JOURNAL OF COMBINATORIAL THEORY, Series A 28, 338–350 (1980)

On Harrington's Partition Relation

MIROSLAV BENDA

Department of Mathematics, University of Washington, Seattle, Washington 98195

Communicated by the Managing Editors

Received January 30, 1979

1

The purpose of this paper is to investigate a Ramsey-like relation among the natural numbers which was recently introduced by L. Harrington. Considering a natural number n as the set of smaller natural numbers $n = \{0, 1, \ldots, n-1\}$ we write

$$n \underset{*}{\rightarrow} (k)_c^e$$

if for any f: $[n]^e \rightarrow c$, where $[n]^e = \{s \subseteq n \mid |s| = e\}$, there is $X \subseteq n$ such that $|X| \geqslant k$, $\min(X) \leqslant |X|$, and f is constant on $[X]^e$.

If we drop the assumption $\min(X) \leqslant |X|$ then we get the usual Ramsey relation. The reason for considering this much stronger relation is the amazing fact that in Peano arithmetic (the usual axiom system for natural numbers) it is impossible to prove

$$(r^*) : (\forall\, k, e, c)(\exists\, n)[n \underset{*}{\rightarrow} (k)_c^e] ;$$

yet this sentence is true of natural numbers. In fact as Paris and Harrington showed the implication "$(r^*) \rightarrow$ arithmetic is consistent" is provable from Peano axioms. A theorem of Gödel says that unless the axioms prove everything the statement "arithmetic is consistent" is unprovable from them. Hence the validity of r^* and provability of the implication is probably the ultimate in the effort toward Hilbert's First Problem.

The work of Harrington and Paris (see [8]) was extended by Solovay, who gave combinatorial arguments on the growth of the function

$$f(n) = \text{least } k[k \underset{*}{\rightarrow} (n+1)_n^n]$$

and showed that it grows spectacularly (see [9]). In this context, the unprovability of (r^*) may be illuminated by saying that n is so large with respect to k, e, c that it can't be defined from them using addition, multiplication, and

0097-3165/80/030338–13$02.00/0

induction and hence it is not forced to be in a (nonstandard) model of arithmetic which contains k, e, and c.

However, given e it is possible to show that

$$(\forall k, c)(\exists n)[n \underset{*}{\rightarrow} (k)_c^e]$$

is provable in arithmetic. This may seem curious to one not familiar with the subtleties of the provability relation. An explanation might run as follows: given e we have a proof but for different e's the proofs are so different that it is impossible to combine their features into an inductive proof. To amplify a proof for $e = 2$ is not of much help when considering $e = 3$.

Our aim in this paper is to give not only a proof of (1) when $e = 2$ but also reasonable bounds for n. We shall be concerned mostly with $c = 2$. To give a preview let us denote the least n such that $n \underset{*}{\rightarrow} (k)_2^2$ by $r^*(k)$. We shall show that

$$r^*(3) = 6\ (= r(3)), \qquad 18 < r^*(4) \leqslant 50, \qquad \text{and} \qquad r^*(5) < 7092.$$

In Section 2, we place the relation completely into graph theory by restating it so that it is independent of the ordering. The rest of the section shows that the statement (r^*) is really true, that it follows from the infinite Ramsey theorem and, moreover, we introduce (R^*), which is equivalent, in arithmetic, to an approximation of the infinite Ramsey theorem. A reader wishing to get a feeling of why (r^*) is unprovable may find it here but the rest of the paper is independent of this part of Section 2. Elaborations in this direction in arithmetic and set theory may be found in Benda [2]. Section 3 discusses the upper and lower bounds and related matters and is purely combinatorial.

Notation. Given f: $[A]^t \rightarrow c$, a set $X \subseteq A$ is called f-homogeneous if $f(s) = f(s')$ for any $s, s' \in [X]^t$. We let h_a for $a \in A$ denote the largest size of an f-homogeneous set to which a belongs. An h-sequence of f: $[A]^t \rightarrow c$ is a sequence $(h_{a_0}, h_{a_1}, \ldots, h_{a_{n-1}})$, where $A = \{a_i \mid i < n\}$, such that $h_{a_i} \leqslant h_{a_j}$ if $i \leqslant j$.

Remark. An h-sequence is similar to the degree sequence of a graph. We do not have, though, a description of those sequences which are h-sequences as is the case for degree sequences (see Erdös and Gallai [4]). The only nontrivial results in this connection we know of are due to Galvin and Krieger (see [7, Corollary 2]).

