The Maximum Independent Set Problem

For Cubic Planar Graphs

by

James E. Burns

Computer Science Department
Indiana University
Bloomington, IN 47405

TECHNICAL REPORT NO. 177

The Maximum Independent Set Problem
For Cubic Planar Graphs

by

James E. Burns

August, 1985

THE MAXIMUM INDEPENDENT SET PROBLEM FOR CUBIC PLANAR GRAPHS

James E. Burns

Indiana University

June 1985

ABSTRACT

A maximum independent set of a graph is a set of vertices with maximum cardinality such that no pair of vertices is connected by an edge. Choukhmane and Franco have presented a polynomial time approximation algorithm for the maximum independent set problem in cubic planar graphs. If M is taken as the ratio of the size of the independent set produced by their algorithm to the size of a maximum independent set of the input graph, then they show that their algorithm gives $M \geq 6/7$ for any cubic planar graph and $M \geq 7/8$ for a triangle-free cubic planar graph. We show that their algorithm gives $M \geq 7/8$ for all cubic planar graphs.

1. INTRODUCTION

The maximum independent set problem, defined formally in the next section, is a problem of wide interest. Unfortunately, it is known to be an NP-complete problem, so there is little hope for an algorithm which can compute an exact solution in polynomial time. (For a comprehensive treatment of NP-complete problems, see Garey and Johnson[7].) It is therefore useful to consider solutions which are approximate, but can be computed efficiently.

For any algorithm W which produces an independent set I for some input graph G, let $M_W(G)$ be the ratio of the size of I to the size of a maximum independent set of G. For any class of graphs $\mathcal{G}$, let $M_W(\mathcal{G})$ be the minimum value of $M_W(G)$ over all $G \in \mathcal{G}$. We define $\mathcal{G}_P = \{\text{planar graphs}\}$, $\mathcal{G}_{3P} = \{\text{cubic planar graphs}\}$, and $\mathcal{G}_{3TFP} = \{\text{cubic triangle-free planar graphs}\}$. Lipton and Tarjan[10] give an $O(n \log n)$ time algorithm, L, which is shown to give an approximation such that $M_L(\mathcal{G}_P) \geq \left(1 - O(1/\sqrt{\log \log n})\right)$, where n is the number of vertices in the input graph. Unfortunately, as shown by Chiba *et al*[2], n must be rather large before $M_L(\mathcal{G}_P) \geq 1/2$. Chiba *et al* then describe an algorithm, C, which executes in

Author's addresses: J.E. Burns, Computer Science Department, 101 Lindley Hall, Indiana University, Bloomington, Indiana 47405.

time $O(n \log n)$ and show $M_C(\mathcal{G}_P) > 1/2$. Baker[1] improve this result by giving a family of algorithms, B_d, for the maximum independent set problem such that $M_{B_d}(\mathcal{G}_P) \geq d/(d+1)$, and which execute in time $O(8^d dn)$. Note that Baker's algorithms are linear for any fixed d.

Choukhmane and Franco[3] describe a polynomial algorithm, A, for the maximum independent set and show that $M_A(\mathcal{G}_{3P}) \geq 6/7$ and $M_A(\mathcal{G}_{3TFP}) \geq 7/8$. Approximation algorithms are relevant for cubic planar graphs because the maximum independent set problem has been shown to be NP-complete even on this subset of graphs[6]. Algorithm A relies on an algorithm of Hadlock[8] which finds a maximum bipartite subgraph of a planar graph. Hadlock's algorithm in turn relies on an algorithm of Edmonds[4] for maximum weighted matching. Since Edmond's algorithm is $O(n^2)$ for planar graphs, Choukhmane and Franco's algorithm is no better than $O(n^2)$. However, their algorithm can be immediately improved to linear if a linear time algorithm can be found for finding a maximum bipartite subgraph in a planar graph, although this appears to be difficult. To achieve an approximation of 7/8, Baker's algorithm requires that $d = 7$, which implies a rather large constant in the linear running time. It is therefore reasonable to consider using Choukhmane and Franco's algorithm.

