COMBINATORICA 4 (1) (1984) 39—45

ON DISJOINTLY REPRESENTABLE SETS

Peter FRANKL and János PACH

Dedicated to Paul Erdős on his seventieth birthday

Received 21 April 1983

A system of sets $E_1, E_2, \ldots, E_k \subset X$ is said to be disjointly representable if there exist $x_1, x_2, \ldots, x_k \in X$ such that $x_i \in E_j \Leftrightarrow i=j$. Let $f(r, k)$ denote the maximal size of an r-uniform set-system containing no k disjointly representable members. In the first section the exact value of $f(r, 3)$ is determined and (asymptotically sharp) bounds on $f(r, k)$, $k>3$ are established. The last two sections contain some generalizations, in particular we prove an analogue of Sauer's theorem [16] for uniform set-systems.

1. Set-systems without k disjointly representable members

We start by recalling two well-known and frequently used definitions in combinatorial set theory. Given an arbitrary set X, $E_1, E_2, \ldots, E_k \subseteqq X$ are said

(a) to form a *Δ-system*, if there exists a subset $K \subseteqq X$ such that $E_i \cap E_j = K$ for every pair (i, j), $1 \leqq i < j \leqq k$ (Cf. [7].)

(b) to have a *system* $\{x_1, x_2, \ldots, x_k\}$ *of distinct representatives*, if all x_i are different and $x_i \in E_i$ for every i, $1 \leqq i \leqq k$. (See e.g. [1].)

In the present note we introduce the following intermediate notion, establishing a natural link between the above two definitions.

Definition 1. *The sets $E_1, E_2, \ldots, E_k \subseteqq X$ are said to be disjointly representable if there exist $x_1, x_2, \ldots, x_k \in X$ such that*

$$x_i \in E_j \Leftrightarrow i = j \quad (1 \leqq i,\ j \leqq k),$$

in other words, no E_i is contained in the union of the others.

If $E_1, E_2, \ldots, E_k$ form a Δ-system, then they are disjointly representable and, consequently, have a system of distinct representatives.

A classical result of Erdős and Rado [7] states that any r-uniform set-system consisting of at least $r!k^r$ sets contains a k member Δ-system. Our next theorem shows that this bound can be essentially improved, if, instead of a Δ-system, the existence of k *disjointly representable* sets is to be guaranteed. Let $T(n, r, k)$, $r<k<n$, denote the Turán number, i.e. the largest integer T such that there exists an r-uniform set-

AMS subject classification (1980): 05 C 35; 05 C 65, 05 A 05

system with T members on an n element set X, which does not contain all r-tuples of a k element subset of X.

Theorem 1. *Let $f(r,k)=\max|\mathcal{H}|$, where the maximum is taken over all r-uniform set-systems $\mathcal{H}$ containing no k disjointly representable members. Then we have*

(i) $$T(r+k-1,k-1,k)\leqq f(r,k)\leqq\binom{r+k-1}{k-1},\quad (r\geqq 1,\ k\geqq 2),$$

(ii) $$f(r,3)=\left\lfloor\frac{r+2}{2}\right\rfloor\cdot\left\lceil\frac{r+2}{2}\right\rceil,\quad (r\geqq 1),$$

and (up to isomorphism) the only r-uniform set-system $\mathcal{H}_{r,3}$ without 3 disjointly representable members, satisfying $|\mathcal{H}_{r,3}|=f(r,3)$, can be constructed as follows. Let A and B be disjoint sets with $|A|=\left\lfloor\frac{r+2}{2}\right\rfloor$ and $|B|=\left\lceil\frac{r+2}{2}\right\rceil$, and let $\mathcal{H}_{r,3}$ $:=\{E\subset A\cup B\,|\,|E|=r$ and $|A\setminus E|=|B\setminus E|=1\}$.

(iii) $$f(2,k)=\binom{k+1}{2}-\left\lceil\frac{k+1}{2}\right\rceil,\quad (k\geqq 2),$$

and the only graph $\mathcal{H}_{2,k}$ with $|\mathcal{H}_{2,k}|=f(2,k)$, which contains no k disjointly representable edges, can be obtained from the complete graph K_{k+1} by deleting $\left\lceil\frac{k+1}{2}\right\rceil$ edges as disjoint as possible.

Proof. (i): Take any $(k-1)$-uniform set-system $\mathcal{F}$ on an $r+k-1$ element set X, which does not contain all $(k-1)$-tuples of any k element subset of X. Then $\mathcal{H}$ $:=\{X\setminus F\,|\,F\in\mathcal{F}\}$ is obviously an r-uniform set-system without k disjointly representable members. This proves the lower bound in (i).