THEOREM 2.1. *The following are equivalent:*

(1) $n \underset{*}{\rightarrow} (k)_t^c$

(2) *If* $|A| \geqslant n$ *and* f: $[A]^t \rightarrow c$ *then for some* m, $k \leqslant m < n$, *all but* m *elements of* A *are in* f*-homogeneous sets of size* $\geqslant m$.

(3) *For any h-sequence* $(h_0, h_1, \ldots, n_{n-1})$ *we have* $m \leqslant h_m$ *for some* m, $k \leqslant m < n$.

Proof. Statements (2) and (3) are clearly equivalent. Note that (2) is a property which in no way depends on any ordering, i.e., is invariant under permutations.

To prove (2) from (1), we take f: $[A]^t \to c$ and order A by $\leqslant$ so that $a \leqslant b$ implies $h_a \leqslant h_b$. We then identify A with n using the order-preserving map and invoke (1) to obtain f-homogeneous relatively large X. If $a = \min(X)$ we have $a \leqslant |X| \leqslant h_a$ so (3) is true.

To go the other way, let f: $[n]^t \to c$. We know by (2) that for some m, $k \leqslant m < n$ all but m points are in f-homogeneous subsets of size $\geqslant m$. Let a be the first such point. Then $a \leqslant m$, so if X contains a and $|X| \geqslant m$, X is relatively large.

We shall now show that r^* is true. We give two proofs. The first proof is short but uses extravagant machinery. The second proof tries to prove r^* by the most elementary means and will enable us to see the difficulty in proving r^* in elementary arithmetic and, therefore, will show what to concentrate on in order to obtain the desired bounds.

THEOREM 2.2 (Harrington). *r^* is true.*

The first proof. Assume that we have k, t, and c and for every $n \in N$ we have A_n and f_n: $[A_n]^t \to c$ which is a counterexample to R^*. Let μ be a finitely additive measure on subsets of N with values 0 or 1 which assigns 0 to all singletons. The existence of this measure requires an essential use of the Axiom of Choice. Take $A = \prod_{n \in N} A_n$ and identify $x, y \in A$ if $\mu(\{n \mid x(n) = y(n)\}) = 1$. Let A/μ be the set of equivalence classes. Define

$$f(x_1/\mu, \ldots, x_t/\mu) = i \qquad \text{if} \quad \mu(\{n \mid f_n(x_1(n), \ldots, x_t(n)) = i\} = 1.$$

Then f: $[A/\mu]^t \to c$ and by the infinitary Ramsey theorem, we know that all but finitely many points of A/μ are in infinite f-homogeneous sets. If m' is the number of such points let $m = \max(k, m')$. It is easy to see from the construction that for some $n > m$ all but m elements are in f_n-homogeneous sets of size $\geqslant m$, so (A_n, f_n) could not have been a counterexample.

The second proof. We shall be informal but up to a certain point this proof could be formalized in elementary arithmetic. We again assume that we have k, t, c such that for any n we have $n \underset{*}{\not\to} (k)_t^c$. Let G_n be the set of counterexamples f: $[n]^t \to c$ in the sense of 2.1(2). Note that G_n is finite.

For $f \in G_n$, $g \in G_m$, $n < m$ define fRg if (n, f) can be embedded into (m, g); that is, for some F: $n \to m$

$$F(f(a_1, \ldots, a_t)) = g(F(a_1), \ldots, F(a_t)).$$

Note that if $g \in G_{n+1}$ there is $f \in G_n$ such that fRg which is obtained (in essense) by taking out an element from $n + 1$ which belongs to the largest g-homogeneous set. This shows that the relation R has arbitrarily long sequences $f_1Rf_2R \cdots Rf_n$. The relation is transitive and we have a situation similar to the König's Tree Theorem. There is, therefore, a sequence $f_n \in G_n$ such that f_nRf_{n+1} for every $n \in N$. Rearanging the natural numbers we can assume that the maps guaranteed by f_nRf_{n+1} are inclusions. The union of the f_n's is then f: $[n]^t \to c$ and it can be seen by inspection of the construction that it is definable from the counterexamples, so we are still within arithmetic.

Now we wish to apply the infinite Ramsey theorem. Its application gives a speedy contradiction to the existence of the counterexamples and finishes the proof of Theorem 2.2. This ending is of course impossible to carry out within elementary arithmetic where we cannot mention infinite sets or functions unless we define them. We shall now try to see how much of the infinite Ramsey theorem can be formulated within arithmetic and whether the resulting statement may enable us to prove that r^* is true. We shall see that even more can be achieved.

