Kragujevac J. Math. 32 (2009) 117–122.

ON THE RULES FOR THE ELIMINATION OF THE NON-CANONICAL MORGAN TREES

Damir Vukičević

Faculty of Natural Sciences and Mathematics, Nikole Tesle 12, HR-21000 Split, Croatia
(e-mail: vukicevi@pmfst.hr)

(Received August 16, 2008)

Abstract. The concept of Morgan tree [6] is shown to be useful in generation of all non-isomorphic trees. Namely, to each tree one can assign canonical Morgan tree. Since, the number of Morgan trees [5, 1] is much larger then number of canonical Morgan trees, it is of interest to create an efficient algorithm that creates only a fraction of Morgan trees not eliminating the single canonical Morgan tree. Then, in the second step, non-canonical trees are eliminated. The rules for the recognition of non-canonical trees are proposed in [4, 3]. However, it seems that Rule 3 in [4] and Rule 1se in paper [3] are not correct. In this paper, we present the counter-examples to these rules.

1. INTRODUCTION

The number of all labeled trees with the prescribed number of vertices is much larger then the number of all non-isomorphic trees with the same number of vertices. Hence, if one is interested in generation of the set of all non-isomorphic graphs, it would be quite inefficient to generate all labeled trees and then to eliminate isomorphic ones. That's why a considerable effort has been put in finding a smaller classes of graphs that contain all non-isomorphic trees (with as little duplicates as possible).

The idea in paper [2] was to generate a subclass of labeled trees named physical trees by assigning labels consecutively in such way that each vertex to be labeled must be adjacent to an already labeled vertex. It can be seen that physical trees are proper subset of labeled trees and that they contain all non-isomorphic trees.

Further restriction was introduced in [6] with the concept of Morgan trees. Labeling follows this procedure:

1) Assign number 1 to any vertex v.

2) Assign numbers 2,…,$d(v)+1$ to its neighbors, where $d(v)$ is degree of vertex v.

3) Consider next vertex possessing the lowest label and having non-labeled neighbors and label them by the following (unused) consecutive numbers.

4) Repeat this procedure till all vertices are labeled.

We illustrate these three concepts by the following figure:

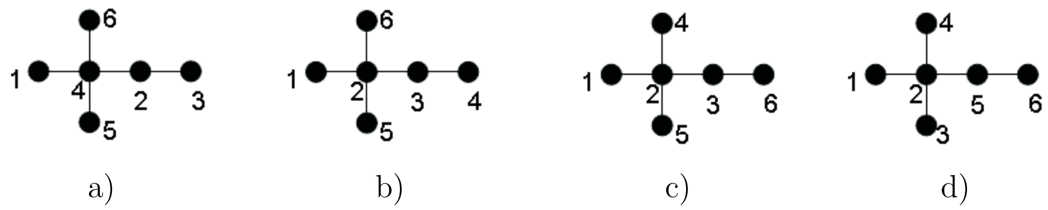


Figure 1. Labeled graphs

Graph on Figure 1a is a chemical tree, but it is not a physical tree since vertex labeled by 2 is not adjacent to previous vertices (i.e. to vertex labeled by 1). Graph of Figure 1b) is a physical tree, but it is not a Morgan tree, because vertex labeled by 4 should not be labeled until all neighbors of 2 got their labels. Graphs on Figures 1c) and 1d) are Morgan trees. For the sake of the simplicity vertex labeled by i will be denoted by v_i.

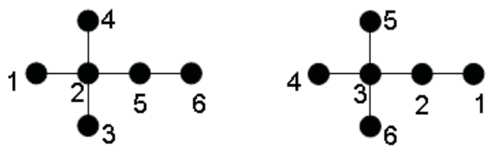
Morgan trees have an interesting property that each vertex (except 1) is labeled to a single vertex with the smaller label [4,3]. This can be utilized to assign to each graph CAM (condensed adjacency matrix) [4, 3] which is an ordered $(n-1)$-tuple (where n is the number of vertices of the graph) such that $CAM_i = j$ if v_j is the neighbor of v_{i+1} such that $j < i+1$. E. g., CAM of Morgan tree presented on Figures 1c) and 1d) are (1, 2, 2, 2, 3) and (1, 2, 2, 2, 5), respectively.

