JOURNAL OF COMBINATORIAL THEORY (A) 16, 199–208 (1974)

A Substitute for Hall's Theorem for Families with Infinite Sets

SAHARON SHELAH

Department of Mathematics, The Hebrew University, Jerusalem, Israel

Communicated by the Managing Editors

Received April 10, 1972

A sufficient condition for the existence of a system of distinct representatives for a family S is that $x \in A \in S$ implies the number of elements of A is not smaller than the number of sets in S to which x belongs.

INTRODUCTION

The celebrated Hall's Theorem [1, 2] gives a necessary and sufficient condition for the existence of a system of distinct representatives for a family S of finite sets: the union of any m sets of S is a set of at least m elements. Synonyms for "a system of distinct representatives" are "a transversal" and "a one-to-one choice function." The problem of finding such conditions for an arbitrary family of sets seems to be difficult (see [5, 6] for partial results and for explanation of the difficulty). So Mirsky suggested looking for sufficient conditions, and successively he, Knight, and Milner formulated such conditions. In [4] it was shown that Knight's condition $A \in S \Rightarrow |\{B: B \in S, A \cap B \neq \varnothing\}| \leqslant |A|$ is sufficient for the family S to have a system of distinct representatives. It was also pointed out in that paper that in order to prove the sufficiency of Milner's weaker condition, stated in the abstract, it is enough to prove it in the case in which S is a denumerable family of sets. This will be proved here. Unlike [4], no set-theory is needed and the proof is combinatorial and computational. Unfortunately it is not elegant. For more information on transversals, see Mirsky [3]. This is a partial answer to Problem 15 of [3].

NOTATION. N will be the set of natural numbers including zero. By a family of sets we mean an indexed one, that is, with possible repetitions. S will be a fixed countable family of non-empty sets, that is,

$S = \{A_n : n \in N\}$. L will denote a subfamily of S consisting of finite sets only. K will denote a finite subfamily of S consisting of finite sets only. For any $S' \subseteq S$, $D(S') = \mathrm{dom}(S') = \bigcup_{A \in S'} A$. X, Y, A, and B are subsets of, and x, y points of, dom(S). The letters $i, j, k, l, m, n, p, q, r$ always denote natural numbers. $|A|$ is the number of elements of A—which is here a natural number or ∞; the same with $|L|$—the number of sets $A \in L$. $v(x) = |\{A: x \in A \in S\}|$. A transversal of S is a one-to-one function F, $F(A) \in A$. Using indexes, we agree $p_i = p(i)$ etc.

MAIN THEOREM. *If S is a countable family of non-empty sets and $x \in A \in S$ implies $v(x) \leqslant |A|$ then S has a transversal.*

Proof. W.l.o.g. for every $x \in \mathrm{dom}(S)$:

(1) $\quad v(x) = \min\{|A|: x \in A \in S\}$.

For otherwise we shall add to S infinite sets A_n' (which may have points outside dom(S)) such that, for $x \in \mathrm{dom}(S)$, $x \in A_n'$ iff $n + 1 \leqslant \min\{|A|: x \in A \in S\} - v(x)$. Let $S_f = \{A: A \in S, A \text{ finite}\}$. W.l.o.g. we may assume that $S_i = S - S_f = \{Y_n: n \in N\}$ is infinite and that the sets Y_n are denumerable (for any infinite member of S can be replaced by a denumerable subset).

A set $B \subseteq \mathrm{dom}(S)$ is *acceptable* if S_f has a transversal F, which avoids B, that is, $F(A) \notin B$ for any $A \in S$. By Hall's Theorem, B is acceptable iff $|D(K) - B| - |K| \geqslant 0$ for any $K \subseteq S$. If equality holds, K is called *B-critical.* It is easy to see that K is B-critical *iff* S has a transversal avoiding B, and any such transversal F satisfies $\{F(A): A \in K\} = D(K) - B$. Hence the union of B-critical sets is a B-critical set, when B is acceptable.

Define

(2) $\quad d(K) = |D(K)| - |K|$,

(3) $\quad d(K, B) = |D(K) - B| - |K|$,

(4) $\quad w(x, L) = \sum_{x \in A \in L} \frac{1}{|A|}$,

(5) $\quad t(x, L) = 1 - w(x, L)$.