Consider the following schema of statements of arithmetic:

For any c, $t \in N$, if F: $[N]^t \to c$ is a (definable) function then some $n \in N$ belongs to arbitrarily large F-homogeneous sets.

This schema may be formulated within arithmetic as infinitely many axioms. Ramsey's theorem implies that all of them are true (in fact much more). However, as we shall presently see, this schema is unprovable in arithmetic for it implies, within arithmetic the statement r^*. The reader may inspect his/her favorite proof of the infinite Ramsey theorem and it will be evident that although the proof constructs an infinite homogeneous set, it does not construct an element of the set. We shall elaborate on this point in the next section.

Before we proceed to show that the Ramsey schema implies r^*, we ask whether the converse is true. It turns out the Ramsey schema is equivalent (in arithmetic) to a natural strengthening of r^* which we define below and denote by R^*:

For any c, $t \in N$ and any (definable) function f: $N \to N$ there is n such that

$$n \to (f)^t_c,$$

where $n \to (f)^t_c$ means: if F: $[n]^t \to c$ then for some $m < n$ all but m points belong to F-homogeneous sets of size $\geqslant f(m)$.

THEOREM 2.3. *R^* implies R.*

Proof. Let F: $[N]^t \to c$ be a definable function and assume that for every $n \in N$ the F-homogeneous sets containing n have size $\leqslant f(n)$. By the finite

Ramsey theorem, for given $k \in N$, we have only finitely many n juch that $f(n) = k$. Hence, by rearranging the natural numbers, we can assume that $n \leqslant m$ implies $f(n) \leqslant f(m)$. We now use R^* for the function $f(x) + 1$. We get n such that $n \to (f)_c^t$. Let g be the restriction of F to $[n]^t$ and $m < n$. We then have a contradiction since the points $0,\ldots, m$ belong to g-homogeneous sets of size $< f(m) + 1$ and there is $m + 1$ of them.

In order to prove the other implication, we need a lemma which is interesting in itself because together with Ramsey theorem, it implies R^*. It is conceivable that some finitary version of it together with the finite Ramsey theorem would imply r^*. We were not able to find an appropriate formulation.

First some terminology: If S is a set of finite subsets of N, we say that $x \in N$ is covered by S if $\{s \in S \mid x \in s\}$ contains sets of arbitrarily large size. For n, $m \in N$ we say that x is m-covered in n if for some $s \in S$ containing x we have $|s \cap n| \geqslant m$. We say S is h-dense, where h: $N \to N$, if for any $n \in N$ and $r \subseteq N$ $|r| \geqslant h(n)$ there is $s \subseteq r$, $|s| \geqslant n$, and $s \in S$.

LEMMA 2.4. *Let S be an h-dense set. Then for any $m \in N$ there are infinitely many n such that if $x < n$ is covered by S then it is m-covered in n.*

Proof. If no x is covered by S there is nothing to prove. Assume that some x is covered and the lemma is false. Then we have n_0 and for $n \geqslant n_0$ we have $g(n) < n$ which is covered by S but not m-covered in n.

Case 1. g is bounded. Then $g(n) = b$ for infinitely many n's and since b is covered, we get $s \in S$ juch that $b \in s$ and $|s| \geqslant m$. Taking n such that $s \subseteq n$ we get a contradiction.

Case 2. g is unbounded. We find a set r in the range of g, $|r| \geqslant h(m)$ and using the h-density of S get $s \in S$, $s \subseteq r$, and $|s| \geqslant m$. Let b be the last element of s. Then for some $n > b$ $g(n) = b$ but b is m-covered in n.

THEOREM 2.5. *R implies R^*.*

Proof. Let c, $t \in N$ and g: $N \to N$ be given such that for no n we have $n \to (g)_c^t$. As in the second proof of Theorem 2.2 we average out the counterexamples and get f: $[N]^t \to c$. Let S be the set of all finite f-homogeneous subsets. It is easy to see the Ramsey schema implies that S not only covers some point but that it in fact covers all but finitely many points. (Reason: we take the set of uncovered points U, if unbounded the function $f \restriction U$ is isomorphic to a definable g on $[N]^t$ showing the Ramsey schema false for g.) Let k be the number of uncovered points. Let $h(x)$ be the least n such that $n \to (x)_c^t$ (from finite Ramsey theorem). Then S is h-dense and the assumptions of Lemma 2.4 are satisfied. Let $m = g(k)$. By Lemma 2.4, we get n such that some covered point is $<n$ and every covered point $<n$ is m-

covered in n. That is, all but k points are in f-homogeneous subsets of n of size $\geqslant g(k)$. Remembering that we rearranged N so that the restriction of f to n would coincide with the supposed counterexample on n we see that we have a contradiction.

