Combinatorics, Probability and Computing (1993) 2, 000–000

Galley Proof

Clique Partitions of Chordal Graphs†

PAUL ERDŐS‡, EDWARD T. ORDMAN§ and YECHEZKEL ZALCSTEIN¶

‡ Mathematical Institute, Hungarian Academy of Sciences ,

§ Memphis State University, Memphis, TN 38152 U.S.A.

¶ Division of Computer and Computation Research, National Science Foundation, Washington, D.C. 20550, U.S.A.

Received 15 March 1993; revised 30 September 1993

Dedicated by the last two authors to Paul Erdős on his 80th birthday

To partition the edges of a chordal graph on n vertices into cliques may require as many as $n^2/6$ cliques; there is an example requiring this many, which is also a threshold graph and a split graph. It is unknown whether this many cliques will always suffice. We are able to show that $(1-c)n^2/4$ cliques will suffice for some $c > 0$.

1. Introduction

We consider undirected graphs without loops or multiple edges. The graph K_n on n vertices for which every pair of distinct vertices induces an edge is called a *complete* graph or a *clique* on n vertices. If G is any graph, we call any complete subgraph of G a *clique* of G (we do not require that it be a maximal complete subgraph). A *clique covering* of G is a set of cliques of G that together contain each edge of G at least once; if each edge is covered exactly once we call it a *clique partition*. The *clique covering number* $cc(G)$ and *clique partition number* $cp(G)$ are the smallest cardinalities of, respectively, a clique covering and a clique partition of G.

The question of calculating these numbers was raised by Orlin [13] in 1977. DeBruijn and Erdős [6] had already proved, in 1948, that partitioning K_n into smaller cliques required at least n cliques. Some more recent studies motivating the current paper include [11, 14, 2, 7, 9].

It is widely known that a graph on n vertices can always be covered or partitioned

† This work was done at Memphis State University.

§ Partially supported by U.S. National Science Foundation Grant DCR-8503922.

¶ Partially supported by U.S. National Science Foundation Grant DCR-8602319 at Memphis State University.

by no more than $n^2/4$ cliques; the complete bipartite graph actually requires this many. Turán's theorem states that if G has more than $n^2/4$ edges, it must contain a clique K_3; if it has more than $n^2(c-2)/(2c-2)$ edges it must contain a K_c. (For a more precise statement and proof, see *e.g.* [3, Chapter 11].)

A subgraph H of a graph G is an *induced* subgraph if for any pair of vertices a and b of H, ab is an edge of H if and only if it is an edge of G.

Two classes of graphs we shall refer to here are chordal graphs and threshold graphs. A graph is *chordal* (or often triangulated; [10, Chapter 4]) if every cycle of size greater than 3 has a chord (no set of more than 3 vertices induces a cycle). A graph G is *threshold* ([10, Chapter 10; 4; 5; 12]) if there exists a way of labelling each vertex A of G with a nonnegative integer $f(A)$ and there is another nonnegative integer t (the threshold) such that a set of vertices of G induces at least one edge if and only if the sum of their labels exceeds t.

A graph is *split* if its vertices can be partitioned into two sets A and B such that the vertices A form a clique and the vertices B induce no edges. (Two vertices, of which one is in A and one is in B, may or may not induce an edge.)

All threshold graphs are split and all split graphs are chordal. In a sense, most chordal graphs are split [1]. Induced subgraphs of chordal graphs are chordal; similar results hold for split graphs and threshold graphs.

2. Preliminary results on split graphs

A *complete matching* in a graph G is a set of edges such that each vertex of G lies on exactly one edge in the set. It is well known that the $t(2t-1)$ edges of K_{2t} can be edge-partitioned by a set of $2t-1$ matchings, each of t edges. By the *join* of two graphs G and H, we mean the graph made by taking the disjoint union of the two graphs and adding all edges of the form gh, where g is a vertex of G and h is a vertex of H.

By the graph $K_n - K_m$, for $n > m$, we mean a graph made by taking K_n and deleting all the edges induced by some particular m of the vertices. Equivalently, this is the join of K_{n-m} with the complement of K_m (a collection of m isolated vertices).

Lemma 2.1. *Let* $G = K_{4t} - K_{2t}$. *Then* $cp(G) \le t(2t+1)$.

Proof. Think of G as a complete graph $A = K_{2t}$ joined completely to an empty graph C on $2t$ vertices. Partition A into $2t-1$ disjoint matchings; join each matching to a different vertex in C, each matching yielding t triangles. The remaining vertex in C lies on $2t$ single edges to A. Thus we partition G by $t(2t-1)$ triangles and $2t$ single edges, a total of $2(2t+1)$ cliques. □

In fact, $cp(G) = t(2t+1)$. See, for example, [7].

