Child}(v) \wedge mwe_u[1] \preceq \ell_v\}$
- RecoverChild( $v$ )  $\equiv mwe_u$  such that FarLcaChild( $v$ ) =  $u$
- FarLca( $v$ )  $\equiv \arg \min_{\prec} \{nca(u, v) : u \in N(v) \setminus \text{Child}(v) \setminus \{p_v\} \wedge mwe_v[1] \neq \emptyset \wedge nca(u, v) \succ mwe_v[1] \wedge nca(u, v) \neq \emptyset\}$
  
- RecoverAdj( $v$ )  $\equiv \begin{cases} (w(u, v), nca(u, v)) \text{ s.t. FarLca}(v) = u & \text{if FarLca}(v) \neq \emptyset \\ (w(u, v), nca(u, v)) & \text{otherwise} \\ \text{s.t. } u = \arg \min_{\prec} \{nca(u, v) : \\ u \in N(v) \setminus \text{Child}(v) \setminus \{p_v\} \wedge nca(u, v) \neq \emptyset\} \end{cases}$
- RecoverEdge( $v$ )  $\equiv \begin{cases} \text{RecoverChild}(v) & \text{if RecoverAdj}(v)[1] \prec \text{RecoverChild}(v) \\ \text{RecoverAdj}(v) & \text{otherwise} \end{cases}$
  
- MinEdge( $v$ )  $\equiv \begin{cases} \text{MergeEdge}(v) & \text{if MergeEdge}(v) \neq \emptyset \\ \text{RecoverEdge}(v) & \text{otherwise} \end{cases}$
  
- NeedReorientation( $v$ )  $\equiv \text{Label}_R(v) \vee (\ell_{p_v} = (\perp, \perp) \wedge p_{p_v} = \text{Id}_v)$
- EndReorientation( $v$ )  $\equiv \ell_{p_v} = (\emptyset, \emptyset) \wedge (\ell_v = (\perp, \perp) \vee \ell_v \neq (\emptyset, \emptyset))$
- TreeMerg( $v$ )  $\equiv \text{NeedReorientation}(v) \vee \text{EndReorientation}(v)$
  
- NewFrag( $v$ )  $\equiv mwe_v = mwe_{p_v} \wedge mwe_v[1] \neq \text{Id}_v \wedge w(v, p_v) > mwe_v[0]$

Figure 3: Predicates used by the MST for the tree correction or the fusion fragments.

in 2 and the fragment rooted in 3. For the case of nodes 10 and 9 the labels are  $\ell_{10} = (2, 3)$  and  $\ell_9 = (2, 1)(9, 0)$  and  $nca(10, 9) = (2, 1)$ . Consequently, 10 and 9 are in the same fragment. The fundamental cycle  $C_e$  with  $e = (9, 10)$  run through the node with the label  $nca(10, 9)$ , in other word the node 5 in Figure 2(b).

The core of the computation of the minimum weight edge is based on the predicate MinEdge (see Figure 3 and  $R_{Min}$ ). The information about the weight of the not-tree edges are sent from the leaves to the root. For the merging phase, our algorithm gives priority to the outgoing-edges via the predicate MinEdge. Namely, an outgoing edge  $e = (u, v)$  is an edge with  $nca(u, v) = \emptyset$ . Therefore, for the tree recovering, the weighted edges are sent only if there exists an unique tree or if a sub-tree of one fragment has no outgoing edge.

Let us explain the rule  $R_{\downarrow}$  which recovers the tree (or some fragment). In this case the information about the non-tree edges are sent to the root as follows. Among all its non-tree edges, a node  $u$  send the edge  $e = (u, v)$  with the  $nca(u, v)$  nearest to the root (see Figure 4(a)). The information about the edge  $e$  is stored in a variable  $mwe_u$ . If the parent  $x$  of the node  $u$  has the same information and the weight of the edge  $w(u, x) > w(e)$  then  $(u, v)$  is removed from the tree (see Figure 4(a-b) for the nodes 6 and 10 ). In the same way, if  $x$  has an information with  $nca$  nearest the root of the  $nca(u, v)$  then  $(u, v)$  is removed from the tree too (see Figure 4(a-b) for the nodes 7 and 4 ). We use the red rule in intensive way, because we remove all the edges with upper weight to  $w(e)$  in fundamental cycle of  $e$ . This interpretation of the red rule allows to insure that after a recover phase there are no invalid edges.

$R_{Min}$ : [ **Minimum computation** ]

**If**  $\neg \text{Cycle}(v) \wedge \text{Label}(v) \wedge [(mwe_v \neq \text{MinEdge}(v) \wedge \text{MergeEdge}(v) \neq \emptyset) \vee (mwe_{p_v} =$