We do not improve Choukhmane and Franco's algorithm, but instead improve the analysis. While their analysis relies on the graph theoretical results of Hopkins and Staton[9] and Staton[12], we provide a more direct proof. More importantly, we show that the bound of 7/8 holds for all cubic planar graphs, not just triangle-free cubic planar graphs. That is, $M_A(\mathcal{G}_{3P}) \geq 7/8$.

Section 2 formally defines the problem and presents a version of Choukhmane and Franco's algorithm. The algorithm is analyzed in Section 3.

2. THE MAXIMUM INDEPENDENT SET PROBLEM

Let G be a graph without loops or multiple edges and with vertices $G(V)$ and edges $E(G) \subseteq V(G) \times V(G)$. G is *r-regular* if every vertex of G has degree r. A *cubic* graph is 3-regular. An *independent set* of G is a subset V' of $V(G)$ such that there is no edge (v, v') in $E(G)$ with both v and v' in $V(G)$. The cardinality of a maximum independent subset of any subset V' of V is denoted by $\mu(V')$. The *maximum independent set problem* requires finding an independent set of G with maximum cardinality.

Let S and $\overline{S} = (V(G) - S)$ be subsets of $V(G)$. The *edge cut of* S, denoted $K(S)$, is the set of all edges of G with one end in S and the other in $\overline{S}$. A set S is a *maximum cut* if $|K(S)|$ is maximum over all subsets of $V(G)$. Finding a maximum cut is equivalent to finding a maximum bipartite subgraph. We let $B(S)$ be the (bipartite) graph with vertex set V and edge set $K(S)$. The algorithm designed by Choukhmane and Franco[3] (given below for completeness) takes advantage of the relationship between the maximum cut problem and the maximum independent set problem and of the known polynomial time algorithm of Hadlock[8] for the maximum bipartite subgraph problem for planar graphs.

Algorithm A [Choukhmane and Franco]. *Given a cubic planar graph G, return an independent set $I \subseteq V(G)$ of G.*

1. Find a maximum cut S of G using the polynominal time algorithm of Hadlock [8]. Let $X \leftarrow S$ and $Y \leftarrow \overline{S}$.
2. For each edge $e = \{x, x'\}$ in $(E(G) - K(S))$ in turn, if both $x, x' \in X$ then $X \leftarrow X - \{x\}$.
3. For each edge $e = \{y, y'\}$ in $(E(G) - K(S))$ in turn, if both $y, y' \in Y$ then $Y \leftarrow Y - \{y\}$.
4. If $|X| \geq |Y|$ then $I \leftarrow X$ else $I \leftarrow Y$.

3. THE ANALYSIS

For any graph G and set $S \subseteq V(G)$, a vertex $v \in V$ is *unstable relative to* S if the number of edges incident with v and in $K(S)$ is less than the number of edges incident with v and in $E(G) - K(S)$. A set S is unstable if any vertex is unstable relative to S, otherwise S is *stable*. The following lemma relates the size of an independent subset of a stable set to the size of its edge cut set. This is the critical lemma for our result.

Lemma 1. *Let G be a cubic graph, $S \subseteq V(G)$ be a stable set, $n = |E(G)|$ and $k = |E(G) - K(S)|$, the number of edges not in the edge cut of S. Then*

$$(n-k)/2 \leq \max(\mu(S), \mu(\overline{S})) \leq \max((3n-2k)/6, (5n-6k)/8).$$

Proof: For the extent of this proof, let the degree of a vertex v be understood as the degree of v in $B(S)$. Let a, b, c, and d be the number of degree two and three vertices in S and degree two and three vertices in $\overline{S}$, respectively. Since S is stable,

$$n = a + b + c + d. \tag{1}$$

Since every edge in $B(S)$ connects S to $\overline{S}$,

$$2a + 3b = 2c + 3d. \tag{2}$$

Every degree two vertex has exactly one incident edge in $E(G) - K(S)$, so

$$2k = c + a. \tag{3}$$

A maximum independent subset of S or $\overline{S}$ will contain every degree three vertex and half of the degree two vertices in S or $\overline{S}$, respectively, so

$$\mu(S) = a/2 + b, \quad \text{and} \quad \mu(\overline{S}) = c/2 + d. \tag{4}$$

