JOURNAL OF COMBINATORIAL THEORY 5, 164–169 (1968)

On Coloring Graphs to Maximize the Proportion of Multicolored k-Edges

PAUL ERDÖS AND DANIEL J. KLEITMAN*

Németvolgit 72^c, Budapest XII, Hungary
and Mathematics Dept. MIT, Cambridge, Massachusetts 02139

Communicated by Gian-Carlo Rota

ABSTRACT

The following results and some generalizations are obtained. Consider all colorings of the n vertices of a k-graph G into l colors. Then, if k is sufficiently large ($k \geqslant k_0(r, l)$), at least a proportion r of the k-edges of G will contain vertices colored in every color for any $r < 1$.

It is possible to color the points of any graph G with two colors so that less than half of the edges in G have endpoints of the same color. Further results of this kind have been obtained by one of the authors [1].

In this note we consider an analogous problem for k-graphs. Let G_k be a k-graph (a collection of k-element sets of points which we will call k-edges) and suppose we color the points of G_k in l colors. We seek the maximum over all colorings of G_k of the proportion of k-edges in G_k which contain at least one point of every color.

Let $p(G_k, l)$ be the maximum just described and let $m(n, k, l)$ and $m(k, l)$ be, respectively, the minimum value of $p(G_k, l)$ over all k-graphs G_k on n points, and the minimum over all finite k-graphs.

Below, we evaluate $m(n, k, l)$ by showing that the graph which minimizes $p(G_k, l)$ for each n is the complete k-graph on n points, $S_{k,n}$; i.e., the k-graph consisting of all k-edges. We also provide a simple direct evaluation of $m(k, l)$.

Our results imply, for example, that

$$\lim_{k\to\infty} m(k, l) = 1$$

for all l, so that for sufficiently large k there exist colorings of any k-graph which make most of its k-edges contain all colors.

* This research was supported in part by NSF Contract GP6165.

Our results can be divided into the following three parts:

THEOREM 1. *For any finite graph*

$$m(n, k, l) > \frac{S_2(k, l)}{l^k}\, l!$$

where $S_2(k, l)$ is a Stirling number of the second kind, i.e., the number of partitions of k elements into exactly l indistinguishable parts.

THEOREM 2. *For any (n, k, l), the minimum value of $p(G_k, l)$ is achieved if G_k is the complete k-graph on n points.*

COROLLARY. *For any (n, k, l) such that l divides n, we have*

$$m(n, k, l) = \sum_{s=0}^{l} (-1)^s \binom{l}{s} \binom{n \frac{(l-s)}{l}}{k} \Big/ \binom{n}{k}$$

and

$$m(k, l) = \frac{S(k, l)\, l!}{l^k}.$$

PROOF OF THEOREM 1: There are l^n distinct colorings of a graph on n points into l colors. Of these, $l^{n-k}R$ colorings will color a given k-tuple in exactly l colors, where R is the number of ways of coloring k points in exactly l colors. Moreover, R is clearly $l!$ times the number of partitions of k into exactly l non-empty parts, which latter number is $S_2(k, l)$. The proportion of all colorings which will give rise to all l-colors in any k-tuple is then

$$\frac{l!}{l^k} S_2(k, l).$$

Let G_k have $q(G_k)$ k-edges, and let $r(C, G_k)$ be the number of edges of G_k colored in all l colors under the coloring C. Let $\theta(C, E)$ be 1 if the edge E is colored in all colors by C and 0 otherwise.

We then have

$$r(C, G_k) = \sum_{E \in G_k} \theta(C, E).$$

By the remarks above, for any edge E, the average value of $\theta(C, E)$ over all colorings is given by

$$\frac{l!}{l^k} S_2(k, l).$$

Averaging the expression above for r over all colorings into l colors then yields

$$\langle r(C, G_k)\rangle_C = \sum_{E \in G_k} \langle \theta(C, E)\rangle$$

$$= q(G_k)\frac{l!}{l^k} S_2(k, l).$$

This equation means that, for any graph G_k, the average proportion of edges colored in all colors, over all possible l-colorings, is

$$\frac{l!}{l^k} S_2(k, l).$$

Since $p(G_k, l)$ is the maximum of $r(C, G_k)/q(G_k)$ over all l-colorings we have

$$p(G_k, l) > \frac{l!}{l^k} S_2(k, l)$$

for any finite graph G_k, and we can conclude that

$$m(k, l) \geqslant \frac{l!}{l^k} S_2(k, l).$$

PROOF OF THEOREM 2: For a coloring C of a graph G_k on n points and an element g of the symmetric group S_n on the n points of G_k, let Cg be the coloring obtained by performing the permutation g before the coloring C. For any k-edge E, let Eg be the k-edge whose elements are the images of the elements of E under g, and let $G_k g$ be the k-graph whose members are of the form Eg for each E in G_k.