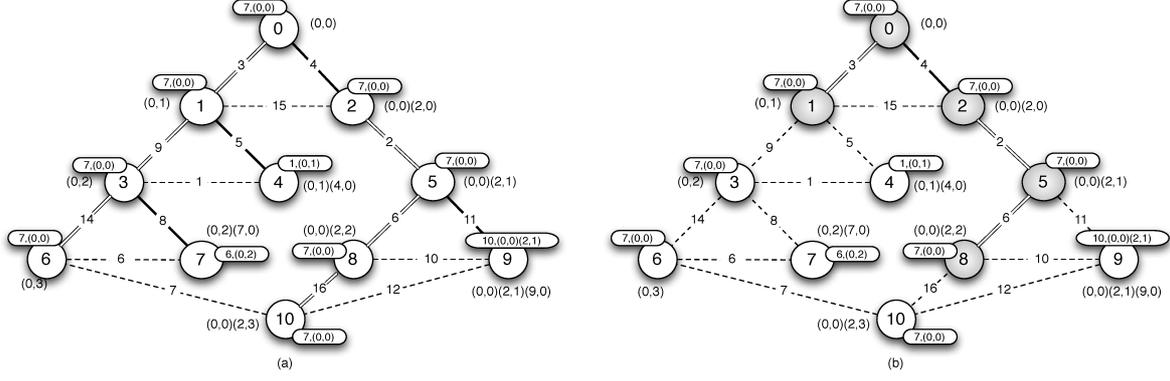


Figure 4: Minimum weighted edge computation and Tree correction

$mwe_v \wedge \text{MergeEdge}(v) = \emptyset]$   
**Then**  $mwe_v := \text{MinEdge}(v);$

$R_{\downarrow}:$  [ **MST Correction** ]  
**If**  $\neg \text{Cycle}(v) \wedge \text{Label}(v) \wedge \text{NewFrag}(v)$   
**Then**  $p_v := \emptyset; \ell_v := (\text{Id}_v, 0);$

To summarize, in this section we explained how to compute the outgoing-edges and the fundamental cycles (rule  $R_{Min}$ ), and how to recover a false tree (Rule  $R_{\downarrow}$ ). The next section addresses the fragments merge operation (rules  $R_{\triangleright\triangleleft}$  and  $R_{\square}$ ).

### Fragments merging

In this phase two rules are executed:  $R_{\triangleright\triangleleft}$  and  $R_{\square}$ . Note that Rule  $R_{Min}$  (described in the previous section) computes from the leaves to the root the minimum outgoing edge  $e = (u, v)$  of the fragment  $F_u$ , with  $u \in F_u$ . The information about  $e$  is stored in the variable  $mwe$  altogether with the weight of the edge and the label of the fragment  $F_v$ . When a root  $r$  of  $F_u$  has stabilized its  $mwe_r$  it starts a merging phase (rule  $R_{\triangleright\triangleleft}$ ). To this end the nodes in the path between  $r$  and  $u$  are reoriented from  $r$  to  $v$ . During this reorientation the labels are locked (for a node  $v$ ,  $\ell_v := (\perp, \perp)$ ). When a node  $u$  becomes the root of the fragment  $F_u$  it can merge with the fragment  $F_v$ . At this moment the process of the labeling is re-started (see  $R_{\square}$ ). The merging phase is repeated until a single fragment is obtained.

$R_{\triangleright\triangleleft}:$  [ **Merging** ]  
**If**  $\text{NeedReorientation}(v) \wedge mwe_v = \text{MergeEdge}(v)$   
**Then If**  $(\exists u \in N(v) \setminus \text{Child}(v) \setminus \{p_v\}, w(u, v) = \text{MergeEdge}(v) \wedge \text{Id}_v > \text{Id}_u)$   
**Then**  $p_v := \min\{\text{Id}_u : u \in N(v) \setminus \text{Child}(v) \setminus \{p_v\} \wedge w(u, v) = \text{MergeEdge}(v)\};$   
 $\ell_v := (\emptyset, \emptyset);$   
**If**  $(\exists u \in N(v) \setminus \text{Child}(v) \setminus \{p_v\}, w(u, v) = \text{MergeEdge}(v) \wedge \text{Id}_v < \text{Id}_u \wedge p_u = \text{Id}_v)$   
**Then**  $\ell_v := (\emptyset, \emptyset);$   
**Else**  $p_v := \min\{\text{Id}_u : u \in \text{Child}(v) \wedge mwe_v = \text{MergeEdge}(v)\}; \ell_v := (\perp, \perp);$

$R_{\square}$ : [ **End Merging** ]  
**If**  $\neg \text{NeedReorientation}(v) \wedge \text{EndReorientation}(v)$   
**Then**  $\ell_v := (\emptyset, \emptyset)$ ;

## 4 Complexity proof

In the following we discuss the complexity issues of our solution. Due to the space limitation the correctness proofs are detailed in the Annexes section.

**Lemma 1** *Algorithms NCA-L and MST have a space complexity of  $O(\log^2 n)$  bits.*

**Proof.** Algorithm NCA-L uses three variables :  $p_v, \ell_v, size_v$ . The first and the last one are respectively a pointer to a neighbor node and a pair of integers, each one needs  $O(\log n)$  bits. However, the variable  $size_v$  is a list of pairs of integers. A new pair of integers is added to the list when a light edge is created in the tree. As noticed in [1], there are at most  $\log n$  light edges on the path from a leaf to the root, i.e., at most  $\log n$  pairs of integers. Thus, the variable  $size_v$  uses  $\log n \times \log n$  bits.

Algorithm MST uses an additional variable  $mwe_v$  which is a pair composed of an integer and the label of a node. The label of a node is stored in variable  $size_v$  which uses  $\log^2 n$  bits. Thus, the variable  $mwe_v$  needs  $\log^2 n$  bits.