We assume, without loss of generality, that $a \leq c$. Let $\alpha \geq 0$ be such that $c = a + 3\alpha$. Then (3) implies that

$$a = k - 3\alpha/2, \quad \text{and} \quad c = k + 3\alpha/2. \tag{5}$$

By (3) we can substitute k for $a+c$ in (1) giving $n = 2k+b+d$. Since $c = a+3/\alpha$, (2) implies that $b = d + 2\alpha$. Then,

$$b = n/2 - k + \alpha, \quad \text{and} \quad d = n/2 - k - \alpha. \tag{6}$$

Therefore, by (4),

$$\mu(S) = (n-k)/2 + \alpha/4, \tag{7}$$

and $\mu(\overline{S}) = (n-k)/2 - \alpha/4$. Since $\alpha \geq 0$,

$$\mu(\overline{S}) \leq (n-k)/2 \leq \mu(S) = \max\big(\mu(S), \mu(\overline{S})\big).$$

Since $a \geq 0$, (5) implies that $\alpha \leq 2k/3$. Similarly, $d \geq 0$ and (6) imply that $\alpha \leq n/2 - k$. Then from (7) we have the final requirement of the lemma:

$$\max\big(\mu(S), \mu(\overline{S})\big) = \mu(S) \leq \max\big((3n-2k)/6, (5n-6k)/8\big)$$

□

The following lemma (in a somewhat different form) is proved by Malle[11] and follows from a result of Erdös[5].

Lemma 2 [Malle]. *If G is a graph with a maximum cut $S \subseteq V(G)$, then S is stable.*

We need two additional lemmas pertaining to sets which contain independent sets.

Lemma 3. *Let G be a cubic graph and $S \subseteq V(G)$. If S is unstable and contains a maximum independent set of G, then S contains a maximum independent set of G and an unstable vertex which is not in this set.*

Proof: Let $P \subseteq S$ be a maximum independent set of G, and let $v \in V(G)$ be an unstable vertex relative to S. If $v \notin P$, the lemma is satisfied by P and v, so assume $v \in P$. Let v' be a neighbor of v in S. If v' is unstable and $v' \notin P$, the lemma is satisfied by P and v'. Suppose instead that v' is stable. Then v can have at most one neighbor in S; in fact, it has exactly one neighbor, v. But then $P' = P - \{v\} \cup v' \subseteq S$ is a maximum independent set of G and the lemma is satisfied by P' and v. □

Lemma 4. *For every cubic graph G, there is a stable set $S \subseteq V(G)$ which contains a maximum independent set of G.*

Proof: Let $S \subseteq V(G)$ be a set containing a maximum independent set of G and such that $|K(S)|$ is maximized. Suppose that S is unstable. Then, by Lemma 3, there is a maximum independent set P of G contained in S and an unstable vertex, v, not in P. Then $S' = (S - \{v\}) \cup (\{v\} - S)$ contains P, a maximum independent set. But, since v is unstable relative to S, $K(S')$ has more edges than $K(S)$, contrary to the choice of S. Thus, S cannot have an unstable vertex. □

We note that the last two lemmas actually hold for any graph with maximum vertex degree three. The main result follows easily from the lemmas.

Theorem. *Let G be a cubic graph, $S \subseteq V(G)$ be a maximum cut of G, $n = |V(G)|$, and $k = |E(G) - K(S)|$. Then*

$$\frac{\max(\mu(S), \mu(\overline{S}))}{\mu(V(G))} \geq \max\left(\frac{3n-3k}{3n-2k}, \frac{4n-4k}{5n-6k}\right) \geq 7/8.$$

Proof: Without loss of generality, assume that $\mu(S) = \max(\mu(S), \mu(\overline{S}))$. By Lemma 2, S is stable. By Lemma 4, there is a set $S' \subseteq V(G)$ which is stable and such that $\mu(S') = \mu(V(G))$. Let $k' = |E(G) - K(S')|$. Since S is a maximum cut, $k' \geq k$. By Lemma 1,

$$\begin{aligned} \frac{\max(\mu(S), \mu(\overline{S}))}{\mu(V(G))} &= \mu(S)/\mu(S') \\ &\geq ((n-k)/2)/\max((3n-2k')/6, (5n-6k')/8) \\ &\geq \max\left(\frac{(3n-3k)}{(3n-2k)}, \frac{(4n-4k)}{(5n-6k)}\right). \end{aligned}$$