Note that

(A) for any B, K $d(K, B) \geqslant d(K) - |B|$,

(B) by definition, K is B-critical *iff* $d(K, B) = 0$,

(C) by Hall's Theorem, B is acceptable *iff*, for any K, $d(K, B) \geqslant 0$.

By the condition in the theorem $x \in A \in S$ implies that $v(x) \leqslant |A|$ and hence that

$$w(x, L) \leqslant \sum_{x \in A \in S} \frac{1}{v(x)} = 1.$$

Thus

(D) $\quad 0 \leqslant w(x, L) \leqslant 1$, and $0 \leqslant t(x, L) \leqslant 1$.

Let us try to connect these numbers:

$$(6) \qquad |K| = \sum_{A \in K} 1$$

$$= \sum_{A \in K} \sum_{x \in A} \frac{1}{|A|}$$

$$= \sum_{x \in D(K)} \sum_{x \in A \in K} \frac{1}{|A|}$$

$$= \sum_{x \in D(K)} w(x, K).$$

So using (2)

$$(7) \qquad d(K) = |D(K)| - |K| = \sum_{x \in D(K)} 1 - \sum_{x \in D(K)} w(x, K)$$

$$= \sum_{x \in D(K)} t(x, K).$$

By (A)

$$(8) \qquad d(K, B) \geqslant d(K) - |B| = \sum_{x \in D(K)} t(x, K) - |B|.$$

Let us prove. We shall define by induction on n distinct $y_n \in Y_n$ such that

$(*)_n$ $\quad S_f$ has a transversal avoiding $B_n = \{y_i : i < n\}$, or equivalently for any K, $d(K, B_n) \geqslant 0$.

Suppose we succeed in defining the y_n so that $(*)_n$ holds. Let $B = \bigcup B_n = \{y_0, y_1, \ldots\}$. Then, for any K, as $D(K)$ is finite, there is n_0 such that $D(K) \cap B = D(K) \cap B_{n_0}$ and hence $d(K, B) = d(K, B_{n_0}) \geqslant 0$. Therefore, by Hall's Theorem, S_f has a transversal F avoiding B. Extend F by putting $F(Y_n) = y_n$ and we are done. All that remains is to define the y_n. By (7) and (D),

$$d(K, \phi) = d(K) \geqslant \sum_{x \in D(K)} t(x, K) \geqslant \sum_{x \in D(K)} 0 = 0,$$

and so $(*)_n$ is satisfied for $n = 0$.

From now on n is a fixed (natural) number, $y_i \in Y_i$ are defined for $i < n$, and $(*)_n$ is satisfied. If there is $y \in Y_n - B_n$ such that, for any B_n-critical K, $y \notin D(K)$ then we define $y_n = y$ and $(*)_{n+1}$ is satisfied. So assume from now that such y does not exist, and we shall get a contradiction. Let $Y_n - B_n = \{x_m : m \in N\}$. So for every m there is a B_n-critical K_m, $x_m \in D(K_m)$. Define K_m inductively so that $| D(\bigcup_{i \leqslant m} K_i)|$ is minimal (for the already chosen K_i, $i < n$). Define

$$\text{(9)} \qquad K^m = \bigcup_{p \leqslant m} K_p, \quad L = \bigcup_{p \in N} K_p, \quad D_m = D(K^m), \quad D = D(L).$$

As the union of B_n-critical families is a B_n-critical family, K^m is B_n-critical.

(E) $d(K^m, B_n) = 0$ for every $m \in N$.

So by (8)

$$0 = d(K^m, B_n) \geqslant \sum_{x \in D_m} t(x, K^m) - | B_n |.$$

Hence, as $| B_n | = n$,

$$n \geqslant \sum_{x \in D_m} t(x, K^m).$$

Since $K^m \subseteq L$, it follows from (4) and (5) that $w(x, K^m) \leqslant w(x, L)$, $t(x, K^m) \geqslant t(x, L)$. Therefore, $n \geqslant \sum_{x \in D_m} t(x, L)$. This inequality holds for any m, and, since $D_m \subseteq D_{m+1}$, $D = \bigcup_m D_m$, it follows that

$$\text{(10)} \qquad n \geqslant \sum_{x \in D} t(x, L).$$

Hence

(F) $\sum_{x \in D} t(x, L)$ converges.