Note that R^* is much stronger than r^*. The first n such that $n \underset{*}{\rightarrow} (k)^t_c$ is the first n such that $n \rightarrow (f_k)^t_c$ where $f_k(x) = \max(k, x)$. Hence the Ramsey schema is unprovable in arithmetic (assuming it consistent). In fact the proof of this is easier than the proof of unprovability of r^* and can be found in [2].

In this section, we shall find bounds for the least numbers satisfying $n \underset{*}{\rightarrow} (k)^2_2$. We start by discussing the situation for small values of k; this will prepare the reader for the general bounds and will give us a feeling of how good the bounds are.

In order to ease the notation, we shall call a function f: $[n]^2 \rightarrow 2$ a colored graph on n; it essentially is the complete graph on vertices whose edges are colored by two colors, 0 and 1. We shall call an f-homogeneous set an f-set, and an i-set ($i = 0, 1$) will be an f-set X such that f is constantly i on $[X]^2$. Our bounds will be expressed in terms of the ordinary Ramsey numbers $r(k, l)$, which is the least number n such that any colored graph on n vertices has a 0-set of size k or it has a 1-set of size l. By $r(k)$, we denote $r(k, k)$ and $r^*(k)$ will stand for the least n satisfying $n \underset{*}{\rightarrow} (k)^2_2$. The basic relationship for $r(k, l)$ is

$$r(k, l) \leqslant r(k-1, l) + r(k, l-1)$$

($k, l \geqslant 2$) and it then follows that $r(k, l) \leqslant \binom{k+l-2}{k-1}$ (see [4, for more information on $r(k, e)$. The best general bounds for $r(k)$ are

$$2^{k/2} \leqslant r(k) \leqslant \binom{2(k-1)}{k-1},$$

the lower bound being due to Erdös. For special numbers, the lower bound may be improved, e.g., $r(2^n + 1) \geqslant 5^n + 1$ (see Abbott [1]).

The case $k = 3$: Since we have $r(k) \leqslant r^*(k)$ we get $6 \leqslant r^*(3)$ and because 3 happens to be a half of 6, we get $r^*(3) = 6$.

In fact, we can do better than this. It is easy to see that any colored graph on 6 vertices has at least two homogeneous triangles so all but two points are in homogeneous sets of size 3. This leads immediately to the question of what happens for larger graphs. In general, we ask how many points can there be in a large, or simply infinite, graph which do not belong to homogeneous sets of size k? Clearly their number is limited by $r(k)$ so we may define a function g where $g(k)$ (for $k > 2$) is the maximal number n such that some infinite colored graph contains n vertices none of which is in a homogeneous set of size k. Our first bounds were obtained using g; however, later we found better bounds may be achieved with g and its variants and that these will be essential for considering $t > 2$. For this reason, we give an evaluation of g.

THEOREM 3.1. $g(k) = r(k, k-1) - 1$ $(k > 2)$.

Proof. First of all, we show that $g(k) \geqslant r(k-1, k) - 1$. The definition of $r(k-1, k)$ implies that it is the least n such that any colored graph f on n vertices which does not have an f-set of size k has 0- and 1-sets of size $k-1$. Therefore, $r(k-1, k) - 1$ is the largest n for which we find $f\colon [n]^2 \to 2$ with say, 0-sets of size $k-1$ but 1-sets of size $<k-1$. Now extend f to N be defining $f(u, v) = 0$ for $u, v \geqslant n$ and $f(u, v) = 1$ for $u < n \leqslant v$. It is easy to see that any $u < n$ is in an f-set of size $k-1$ at most.

To show $g(k) < r(k-1, k)$ assume we have $f\colon [N]^2 \to 2$ with $n = r(k-1, k)$ vertices not in f-sets of size k. Assume that the points are the numbers $<n$. Considering f on $[n]^2$, we see that there are i-sets of size $k-1$ included in n for $i = 0$ and $i = 1$. For $S \subseteq n$ define

$$C_S = \{u \mid u \geqslant n \wedge S = \{m \mid m < n \text{ and } f(m, u) = 0\}\}.$$

Note that $C_\varnothing = \varnothing$ because if $u \in C_\varnothing$ then taking a 1-set $X \subseteq n \mid X \mid = k-1$ we would get an $m < n$ belonging to the 1-set $X \cup \{u\}$ of size k. Similarly $C_n = \varnothing$. Now if $0 \subset S \subsetneq n$ then

$$|C_S| < r(k-1, k-1)$$

for if not C_S would include, say 1-set X of size $k-1$ and then taking $X \cup \{m\}$, where $m < n$ and $m \notin S$, we would get a 1-set of size k with m in it As the set $\{C_S \mid S \subseteq n\}$ covers an infinite set one of them should be infinite.

