MINIMUM SUB-TREE COVERING EVERY COLOR

THINH D. NGUYEN

ABSTRACT. We show that constructing a sub-tree of a given vertex-colored graph that covers every color is a hard task.

1. PROBLEM STATEMENT

Let us formalize our problem as follows.

Definition 1. COLORED NODE COVERING (**CNC**) problem:

Input: An instance consists of the following:

- a connected graph $G = (V, E)$
- a positive integer $k \leq |V|$ and a map $C : V \rightarrow \{1, 2, \dots, k\}$, called coloring map with $|C(V)| = k$;
- a positive integer m

Question: Is there a subtree $T = (V', E')$ of G that $|E'| \leq m$ and $|C(V')| = k$?

We reduce STEINER TREE (ST) problem to ours. The proof of the hardness of STEINER TREE can be found in [1].

Claim 2. *We have that $ST \leq_p CNC$*

2. REDUCING ST TO CNC

Let us recall ST can be defined in decisional form as follows:

Instance:

- a connected graph $G = (V, E)$;
- a subset of the nodes $R \subseteq V$, called terminal nodes;
- a positive integer m .

Question: Is there a subtree of G that includes all the nodes of R and that contains at most m edges?

Let us show a polynomial-time reduction from ST to CNC by building an instance of CNC starting from a generic instance of ST. Given a generic instance s of ST, defined by a graph $G = (V, E)$, the set of terminals $R = \{v_1, v_2, \dots, v_k\} \subseteq V$, and the upper-bound m , define an instance c of CNC as the following.

- the same graph $G = (V, E)$
- the coloring $C : V \rightarrow \{1, 2, \dots, k, k+1\}$, $C(v_i) = i$ for $1 \leq i \leq k$ and $C(v) = k+1$ otherwise
- the same upper bound m

Key words and phrases. cover, graph, tree.
Perebor.

If the induced graph $G[R]$ is connected, then both s and c can be determined in polynomial time in $|V|$ (in fact, constant time), since the spanning tree of $G[R]$ is a tree with the minimal number, $k - 1$ of edges that can include all vertices of R .

Otherwise let us assume $G[R]$ is not connected. Then any subtree of G that contains all nodes of R must contain at least one node that is not in R , which must be colored $k + 1$. For any subtree H of G , H includes all the nodes of R and contains at most m edges if and only if H includes nodes of every possible color under C and contains at most m edges. That is, s has a YES answer if and only if c has a YES answer.

It is easy to see that the transformation from s to c can be done in polynomial time (in fact, constant time). So we have reduced ST to CNC by a polynomial-time reduction.

3. CONCLUSION

Garey and Johnson [1] shapes their theory based on previous primal works of Cook, Levin and Karp. Johnson [2] moves on with the guide to this theory. As long as we study a mathematical conjecture, we should encourage ourselves of having enough labouring hours on popular maths books like these. Then, reading some articles on theory of computing like [3] is a good practice. Only after that, could we think of the ultimate final for all mathematics sciences.

REFERENCES

1. Michael R. Garey, David S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*
2. David S. Johnson, *The NP-Completeness Column: An Ongoing Guide*. pp.393-405
3. Phan Dinh Dieu, Le Cong Thanh, Le Tuan Hoa, *Average Polynomial Time Complexity of Some NP-Complete Problems*, Theor. Comput. Sci. 46(3): pp.219-237 (1986)
Current address: Department of Mathematics, Moscow State University
Email address: kosmofarmer@yandex.com