Therefore, Algorithm NCA-L and MST use  $O(\log^2 n)$  bits of memory at each node.  $\square$

**Lemma 2** *Starting from any configuration, all cycles are removed from the network in at most  $O(n^2)$  rounds, with  $n$  the number of nodes in the network.*

**Proof.** As explained in the proof of Lemma 6, to break a cycle  $C_k$  a part of the nodes in  $C_k$  must compute their new labels, that is a label computation must be initiated from one node and then this process must cross  $C_k$ . Thus, the worst case is a configuration in which all the nodes in  $C_k$  have to compute their new labels using Rule  $R_{\ell}$  to detect the presence of cycle  $C_k$ . Therefore, at most  $O(n)$  rounds are needed to compute the new label of the nodes in  $C_k$  based on the label of one node  $x$  in  $C_k$ . According to Lemma 6, when this computation is done the cycle  $C_k$  is detected and removed by the node  $x$ . At most  $O(n)$  additional rounds are needed to break the cycle  $C_k$ .

Since there is at most  $n/2$  cycles in a network, at most  $O(n^2)$  rounds are needed to remove all the cycles from the network.  $\square$

**Lemma 3** *Starting from a configuration which contains a tree  $T$ , using Algorithm NCA-L any node  $v \in V_T$  has a correct label in at most  $O(n)$  rounds.*

**Proof.** As described in proof of Lemma 7, the correction of node labels is done using a bottom-up computation followed by top-down computation in the tree  $T = (V_T, E_T)$ .

The bottom-up computation is started by the leaves of  $T$ , when the leaf nodes  $v \in V_T$  have corrected their variable  $size_v$  to  $(1, \perp)$  then internal nodes  $u \in V_T$  can start to correct their variable  $size_u$ . A internal node  $v \in V_T$  computes a correct value in its variable  $size_v$  using Rule  $R_{\ell}$  when all its children  $u$  have a correct value in their variable  $size_u$ . Since the computation is done in a tree sub-graph then in at most  $O(n)$  rounds each node  $v \in V_T$  has corrected its variable  $size_v$ .

The top-down computation is started by the root of the tree  $T$ . When the root  $v$  has a correct value in variable  $size_v$  then the computation of correct labels can start. Thus, if the parent of a node  $v$  has a correct value in its variable  $size_{p_v}$  and  $\ell_{p_v}$  then  $v$  can compute its

correct label in  $\ell_v$  using Rule  $R_\ell$ . As for the bottom-up computation, the top-down computation is done in at most  $O(n)$  rounds since it is done in a tree sub-graph.

Therefore, in at most  $O(n)$  rounds each node  $v$  in the tree  $T$  has a correct label stored in variable  $size_v$ .  $\square$

**Lemma 4** *Starting from any configuration, Algorithm NCA-L reaches a legitimate configuration in at most  $O(n^2)$  rounds.*

**Proof.** The initial configuration  $\mathcal{C}$  could contain one or more cycles, so according to Lemma 2 in at most  $O(n^2)$  rounds the system reaches a new configuration  $\mathcal{C}'$  which contains no cycle. Moreover according to Lemma 3, the nodes  $v$  in each tree  $T$  in the configuration  $\mathcal{C}'$  have a correct label in at most  $O(n)$  rounds. Therefore, starting from an arbitrary configuration each node  $v \in V$  computes its correct label in at most  $O(n^2)$  rounds.  $\square$

**Lemma 5** *Starting from any configuration, Algorithm MST reaches a legitimate configuration in at most  $O(n^2)$  rounds.*

**Proof.** According to Lemma 2, starting from any configuration in at most  $O(n^2)$  rounds all the cycles are removed from the network, i.e., it remains a forest of trees in at most  $O(n^2)$  rounds. Moreover, according to Lemma 3 in at most  $O(n)$  additional rounds each node  $v \in V$  has a correct label since each node belongs to an unique tree.

As described in Lemma 10, while it is possible to make a merging between two distinct trees in the forest a merging phase is started. This merging phase is done in three steps: (1) information (edge weight, nearest common ancestor label) corresponding to the minimum outgoing edge is propagated in a bottom-up fashion in each tree, (2) the orientation is reversed from the root of a tree until reaching the node in the tree adjacent to the minimum outgoing edge, and (3) the node labels are changed to inform of the end of the merging phase, followed by a propagation of the new correct node labels in the new tree resulting from the merging phase.

The first step is propagation of information in a bottom-up fashion in a tree which is done in at most  $O(n)$  rounds. The second step reverse and propagate new node labels on a part of the tree (between the root and the node adjacent to minimum outgoing edge) which at most  $O(n)$  rounds too. Step three modifies the label of the nodes which have change their parent pointer in step 2, so this last step takes also at most  $O(n)$  rounds and the relabeling of the nodes in the new tree is done in at most  $O(n)$  rounds according to Lemma 3. Thus, a merging phase is accomplished in at most  $O(n)$  rounds and as there is in the worst case  $n$  trees then in at most  $O(n^2)$  rounds a spanning tree is constructed.  $\square$

## 5 Conclusions

We extended the Gallager, Humblet and Spira (GHS) algorithm, [7], to self-stabilizing settings via a compact informative labeling scheme. Thus, the resulting solution presents several advantages appealing for large scale systems: it is compact since it uses only logarithmic memory in the size of the network, it scales well since it does not rely on any global parameter of the system, it is fast — its time complexity is the better known in self-stabilizing settings. Additionally, it self-recovers from any transient fault. The time complexity is  $O(n^2)$  rounds and the space complexity is  $O(\log^2 n)$ .

