THE MAXIMUM LEAF SPANNING TREE PROBLEM FOR CUBIC GRAPHS IS NP–COMPLETE

By

Paul Lemke

IMA Preprint Series # 428

July, 1988

THE MAXIMUM LEAF SPANNING TREE PROBLEM FOR CUBIC GRAPHS IS NP-COMPLETE

PAUL LEMKE*

Abstract. An NP-completeness proof is given for the MAXIMUM LEAF SPANNING TREE problem for cubic graphs and an extension of the proof to regular graphs of degree d, for all fixed $d \geq 5$.

The problem:

INSTANCE: A graph G which is regular of degree d, and an integer K.

QUESTION: Does G posess a spanning tree with at least K vertices of degree 1?

is known to be NP-complete for $d = 4$ ([1]). I will prove in this paper that it is also NP-complete for $d = 3$, and will give a sketch of an extension of the proof to all fixed $d \geq 5$. For $d = 3$ I will show NP-completeness by proving that the following more specific problem is NP-complete:

INSTANCE: A cubic graph G.

QUESTION: Does G posess a spanning tree with no vertices of degree 2?

The proof is by reduction from EXACT COVER BY 3-SETS:

INSTANCE: Positive integers n and m, subsets $S_1, S_2, \ldots, S_m$ of $\{1, 2, \ldots, n\}$, with $|S_i| = 3 \quad \forall i$.

QUESTION: Is there a subset $Q \subseteq \{1, 2, \ldots, m\}$ such that $\bigcup_{i \in Q} S_i = \{1, 2, \ldots, n\}$ and $S_{i_1} \cap S_{i_2} = \phi \quad \forall i_1, i_2 \in \{1, 2, \ldots, m\},\ i_1 \neq i_2$?

which is NP-complete ([2]).

Proof. Given an instance of EXACT COVER BY 3-SETS we construct a graph G as follows:

Define the numbers $a_0, a_1, \ldots, a_n$ by:

$$a_0 = 2m$$
$$a_j = |\{i | j \in S_i\}| \quad \text{for } j = 1, 2 \ldots, n.$$

For the construction given below we need $a_j \geq 3 \quad \forall j$. If this is not the case, we can always make it so by adding duplicate sets to $\{S_1, S_2, \ldots, S_m\}$.

*Institute for Mathematics and its Applications, 514 Vincent Hall, 206 Church St. S.E., University of Minnesota, Minneapolis, MN 55455

The vertex set V of G is defined to be the union of the pairwise disjoint sets $U_0, U_1, \ldots, U_n, W$, and X, where:

$$
\begin{aligned}
|U_j| &= 10a_j - 18 \quad \text{for } j = 0, 1, \ldots, n \\
W &= \{w_1, w_2, \ldots, w_m\} \\
X &= \{x_1, x_2, \ldots, x_m\}.
\end{aligned}
$$

The edge set E of G consists of those edges in the subgraph induced on G by U_j (call it G_j) for each j, as shown in Figure 1a, and those in the subgraphs induced by the sets $\{u_{0,2i-1}, u_{0,2i}, w_i, x_i, v_{i1}, v_{i2}, v_{i3}\}$ (call them H_i) for $i = 1, 2, \ldots, m$, as shown in Figure 2a. The vertices $\{u_{j\ell}\}$, $\ell = 1, 2, \ldots, a_j$ are members of U_j for $j = 1, 2, \ldots, n$ as indicated in Figure 1a. The vertices $\{v_{ik}\}$, $i = 1, 2, \ldots, m$, $k = 1, 2, 3$ are the $3m$ members of $\bigcup_{j=1}^{n} \bigcup_{\ell=1}^{a_j} \{u_{j\ell}\}$ re-labelled so that:

$$
(1) \qquad |\{v_{i1}, v_{i2}, v_{i3}\} \cap \bigcup_{\ell=1}^{a_j} \{u_{j\ell}\}| = |S_i \cap \{j\}| \quad \text{for } i = 1, 2, \ldots, m, \quad j = 1, 2, \ldots, n.
$$

Because of the way the numbers $\{a_j\}$ and the vertices $\{u_{j\ell}\}$ are defined, this is always possible.

