A correction on Shiloach's algorithm for minimum linear arrangement of trees

J. L. Esteban[1] and R. Ferrer-i-Cancho[2]

[1] Logic and Programming, LOGPROG Research Group, Departament de Ciències de la Computació, Universitat Politècnica de Catalunya, Campus Nord, Edifici Omega. Jordi Girona Salgado 1-3. 08034 Barcelona, Catalonia, Spain

[2] Complexity & Qualitative Linguistics Lab, LARCA Research Group, Departament de Ciències de la Computació, Universitat Politècnica de Catalunya, Campus Nord, Edifici Omega. Jordi Girona Salgado 1-3. 08034 Barcelona, Catalonia, Spain

Abstract

More than 30 years ago, Shiloach published an algorithm to solve the minimum linear arrangement problem for undirected trees. Here we fix a small error in the original version of the algorithm and discuss its effect on subsequent literature.

Keywords: minimum linear arrangement, trees, Shiloach's algorithm.
AMS subject classification: 05C05, 68R10.

Introduction

More than 30 years ago, Shiloach published an $\mathcal{O}(n^{2.2})$ algorithm to solve the minimum linear arrangement problem for undirected trees [1]. A few years after, Chung published two different algorithms for solving the same problem [2]. The first one has cost $\mathcal{O}(n^2)$ and it is quite similar to Shiloach's algorithm. The second one has cost $\mathcal{O}(n^\lambda)$, where $\lambda > \log 3/\log 2$. To our knowledge, Chung's second algorithm is still the most efficient algorithm for undirected trees. This is corroborated by surveys [3, 4, 5]. As far as we know, these algorithms have not been implemented and tested. We implemented Shiloach's algorithm and found an error, which is the subject of this note.

The error and its correction

At the bottom of p. 18, Shiloach defines

$$S_0 = ... + p_0(n_* + 1) \quad (1)$$

$$S_1 = ... + p_1(n_* + 1) - 1. \tag{2}$$

An accurate derivation of S_0 and S_1 (see below) indicates that these definitions should read as

$$S_0 = ... + p_0(n_* + n_0 + 1) \tag{3}$$
$$S_1 = ... + p_1(n_* + n_0 + 1) - 1. \tag{4}$$

This little mistake implies that part b) of Theorem 3.1.2 is wrong. The error concerns steps 6-7 of Shiloach's algorithm (Section 3.2 of his article, pp. 19-20). Shiloach omitted the proof of that theorem arguing that it "is by a straightforward calculation which follows from elementary definitions" (pp. 19).

We discovered this mistake trying to understand why our implementation of Shiloach's algorithm failed for complete binary trees of k levels with $k \geq 5$. For trees with $k \geq 1$, the solution of the m.l.a. is [6]

$$D_{min} = 2^k \left(\frac{k}{3} + \frac{5}{18} \right) + (-1)^k \frac{2}{9} - 2. \tag{5}$$

For $k = 5$, Eq. 5 gives $D_{min} = 60$ while our original implementation of Shiloach's algorithm gave $D_{min} = 46$. Once we corrected the definitions of S_0 and S_1, our implementation of Shiloach's algorithm ceased to give wrong results.

To understand the little errors in the definitions of S_0 and S_1, notice that a type B arrangement, as defined in Theorem 3.1.1 a), of a tree $T(\alpha)$ depends on a certain calculated parameter p_α (where α is either 0 or 1), and consists of placing the tree $T_* = T(\alpha) - (T_1, ..., T_{2p_\alpha - \alpha})$ at the center surrounded by subtrees $T_1, ..., T_i, ..., T_{2p_\alpha - \alpha}$ as indicated in Figs. 4(a) and 4(b) of Shiloach's article. $\alpha = 0$ indicates that the tree T is a free tree (Fig. 4(a)) and $\alpha = 1$ indicates that T is an anchored tree (Fig. 4(b)). Part b) of Theorem 3.1.2 defines the cost of an arrangement of type B for a tree T as

$$C_\alpha = C[\pi, T(\alpha)] = \sum_{\substack{i=1 \\ i \text{ is odd}}}^{2p_\alpha - 1} C[\pi_i, \overrightarrow{T}(v_i)] + \sum_{\substack{i=1 \\ i \text{ is even}}}^{2p_\alpha - 2\alpha} C[\pi_i, \overleftarrow{T_i}(v_i)] + C[\pi_*, T_*] + S_\alpha, \tag{6}$$

Let us consider the subtree $\overrightarrow{T}(v_i)$. In this case, $C[\pi_i, \overrightarrow{T}(v_i)]$ includes the cost (or length) of the anchor of $\overrightarrow{T}(v_i)$, but notice that the cost of the anchor is only a part of the cost of the edge joining node v_i to T_*. Thus, the couple of summations in Eq. 6 comprise the cost of the anchors of every anchored subtree T_i, but the edge joining any T_i to T_* is longer that the anchor of T_i. S_α is added to account for the missing part of the cost, which in case that $\alpha = 1$ also has to account for the cost of the anchor of T. When $\alpha = 0$, T_* has p_0 subtrees to its left and p_0 subtrees to its right. In contrast, when $\alpha = 1$, T_* has p_1 subtrees to its left and $p_1 - 1$ subtrees to its right. With this background in mind, a derivation of S_0 and S_1 is straightforward.