## References

- [1] Alstrup Stephen and Gavoille Cyril and Kaplan Haim and Rauhe Theis. Nearest common ancestors: a survey and a new algorithm for a distributed environment. Theory of Computing Systems, 37(3):441–456, 2004.
- [2] Lélia Blin, Maria Potop-Butucaru, Stephane Rovedakis, Sébastien Tixeuil. A New Self-stabilizing Minimum Spanning Tree Construction with Loop-Free Property. DISC, volume 5805 of Lecture Notes in Computer Science, pages 407–422. Springer 2009.
- [3] Jungho Park, Toshimitsu Masuzawa, Ken’ichi Hagihara, Nobuki Tokura. Efficient distributed algorithm to solve updating minimum spanning tree problem. Systems and Computers in Japan, 23(3):1–12, 1992.
- [4] Doina Bein, Ajoy Kumar Datta, Vincent Villain. Self-Stablizing Pivot Interval Routing in General Networks. ISPAN, pages 282–287, 2005.
- [5] Edsger W. Dijkstra. Self-stabilizing systems in spite of distributed control. Commun. ACM, 17(11):643–644, 1974.
- [6] Shlomi Dolev. Self-Stabilization. MIT Press, 2000.
- [7] Robert G. Gallager, Pierre A. Humblet, and Philip M. Spira. A distributed algorithm for minimum-weight spanning trees. ACM Trans. Program. Lang. Syst., 5(1):66–77, 1983.
- [8] D. Harel and R. E. Tarjan. Fast algorithms for finding nearest common ancestors. SIAM Journal Computing, 13(2):338-355, 1984.
- [9] Lisa Higham and Zhiying Liang. Self-stabilizing minimum spanning tree construction on message-passing networks. In DISC, pages 194–208, 2001.
- [10] S Katz and KJ Perry. Self-stabilizing extensions for message-passing systems. Distributed Computing, 7:17–26, 1993.
- [11] Sandeep K. S. Gupta and Pradip K. Srimani. Self-stabilizing multicast protocols for ad hoc networks. J. Parallel Distrib. Comput., 63(1):87–96, 2003.
- [12] Joseph B. Kruskal. On the shortest spanning subtree of a graph and the travelling salesman problem. Proc. Amer. Math. Soc., 7:48–50, 1956.
- [13] R.C. Prim. Shortest connection networks and some generalizations. Bell System Tech. J., pages 1389–1401, 1957.

## Annexes: Correctness proof

**Lemma 6** *Let  $\mathcal{C}$  a configuration where the set of variables  $p_v, v \in V$ , form at least one cycle in the network. In a finite time, Algorithm NCA-L removes all the cycles from the network.*

**Proof.** If a node  $v$  has a parent which is not in its neighborhood or if  $v$  has no parent then the parent and the label variable of  $v$  is modified to  $\emptyset$  and  $(\text{ld}_v, 0)$  respectively with Rule  $R_{\odot}$ .

A node  $v$  identifies a cycle with Predicate  $\text{Cycle}(v)$  which uses  $v$ 's label and the label of its parent. In a legitimate configuration,  $v$ 's label is smaller than the label of its parent and is constructed using the label of its parent, i.e., the label of the parent of  $v$  is included to  $v$ 's label. Thus, if the label of  $v$  is included or is smaller than the label of its parent then a cycle is detected and Predicate  $\text{Cycle}(v)$  is true. In this case,  $v$  reinitiates its parent and label variable using Rule  $R_{\odot}$ .

In order to detect a cycle the label computation process must cross a part or all the nodes of the cycle. However, since we consider a distributed scheduler then all the nodes in a cycle can be activated and we can have a rotation of the labels of the nodes in the cycle. This may lead to a new configuration in which the labels cannot be used to detect a cycle, because the label of one node is not used to compute some other labels and to detect a cycle. To break this symmetry, we use the node identifiers with Predicate  $\text{MinEnabled}(v)$ . This predicate allows to activate the node  $v$  iff  $v$  has no neighbor  $u$  such that  $u$  is activated and  $u$ 's identifier is lower than  $v$ . Therefore, there is at least a node  $x$  in the cycle which is not activated and when the label of  $x$  is used by some other nodes to compute their labels then Predicate  $\text{Cycle}(x)$  is true and  $x$  breaks the cycle using Rule  $R_{\odot}$ .  $\square$

According to Lemma 6, if the system starts from a configuration which contains at least one cycle then all the cycles are removed from the network in a finite time. Therefore, in the following we consider only configurations containing no cycle.

**Definition 2 (Legitimate state of NCA-L)** *Let  $\mathcal{C}$  a configuration with no cycle in the network, i.e., which contains a forest of trees  $T = (V_T \subseteq V, E_T \subseteq E)$ . The configuration  $\mathcal{C}$  is legitimate for Algorithm NCA-L iff each node  $v \in V_T$  satisfies one of the following conditions:*

1. the label  $\ell_v$  of  $v$  is equal to  $\ell_{p_v} \cdot (\text{ld}_v, 0)$ , if the edge between  $v$  and  $v$ 's parent in  $T$  is a light edge or  $v$  is the root of the tree;
2. the label  $\ell_v$  of  $v$  is equal to  $\ell_{p_v}$  and  $\text{last}(\ell_v)[1] = \text{last}(\ell_{parent})[1] + 1$ , if the edge between  $v$  and  $v$ 's parent in  $T$  is a heavy edge.

