ON A GENERALIZATION OF RAMSEY NUMBERS

Paul ERDÖS

M.I.T. Department of Mathematics, Cambridge, Mass., USA

Patrik E. O'NEIL*

M.I.T. Department of Electrical Engineering, Cambridge, Mass., USA

Received 20 December 1971

Abstract. Given the integers l_1, k_1, l_2, k_2, r, which satisfy the condition $l_1, l_2 \geq r \geq k_1, k_2 > 0$, we define $m = N(l_1, k_1; l_2, k_2; r)$ as the smallest integer with the following property: if S is a set containing m points and the r-subsets of S are partitioned arbitrarily into two classes, then for $i = 1$ or 2 there exists an l_i subset of S each of whose k_i-subsets lies in some r-subset of the i^{th} class. The integers defined in this way form a collection of which the usual Ramsey numbers are a special case: i.e., the Ramsey number $N(l_1, l_2; r)$ is represented as $N(l_1, r; l_2, r; r)$. We derive two major results concerning the values of these generalized Ramsey numbers. If $k_1 + k_2 = r + 1$ then $N(l_1, k_1; l_2, k_2; r) = l_1 + l_2 - k_1 - k_2 + 1$, corresponding to the "pigeonhole principle". For $k_1 + k_2 \leq r$, we show that $N(l_1, k_1; l_2, k_2; r) = \max(l_1, l_2)$. The next interesting case occurs for $k_1 + k_2 = r + 2$, where we show that there are constants c_1 and c_2 such that for sufficiently large l, $2^{c_1 l} < N(l, k_1; l, k_2; r) < 2^{c_2 l}$.

Given integers $l_i, k_i, i = 1, ..., n$, and r, which satisfy the properties $l_i \geq r \geq k_i > 0$, for $i = 1, ..., n$, we may define an integer $N(l_1, k_1; l_2, k_2; ...; l_n, k_n; r) = m$ as the smallest integer with the following property: If S is a set containing m points and the r-subsets of S are partitioned arbitrarily into n classes, then for some i, $1 \leq i \leq n$, there exists an l_i-subset of S each of whose k_i-subsets lies in some r-subset of the i^{th} class. The fact that such an integer exists follows immediately from the existence of the Ramsey number $N(l_1, l_2, ..., l_n; r)$, for if the set S contains this many points, there is some i, $1 \leq i \leq n$, such that all the r-subsets of some l_i-set are of the i^{th} class [3]. Then certainly each k_i subset of this l_i subset lies in such an r-set, since $k_i \leq r$. In what follows we shall be con-

* This work was supported by the Office of Naval Research Grants N00014-67-A-0204-0034 and N00014-70-A-0362-0002.

cerned mainly with the case where there are only two classes of r-subsets ($n = 2$). The proof of the following remarks are entirely analogous to those found in [3] and will be omitted.

Remark 1. $N(r, k_1; l, k_2; r) = N(l, k_1; r, k_2; r) = l$.

Remark 2. $N(l_1, k_1; l_2, k_2; r) \le N(N(l_1-1, k_1; l_2, k_2; r), k_1-1; N(l_1, k_1; l_2-1, k_2; r), k_2-1; r-1) + 1$.

The following remark has no counterpart in any theorem on Ramsey numbers, but is elementary.

Remark 3. If $k_1' \le k_1$ and $k_2' \le k_2$, then $N(l_1, k_1'; l_2, k_2'; r) \le N(l_1, k_1; l_2, k_2; r)$.

Proof. Let $m < N(l_1, k_1; l_2, k_2; r)$. Then there exists a partition of the r-subsets of the m-set S into two classes such that every l_i subset contains a k_i subset all of whose containing r-subsets are class j, $j \neq i$, $i, j = 1, 2$. Since $k_1' \le k_1$, $k_2' \le k_2$, the above property is inherited and $m < N(l_1, k_1'; l_2, k_2'; r)$. Thus $N(l_1, k_1'; l_2, k_2'; r) \le N(l_1, k_1; l_2, k_2; r)$.