Remark. In particular, we have $g(3) = r(2, 3) - 1 = 2$, therefore any infinite colored graph has at most two points not in homogeneous triangles and, in fact, this is true for any graph on at least 6 points. Also $g(4) = r(3, 4) - 1 = 9 - 1 = 8$.

The proof also gives a bound on the size of a graph which contains $m \geqslant r(k-1, k)$ vertices not in homogeneous sets of size k. The bound is $(2^m - 2)r(k-1, k-1)$. So the largest graph with vertices not in homogeneous sets of size 4 has at most 6510 vertices. Actually, with a more detailed work, one can get a bound which is about a tenth of this.

Inspecting the proof of $g(k) \geqslant r(k-1, k) - 1$ we see that we have introduced a large homogeneous set, i.e., $\{u \mid u \geqslant n\}$. This is in fact typical, e.g., if we have a graph on $n \geqslant 6$ vertices which has two points not in komogeneous triangles then it will have a homogeneous set of size $\geqslant n - 3$. This is illustrated in general below.

THEOREM 3.2. $19 \leqslant r^*(4) \leqslant 50$.

Proof. We know 18 is the smallest n such that $n \to (4)^2_2$. An example that $17 \nrightarrow (4)^2_2$ may be obtained by taking

$$f(u, v) = 0 \quad \text{if } u - v \text{ is a quadratic residue mod 17,}$$
$$= 1 \quad \text{otherwise,}$$

for $0 \leqslant u, v < 17$. Extending f by defining $f(u, 17) = 0$ iff $u \equiv 0 \pmod 3$, we see that the points 0, 3, 6, 9, 12, 15 are in no f-homogeneous sets of size 4 and there are 6 of them. As there are no f-sets of size >4, we see that $18 \nrightarrow_* (4)^2_2$.

To find the upper bound, we take f: $[n]^2 \to 2$ and assume that we have four points not in f-sets of size 4. We could assume to have five of them but we do not see a way of using this information. Then there will be three among these a, a_0, and a_1 say, such that $f(a, a_0) = 0$ and $f(a, a_1) = 1$. Let

$$A_i = \{x < n \mid x \neq a, f(a, x) = i\}.$$

Note that $a_i \in A_i$ for $i = 0$ or 1 and define for $i, j \in \{0, 1\}$,

$$A_{ij} = \{x \in A_i \mid x \neq a_i \text{ and } f(a_i, x) = j\}.$$

We then have: A_{00} is a 1-set, A_{11} is a 0-set, A_{01} cannot contain an f-set of size 3 so $|A_{01}| \leqslant 5$, and similarly for A_{10}. We thus have two f-sets A_{00} and A_{11} of sizes k_0 and k_1 resp. and the rest of n has at most 13 points. Now we need a condition on n which would allow us to conclude that if $n = 13 + k_0 + k_1$ then either $k_0, k_1 \geqslant 13$ or if, say, $k_0 < 13$ then $k_1 \geqslant k_0 + 13$. Then we are done for in the former case all the points of $A_{00} \cup A_{11}$ are in f-sets of size $\geqslant 13$ and in the latter case all but $k_0 + 13 \leqslant k_1$ points are in f-sets of size k_1 (i.e., in A_{11}). The smallest n satisfying this is $4 \cdot 13 - 2 = 50$.

The argument has not been typical. A more typical argument, though we shall still use information not available in general is:

THEOREM 3.3. $r^*(5) \leqslant 7\,092$.

Proof. Let f: $[n]^2 \to 2$ be such that we have six points not in f-sets of size 5. We can then find a, a_0, a_{00}, and a_1 among these such that $\{a, a_0, a_{00}\}$ is a 0-set and $f(a, a_1) = 1$ (we use the fact that $6 \to (3)^2_2$ here). Let

$$A_i = \{x \mid x \neq a \text{ and } f(a, x) = i\}$$

so $a_0, a_{00} \in A_0$, and $a_1 \in A_1$. For $i, j \in \{0, 1\}$ define

$$A_{ij} = \{x \in A_i \mid x \neq a_i \text{ and } f(a_i, x) = j\}.$$

Then $a_{00} \in A_{00}$ and we define

$$A_{00i} = \{x \in A_{00} \mid x \neq a_{00} \text{ and } f(a_{00}, x) = i\}.$$

It can be easily seen that A_{000} is a 1-set, A_{01} and A_{10} have no f-sets of size 4 thus they have at most 17 elements and A_{11} has no 1-set of size 3. Finally, A_{001} has no 0-set of size 3 and no 1-set of size 4 so $|A_{001}| < r(3, 4) = 9$. Consequently,