If T is a spanning subgraph of G, then Figure 1b shows how it is possible to assign edges of G_i to T so that T spans G_j, is connected on G_j, has no cycles on G_j, and all its vertices on G_j are of odd degree, for each j. If the answer to the EXACT COVER BY 3-SETS question is "yes", so that the subset Q of $\{1, 2, \ldots, m\}$ exists with the properties indicated, then we additionally assign edges in H_i to T for each i as indicated on the left side of Figure 2b if $i \notin Q$, and as indicated on the right side of Figure 2b if $i \in Q$. Clearly all of T's vertices will be of odd degree. Also, given any $j \in \{1, 2, \ldots, n\}$ we know that there exists an $i \in Q$ such that $j \in S_i$, and hence from (1) we know that there is a $k \in \{1, 2, 3\}$ such that $v_{ik} \in \bigcup_{\ell=1}^{a_j} \{u_{j\ell}\} \subset U_j$. Thus there is a path in T between any vertex in U_j and any vertex in U_0. Also, $\{w_i\}$ and $\{x_i\}$ are clearly connected by a path to a vertex in U_0, so that T is connected. Finally, since T had exactly $n + 1 + 2m$ components before assigning to it the edges from the subgraphs $\{H_i\}$, and since we assigned $2(m - |Q|) + 5|Q|$ edges to it from these subgraphs (and since we must have $3|Q| = n$ to make an exact 3-covering possible), we see that T has just enough edges to make it connected, so that T has no cycles and must be a tree. We conclude that if the answer to the EXACT COVER BY 3-SETS question is "yes", the answer to whether G posesses a spanning tree with all vertices of odd degree is also "yes".

Conversely, assume that the answer to the above graph problem is "yes", so that G has a spanning tree T with all vertices of odd degree. With a little trial and error it can be seen that the only ways to assign edges from the repeating units of the graphs $\{G_j\}$ to T so that T is a spanning tree of G with no vertices of degree 2 are as shown in Figure 1b.

The subgraph of T induced by U_j must therefore be connected and hence a tree for each j. It is also easy to see that the only way to assign edges of the subgraphs $\{H_i\}$ to T are the two ways shown in Figure 2b. (given that the 5 edges incident from the outside of H_i are in T, which they must be). Let Q be the set of $i \in \{1, 2, \dots, m\}$ for which T contains 5 edges from H_i (as in the left side of Figure 2b). Then as before we must have $|Q| = \frac{n}{3}$ in order to give T the right number of edges to be a tree. If $(\bigcup_{i \in Q}\{v_{i1}, v_{i2}, v_{i3}\}) \cap \bigcup_{\ell=1}^{a_j}\{u_{j\ell}\} = \phi$ for any $j \in \{1, 2, \dots, n\}$ then there can be no path in T between a vertex of G_j and any vertex outside G_j, so that T would be disconnected. We must therefore assume the contrary. With (1) this implies $\sum_{i \in Q}|S_i \cap \{j\}| \geq 1$ for each $j \in \{1, 2, \dots, n\}$ so that $\{1, 2, \dots, n\} \subseteq \bigcup_{i \in Q} S_i$, so that we have a covering for $\{1, 2, \dots, n\}$, and because $|Q| = \frac{n}{3}$ and $|S_i| = 3$ for each i, it is an exact covering. Thus the answer to the EXACT COVER BY 3 SETS question is "yes". □ □

COROLLARY. *The MAXIMUM LEAF SPANNING TREE problem on cubic graphs is NP-complete.*

Proof. If T is a spanning tree of a graph G with vertex set V, and if d_k is the number of vertices of degree k in T for each k, then if G is a cubic graph we have:

$$d_1 + d_2 + d_3 = |V|$$

$$\text{and} \qquad d_1 + 2d_2 + 3d_3 = 2|V| - 2,$$

from which we derive:

$$2d_1 + d_2 = |V| + 2,$$

so that the number of leaves is maximized when the number of vertices of degree 2 is minimized. In light of the previous result, therefore, the MAXIMUM LEAF SPANNING TREE problem for cubic graphs is NP-complete when $K = \frac{1}{2}|V| + 1$.