Note that different CAM, may correspond to the isomorphic graphs. In order to overcome this difficulty, the concept of lexicographical order is needed. On the set of all sequences of integers of length n, lexicographical order is defined by $a < b$ if and only if there is $j \in \{1, ..., n\}$ such that $a_i = b_i$ for each $i < j$ and $a_j < b_j$. The name lexicographical order is used since this is the ordering used in lexicons.

It is defined that Morgan tree M corresponding to unlabeled tree T is canonical if for each Morgan tree M' corresponding to the same unlabeled tree T, it holds $CAM\,(M) \geq CAM\,(M')$. In this case, we say that $CAM\,(M)$ is canonical CAM of T. Canonical CAMs uniquely determine non-isomorphic trees [4, 3]. Since, the number of Morgan trees is still much larger then the number of non-isomorphic trees, it is of interest to further restrict the class of Morgan trees.

LDF-Morgan trees are trees in which we use the same methodology of labeling as in Morgan trees with one additional rule – we always first assign the labels to the vertices with the smaller degree. Note that graph on the Figure 1c) is not LDF-Morgan tree, because during the labeling of neighbors of v_2 label 3 is given to the vertex of higher degree then vertices labeled by 4 and 5. The graph on Figure 1d) is LDF-Morgan tree.

It can be shown that each canonical Morgan tree is LDF-Morgan tree [4, 3]. Hence, indeed we can restrict ourselves to observing only LDF-Morgan trees. However, different LDF-Morgan trees can still correspond to the same graph. In Figure 2, the isomorphic LDF-Morgan trees are presented:



a) b)

Figure 2. LDF-Morgan trees

CAMs of the isomorphic graphs presented on the Figures 2a) and 2b) are (1, 2, 2, 2, 5) and (1, 2, 3, 3, 3), respectively. Therefore further rules which eliminate non-canonical LDF-Morgan trees are needed. In papers [4, 3] several such rules are presented.

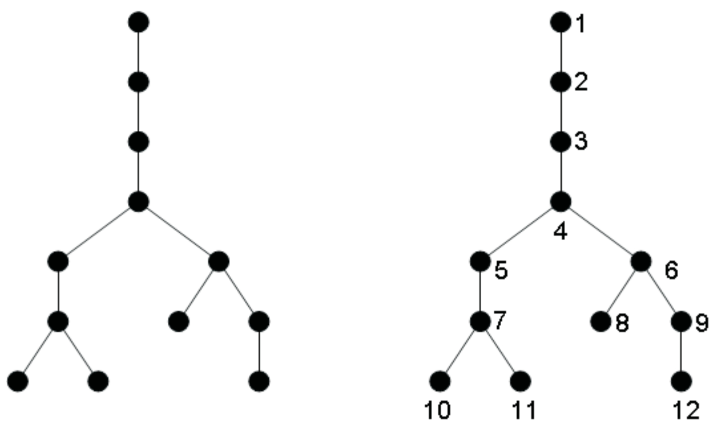
However, we claim that these rules are not all correct. Namely, we claim that the following rules are not correct:

Rule A (*Rule 3* in [4]) Delete any $CAM = (1, 2, ..., X, ..., X, Z)$ if $Z < n - 1$ (in words [4]: if the last digit is not equal to $n - 1$ and it is preceded by at least two identical digits, then the CAM has to be deleted).

Rule B (*Rule 1se* in [3]) Consider paths connecting vertex 1 (an end point) and any other end point x of T in turn, and write down the valences belonging to this path starting with the degree of vertex 1, then with the degree of its first neighbor, then with the degree of its second neighbor, etc. In this way a valence code has been created. Repeat the same procedure starting with a vertex x of the same path if the "reversed" valence code is less then the original code, delete the underlying CAM.