We will show:

Theorem 1. *If* $k_1 + k_2 = r + 1$, *then* $N(l_1, k_1; l_2, k_2; r) = l_1 + l_2 - k_1 - k_2 + 1$. *Further, if* $k_1 + k_2 \le r$, *then* $N(l_1, k_1; l_2, k_2; r) = \max(l_1, l_2)$.

Proof. Let us first dispose of the simpler case where $k_1 + k_2 \le r$. We may assume $l_1 \le l_2$. Clearly, $N(l_1, k_1; l_2, k_2; r) \ge l_2$; merely consider the set S containing $l_2 - 1$ points all of whose r-sets lie in class 2; S has no l_2 subset at all and no l_1 subsets with k_1-subsets contained in class 1 r-sets. Now assume S contains l_2 points. If every k_2 subset lies in an r-set of class 2, we are finished; therefore assume there is a k_2-subset S_1 all of whose containing r-sets are class 1. But now all k_1-subsets $S_2 \subseteq S$ lie in an r-subset of class 1, since $|S_1 \cup S_2| \le r$.

Now let $k_1 + k_2 = r + 1$. Assume $S = S_1 \cup S_2$, S_1 and S_2 disjoint, with $|S_1| = l_1 - k_1$ and $|S_2| = l_2 - k_2$. We construct a partition of the r-set of S as follows: place r-sets in class 1 which intersect S_1 in $\ge k_2$ points and all other r-sets in class 2, i.e., those that intersect S_2 in at least k_1 points, since $k_1 + k_2 = r + 1$. Note that any l_1-set must contain at least k_1 points in S_2 and those k_1 points are contained only in class 2 r-sets. The situation is entirely symmetric for l_2-sets. Thus

$$(1) \qquad N(l_1, k_1; l_2, k_2; r) > |S| = l_1 - k_1 + l_2 - k_2 .$$

Now assume that we have a set S whose r-sets are partitioned into two classes such that there is no l_i-set each of whose k_i-sets lies in an r-set of class i, $i = 1, 2$. We may assume that S is of maximal cardinality with the property

$$(2) \qquad |S| = N(l_1, k_1; l_2, k_2; r) - 1 .$$

Let T_1 be a maximal subset of S such that each of its k_1-subsets is contained in an r-set of class 1. Then $|T_1| \leq l_1 - 1$. Let $T_2 = S \backslash T_1$, and choose any point p in T_2. If all r-subsets containing p and intersecting T_1 in $r-1$ points were in class 1, then T_1 would not be maximal since the point p could be adjoined. Therefore there is an r-set U of class 2 which intersects T_1 in $r-1$ points. We now show that $T_2 \cup U$ has the property that each of its k_2-subsets is contained in an r-set of class. Obviously any k_2-set lying in $U \setminus T_2$ is contained in the set U which is of class 2. Now take any k_2-set V in $T_2 \cup U$ such that $V \cap T_1 \subset U$ and $k_2 \cap T_2 = W \neq \emptyset$ (this is the only remaining case). We assume that V lies only in r-subsets of class 1 and arrive at a contradiction. For take any k_1 subset V' lying in $T_1 \cup W$. If V' lies totally in T_1, it is contained in an r-set of class 1; but if $V' \cap W \neq \emptyset$, then $|V' \cup V| \leq r$ and since all r-subsets containing V are of class 1, V' is contained in an r-set of class 1. Therefore any k_1-subset of $T_1 \cup W$ lies in an r-set of class 1 which contradicts the maximality of T_1. Thus the arbitrarily chosen k_2-set V in $T_2 \cup U$ must lie in some r-subset of class 2. But $|T_2 \cup U| = |T_2| + r - 1$, and we must have $|T_2| + r - 1 \leq l_2 - 1$ by the definition of S. Since $|T_1| \leq l_1 - 1$, we have

$$|S| = |T_1| + |T_2| \leq l_1 + l_2 - r + 1 = l_1 + l_2 - k_1 - k_2 .$$

or by equation (2)

$$N(l_1, k_1; l_2, k_2; r) \leq l_1 + l_2 - k_1 - k_2 + 1 .$$

Combining this result with inequality (1), we see that

$$N(l_1, k_1; l_2, k_2; r) = l_1 + l_2 - k_1 - k_2 + 1 , \tag{3}$$

which proves the theorem.