The rest of the proof is dedicated to contradicting (F). For any m, $x_m \in D_m \subseteq D$, $x_m \in Y_n \notin L$, and so

$$w(x_m, L) = \sum_{x_m \in A \in L} \frac{1}{| A |} \leqslant \sum_{x_m \in A \in L} \frac{1}{v(x_m)} \leqslant [v(x_m) - 1] \frac{1}{v(x_m)}.$$

Therefore

$$\text{(11)} \qquad t(x_m, L) = 1 - w(x_m, L) \geqslant \frac{1}{v(x_m)}.$$

Sh:29

As $\sum_{x\in D} t(x, L)$ converges, $\{x \in D: t(x, L) \geqslant 1/p\}$ is finite for any p, hence

(G) $\lim_{m\to\infty} v(x_m) = \infty.$

Notice that

$$\frac{1}{m} = \sum_{p=m}^{\infty} \left(\frac{1}{p} - \frac{1}{p+1}\right)$$

and

$$\frac{1}{m} - \frac{1}{m_1} = \sum_{p=m}^{m_1-1} \left(\frac{1}{p} - \frac{1}{p+1}\right)$$

for $0 < m \leqslant m_1$. Using these identities, we obtain from (4), (5), and (1) that

$$(12) \qquad t(x, L) = 1 - \sum_{x\in A\in L} \frac{1}{|A|}$$

$$= \sum_{x\in A\in S} \frac{1}{v(x)} - \sum_{x\in A\in L} \frac{1}{|A|}$$

$$= \sum_{x\in A\in S-L} \frac{1}{v(x)} + \sum_{x\in A\in L} \left(\frac{1}{v(x)} - \frac{1}{|A|}\right)$$

$$= \sum_{x\in A\in S-L} \sum_{p=v(x)}^{\infty} \left(\frac{1}{p} - \frac{1}{p+1}\right)$$

$$+ \sum_{x\in A\in L} \sum_{p=v(x)}^{|A|-1} \left(\frac{1}{p} - \frac{1}{p+1}\right).$$

Let

$$(13) \qquad c(p, x) = \begin{cases} 0, & \text{if } p < v(x), \\ v(x) - |\{A: x \in A \in L, (v(x) \leqslant) |A| \leqslant p\}|, & \text{if } p \geqslant v(x); \end{cases}$$

thus

$$(14) \qquad t(x, L) = \sum_{p=1}^{\infty} \left[\frac{1}{p} - \frac{1}{p+1}\right] c(p, x).$$

Let

$$(15) \qquad c_p = \sum_{x\in} c(p, x).$$

The convergence of these sums follows from

$$\text{(16)} \qquad \infty > \sum_{x\in D} t(x, L) = \sum_{x\in D} \sum_{p=1}^{\infty} \left(\frac{1}{p} - \frac{1}{p+1}\right) c(p, x)$$

$$= \sum_{p=1}^{\infty} \left(\frac{1}{p} - \frac{1}{p+1}\right)\left(\sum_{x\in D} c(p, x)\right)$$

$$= \sum_{p=1}^{\infty} \left(\frac{1}{p} - \frac{1}{p+1}\right) c_p .$$

So clearly by (15) and (16)

(H) each c_p is a natural number and

$$\sum_{p=1}^{\infty} \left(\frac{1}{p} - \frac{1}{p+1}\right) c_p$$

converges.

Also

(I) $\lim_{p\to\infty} c_p = \infty.$

As $x_m \in D$, $x_m \in Y_n \in S - L$, we have $c(p, x_m) \geqslant 1$ if $v(x_m) \leqslant p$. Therefore, $c_p \geqslant |\{m: v(x_m) \leqslant p\}| \to \infty$.

We next prove that for any p

(*) $L_p = \{A \in L: |A| = p\}$ is finite.

Call $x \in D$ *balanced* if $x \in A \in S \Rightarrow |A| = v(x)$, $A \in L$ and let $A^* = \{x \in D: v(x) \leqslant p, x \text{ not balanced}\}$. For $x \in A^*$ we have

$$w(x, L) = \sum_{x\in A\in L} \frac{1}{|A|} \leqslant \frac{v(x) - 1}{v(x)} + \frac{1}{v(x) + 1},$$

$$t(x, L) \geqslant \frac{1}{v(x)(v(x) + 1)} \geqslant \frac{1}{p(p + 1)}.$$

Therefore A^* is finite by (F). Hence there are m_0, m_1 so that $A^* \subset D_{m_0}$ and $A \in K^{m_1}$ whenever $A \in L_p$ and $A \cap A^* \neq 0$. To prove (*) it will be enough to prove that $L_p \subseteq K^{m_1}$. Suppose this is false. Thus there is a least integer $q > m_1$ such that $L_p' = (K^q - K^{m_1}) \cap L_p \neq \varnothing$. By the definition of A^*, if $x \in A \in L_p'$, then x is balanced and $A \in K^q - K^{q-1}$.