Lemma 2.2. *In the graph G of the previous lemma, suppose r edges are deleted. Then this new graph has clique partition number not exceeding* $t(2t+1)+r$.

Proof. Start with the same partition as above. Each edge deletion at worst demolishes one triangle, requiring it to be replaced in the partition by two edges. □

3. Preliminary results on chordal graphs

We will rely heavily on the following lemma of Bender, Richmond, and Wormald, which gives a means of constructing an arbitrary chordal graph.

Lemma 3.1. *[1, Lemma 1.] For each chordal graph G and each clique R of G there is a sequence*

$$R = G_r, G_{r+1}, \ldots, G_n = G$$

of graphs such that G_{i+1} is obtained from G_i by adjoining a new vertex to one of its cliques.

Corollary 3.2. *If G is a chordal graph on n vertices with largest clique of size r, then G can be covered by at most $n - r + 1$ cliques.*

It is easy to see that the bound in the corollary cannot be improved; $K_n - K_{n-r+1}$ is an example requiring $n - r + 1$ cliques to cover.

Covering G may require less than $n - r + 1$ cliques. If G consists of two copies of K_t with a single vertex in each identified, G has $2t - 1$ vertices, the largest clique is of size t, this corollary produces a covering by $(2t - 1) - t + 1 = t$ cliques, but obviously there is a covering (and for that matter a partition) by two cliques.

We now utilize this construction with one additional specialization: we begin with a clique of maximum possible size in G. Supposing this clique to be of size r, each subsequently added vertex will add, at the time it is adjoined, at most $r - 1$ edges (or it would form a clique of more than r vertices).

Corollary 3.3. *A chordal graph on n vertices with a largest clique having r vertices has at most $(n-r)(r-1)$ edges outside that clique.*

Theorem 3.4. *Let G be a chordal graph on n vertices and $1/4 > d > 0$. Suppose G has at least dn^2 edges. Then G contains a clique with at least $(1 - \sqrt{1 - 2d})n > dn$ vertices.*

Proof. If the largest clique in G contains cn vertices, then that clique contains $cn(cn - 1)/2$ edges and each of the remaining $n - cn$ vertices of G can be added to G adding at most $cn - 1$ edges at each stage. Hence the total number of edges of G is at least dn^2 and at most $cn(cn - 1)/2 + (cn - 1)(n - cn)$, so $dn \leq (2c - c^2)(n/2) + (c - 2)/2$ and $dn < (2c - c^2)(n/2)$ since $c \leq 1$. Hence $d < (2c - c^2)/2$ and $c > 1 - \sqrt{1 - 2d} > d$ as needed. □

The result of this theorem turns out to be essentially best possible, not only for chordal graphs, but for split graphs and threshold graphs as well.

Example 3.5. Let $0 < c < 1$. Consider the graph $K_n - K_k$ where $k = n - cn + 1$, that is, the base clique has $cn - 1$ vertices and forms a clique on cn vertices with each other vertex. Clearly there are

$$(cn-1)(cn-2)/2 + (n - cn + 1)(cn - 1) = (c - c^2/2)n^2 - (1 - c/2)n$$

edges. So a graph can be threshold (hence split and chordal) and have almost $(c - c^2/2)n^2$ edges and no clique on more than cn vertices.

4. Clique partitions of chordal graphs

An arbitrary graph on n vertices may require $n^2/4$ cliques to cover or partition it [8]. We saw above that a chordal graph on n vertices may always be covered by fewer than n cliques. It may, however, still require a large number of cliques to partition it. The examples in [7] with high clique partition numbers are chordal graphs.

Example 4.1. [7] The graph $K_n - K_{2n/3}$ requires $n^2/6 + n/6$ cliques to partition it and $2n/3$ cliques to cover it. Thus for a chordal graph, both $cp(G)$ and $cp(G) - cc(G)$ can be approximately $n^2/6$.

We note that for a different example, the ratio of $cp(G)$ to $cc(G)$ may be larger.

Example 4.2. [7] The graph G_n composed of 3 cliques $K_{n/3}$, with all vertices of the first clique attached by edges to all vertices of the second and third, is a chordal graph (but not a split graph or threshold graph). As n increases, $cp(G_n)/cc(G_n)$ grows at least as fast as cn^2 for some $c > 0$.

We do not know if $cp(G)$ can significantly exceed $n^2/6$ for a chordal graph, or even for a split graph or a threshold graph.