Theorem 1 is a generalization of the pigeonhole principle, the simplest Ramsey result which states that the Ramsey number $N(l_1, l_2; 1)$ (equivalent in our notation to $N(l_1, 1; l_2, 1; 1)$) is given by $l_1 + l_2 - 1$.

We now consider the numbers $N(l_1, k_1; l_2, k_2; r)$, with the condition that $k_1 + k_2 = r + 2$. These are analogous to the Ramsey numbers $N(l_1, l_2; 2)$ ($N(l_1, 2; l_2, 2; 2)$, in our notation). To get an exact formula for these numbers would be too much to expect since this has not been possible with the usual Ramsey numbers even in very restricted cases. The numbers are so highly variable if both l_1 and l_2 are allowed to range that we shall restrict ourselves to studying the asymptotic behavior of $N(l, k_1; l, k_2; r)$. We shall prove:

Theorem 2. *If $k_1 + k_2 = r + 2$, then there exist constants c_1 and c_2 such that for sufficiently large l, $2^{c_1 l} < N(l, k_1; l, k_2; r) < 2^{c_2 l}$.*

Proof. We first show that $N(l, k_1; l, k_2; r) < 2^{c_2 l}$. Let S be a set containing $N(l, l; 2)$ points. This assures that if the edges defined by pairs of points in S are partitioned into two classes, there will be an l-gon (a complete l-graph) all of whose edges are in one class.

Now partition the r-tuples of S into class 1 and class 2 in any manner. We say that a k_1-set is of class 1 if all r-tuples containing it are of class 1, and a k_2-set is of class 2 if all r-tuples containing it are of class 2. Note that a k_1-set of class 1 and a k_2-set of class 2 intersect in at most one point, for if they intersected at two points, then the union of the k_1-set and k_2-set would be an r-set which would have to be in class 1 and class 2 simultaneously.

We define two edge disjoint graphs as follows: an edge is in G_1 if it is contained in a k_1-set of class 1 and is in G_2 if it is contained in a k_2-set of class 2. Since $|S| = N(l, l; 2)$, there is an l-set where either G_1 or

G_2 has no edge, say G_1. But then every k_2-subset of this l-set is contained in an r-set of class 2. Thus, $N(l, k_1; l, k_2; r) \leq N(l, l; 2)$. Now it has been shown in [2] that $N(l, l; 2) \leq \binom{2l-2}{l-1} < 2^{2l-2}$, and for $c_2 = 2$,

$$N(l, k_1; l, k_2; r) < 2^{c_2 l} . \tag{4}$$

We will now show that $N(l, k_1; l, k_2; r) > 2^{c_1 l}$, for some constant c_1, with l sufficiently large. To do this we shall need the following result:

Lemma 1. *Let $F_k(2, l)$ be the largest integer for which there is a graph G on $F_k(2, l)$ vertices so that every set of l vertices in it contains a complete k-gon and a set of k independent points (no two joined by an edge). There is a constant c_k depending only on k such that for l sufficiently large, $F_k(2, l) \geq 2^{c_k l}$.*

Consider a set of $S = F_k(2, l)$ points and let $k = \max(k_1, k_2) \leq r$, $k_1 + k_2 = r + 2$. We partition the r-sets of S as follows: place an r-set in class 2 if it contains a k_1-gon in the graph G and in class 1 if it contains an independent k_2-set. Note that an r-set cannot contain both a k_1-gon and an independent k_2-set as they would intersect at two points; thus the partition is well defined if we add that r-sets not containing either a k_1-gon or an independent k_2-set are placed arbitrarily in either class.

Now with $|S| = F_k(2, l)$ points we have constructed a partition of the r-tuples such that every l-set contains a k_1-set all of whose containing r-sets are class 2 and a k_2-set all of whose containing r-sets are class 1. Thus

$$N(l, k_1; l, k_2; r) > F_k(2, l), \quad k = \max(k_1, k_2) . \tag{5}$$

This shows that the definition of $F_k(2, l)$ as the largest integer with the given property is proper, since we know that $N(l, k_1; l, k_2; r)$ is bounded above. Furthermore, given the result of Lemma 1 and eq. (5), we will know that there is an integer c_1 depending only on $\max(k_1, k_2)$, such that, for l sufficiently large,

$$N(l, k_1; l, k_2; r) > 2^{c_1 l} . \tag{6}$$

This taken with eq. (4) will give Theorem 2. It is only necessary to prove Lemma 1 therefore.