COMMENT. It is possible to economize in the construction of G to use fewer vertices for a given instance of EXACT COVER BY 3-SETS, especially if one is willing to sacrifice the 3-connectedness of G (as constructed above G is 3-connected under fairly general conditions, or can be made so just by adding some simple structures to G).

COMMENT. As a by-product, it follows that the following two problems are NP-complete:

INSTANCE: A graph G.

QUESTION: Does G posess a spanning tree with no vertices of degree 2?

INSTANCE: A graph G

QUESTION: Does G posess a spanning tree all of whose vertices have odd degree?

The above proof can be extended to prove that the problem:

INSTANCE: A graph G which is regular of degree d.

QUESTION: Does G posess a spanning tree whose vertices all have either degree 1 or degree d?

is NP-complete for all fixed $d \geq 5$. This is done by constructing a graph G from structures like those which are illustrated in Figures 3 and 4 for $d = 5$. NP-completeness is then shown by reduction from EXACT COVER BY $(d-1)$-SETS, which is NP-complete (this can easily be shown by reduction from EXACT COVER BY 3-SETS). It then follows that the MAXIMUM LEAF SPANNING TREE problem on regular graphs of degree d is NP-complete for all fixed d.

Although the structure in Figure 3 is only 2-connected, it can probably be modified to make it d-connected in general.

REFERENCES

[1] David S. Johnson, unpublished manuscript.

[2] R.M. Karp, *Reducibility among combinatorial problems*, R.E. Miller and J.W. Thatcher (eds.), in *Complexity of Computer Computations*, Plenum Press, New York, 1972, pp. 85–103.

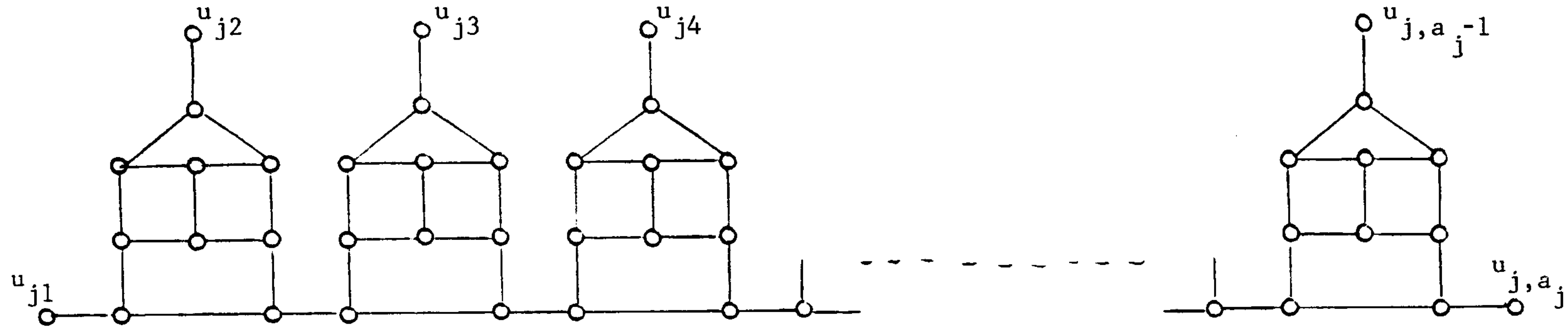

Figure 1a. The subgraph G_j induced on G by U_j for $j = 1, 2, \dots, n$.

