On the Covering of the Vertices of a Graph by Cliques*

Paul Erdös

(Magyar Tudományos Akadémia, Budapest)

In conversation I was told by Professor R. Brigham the following conjecture [1]. Let $G(n)$ be a graph of n vertices. Denote by $f(G(n))=t$ the smallest integer for which the vertices of $G(n)$ can be covered by t cliques. Denote further by $h(G(n))=l$ the largest integer for which there are l edges of our $G(n)$ no two of which are in the same clique. Clearly $h(G(n))$ can be much larger than $f(G(n))$ e.g. if $n=2m$ and $G(n)$ is the complete bipartite graph of m white and m black vertices. Then $l(G(n))=m$ and $l(G(n))=m^2$. It was conjectured that if $G(n)$ has no isolated vertices then

$$f(G(n))\leqslant h(G(n)) \tag{1}$$

holds for all graphs. R. Brigham showed me that (1) is true and easy if $h(G(n))\leqslant 2$;

A simple application of the probability method shows that (1) fails for almost all graphs. In fact we prove

Theorem 1. There are positive absolute constants c_1 and c_2 for which for $n>m_0(c_1,c_2)$

$$c_1\frac{n}{(\log n)^3}<\max_{G(n)}\frac{f(G(n))}{h(G(n))}<c_2\frac{n}{(\log n)^3}\ . \tag{2}$$

In fact we will show that the lower bound in (2), holds for almost all graphs $G(n)$, i.e. it holds for all but $\sigma(2)^{\binom{n}{2}})$ labelled graphs of n vertices. We do not give the details of the proof of the upper bound.

* Received June 1, 1981.

Observe that $f(G(n))$ is the chromatic number of $\bar{G}(n)$, the complement of $G(n)$.

I proved that for almost all graphs $G(n)$ the chromatic number of $G(n)$ is between $c_1 n/\log n$ and $c_2 n/\log n$. Thus for almost all $G(n)$ [2]

$$(3)\qquad c_1 \frac{n}{\log n} < f(G(n)) < c_2 \frac{n}{\log n}.$$

It seems certain that for almost all $G(n)$

$$(4)\qquad f(G(n)) = (c + o(1)) \frac{n}{\log n}$$

for a certain absolute constant c, but I have never been able to prove (4).

Next we prove that for almost all graphs $G(n)$

$$(5)\qquad c_3(\log n)^2 < h(G(n)) < c_4(\log n)^2.$$

Let $G(n)$ be a random graph of n vertices and let $e_1,\dots,e_m$ be the largest family of pairwise independent edges of $G(n)$ (i.e. no two e's are in a clique). First observe that we can assume that for every vertex x of $G(n)$ the number of e's incident to x is less than $c_3 \log n$. This remark follows immediately from the fact that the other endpoints of the e's incident to x must form an independent set in $G(n)$ (for if not then two e's incident to x are contained in a triangle which is impossible). Now it is well known and easy to see that the largest independent set in the random graph $G(n)$ is less than $c \log n$ [3].

Next observe that if $e_1 \dots, e_t$ is a set of edges without a common vertex no two of which are contained in a clique (which here is of size 4), then for almost all $G(n)$

$$(6)\qquad t < c_0 \log n.$$

To prove (6) observe that the probability that two edges e_1 and e_2 (not having a common vertex) are not contained in a clique is $1-\frac{1}{16}$ (since all four edges joining the endpoints of e_1 and e_2 must be in $G(n)$ if e_1 and e_2 are contained in a clique). The $\binom{t}{2}$ events e_i and e_j $(1 \leqslant i < j \leqslant t)$ are not contained in a clique are clearly independent and thus the probability that no two of the edges $e_i, e_j (1 \leqslant i < j \leqslant t)$ are in a clique is $(1-\frac{1}{16})^{\binom{t}{2}}$. For $e_1,\dots,e_t$ there are $\binom{\binom{n}{2}}{t} < n^{2t}$ choices. Thus the probability that our $G(n)$ has t independent edges is less than $n^{2t}(1-\frac{1}{16})^{\binom{t}{2}}$ which tends to 0 if $t>c_6\log n$, which proves (6). Now (6) and the fact that for almost all $G(n)$ each vertex is incident to fewer than $c_5\log n$ of the e's gives that for almost all $G(n)$ $h(G(n))<2c_3c_6\log n$, which proves the upper bound of (5). The upper bound of (5) and the lower bound of (3) give the lower bound of (2).