Notice that $A \in L_p$, $B \in L - L_p$ implies $A \cap B \subseteq A^* \subseteq D_{m_0}$ and $x_m \in \bigcup \{A: A \in L_p\}$ implies $x_m \in A^*$. Considering any transversal F of K^q

avoiding B_n, it covers $D(K^q) - B_n$, hence its restriction F_1 to $K^q - L_p'$ covers $D(K^{m_1}) - B_n$, as K^{m_1} is B_n-critical. Also F_1 covers $D(K^q - L_p') - D(K^{m_1}) - B_n$, so F_1 is a transversal of $K^q - L_p'$ which covers $D(K^q - L_p') - B_n$. It follows that $K^q - L_p'$ is B_n-critical. By the minimality of $D(K^q)$ it follows that L_p' is empty and this is a contradiction.

We now prove that

$$\text{(J)} \qquad c_p \equiv c_{p+1} \pmod{p+1}.$$

All the congruences in the following will be modulo $p + 1$. By (13) we have

$$c(p, x) - c(p+1, x) = \begin{cases} 0, & \text{if } \ v(x) > p+1, \\ -c(p+1, x), & \text{if } \ v(x) = p+1, \\ |\{A: x \in A \in L, \ |A| = p+1\}|, & \text{if } \ v(x) < p+1. \end{cases}$$

Therefore, since $c(p+1, x) \equiv -|\{A: x \in A \in L, v(x) \leqslant |A| \leqslant p+1\}|$ for $v(x) = p+1$, it follows that

$$c(p, x) - c(p+1, x) \equiv |\{A: x \in A \in L, |A| = p+1\}|.$$

Now

$$\begin{aligned} c_p - c_{p+1} &= \sum_{x \in D} (c(p, x) - c(p+1, x)) \\ &\equiv \sum_{\substack{x \in D \\ v(x) \leqslant p+1}} |\{A: x \in A \in L, |A| = p+1\}| \\ &\equiv \sum_{\substack{A \in L \\ |A| = p+1}} |\{x: x \in A\}| \\ &\equiv 0, \end{aligned}$$

since the summation contains only finitely many terms by (*). So (J) was proved.

It remains to prove only

LEMMA 1. *If $c_p, p \geqslant 1$ is a sequence of natural numbers,* $\lim_{p \to \infty} c_p = \infty$, *and $c_p = c_{p+1}$ (mod $p+1$), then*

$$\sum_{p=1}^{\infty} \left(\frac{1}{p} - \frac{1}{p+1}\right) c_p = \infty.$$

Proof of Lemma 1. We shall assume it converges and get a contradiction.

Sh:29

Choose ϵ, $0 < \epsilon < 0.1$. Define

(17) $$C_q{}^r = \sum_{p=q}^{r-1} \left(\frac{1}{p} - \frac{1}{p+1}\right) c_p \quad (r \text{ may be } \infty).$$

Notice

(K) If $p(k)$ is an increasing sequence $k < l$ then

$$C_{p(k)}^{p(l)} = \sum_{j=k}^{l-1} C_{p(j)}^{p(j+1)}.$$

Choose n_0 such that $\epsilon > C_{n_0}^{\infty}$. Let

(18) $$c_p = c_{p+1} - l_{(p+1)}(p+1),$$

(19) $$l(p, q) = \sum_{i=p+1}^{q} l(i).$$

So

$$c_p\left(\frac{1}{p} - \frac{1}{p+1}\right) = \frac{c_p}{p} - \frac{c_p}{p+1} = \frac{c_p}{p} - \frac{c_{p+1} - l_{(p+1)}(p+1)}{p+1}$$
$$= \frac{c_p}{p} - \frac{c_{p+1}}{p+1} + l_{(p+1)}.$$