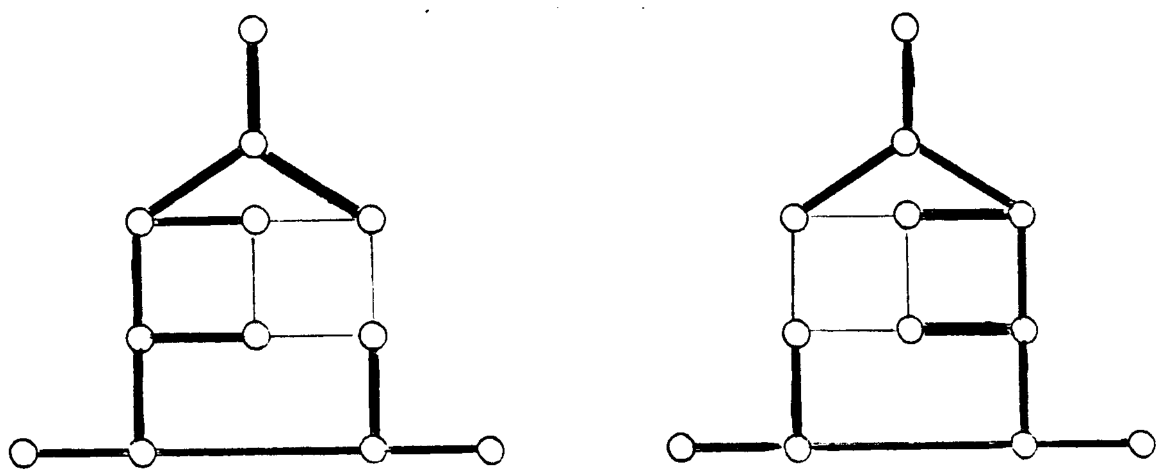

Figure 1b. The only two ways to assign the edges of a repeating sub-unit of the graphs G_j to a spanning tree T of G so that its nine interior vertices all have odd degree in T.

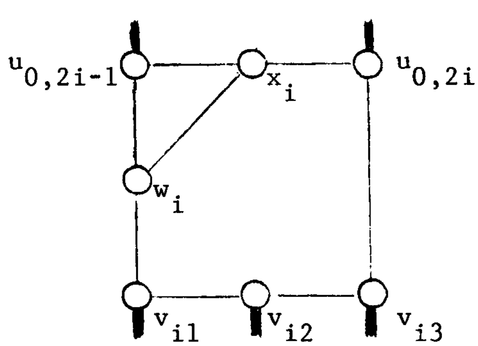

Figure 2a. The subgraph H_i induced on G by $\{u_{0,2i-1}, u_{0,2i}, v_{i1}, v_{i2}, v_{i3}, w_i, x_i\}$, along with part of the five edges of G incident on H_i from outside H_i.

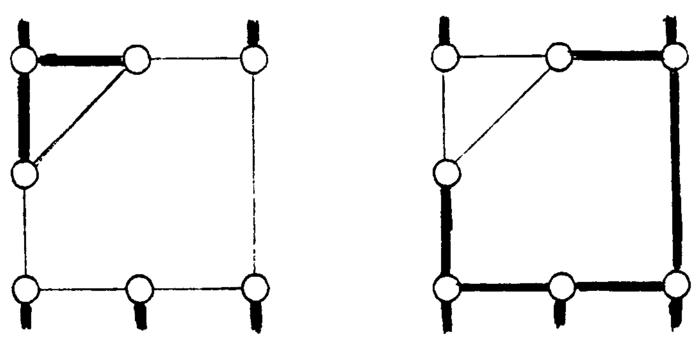

Figure 2b. The only two ways to assign the edges of H_i to a spanning tree T of G so that the vertices of T all have odd degree.