**Lemma 7 (Convergence for NCA-L)** *Starting from an illegitimate configuration for Algorithm NCA-L, eventually Algorithm NCA-L reaches in a finite time a legitimate configuration.*

**Proof.** To compute correct node labels, heavy and light edges in each tree  $T = (V_T, E_T)$  of the forest in the network must be identified. To this end, each node  $v \in V_T$  maintains in the variable  $\text{size}_v$  two information: the size of its subtree in  $T$  and the identifier of the child with the subtree of maximum number of nodes. Based on the first information given by its children  $u \in V_T$  stored in variable  $\text{size}_u$ , each node  $v$  compute the size of its subtree in  $T$  and informs its child  $u$  if the edge  $(u, v) \in E_T$  is a light or heavy edge. The edge  $(u, v) \in E_T$  is a heavy edge if the second information stored in  $\text{size}_v$  is equal to  $\text{ld}_u$  the identifier of  $u$ , a light edge otherwise. Therefore, each node  $v \in V_T$  can detect if its label is correct according to its parent label. Note that Predicate  $\text{MinEnabled}(v)$  is used at node  $v \in V$  to help to break cycle if we are in the case of a configuration described in proof of Lemma 6. Moreover, since we use the

labeling scheme with minimum spanning tree computation rules Predicate TreeMerg( $v$ ) is used to forbid the label correction when it has been modified by Rule  $R_\ell$ ,  $R_{\triangleright\triangleleft}$  and  $R_\odot$ .

The computation of the information stored in the variable  $size_v$  at each node  $v \in V_T$  is done via bottom-up fashion in the tree  $T$ . According to Predicate SizeC( $v$ ), if the variable  $size_v$  is not equal to  $(1, \perp)$  at a leaf node  $v$  in  $T$  then Predicate SizeC( $v$ ) = *false* and  $v$  can execute Rule  $R_\ell$  to correct its variable  $size_v$ . Otherwise according to Predicate SizeC( $v$ ), for any internal node  $v \in V_T$  the first information of  $size_v$  must be equal to one plus the sum of the size of the children subtrees and the second one to the identifier of its child with the maximum subtree size. Thus, if variable  $size_v$  is not correct (i.e., Predicate SizeC( $v$ ) = *false*) then  $v$  can execute Rule  $R_\ell$  to correct its variable  $size_v$ . Using the same argument, one can show by induction that for any internal node  $v \in V_T$  we have  $size_v = (1 + \sum_{u \in \text{Child}(v)} size_u, \arg \max\{size_u : u \in \text{Child}(v)\})$ .

The computation of the node labels in a tree  $T$  is done via top-down fashion starting from the root of  $T$ . For convenience, a path is called heavy (resp. light) if it contains only heavy (resp. light) edges. Moreover, a node is called heavy (resp. light) if the edge between its parent and itself is a heavy (resp. light) edge.  $v$  is informed by its parent with  $size_{p_v}$  if  $v$  is a heavy (i.e.,  $size_{p_v}[1] = \text{Id}_v$ ) or light (i.e.,  $size_{p_v}[1] \neq \text{Id}_v$ ) node. The root node  $v$  of  $T$  is also the root of a heavy path and its label must be equal to  $(\text{Id}_v, 0)$ . According to Predicate Label( $v$ ) and Label $_R$ ( $v$ ),  $v$  can execute Rule  $R_\ell$  to correct its label. Otherwise, we have two cases: heavy or light nodes. When the root have a correct label then all its children can compute their correct label. If a heavy (resp. light) node has a label different from  $\ell_{p_v}$  and  $last(\ell_v)[1] = last(\ell_{parent})[1] + 1$  (resp.  $\ell_{p_v} \cdot (\text{Id}_v, 0)$ ) then we have Predicate Heavy( $v$ ) = *false* (resp. Light( $v$ ) = *false*), Label $_{Nd}$ ( $v$ ) = *false* and Label( $v$ ) = *false*. Therefore,  $v$  can execute Rule  $R_\ell$  to correct its label  $\ell_v$  accordingly with its parent. Using the same argument, one can show by induction that for any internal node  $v \in V_T$  we have  $\ell_{p_v}$  and  $last(\ell_v)[1] = last(\ell_{parent})[1] + 1$  (resp.  $\ell_{p_v} \cdot (\text{Id}_v, 0)$ ) for heavy (resp. light) nodes.  $\square$

**Lemma 8 (Closure for NCA-L)** *The set of legitimate configurations for NCA-L is closed.*

**Proof.** According to Algorithm NCA-L, the labeling procedure is done using only Rule  $R_\ell$ . In any legitimate configuration for Algorithm NCA-L, for any node  $v \in V$  Predicate SizeC( $v$ ) and Label( $v$ ) are true and Rule  $R_\ell$  cannot be executed by a node  $v$ . So, starting from a legitimate configuration for Algorithm NCA-L the system remains in a legitimate configuration.  $\square$

**Definition 3 (Legitimate state of MST)** *A configuration is legitimate for Algorithm MST iff each node  $v \in V$  satisfies the following conditions:*

1. a tree  $T$  spanning the set of nodes in  $V$  is constructed;
2.  $T$  is of minimum weight among all spanning trees.