Conjecture 1. *The clique partition number of a chordal graph, split graph, or threshold graph on n vertices cannot exceed $n^2/6$ (except by a term linear in n).*

It is even possible that $K_n - K_{2n/3}$ is literally the best example. (Some very minor adjustments to $n^2/6 + n/6$ may be needed because of round-off error). However, it is unclear how one would go about proving the following:

Conjecture 2. *No chordal, threshold, or split graph on n vertices requires more than $cp(K_n - K_{2n/3})$ cliques to partition it.*

For chordal graphs in general, we are very far from proving that $n^2/6$ cliques will suffice for a partition. In fact, we can improve only slightly on $n^2/4$.

Theorem 4.3. *There is a constant $c > 0$ such that if G is a chordal graph with n vertices, G may be partitioned into no more than $(1-c)n^2/4$ cliques.*

Proof. As the details are messy, we first give an outline; we follow this by some indication of more precise calculations, which the reader may choose to ignore, and a few numeric indications. As the result is clear for $n < 5$, we assume $n \geq 5$ in the proof. Let the largest clique in G have $(1+a)n/2$ vertices (a may be negative). Pick such a clique and call it A. Let C denote the subgraph of G induced by those vertices not in A; the set of edges not in A or C will be denoted B.

In case 1, The large clique is larger or smaller than half the vertices by a reasonable amount ($a^2 > c$). By Corollary 3.3, there are so few edges outside A that we can cover them by single edges. In case 2, A has close to half the vertices, and C has a significant number of edges. By Theorem 3.4, C contains a large clique C'; we can cover by A, C', and single edges. In case 3, A has close to half the vertices and C has few edges; in this case the graph must be very similar in form to $K_n - K_{n/2}$ and Lemma 2.2 can be used to construct a partition with 'little more than' $n^2/8$ triangles and edges.

We now give somewhat more precise calculations.

1 If $a^2 > c$, we can cover A with one clique and each edge not in A by a single edge. The number of edges outside A is at most

$$(1-a)(n/2)((1+a)n/2-1) < (1-a^2)n^2/4 < (1-c)n^2/4$$

as desired. Hereafter, we suppose $a^2 \leq c$.

2 If C has very many edges, we can cover A with a clique, the largest clique in C with a clique, and all other edges singly. Suppose C has dn^2 edges. Then, since C is an induced subgraph of G, it is a chordal graph with $v = (1-a)n/2$ vertices and $dn^2 = (dn^2/((1-a)n/2)^2)v^2$ edges; so by Theorem 3.4 it contains a clique with at least $(dn^2/((1-a)n/2)^2)v = 2dn/(1-a)$ vertices and $(2(dn)^2 - dn(1-a))/(1-a)^2$ edges. Covering this clique by itself, A by a clique, and each remaining edge with an edge, we get a number of cliques guaranteed to be less than

$$2 + (1-a^2)n^2/4 - (1-a)n/2 - (2(dn)^2 - dn(1-a))/(1-a)^2$$

$$= (1 - a^2 - 8d^2/(1-a)^2)n^2/4 + 2 - (1-a)n/2 + dn/(1-a)$$

Now supposing $c < .01$, $|a| < .1$, $n > 4$, and $d < .04$, we see that

$$2/n + d/(1-a) + a/2 < 1/2,$$

so

$$2 - (1-a)n/2 + dn/(1-a) < 0$$

and we need only have

$$1 - a^2 - 8d^2/(1-a)^2 < 1 - c$$

to finish, which is clearly true if $d^2 > (c - a^2)(1-a)^2/8$. If that condition is met, we are done. Hereafter, we assume that $d^2 \leq (c-a^2)(1-a)^2/8$, and hence that $d^2 < c(1+\sqrt{c})^2/8$. In particular, as c nears 0, so does d.

3 In the remaining case, we will cover the edges in C by single edges, and cover the edges in B and A by triangles and single edges using the technique of Lemma 2.2. Consider the number of edges in B. Since B and C together must have at least $(1-c)n^2/4$

edges and C has no more than dn^2 edges, we see that B has at least $(1-c-4d)n^2/4$ edges. In the 'complete' graph $H = K_n - K_{(1-a)n/2}$ there are $(1-a^2)n^2/4$ edges in B, so we see that if we can partition H by edges and triangles, we can partition G with only a few extra cliques: dn^2 for the edges in C and an allowance of at most $(1-a^2)(n^2/4) - (1-c-4d)(n^2/4) = (c-a^2+4d)n^2/4$ for the 'missing' edges of B. We now set out to clique-partition H. We neglect some constant multiples of n to reduce the bulk of the expressions below. As in Lemma 2.1, partition $A = K_{(1+a)n/2}$ into $(1+a)(n/2)-1$ matchings of $(1+a)(n/4)$ edges each (if $(1+a)n/2$ is odd, there is an extra linear factor in n neglected below). We must consider two subcases, $a \geq 0$ and $a < 0$.