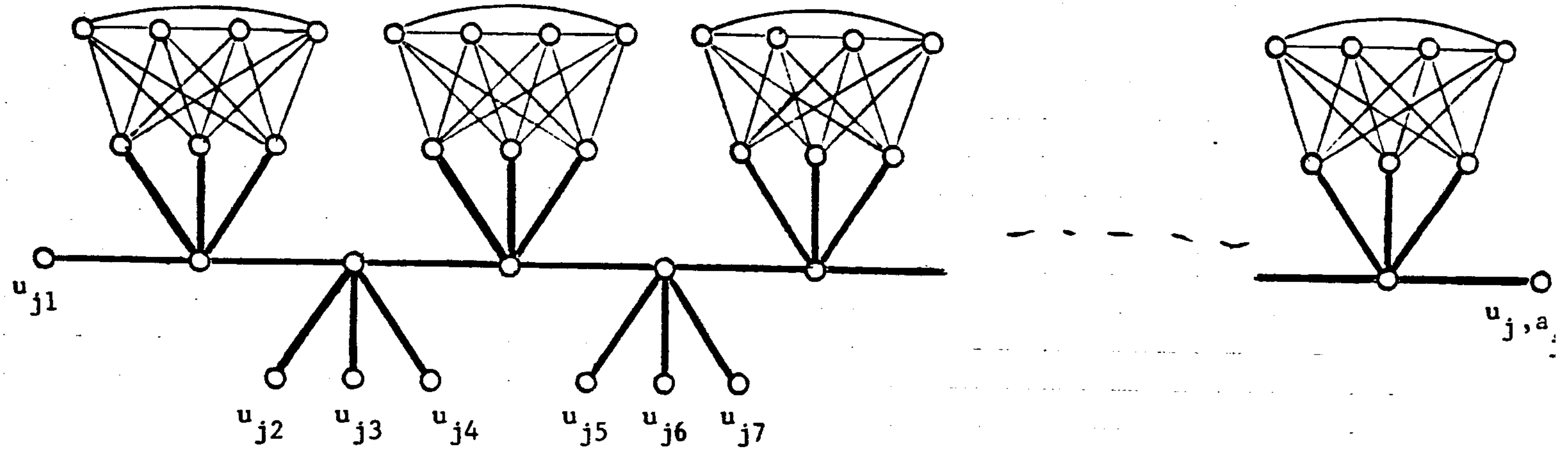

Figure 3. A structure for the case $d = 5$ anologous to the structure of Figure 1a. Edges drawn in dark are edges which **must** be contained in any spanning tree of G whose edges all have degree 1 or d.

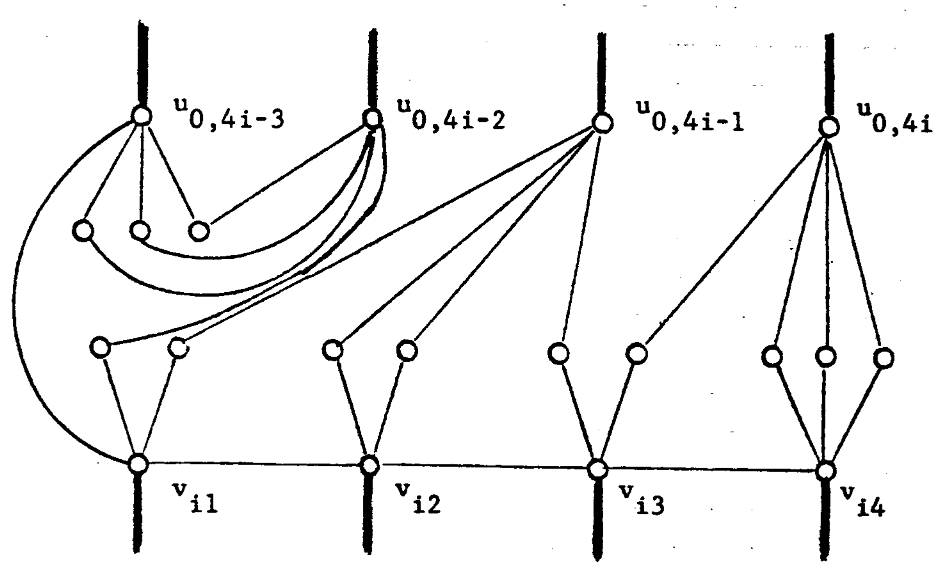

Figure 4. The structure for the case $d = 5$ anologous to the structure of Figure 2a. The subgraph induced by the 20 vertices shown contains other edges not shown between the unlabelled vertices, so as to give each of them degree $d = 5$. In the desired spanning tree, we must either have vertices $u_{0,4i-2}$,$u_{0,4i-1}$ and $u_{0,4i}$ be of degree 5 and all of the other 17 vertices shown above be of degree 1, or else have $u_{0,4i-3}$, v_{i1}, v_{i2}, v_{i3} and v_{i4} be of degree 5 and the others be of degree 1.