2. COUNTER-EXAMPLE TO RULE A

Let T be a tree given by Figure 3a). It can be easily checked manually (or by computer) that its canonical Morgan tree is given by Figure 3b).



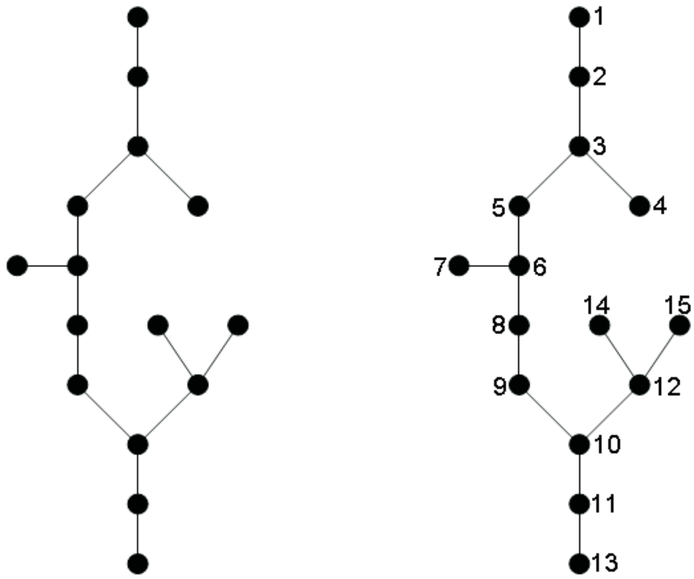
a) b)

Figure 3. Unlabeled tree and corresponding canonical Morgan tree

CAM that corresponds to canonical Morgan tree reads as (1, 2, 3, 4, 4, 5, 6, 6, 7, 7, 9). It can be easily seen that digits on the places $n - 2$ and $n - 3$ are the same, but never the less the last digit is different then $n - 1$ which is contradiction.

3. COUNTER-EXAMPLE TO RULE B

Let T be a tree given by Figure 4a). It can be easily checked manually (or by computer) that its canonical Morgan tree is given by Figure 4b).



a) b)

Figure 4. Unlabeled tree and corresponding canonical Morgan tree

Figure 4b) corresponds to CAM (1, 2, 3, 3, 4, 5, 5, 6, 9, 9, 11, 12, 12, 14). Now, let us observe the valence sequence from v_1 to v_{13}. This sequence reads as: (1,2,3,2,3,2,2,3,2,1). The reverse of this sequence reads as: (1,2,3,2,2,3,2,3,2,1). Note that reversed sequence is smaller then the original one. Hence, this Morgan tree according to the rule B should be eliminated. But, this is not true – this is canonical tree.

Acknowledgements: Partial support of Ministry of Science, Education and Sport of the Republic of Croatia (grants no. 037-0000000-2779 and 177–0000000-0884) is gratefully acknowledged.

References

[1] T. Došlić, *Morgan Trees and Dyck Paths*, Croat. Chem. Acta. **75** (2002), 881-889.

[2] J. V. Knop, W. R. Müller, K. Szymanski, H. W. Kroto and N. Trinajstić, *Computer Enumeration and Generation of Physical Trees*, J. Comput. Chem. **8** (1987), 549-554

[3] I. Lukovits, *Isomer Generation: Semantic Rules for Detection of Isomorphism*, J. Chem. Inf. Comput.Sci. **40** (2000) 361-366.

[4] I. Lukovits, *Isomer Generation: Syntatic Rules for Detection of Isomorphism,* J. Chem. Inf. Comput.Sci. **39** (1999) 563-568.

[5] I. Lukovits and I. Gutman, *On Morgan Trees,* Croat. Chem. Acta. **75 (2)** (2002) 563-576.

[6] H. L. Morgan, *The Generation of a Unique Machine Description for Chemical Structures,* J. Chem. Doc. **5** (1965) 107-113.