With the help of Fig. 4(a) and the definition of C_α, one obtains

$$S_0 = (n_3 + n_4) + 2(n_5 + n_6) + (p_0 - 1)(n_{2p_0-1} + n_{2p_0}) + p_0(Z + 1), \tag{7}$$

where Z is the number of vertices of the tree T_*. If $Z = n_*$ we get exactly Shiloach's definition of S_0. The problem is that the tree T has been decomposed into the subtrees $T_0, T_1, ..., T_i, ..., T_{2p_\alpha - \alpha}$ and theorem 3.1. (p. 18) defines n_* as

$$n_* = n - \sum_{i=0} n_i, \tag{8}$$

where n_i is the size of the i-th subtree. Thus T_* includes T_0 and then $Z = n_* + n_0$. Recalling that S_1 comprises the length of the anchor of $T(\alpha)$, Fig. 4(b) helps one to see that

$$S_1 = (n_2 + n_3) + 2(n_4 + n_5) + (p_1 - 1)(n_{2p_1-2} + n_{2p_1-1}) + p_1(Z + 1) - 1. \tag{9}$$

where $Z = n_* + n_0$ again.

Alternatively, the error could be in Shiloach's definition of n_* (recall Eq. 8), which should be replaced by

$$n_* = n - \sum_{i=1} n_i. \tag{10}$$

If that was the case, Theorem 3.1.2 would be correct but then Theorem 3.1.1, where n_* is defined, could be wrong. We discarded this alternative because our implementation with the new definition of n_* (Eq. 8) fails with other kinds of trees.

We note also that the version of Eq. 6 in Shiloach's article (middle of p. 19) has a couple of typos: $C[\pi, \overrightarrow{T_i}(v_i)]$ should read $C[\pi_i, \overrightarrow{T_i}(v_i)]$ and $C[\pi, T_*]$ should read $C[\pi_*, T_*]$.

Discussion

Chung's first algorithm is similar to Shiloach's: for certain values p and q, which play a role similar to Shiloach's p_α, Chung's first algorithm arranges vertices placing the tree $T_* = T - (T_{i_1}, ..., T_{i_{2p+1}})$ at the center surrounded by subtrees $T_{i_1}, ..., T_{i_{2p+1}}$, or placing the tree $T_* = T - (T_{i_1}, ..., T_{i_{2q}})$ at the center surrounded by subtrees $T_{i_1}, ..., T_{i_{2q}}$. In Chung's first algorithm, calculations that are equivalent to Shiloach's S_0 and S_1 appear within Properties 12 and 13 (p. 46). In particular, the bit

$$ns - \sum_{j=1}^{s} (s - j + 1)(t_{i_j} + t_{i_{2s-j+1}}) \tag{11}$$

in Property 12 corresponds to S_0. The bit

$$n(s+1) - \sum_{j=1}^{s} (s-j+1)(t_{i_j} + t_{i_{2s-j+1}}) - (s+1)t_{i_{2s+1}} \tag{12}$$

in Property 13 corresponds to S_1. Properties 12 and 13 are used, respectively, in Step 4 and Step 7 of Chung's first algorithm. While Shiloach calculates S_0 and S_1 by summation and omits one number in each, Chung proceeds by substraction from a maximum and omits no number. Furthermore, we have checked both properties and we find them correct. Chung's second algorithm (the one with subquadratic cost) also uses Properties 12 and 13, which are correct. Therefore, we conclude that Chung's algorithms are not affected by the error in Shiloach's algorithm reported above.

Beyond Shiloach's and Chung's algorithm, we expect that the error in Shiloach's algorithm does not affect or can be easily fixed because it concerns a very specific component of the algorithm. For instance, Shiloach's algorithm was parallelized by Díaz and colleagues [7]. The error does not affect their conclusions. Just correcting the formulae as indicate above suffices.

Acknowledgements

We thank M. Serna and J. Petit for their advice. JLE is funded by the project TASSAT2 (TIN2013-48031-C4-1-P) from MINECO (Ministerio de Economia y Competitividad). RFC is funded by the grants 2014SGR 890 (MACDA) from AGAUR (Generalitat de Catalunya) and also the APCOM project (TIN2014-57226-P) from MINECO.

References

[1] Y. Shiloach. A minimum linear arrangement algorithm for undirected trees. *SIAM J. Comput.*, 8(1):15–32, 1979.

[2] F. R. K. Chung. On optimal linear arrangements of trees. *Comp. & Maths. with Appls.*, 10(1):43–60, 1984.

[3] J. Petit. Addenda to the survey of layout problems. *Bulletin of the European Association for Theoretical Computer Science*, 105:177–201, 2011.

[4] Y.-L. Lai and K. Williams. A survey of solved problems and applications on bandwidth, edgesum, and profile of graphs. *Journal of Graph Theory*, 31(2):75–94, 1999.

[5] J. Díaz, J. Petit, and M. Serna. A survey of graph layout problems. *ACM Computing Surveys*, 34:313–356, 2002.

[6] F. R. K. Chung. A conjectured minimum valuation tree. *SIAM Review*, 20:601 604, 1978.

[7] J. Díaz, A. Gibbons, G. E. Pantziou, M. J. Serna, P. G. Spirakis, and J. Toran. Parallel algorithms for the minimum cut and the minimum length tree layout problems. *Theoretical Computer Science*, 181(2):267 – 287, 1997.