$$|\{a, a_0, a_{00}\} \cup A_{01} \cup A_{10} \cup A_{001}| \leqslant 44.$$

The points listed in the set above are considered useless because we do not know whether they belong to large homogeneous sets and we intend to throw them out of the graph. The points in A_{000} are useful but only if $|A_{000}| \geqslant 44$ because we already intend to throw out 44 points so the rest should be in f-sets of size at least 44. We therefore assume that $|A_{000}| \leqslant 43$ and go to A_{11} to look for a large homogeneous set. If all points of A_{11} are in f-sets of size $43 + 44 = 87$ we are done. If not, let $b \in A_{11}$ be a point not in an f-set of size 87. Define

$$A_{11i} = \{x \in A_{11} \mid x \neq b \text{ and } f(c, x) = i\}$$

Then A_{110} has no 1-sets of size 3 (because $A_{110} \subseteq A_{11}$) and no 0-sets of size 86 so $|A_{110}| < r(3, 86)$. The set A_{111} is 0-homogeneous. There is a bound for the numbers $r(3, k)$ namely $r(3, k) \leqslant (k^2 + 3)/2$ so $|A_{110}| \leqslant (86^2 + 3)/2 = 3458$. Thus if $|A_{000}| \leqslant 43$ we throw out at most $3459 + 87 = 3546$ points and in order to make A_{111} of size 3546 we have to make $n \geqslant 7092$. In the case when $|A_{000}| \geqslant 44$ we proceed as follows: we assume we have a point $c \in A_{11}$ which is not in a homogeneous set of size 44, for otherwise we are done. Forming the sets A_{11i} as before, we find that $|A_{110}| < r(3, 43) \leqslant 924$. We are now in the same situation as in Theorem 3.2. We have $925 + 44 = 969$ definitely useless points and two homogeneous sets of size k_0 and k_1. In order to make sure that $k_0, k_1 \geqslant 969$, or $k_0 \geqslant 969 + k_1$, or $k_1 \geqslant 969 + k_2$ whenever $n = 969 + k_0 + k_1$ we have to take $n = 4 \cdot 969 - 2 = 3874$ in this case which is smaller than the former bound.

This process may be applied to a general graph where, however, we shall not be able to use any special information and so the bounds will be much looser. Even in the case $k = 5$ the bound is probably too high because we in effect show that for any graph on 7092 points with 4 points as described above a half of it is a homogeneous set or two thirds of it consist of two homogeneous sets of different colors.

We define the bound for $k \geqslant 3$, b_k, by induction as follows: For $i \leqslant k - 2$, let $e_{i+1} = e_i + 2 \cdot r(k - i, e_i - 1)$, with $e_0 = k$. Then for $1 \leqslant i \leqslant k - 2$ let $f_{i+1} = f_i + 2 \cdot r(k - i, f_i - 1)$ with $f_1 = e_{k-2}$. The number b_k is f_{k-2}. Thus $b_3 = 3 + 2 \cdot r(2, 2) = 7$ (pretty good) but $b_4 = 144$.

THEOREM 3.4. $r^*(k) \leqslant b_k$.

Proof. Assume $|A| \geqslant b_k$ and $f: [A]^2 \to 2 = \{0, 1\}$. We define by induction sets A_i^c, U_i, and points a_i^c where $c = 0$ or 1 and as far as the i's are concerned, we indicate below when the induction terminates.

We assume we have a point $a \in A$ which is not in f-sets of size $\geqslant k$. Let, for $c < 2$,

$$A_0^c = \{x \in A \mid x \neq a \text{ and } f(x, a) = c\}.$$

We define $U_0 = \{a\}$ (the elements of some U_i are to be thought of as useless) and let $k_0 = k$. In general k_i will be $\max(|U_i|, k)$. Assume now that for $i \leqslant n$ the sets A_i^c, U_i have been constructed and the points a_i^c have been constructed for $i < n$.

Case 1. Every point $x \in A_n^c$ ($c = 0$ or 1) is in an f-set of size $\geqslant k_n$. In this case, the definition is terminated.