For any coloring and any edge we have

$$\theta(C, E) = \theta(Cg, Eg)$$

hence

$$r(C, G_k) = r(Cg, G_k g).$$

Using these facts, we find, upon averaging $r(Cg, G_k)$ over all g in S_n, that

$$\langle r(Cg, G_k)\rangle_{g \in S_n} = \langle r(C, G_k g^{-1})\rangle_{g \in S_n}$$

$$= \frac{1}{n!}\sum_E \theta(C, E)\left\{n!\, q(G_k)\Big/\binom{n}{k}\right\},$$

since each edge E lies in $G_k g^{-1}$ for exactly $n! q(G_k)/\binom{n}{k}$ elements g of S_n.

This latter expression tells us that

$$\langle r(Cg, G_k)\rangle_{g\in S_n}/q(G_k)$$

is independent of G_k and in fact is given by $r(C, S_{k,n})/\binom{n}{k}$ since, for $G_k = S_{k,n}$, the averaging is trivial.

Now, if $p(G_k, l) \leqslant p(S_{k,n}, l)$, there must be a coloring C such that

$$r(C, G_k)/q(G_k) \leqslant r(C, S_{k,n})\Big/\binom{n}{k}.$$

But by the result immediately above there must then be some $g \in S_n$ such that

$$r(Cg, G_k)/q(G_k) \geqslant r(C, S_{k,n})\Big/\binom{n}{k};$$

thus, for all G_k, $p(G_k, l) \geqslant p(S_{k,n}, l)$.

Proof of Corollary: For the complete graph one can easily verify that any coloring C which assigns exactly $[n/l]$ or $[n/l] + 1$ points to each color will satisfy $r(C, S_{k,n}) = m(n, k, l)$. The value of $m(n, k, l)$ can immediately be deduced from this fact; by the principle of inclusion-exclusion, we obtain

$$m(n, k, l) = \sum_{s=0}^{l} (-1)^s \binom{l}{s}\binom{n(l-s)/l}{k}\Big/\binom{n}{k}.$$

For large values of n this expression is asymptotic to and always greater than

$$\sum_{s=0}^{l} (-1)^s \binom{l}{s}(l-s)^k/l^k.$$

An elementary identity for the Stirling numbers,

$$S_2(k, l) = (-1)^l \sum_{t=0}^{l} \frac{(-1)^t\, t^k}{t!(k-t)!},$$

implies that the upper limit for $m(k, l)$ obtained here is the same as the bound obtained in Theorem 1.

Our Theorem 2 above is capable of wide generalization. In fact, the method of proof used above applies to any situation in which we seek to minimize over a class of graphs the maximum over any class of colorings that is symmetric under point permutation, the proportion of the edges possessing any property which is invariant under point permutation. This generalization can be expressed as the following theorem.

THEOREM 3. *Let G_k be a k-graph and C_l a class of colorings of the points or, for $m \leqslant k$, of the m edges of G_k. Let π be a property of colored k-edges. (We say an edge E has property π for coloring C if $\theta(E, C) = 1$.) Let $p(G_k, C_l)$ be the maximum over all colorings in C_l of the proportion of k-edges in G_k which have property π. Let $m(n, k, C_l)$ be the minimum value of $p(G_k, C_l)$ over all k-graphs on n points. Then, if the class C_l is symmetric under point interchange, and the property π is invariant under simultaneous permutation of coloring and edge (so that S_n $g \in S_n$ $\theta(E, C) = \theta(Eg, Cg)$),*

$$m(n, k, C_l) = p(S_{k,n}, C_l)$$

where $S_{k,n}$ is the complete k-graph on n points.

The proof of this theorem is the same as that of Theorem 2.

We can apply the generalization to obtain answers to the following modifications of our original problem.

(a) Let

$$\begin{Bmatrix} M(n, k, l, r) \\ m(n, k, l, r) \end{Bmatrix}$$

be the minimax, that is, the minimum over k-graph of the maximum over all colorings of points in l colors, of the proportion of edges colored in

$$\begin{Bmatrix} \text{at least} \\ \text{exactly} \end{Bmatrix} r \text{ colors.}$$

(b) Let $m(n, k, l, s_1, ..., s_l)$ be the minimax over graph colorings of the proportion of edges such that s_j points per edge are colored in color j.

(c) Let $m(n, k, l, s_1, ..., s_l, t_1, ..., t_l)$ be the minimax over graph colorings in n points for which t_j edges are colored in color j, of the proportion of edges such that s_j points per edge are colored in color j.

The possibilities of further variation in this problem are obviously great. In each case above we can conclude that the minimum-maximum occurs for the complete graph $S_{k,n}$, for which to answer to each of these questions becomes a routine computation.

The conclusion that $\lim_{k\to\infty} m(k, l) = 1$ has been used by C. Jockusch [2] in proving the following result in recursion theory, nomenclature for which will not be defined here.

The sets with the property that some set of lesser degree of unsolvability is strongly hyper-hyperimmune have zero measure. This result implies that the measure of the family of sets, each one of which is Turing equivalent to some member of a given downward closed family C, is the same as the measure of the family whose members each have some member of C recursive in it. A downward closed family here means a family of infinite sets such that every infinite subset of every member is also a member.

References

1. P. Erdös, On The Structure of Linear Graphs, *Israel Journal of Math* 1 (1963), 56–60.
6. C. Jockusch, (to be published).