Proof of Lemma 1. By a theorem of Erdös and Hanani [1], for fixed k and l large, a set of l elements contains $(1+\varphi(1))\, l^2/k^2 = L(k,l)$ k-subsets, every two of which have at most one element in common, asymptotically in l. We shall disregard the $\varphi(1)$ term, since we shall see that it only affects the value of the constant c_k.

Let $m \le 2^{c_k l}$. (We shall indicate the value of c_k later). There are $2^{\binom{m}{2}}$ graphs on m labelled vertices. We first estimate the number of graphs G on m points for which a given l-subset of points does not contain both a complete k-gon and k independent points. Consider our $L(k,l)$ k-sets. Let us say that we do not permit k-gons in this l-subset of G. Then there are $2^{\binom{k}{2}}-1$ ways in which the edges of the graph G may be placed in each of the $L(k,l)$ k-sets, and since the $L(k,l)$ k-sets are edge disjoint, the colorings are independent. The number of graphs on l points which do not contain k-gons is therefore at most

$$(7) \qquad 2^{\binom{l}{2}}\,(1-1/2^{\binom{k}{2}})^{L(k,l)}\,.$$

Since we could just as well have permitted k-gons and forbidden independent k-sets, the number of graphs on l points becomes at most twice the number. All the remaining edges among the m points, $\binom{m}{2}-\binom{l}{2}$ in number may be included or not included in the graph G arbitrarily and it will remain a graph for which a given l-subset of points does not contain both a complete k-gon and k independent points. The number of such graphs G is

$$2\cdot 2^{\binom{m}{2}}\,(1-1/2^{\binom{k}{2}})^{L(k,l)}\,.$$

Since there are $\binom{m}{l}$ l-subsets, the total number of graphs with some l-subset which does not contain a complete k-gon or k independent points is not greater than

$$(8) \qquad \binom{m}{l}\,2^{\binom{m}{2}+1}\,(1-1/2^{\binom{k}{2}})^{L(k,l)} \le 2\cdot m^l(1-1/2^{\binom{k}{2}})^{L(k,l)}\,2^{\binom{m}{2}}\,,$$

which we may prove is less than $2^{\binom{m}{2}}$, for l sufficiently large. We need merely show that

$$2 \cdot m^l \left(1-1/2^{\binom{k}{2}}\right)^{l^2/k^2} < 1 .$$

Cancelling, we get

$$m(1-c_3)^{c_4 l} < 1 ,$$

where c_3 and c_4 depend only on k, and for a proper choice of c_k, $m \leq 2^{c_k l}$ guarantees this. But this means that among the $2^{\binom{m}{2}}$ graphs on $m \leq 2^{c_k l}$ points there are some all of whose l-subsets contain both a complete k-gon and k independent points. Since $F_k(2,l)$ is the largest cardinality for such graphs

$$F_k(2,l) \geq 2^{c_k l} ,$$

and the lemma is proved.

As a final remark, we note that using essentially the same technique as above, we may show that if $k_1 + k_2 = r + 3$, then for l sufficiently large

$$N(l, k_1; l, k_2; r) > 2^{c_1 l^2} ,$$

where c_1 depends only on max (k_1, k_2). This bound is probably very poor, however. By somewhat more complicated methods, we can prove that

$$N(l, k_1; l, k_2; r) < 2^{c_r l} ,$$

for $r > r(\epsilon)$, $c_r < \epsilon$ if $k_1 + k_2 = r + 2$. We hope to return to this and other related questions in another paper.

References

[1] P. Erdös and H. Hanani, On a limit theorem in combinatorial analysis, Publ. Math. Del. 10 (1963) 10–13.
[2] P. Erdös and G. Szekeres, A combinatorial problem in geometry, Compositio Math. 2 (1935) 463–470.
[3] H.J. Ryser, Combinatorial mathematics, Carus Math. Monographs (Wiley, New York).