If $a \geq 0$ we join $(1-a)n/2$ of these matchings to distinct points in C to form $(1-a^2)n^2/8$ triangles consuming all the connecting (B) edges of H; this leaves $(2a)(1+a)n^2/8$ edges of A unused and we cover them with single edges. Thus we partition H with $(1-a^2)(n^2/8) + a(1+a)(n^2/4)$ triangles and edges. This means we obtain a clique partition of G using no more cliques than

$$(1-a^2)(n^2/8) + a(1+a)(n^2/4) + dn^2 + (c-a^2+4d)n^2/4$$
$$= (n^2/4)[(1/2)(1-a^2) + a(1+a) + 4d + (c-a^2+4d)].$$

But it is easy to see that as c approaches 0 so that a and d also approach 0, this expression approaches $(n^2/4)[1/2+0+0+0]$, so it can clearly be made less than $(n^2/4)[1-c]$ as required.

If $a < 0$ we are able to join all the $(1+a)(n/2)-1$ matchings in A to distinct points in C. The resulting $(1+a)^2n^2/8$ triangles consume all (except a constant multiple of n) of the edges of A but only $(1+a)^2n^2/4$ edges of B, leaving as many as $(1-a^2)n^2/4 - (1+a)^2n^2/4$ to cover with single edges. Thus we partition H into

$$(1+a)^2(n^2/8) + (1-a^2)(n^2/4) - (1+a)^2(n^2/4)$$

cliques (which approaches $(1/2)n^2/4$ as c approaches 0), and the rest of the argument goes exactly as in the prior paragraph.

A somewhat more careful calculation suggests that letting $c = 1/400$ will easily suffice for $n \geq 5$, forcing $|a| < .05$ by case 1 and $d < .02$ by case 2. Unfortunately, linear terms neglected here, such as $(1+a)n/4$, complicate the actual calculation of c badly for low values of n. □

If we require G to be threshold, or split, the situation simplifies somewhat, since C will contain no edges and case (2) becomes unnecessary. Still, this method appears to produce only a marginal improvement in the c in these cases. The first two authors and Guan-Tao Chen have made some further progress in the case that G is a split graph, but are still not close to $n^2/6$; this will be pursued elsewhere.

References

[1] Bender, E. A., Richmond, L. B. and Wormald, N. C. (1985) Almost all chordal graphs split. *J. Austral. Math. Soc.* (A) **38** 214–221.

[2] Caccetta, L., Erdős, P., Ordman, E. and Pullman, N. (1985) The difference between the clique numbers of a graph. *Ars Combinatoria* **19** A 97–106.

[3] Chartrand, G. and Lesniak, L. (1986) *Graphs and Digraphs*, 2nd. Edition, Wadsworth, Belmont, CA.

[4] Chvátal, V. and Hammer, P. (1973) *Set packing and threshold graphs*, Univ. of Waterloo Research Report CORR 73-21.

[5] Chvátal, V. and Hammer, P. (1977) Aggregation of inequalities in integer programming. *Ann. Discrete Math* **1** 145–162.

[6] DeBruijn, N. G. and Erdős, P. (1948) On a combinatorial problem. *Indag. Math.* **10** 421–423.

[7] Erdős, P., Faudree, R. and Ordman, E. (1988) Clique coverings and clique partitions. *Discrete Mathematics* **72** 93–101.

[8] Erdős, P., Goodman, A. W. and Posa, L. (1966) The representation of a graph by set intersections. *Canad. J. Math.* **18** 106–112.

[9] Erdős, P., Gyárfás, A., Ordman, E. T. and Zalcstein, Y. (1989) The size of chordal, interval, and threshold subgraphs. *Combinatorica* **9** (3) 245–253.

[10] Golumbic, M. (1980) *Algorithmic Graph Theory and Perfect Graphs*, Academic Press, New York.

[11] Gregory, D. A. and Pullman, N. J. (1982) On a clique covering problem of Orlin. *Discrete Math.* **41** 97–99.

[12] Henderson, P. and Zalcstein, Y. (1977) A graph theoretic characterization of the PVchunk class of synchronizing primitives. *SIAM J. Comp.* **6** 88–108.

[13] Orlin, J. (1977) Contentment in Graph Theory: covering graphs with cliques. *Indag. Math.* **39** 406–424.

[14] Wallis, W. D. (1982) Asymptotic values of clique partition numbers. *Combinatorica* **2** (1) 99–101.