Case 2. Every $x \in A_n^1$ is in an f-set of size $\geqslant k_n$ but some point in A_n^0 is not in such set

In this case, we let $a_n^0 \in A_n^0$ to be a point not in an f-set of size $\geqslant k_n$ and let $a_n' = a_{n-1}'$. We let

$$A_{n+1}^0 = \{x \in A_n^0 \mid x \neq a_n^0 \text{ and } f(a_n^0, x) = 0\},$$

$$B_{n+1} = \{x \in A_n^0 \mid x \neq a_n^0 \text{ and } f(a_n^0, x) = 1\},$$

$$U_{n+1} = U_n \cup B_{n+1} \cup \{a_n^0\}.$$

We define $A_{n+1}^1 = A_n^1$.

Case 3. This is like Case 2 except that the roles of 0 and 1 are interchanged in the statement of the case as well as in the definitions.

Case 4. Neither of the cases above. We then have for $c < 2$ points $a_n^c \in A_n^c$ which do not belong to f-sets of size $\geqslant k_n$ and we define, for $c < 2$

$$A_{n+1}^c = \{x \in A_n^c \mid x \neq a_n^c \text{ and } f(x, a_n^c) = c\}$$

$$B_{n+1}^c = \{x \in A_n^c \mid x \neq a_n^c \text{ and } f(x, a_n^c) = 1 - c\}$$

$$U_{n+1} = B_{n+1}^0 \cup B_{n+1}^1 \cup \{a_{n+1}^0, a_{n+1}^1\} \cup U_n.$$

It is clear that the definition will terminate in finitely many steps because we remove at every step a point from A_i^0 or from A_i^1 and A is finite. Actually, we now show that it will terminate in less than $2k$ steps.

Let $r_n^c = |\{a_i^c \mid i < n\}| + 1$.

Fact 1. If $X \subsetneq A_n{}^c$, $n > 0$, is a c-set then $|X| < k - r_n{}^c$. The reason for this is that if $X \subsetneq A_n{}^c$ is a c-set then so is

$$X \cup \{a_i{}^c \mid i < n\} \cup \{a\},$$

but the point a is not in any homogeneous set of size $\geqslant k$. Hence $|X| < k - r_n{}^c$.

Now the sequences r^0 and r^1 are nondecreasing and have the property that if, e.g., $r_n{}^0 = r_{n+1}^0$ then $r_n{}^1 < r_{n+1}^1$. Second, by Fact 1, for any n $r_n{}^c \leqslant k - 1$. The longest possible number of steps can, therefore, be $(k - 1) + (k - 2) = 2k - 3$. This might be realized, for example, if for the first $k - 2$ steps we were in Case 2, on the $(k - 1)$ step in Case 4 and then in Case 3 for $k - 2$ steps more.

On the last step, call it l when we enter Case 1, the sets $A_l{}^0$ and $A_l{}^1$ might be empty. If not then any element of $A_l{}^0 \cup A_l{}^1$ is in an f-set of size $\geqslant k_l < |A|$, i.e., all but k_l elements (in fact, possibly less) are in f-sets of size $\geqslant k_l$. We thus have to show the procedure will not turn A into U_l.

To do this, we need a record of which case we considered on a particular step. Let $z(0) = (1, 1)$ and $z(n) = (r_{n+1}^0 - r_n{}^0, r_{n+1}^1 - r_n{}^1)$ for $n > 0$. Thus if $z(n) = (0, 1)$ then we know we considered Case 3 on step n because we had $a_n{}^0 = a_{n-1}^0$ and $a_n{}^1 \neq a_{n-1}^1$. Given z we define a sequence of numbers of the same length by induction as follows:

Let $m_0 = k$, and assume m_i has been defined for $i \leqslant n$, $n < l - 1$.

(a) If $z(n + 1) = (0, 1)$ let $m_{n+1} = m_n + r(k - r_n{}^1, k_n - 1)$.

(b) If $z(n + 1) = (1, 0)$ let $m_{n+1} = m_n + r(k - r_n{}^0, k_n - 1)$.

(c) If $z(n + 1) = (1, 1)$ let $m_{n+1} = m_n + r(k - r_n{}^0, k_n - 1) + r(k - r_n{}^1, k_n - 1)$.