Since S is stable, every edge in $E(G) - K(S)$ is adjacent to a pair of distinct vertices in $V(G)$, which implies that $k \leq n/2$. Over the range $0 \leq k \leq n/2$, the first term of the maximum is strictly decreasing with increasing k, while the second term is strictly increasing, so the maximum is minimized when $(3n-3k)/(3n-2k) = (4n-4k)/(5n-6k)$, which implies $k = 3n/10$. By substitution,

$$\frac{\max(\mu(S), \mu(\overline{S}))}{\mu(V(G))} \geq \max\left(\frac{(3n-3k)}{(3n-2k)}, \frac{(4n-4k)}{(5n-6k)}\right) \geq 7/8,$$

as required. □

Let G_{20} be the planar projection of a dodecahedron. Choukhmane and Franco[3] have shown that there is a set $S \subseteq V(G_{20})$ such that $B(S)$ is a maximum cut of G_{20} and $\mu(S) = \mu(\overline{S}) = 7$, while $\mu(V(G_{20})) = 8$. Thus, the bound of the Theorem cannot be improved, in general.

Note that the Theorem applies to all cubic graphs, not just planar ones. The following corollary applies the Theorem to Choukhmane and Franco's algorithm.

Corollary. *Let G be any cubic planar graph, $n = |V(G)|$, and $k = |E(G) - E(B)|$, where B is any maximum bipartite subgraph of G. Then,*

$$M_A(G) \geq \max\left(\frac{3n-3k}{3n-2k}, \frac{4n-4k}{5n-6k}\right) \geq 7/8.$$

Note that while the algorithm is guaranteed to achieve a performance of at least 7/8ths of optimal, in many cases (when k differs from $3n/10$) the algorithm will give an answer which is known to be closer to optimal than 7/8. For example, if $k = n/10$ then Algorithm A will give an answer which is no worse than 27/28 of optimal. Since k is easily calculated during execution of the algorithm, it is possible to report not only the approximation to the maximum independent set, but also the maximum that it underestimates the true size of the actual maximum independent set.

ACKNOWLEDGMENT

I would like to thank John Franco for interesting me in this problem, for many valuable discussions, and for a careful reading of the manuscript.

REFERENCES

1. BAKER, B.S. Approximation algorithms for NP-complete problems on planar graphs. (Preliminary version). *Proceedings of the 24th Annual Symposium on Foundations of Computer Science.* Tuscon, Arizona (1983) 265–273.
2. CHIBA, N., NISHIZEKI, T., AND SAITO, N. An approximation algorithm for the maximum independent set problem on planar graphs. *SIAM J. Computing 11* (1982), pp. 663–675.
3. CHOUKHMANE, E., AND FRANCO, J. An approximation algorithm for the maximum independent set problem in cubic planar graphs. Accepted for publication in *Networks.*
4. EDMONDS, J. Paths, trees, and flowers. *Canadian J. Math. 17* (1965) 449–467.
5. ERDŐS, P. On some extremal problems in graph theory. *Isr. J. Math. 3* (1965) 113–116.
6. GAREY, M.R., AND JOHNSON, D.S. The rectilinear Steiner tree problem is NP-complete. *SIAM J. Appl. Math. 6* (1977) 826–834.
7. GAREY, M.R., AND JOHNSON, D.S. *Computers and Intractability: A Guide to the Theory of NP-completeness* (1979) W.H. Freeman, San Francisco.
8. HADLOCK, F.O. Finding a maximum cut of a planar graph in polynomial time. *SIAM J. Computing 4* (1975), pp. 221–225.
9. HOPKINS, G., AND STATON, W. Extremal bipartite subgraphs of cubic triangle-free graphs. *J. of Graph Theory 6* (1982) 115–121.
10. LIPTON, R.J., AND TARJAN, R.E. Applications of a planar separator theorem. *SIAM J. Computing 9* (1980) 615–627.
11. MALLE, G. On maximum bipartite subgraphs. *J. of Graph Theory 6* (1982) 105–113.
12. STATON, W. Edge deletion and the chromatic number. *Ars Combinatorica 10* (1980) 103–106.