Now we prove the lower bound of (5) (we will not need it for the proof of Theorem 1). Observe that almost all graphs $G(n)$ contain a set of independent vertices of size $t>c\log n$ i.e. there are vertices $x_1,\ldots,x_t$ no two of which are joined by an edge. It is well known and easy to see [3] that for almost all $G(n)$ all the vertices have valency (or degree) $(1+o(1))\frac{n}{2}$. A simple computation now shows that there is a constant c' so that fot every vertex x there are $c'\log n$ vertices which are all joined to x and which are independent in $G(n)$. Thus we obtain $cc'(\log n)^2$ edges no two of which are on the same clique. This completes the proof of (5). It would be easy to insist that these independent edges should be vertex disjoint except for $x_1,\ldots,x_t$.

My proof of the upper bound of (2) is surprisingly complicated. By repeated application of known inequalities for Ramsay numbers [4] I can prove that $f(G(n))/h(G(n))$ can be of the order of magnitude $\frac{n}{(\text{leg}\, n)^3}$ only if $f(G(n))$ is of the order $n/(\log n)$ and $h(G(n))$ of the order $(\log n)^2$. I suppress the details because perhaps a much simpler proof can be found. If nobody finds a simpler proof I will publish my complicated proof.

It would be of interest to prove that there is a c for which

$$(7)\qquad \lim_{n\to\infty}\frac{1}{n}\max_{G(n)}\frac{f(G(n))}{h(G(n))}=c.$$

I expect that the proof of (7) will be difficult.

It would be interesting to know the largest t for which $h(G(n))\leqslant t$ implies $f(G(n))\leqslant h(G(n))$.

I can prove that there is a t_0 so that for every $t>t_0$ there is a $G(n)$ satisfying

$$(8)\qquad h(G(n))\geqslant t \text{ and } f(G(n))>n^{c_t}.$$

It would be interesting to determine t_0 and the best possible value of c_t. I do not expect this to be easy. Finally it is not hard to prove that

$$\max_{G(n)} \frac{h(G(n))}{f(G(n))} = \left[\frac{n^2}{4}\right]\left[\frac{n+1}{2}\right]^{-1},$$

in other words the example mentioned in the introduction is optimal.

References

1. Parthasarathy, K. R. and Choudum, S. A., The edges clique cover number of products of graphs, *Jour. Math. Phys. Sci.* 10 (1976), 3 255—261.
2. Erdös, P., Some remarks on chromatic graphs, *Colloquium Math.* 16 (1967), 253—256. See also Erdös, P., The Art of Counting, *Selected Writings MIT Press* 1973, 103—106.
3. Erdös, P. and Renyi, A., On the evolution of random graphs, *Publ. Math. Inst. Hung. Acad. Sci.* 5 (1960), 17—61. See also The Art of Counting, 569—617.
4. Erdös, P. and Szekeres, G., A combinatorial problem in geometry, *Composition Math.* 2(1935), 463—470. See also The Art of Counting, 5—12.

本刊编辑部启事

一、为加强对外交流，本刊对刊出的每篇论文希望能有较详粗细的英文摘要，但每篇摘要以不超过印刷面一页为限，这种摘要必须用印刷体书写，最好是间行打字．又为了缩短审稿时间，要求来稿一式二份．

二、本刊为中英文合刊，每期将刊登少量英文稿件．外文稿也须用印刷体字书写或间行打字。

三、本刊出版以来受到各方面的支持．现今已出现稿件拥挤现象．为了能多刊登一些优秀文章和短篇创作，对“专题论文”稿要求以不超过6000字为限(包括摘要在内)．对思想新颖、成果突出的短篇创作尤为欢迎，将尽可能优先考虑及早发表．

四、本刊鼓励学术上不同观点不同见解的讨论和争鸣，特别欢迎数学领城内的实事求是的学术讨论与评述性文章．凡有不同意见，都可以保留，以待实践检验．

五、对于半年至一年内未能收到刊登消息的稿件，作者可以来函询问稿件处理情况，或声明将稿件改投别处．