Fact 2. For $n \leqslant l$ $k_n \leqslant m_n$: If $n = 0$ then $k_0 = k = m_0$. Assume the fact for n and, so as not to consider all cases, suppose that $z(n + 1) = (1, 1)$, that is on the $(n + 1)$-step we were in Case 4. The other cases are treated similarly. We should estimate the size of the sets B_{n+1}^c. Because $B_{n+1}^c \subsetneq A_n{}^c$ then we cannot have c-set $X \subsetneq B_{n+1}^c$ of size $k - r_n{}^c$ as follows from Fact 1. If $x \in B_{n+1}^c$ then $f(a_n{}^c, x) = 1 - c$ and the point $a_n{}^c$ is not in an f-set of size k_n, so the $(1 - c)$-sets $\subsetneq B_{n+1}^c$ have size $< k_n - 1$. Hence $|B_{n+1}^c| < r(k - r_n{}^c, k_n - 1)$ and therefore,

$$|U_{n+1}| \leqslant |U_n| + r(k - r_n{}^0, k_n - 1) - 1 + r(k - r_n{}^1, k_n - 1) - 1 + 2 \leqslant m_{n+1}.$$

All we need to check now is that given any sequence z which corresponds to an actual construction the last number of the sequence m_l connected with z as above is smaller than the number b_k defined previously. It may be seen that we have been more than generous, for b_k is defined from the sequence $((1, 1), \ldots, (1, 1))$ ($2k - 2$ times) and thus majorizes term by term any m_i which occurs above.

Remark and Conjecture. When we were doing preliminary estimates for $r^*(k)$ we supposed we were in Case 2 all the time and the construction eneded in $k - 2$ steps with A^0_{k-2} homogeneous. The bound for the useless part then came out

$$| B_n | \leqslant r(k - n, r(k - n + 1, \ldots, r(k - 1, k - 1) \cdots))$$

and so the largest set thrown into the useless part has size $\leqslant r(2, r(3, \ldots, r(k - 1, k - 1 \cdots))$. This number cannot serve as a bound for small values of k, however, the number $r(2, r(3, \ldots, r(k - 1, r(k, k)) \cdots))$ is large enough, has natural combinatorial interpretation and we conjecture that it is indeed a bound for $r^*(k)$.

The proof of Theorem 3.4 has the advantage that it readily generalizes to the relation $n \to (g)^2_2$, where $g\colon N \to N$ for which we assume to avoid trivialities that $g(k) \geqslant \max(k, 3)$. Define $e_0 = g(0)$ and $e_{i+1} = e_i + 2 \cdot r(k - i, g(e_i) - 1)$ for $i \leqslant k - 2$ and $f_{i+1} = f_i + 2 \cdot r(k - i, g(f_i) - 1)$ for $1 \leqslant i \leqslant k - 2$ with $f_1 = e_{k-2}$. Let $b_g = f_{k-2}$.

THEOREM 3.5. *The least n such that $n \to (g)^2_2$ is at most b_g.*

Proof. The proof differs from the proof of 3.4 in the formulation of the cases. For example, Case 2 is modified as follows: Every $x \in A_n{}^1$ is an f-set of size $g(k_n)$ but some $x \in A_n{}^0$ is not in such set. The sets $A_i{}^c$, U_i, and the points $a_i{}^c$ are defined as before and when it comes to the bounds, we first notice that Fact 1 does not depend on g at all and in Fact 2 the only alteration to be made is to replace k_n by $g(k_n)$. This, of course, applies to the definition of the sequence m_n as well.

Having 3.5, we can also extend the bounds to graphs colored by more than two colors. What is required is a further modification of the construction 3.4 and the knowledge of the bounds b_g for functions $g(i) = i + b_k$ and $h(i) = i + b_g$. For three colors, i.e., $f[A]^2 \to \{0, 1, 2\} r_3^*(k) \leqslant b_k + b_g + b_h$. The argument is not straightforward but we omit it here.

REFERENCES

1. H. L. ABBOTT, Lower bounds for some ramsey numbers, *Discrete Math.* **2** (1972), 289–293.
2. M. BENDA, On strong axioms of induction in arithmetic and set theory, *in* "Proceedings, IV Symposium on Logic, Chile 1978," to appear.

3. J. A. Bondy and U. S. R. Murty, "Graph Theory With Applications," Elsevier, New York, 1976.
4. P. Erdös and G. Szekeres, A combinatorial problem in geometry. *Comp. Math.* **2** (1935), 463–470.
5. P. Erdös, Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.* (1947), 292–294.
6. P. Erdös and T. Gallai, Graphs with prescribed degrees of vertices, *Mat. Lapok* **11** (1960), 264–274.
7. F. Galvin and M. M. Krieger, The minimum number of cliques in a graph and its complements, *in* "Proceedings, 2nd Louisiana Conference on Combinatorics, Baton Rouge 1971," pp. 345–352.
8. L. Harrington and J. Paris, A mathematical incompleteness in peano arithmetic, *in* "Handbook of Mathematical Logic," (K. J. Barwise, Ed.), Elsevier, New York, 1977.
9. R. Solovay, Rapidly growing Ramsey functions, manuscript.