**Lemma 9** *Let  $T_i = (V_{T_i}, E_{T_i})$  a tree (or fragment). Eventually for each node  $v \in V_{T_i}$  the variable  $mwe_v$  contains a pair of values: the weight of the minimum outgoing edge  $(u, v)$  of the fragment  $T_i$  and  $\emptyset$ , if a merging is possible between two fragments  $T_j$  and  $T_i$ , ( $j \neq i$ ).*

**Proof.** We assume that  $T_i$  and  $T_j$ , ( $j \neq i$ ) are two distinct coherent trees (different root and correct labels otherwise Rule  $R_\odot$  and  $R_\ell$  are used to correct the trees) in the network and that a merging is possible between  $T_i$  and  $T_j$ . The computation of the minimum outgoing edge of  $T_i$  (resp.  $T_j$ ) is done in a bottom-up fashion. We consider the tree  $T_i$  but the computation in  $T_j$  is done in a same way. A leaf node  $v$  can compute and store in its variable  $mwe_v$  its local adjacent minimum outgoing edge leading to another tree. Macro MinEdge( $v$ ) returns the local minimum outgoing edge if Macro MergeEdge( $v$ )  $\neq \emptyset$ . To this end, if there is an adjacent

outgoing edge  $(u, v)$  leading to another tree  $T_j$  (i.e.,  $\text{MergeAdj}(v) \neq \emptyset$  and  $\text{MergeEdge}(v) \neq \emptyset$ ) and we have  $mwe_v \neq \text{MinEdge}(v)$  then  $v$  can execute Rule  $R_{Min}$  to compute its local outgoing edge stored in the variable  $mwe_v$ . A internal node  $v$  must use the local outgoing edges computed by its children and selects the edge of minimum weight among these edges, then it compares this value with the weight of its adjacent local outgoing edge and again it holds the edge of minimum weight. The selection of its children minimum outgoing edge is done by Macro  $\text{MergeChild}(v)$  and the computation of its local outgoing edge is done by Macro  $\text{MergeAdj}(v)$  as for a leaf node. Thus, if there is an outgoing edge which can be used to make a merging with another tree (i.e.,  $\text{MergeEdge}(v) \neq \emptyset$ ) and we have  $mwe_v \neq \text{MinEdge}(v)$  for a internal node  $v$ , then  $v$  can execute Rule  $R_{Min}$  to compute in the variable  $mwe_v$  its local minimum outgoing edge. Using the same argument, one can show by induction that for any internal node  $v \in V_{T_i}$  we have  $mwe_v = \text{MergeEdge}$ . Therefore, the local outgoing edge computed by the root node  $v$  is the minimum outgoing edge of  $T_i$ .  $\square$

**Lemma 10** *Let  $T_i = (V_{T_i}, E_{T_i})$  a tree (or fragment). Eventually for each node  $v \in V_{T_i}$  the variable  $mwe_v$  contains a pair of values: the weight of an edge  $(u, v)$  and the label of the nearest common ancestor of  $u \in V_{T_i}$  and  $v$ , if the edge  $(u, v) \in E_{T_i}$  is not part of a minimum spanning tree of the network. Moreover, eventually all local internal edges of  $v$  are computed.*

**Proof.** We assume that  $T_i = (V_{T_i}, E_{T_i})$  is a coherent tree, otherwise Rule  $R_{\odot}$  and  $R_{\ell}$  are used to break the cycles and to correct the labels. Each node  $v \in V_{T_i}$  starts to compute local internal edges (i.e., edges  $(x, y)$  such that  $x, y \in V_{T_i}$  and  $x = v$  or  $x$  or  $y$  is in the subtree of  $v$ ) when it has no local outgoing edge (adjacent outgoing edge or outgoing edge given by a child) leading to another tree. In this case, a node  $v$  informs its parent of its local internal edges using its variable  $mwe_v$ . Macro  $\text{RecoverEdge}(v)$  returns the local internal edge of  $v$  which has the common ancestor nearest from the root among the internal edges that were not taken into account by its parent using the node labels. Each node  $v \in V_{T_i}$  which is in a recover phase sends all its local internal edges. To this end,  $v$  compute its next local internal edge when its parent has taken into account  $v$ 's current local internal edge (i.e.,  $mwe_{p_v} = mwe_v$ ). Thus, the information of internal edges are put back up in the tree until reaching the nearest common ancestor and  $v$  do not wait an acknowledgement from the nearest common ancestor to send to compute its next internal edge. Moreover, Macro  $\text{FarLca}(v)$  compute the next internal edge adjacent to  $v$  such that the label of the nearest common ancestor associated to  $(u, v)$  with  $u \in N(v)$  is greater (according to operator  $>_{\ell}$ ) than  $mwe_v[1]$  which is used to define an order on the internal edges. Otherwise, if  $\text{FarLca}(v) = \emptyset$  then according to Macro  $\text{RecoverAdj}(v)$  the node  $v$  reset the computation of its adjacent internal edge to assure that every internal edge is taken into account. The recover phase is started at node  $v$  if  $v$  has no local outgoing edge (i.e.,  $\text{MergeEdge}(v) = \emptyset$ ) and  $v$ 's parent has taken into account the information associated to its current internal edge and stored in variable  $mwe_v$  (i.e.,  $mwe_{p_v} = mwe_v$ ). In this case, the guard of Rule  $R_{Min}$  is satisfied and  $v$  can execute Rule  $R_{Min}$  to update its variable  $mwe_v$  with the information of its next local internal edge. The information given by a child are stopped at the nearest common ancestor  $v$  of the corresponding internal edge since Macro  $\text{FarLcaChild}(v)$  selects only  $mwe_u$  from a child  $u$  such that  $mwe_u[1] \leq_{\ell} \ell_v$ .  $\square$