(20) $$C_p{}^q = \sum_{i=p}^{q-1} c_p\left(\frac{1}{p} - \frac{1}{p+1}\right)$$
$$= \frac{c_p}{p} - \frac{c_q}{q} + \sum_{i=p}^{q-1} l_{i+1}$$
$$= \frac{c_p}{p} - \frac{c_q}{q} + l(p, q).$$

If there is q_0 such that, for $q \geqslant q_0$, $c_q \geqslant \epsilon q$ then

$$\sum_{p=q_0}^{\infty} c_p\left(\frac{1}{p} - \frac{1}{p+1}\right) \geqslant \epsilon \sum_{p=q_0}^{\infty} p\left(\frac{1}{p} - \frac{1}{p+1}\right) = \epsilon \sum_{p=q_0}^{\infty} \frac{1}{p+1} = \infty,$$

a contradiction. So

(L) for arbitrarily large q, $c_q < \epsilon q$.

Choose $q_0 \geqslant n_0$, $c_{q_0} < \epsilon q_0$. Now for any $r \geqslant q_0$ by (20)

$$C_{q_0}^{r} = \frac{c_{q_0}}{q_0} - \frac{c_r}{r} + l(q_0, r)$$

or

$$\frac{c_r}{r} - l(q_0 r) = \frac{c_{q_0}}{q_0} - C^r_{q_0}.$$

$$(21) \qquad | c_r - rl(q_0, r)| = r \left| \frac{c_{q_0}}{q_0} - C^r_{q_0} \right| \leqslant r \left(\frac{\epsilon q_0}{q_0} + \epsilon \right) = 2\epsilon r,$$

as c_r, r, and $l(q_0, r)$ are integers, $c_r \geqslant 0$, and $2\epsilon < 1$ necessarily

(M) $\quad l(q_0, r) \geqslant 0$, for $r \geqslant q_0$.

Define

$$(22) \qquad k(q_0, r) = \sum_{i=q_0}^{r-1} l(q_0, i).$$

By (M)

(N) $\quad k(q_0, r) \geqslant 0$, for $r \geqslant q_0$.

We now prove by induction on r that

$$(23) \qquad c_r = c_{q_0} + rl(q_0, r) - k(q_0, r), \qquad \text{for} \quad r \geqslant q_0.$$

For $r = q_0$, $l(q_0, r) = 0$ and $k(q_0, r) = 0$ so (23) holds. Assume it holds for r.

$$\begin{aligned} c_{r+1} &= c_r + l_{r+1}(r+1) \\ &= c_{q_0} + rl(q_0, r) - k(q_0, r) + (r+1)\, l_{r+1} \\ &= c_{q_0} + (r+1)\, l(q_0, r) + (r+1)\, l_{r+1} - k(q_0, r) - l(q_0, r) \\ &= c_{q_0} + (r+1)[l(q_0, r) + l_{r+1}] - [k(q_0, r) + l(q_0, r)] \\ &= c_{q_0} + (r+1)\, l(q_0, r+1) - k(q_0, r+1). \end{aligned}$$

So (23) holds.

From (23) and (N) we can deduce $c_r \leqslant c_{q_0} + rl(q_0, r)$. But we noticed in (L) that, for arbitrarily large r, $c_r \leqslant \epsilon r$. Hence by (21) $l(q_0, r) = 0$, hence $c_r \leqslant c_{q_0}$, contradicting $\lim_{p\to\infty} c_p = \infty$.

Remark. Instead of $\lim_{p\to\infty} c_p = \infty$, it suffices to demand that c_p is not eventually constant.

Acknowledgments

I would like to thank Eric Milner for pointing out a gap in the proof; which was filled by adding the demand of minimality in the definition of the K_m, and the proof that L_p is finite.

I would like also to thank the referee for detecting many errors and suggesting improvements in the presentation.

References

1. M. Hall, Jr., Distinct representatives of subsets, *Bull. Amer. Math. Soc.* **54** (1948), 922–926.
2. P. Hall, On representatives of subsets, *J. London Math. Soc.* **10** (1935), 26–30.
3. L. Mirsky, "Transversal Theory," Academic Press, New York/London, 1971.
4. E. Milner and S. Shelah, in preparation.
5. R. A. Brualdi and E. B. Scrimger, Exchange systems, matching and transversals, *J. Combinatorial Theory* **5** (1968), 224–257.
6. J. Folkman, Transversal of infinite families with finitely many infinite members, Rand Corporation, Memorandum, RM-5676-PR, 1968.