**Lemma 11** *Let  $T = (V_T, E_T)$  a tree and any edge  $(x, y) \in E_T$ . Eventually, if  $(x, y)$  is not part of a minimum spanning tree of the network then  $(x, y)$  is removed from  $T$ .*

**Proof.** We assume that  $T = (V_T, E_T)$  is a coherent tree, otherwise Rule  $R_{\odot}$  and  $R_{\ell}$  are used to break the cycles and to correct node labels. Let an edge  $(x, y) \in E_T$  (w.l.o.g.  $p_y = x$ ) which is not part of a minimum spanning tree. As the network has a finite size then there is a time after

which there exists no merging between two trees in the network. Thus according to Lemma 10 each internal edge  $e$  of  $T$  is put back up in  $T$  until reaching the nearest common ancestor associated to  $e$ . Since  $(x, y)$  is not in a minimum spanning tree, there is an edge  $(u, v)$  such that  $w(u, v) < w(x, y)$  and  $(x, y)$  is on the path between  $u$  and  $nca(u, v)$  or between  $v$  and  $nca(u, v)$ . So, there is a time such that  $mwe_y = (w(u, v), nca(u, v))$  and  $mwe_x = (w(u, v), nca(u, v))$  according to Lemma 10. Then Predicate  $NewFrag(y)$  returns true because we have  $mwe_y = mwe_x, mwe_y[1] \neq ld_y$  and  $w(x, y) > w(u, v)$ . Therefore,  $y$  can execute Rule  $R_{\frac{1}{2}}$  to create a new tree rooted at  $y$  and as a consequence the edge  $(x, y)$  is removed from  $T$ .  $\square$

**Lemma 12** *Let two distinct trees  $T_i = (V_{T_i}, E_{T_i})$  and  $T_j = (V_{T_j}, E_{T_j})$  with  $i \neq j$ . Let an edge  $(x, y)$  such that  $(x, y)$  is part of a minimum spanning tree and  $x \in T_j$  and  $y \in T_i$ . Eventually  $(x, y)$  is used to merge the trees  $T_i$  and  $T_j$ .*

**Proof.** We assume that  $T_i = (V_{T_i}, E_{T_i})$  and  $T_j = (V_{T_j}, E_{T_j})$  are coherent trees, otherwise Rule  $R_{\odot}$  and  $R_{\ell}$  are used to break the cycles and to correct node labels. According to Lemma 9, the merging edge  $(x, y)$  is computed by the root and it starts the merging phase since only the root can choose the edge of minimum weight leading to another tree to use in order to make a merging. In the remainder, we focus on tree  $T_i$  but the same arguments are also true for  $T_j$ .

When the root  $v$  has finished to compute its minimum outgoing edge from its fragment (i.e., we have  $mwe_v = MergeEdge(v)$ ) then  $v$  can execute Rule  $R_{\triangleright\triangleleft}$  because Predicate  $NeedReorientation(v)$  is satisfied since  $v$  has a coherent label. Note that we permit the creation of cycles of length two only if at least one node has a label equal to  $(\perp, \perp)$ . Indeed, during the merging phase the orientation is reversed on the path between the root of  $T_i$  and the node adjacent to the edge used for the merging, that is why a cycle is detected in Rule  $R_{\odot}$  if Predicate  $NeedReorientation(v)$  is not satisfied. Thus,  $v$  can change its variables  $p_v$  and  $\ell_v$  as following. If there is an edge  $(x, y)$  adjacent to  $v$  such that  $w(x, y) = mwe_v[0]$  and  $y = v$  then  $v$  selects  $x$  as its new parent (only if  $ld_v > ld_x$ ) and  $v$  changes its label to  $(\emptyset, \emptyset)$  to inform its subtree that the merging is done. Otherwise,  $v$  selects its child  $u$  such that  $mwe_u = mwe_v$  as its new parent and  $v$  changes its label to  $(\perp, \perp)$  to inform  $u$  that a merging is started.

Any other node  $v$  on the path between the root and the node adjacent to the merging edge take part in the merging phase when its parent has selected  $v$  as its new parent (i.e.,  $pp_v = ld_v$ ) and  $v$ 's parent label is equal to  $(\perp, \perp)$ . Thus, Predicate  $NeedReorientation(v)$  is satisfied and  $v$  can execute Rule  $R_{\triangleright\triangleleft}$  since  $mwe_v = MergeEdge(v)$  (otherwise Rule  $R_{Min}$  is executed to update its variable  $mwe_v$ ). So,  $v$  changes its variables  $p_v$  and  $\ell_v$  as described above for the root. Since the merging phase is done on a path, using the same argument one can show by induction that for any internal node  $v \in V_{T_i}$  on the path between the root and the node  $y$  adjacent to the merging edge (except for  $y$ ) we have  $p_v = \min\{ld_u : u \in Child(v) \wedge mwe_v = MergeEdge(v)\}$  and  $\ell_v = (\perp, \perp)$  and for the node  $y$  we have  $p_y = x$  and  $\ell_y = (\emptyset, \emptyset)$ .  $\square$

**Lemma 13** *Eventually all the nodes have a correct label in the new fragment resulting from a merging phase.*

**Proof.** According to Lemma 12, a merging phase is done using the minimum outgoing edge  $(x, y)$  between two distinct trees  $T_i$  and  $T_j$  if it is possible. Moreover, when the edge  $(x, y)$  is added by the extremity of minimum identifier, w.l.o.g. let  $y \in T_i$ , then  $y$ 's label is equal to  $(\emptyset, \emptyset)$  and the end of the merging phase is propagated in the resulting fragment  $T$ . In the reminder we focus on tree  $T_i$  but the same arguments are true for tree  $T_j$ .

Let the node  $v \in T_i$  such that  $p_v = y$  and  $\ell_v = (\perp, \perp)$  (i.e.,  $v$  is the child of  $y$  on the path between  $y$  and the old root of  $T_i$ ). Predicate  $NeedReorientation(v)$  is false because  $v$  is not a root node and the label of its parent  $y$  is not equal to  $(\perp, \perp)$ . Moreover, Predicate

$\text{EndReorientation}(v)$  is true since  $y$ 's label is equal to  $(\emptyset, \emptyset)$  and  $v$ 's label to  $(\perp, \perp)$ . Thus,  $v$  can execute Rule  $R_{\square}$  to modify its label to  $(\emptyset, \emptyset)$ . Using the same argument, one can show by induction that every node  $v$  on the path between  $y$  and the old root of  $T_i$  can execute Rule  $R_{\square}$ , thus there is a time such that we have  $\ell_v = (\emptyset, \emptyset)$ .

Now we show that the other nodes in  $T_i$  can execute Rule  $R_{\square}$ . Consider the node  $r \in V_{T_i}$  such that  $r$  is the old root of  $T_i$  and  $\ell_z = (\emptyset, \emptyset)$ . Let a node  $v \in V_{T_i}$  such that  $p_v = \text{ld}_z$ . Predicate  $\text{NeedReorientation}(v)$  is false because  $v$  is not a root node and the label of its parent is not equal to  $(\perp, \perp)$ . Moreover, Predicate  $\text{EndReorientation}(v)$  is true because  $\ell_z = (\emptyset, \emptyset)$  and  $v$ 's label is not equal to  $(\perp, \perp)$  since  $v$  is not on the path between  $y$  and the old root of  $T_i$ , and  $v$ 's label is different from  $(\emptyset, \emptyset)$ . Note that since the start of the merging phase, Predicate  $\text{TreeMerg}(v)$  is true because Predicate  $\text{NeedReorientation}(v)$  or  $\text{EndReorientation}(v)$  is true. So, Rule  $R_{\ell}$  cannot be executed by  $v$  and  $v$ 's label has not changed. Thus,  $v$  can execute Rule  $R_{\square}$  to modify its label to  $(\emptyset, \emptyset)$ . Using the same argument, one can show by induction that every node  $v$  on the path between  $z$  and a leaf node can execute Rule  $R_{\square}$ , thus there is a time such that we have  $\ell_v = (\emptyset, \emptyset)$ .

Every node  $v$  in the resulting fragment  $T$  can execute Rule  $R_{\ell}$  when the edge  $(x, y)$  is added in  $T$  and  $y$ 's label has been modified from  $(\emptyset, \emptyset)$  to its new label based on  $x$ 's label. Indeed, in this case for every node  $v$ , with  $v \neq y$  and  $v \in V_{T_i}$ , Predicate  $\text{TreeMerg}(v)$  is false and  $v$  can execute Rule  $R_{\ell}$ . Therefore, there is a time such that every node  $v$  in  $T$  has a correct label.  $\square$

**Lemma 14 (Convergence for MST)** *Starting from an illegitimate configuration for Algorithm MST, eventually Algorithm MST reaches in a finite time a legitimate configuration.*

**Proof.** We assume that there is a forest of trees  $T_i, 1 \leq i \leq n$ , in the network, otherwise according to Lemma 6 Rule  $R_{\square}$  is executed to remove the cycle from the network. Moreover, we assume also that the node's label are correct in tree  $T_i$ , otherwise according to Lemma 7 and 8 there is a time such that the node's label are corrected.

According to Lemma 10 and 11, if an edge  $(u, v) \in E_{T_i}$  and  $(u, v)$  is part of no minimum spanning tree of the network then  $(u, v)$  is removed from tree  $T_i$ . Thus, there is a time such that the existing fragments in the network are part of a minimum spanning tree. Moreover according to Lemma 9,12 and node labels are corrected after a merging phase according to Lemma 13. Since the size of the network is finite there is a finite number of merging. Therefore, in a finite time a spanning tree of minimum weight is computed by Algorithm MST.  $\square$

**Lemma 15 (Closure for MST)** *The set of legitimate configurations for MST is closed.*

**Proof.** Let  $\mathcal{C}$  a legitimate configuration such  $T = (V_T, E_T)$  is a minimum spanning tree of the network and an edge  $(x, y) \in E_T$ . To be illegitimate, the configuration  $\mathcal{C}$  must contain an edge  $(x, y)$  such that it exists an edge  $e = (u, v)$  with  $w(u, v) < w(x, y)$  and  $(x, y)$  and  $(u, v)$  are included in the same fundamental cycle  $C_e$ . Thus, this imply that the edge  $e$  is not used to verify if it is possible to replace an edge of  $C_e$  with  $(u, v)$  which contradicts Lemma 10 and 11. Moreover, since  $T$  is a spanning tree then no merging is done in the network. Therefore, starting from a legitimate configuration for Algorithm MST a legitimate configuration is preserved.  $\square$