Let $\mathcal{H}=\{E_1,E_2,\dots,E_m\}$ be an arbitrary r-uniform setsystem containing no k disjointly representable sets. For each i $(1\leqq i\leqq m)$, choose a minimal subset $F_i\subseteq\left(\bigcup_{j=1}^{m}E_j\right)\setminus E_i$ so that $F_i\cap E_j\neq\emptyset$ for all $j\neq i$. Observe that $|F_i|\leqq k-1$ $(1\leqq i\leqq m)$, otherwise, by minimality, F_i would be a representative system of $|F_i|\geqq k$ disjointly representable members of $\mathcal{H}$. The upper bound is now an immediate consequence of the following lemma due to Bollobás [2] (see also Jaeger—Payan [10], Katona [12], Lovász [14]):

Lemma 1. *Let $\mathcal{H}=\{E_1,E_2,\dots,E_m\}$, $\mathcal{H}'=\{F_1,F_2,\dots,F_m\}$ be two set-systems satisfying $|E_i|\leqq r$ and $|F_i|\leqq k-1$ for all i, $1\leqq i\leqq m$. Suppose further that $E_i\cap F_j$ $=\emptyset\Leftrightarrow i=j$. Then $m\leqq\binom{r+k-1}{k-1}$.* ∎

(ii): We proceed by induction on r. In case $r=1$ the assertion trivially holds. Assume now $r>1$, and let $\mathcal{H}=\{E_1,E_2,\dots,E_m\}$, $\mathcal{H}'=\{F_1,F_2,\dots,F_m\}$ denote the same set-systems as in the previous section (with $k=3$). Suppose first that $|F_i|<2$ for some i. Then, applying the induction hypothesis to the system $\{E_j\setminus F_i\,|\,1\leqq j\leqq m,$ $j\neq i\}$, we obtain $m\leqq f(r-1,3)+1<\left\lfloor\frac{r+2}{2}\right\rfloor\cdot\left\lceil\frac{r+2}{2}\right\rceil$ and we are done.

Hence, we may assume $|F_i|=2$ for all i. In other words, $\mathscr{H}'$ is a simple graph. Set $\mathscr{H}_1':=\{F_i\in\mathscr{H}'|F_i\subseteq E_1\}$ and $\mathscr{H}_2':=\mathscr{H}'\setminus\mathscr{H}_1'$. The following three properties can readily be checked by the definitions, using the fact that $\mathscr{H}$ does not contain 3 disjointly representable sets:

(a) $\mathscr{H}'$ is a triangle-free graph;

(b) $E_1\cap F_i\neq\emptyset$ for every i, $1<i\leq m$;

(c) $|\{F_i\colon x\in F_i\in\mathscr{H}_2'|F_i\in x\}|\leq 1$ for every $x\in E_1$.

From here, applying the Turán theorem [18] to $\mathscr{H}_1'$, it follows that

$$m=|\mathscr{H}'|=|\mathscr{H}_1'|+|\mathscr{H}_2'|\leq\left\lfloor\frac{r}{2}\right\rfloor\cdot\left\lceil\frac{r}{2}\right\rceil+r+1=\left\lfloor\frac{r+2}{2}\right\rfloor\cdot\left\lceil\frac{r+2}{2}\right\rceil,$$

as required. The verification of the uniqueness of the extremal construction described in the theorem is left to the reader.

(iii): Let $\mathscr{H}$ be an arbitrary graph on an n-element vertex set. Observe that k edges of $\mathscr{H}$ are disjointly representable if and only if they form a number of vertex-disjoint stars. Thus, the following two properties are equivalent:

(a) $\mathscr{H}$ does not contain k disjointly representable edges;

(b) every set of $n-k$ vertices has a common neighbour in $\overline{\mathscr{H}}$ (i.e. in the complement of graph $\mathscr{H}$).

Hence, (iii) follows from the next result due to P. Erdős and L. Moser [6]:

Lemma 2. *Given two natural numbers t, n $(t<n)$, let G be a graph of n vertices with the property that any t of them have a common neighbour. Then G has at least $(n-t+1)\cdot(t-1)+\binom{t-1}{2}+\left\lceil\frac{n-t+1}{2}\right\rceil$ edges, and equality holds here if and only if G has $t-1$ vertices joined to every other vertex, and $\left\lceil\frac{n-t+1}{2}\right\rceil$ further edges as disjoint as possible.* ∎

Remarks 1. The gap between the lower and upper bounds in part (i) of Theorem 1 is not too big, if $r\gg k$ are large enough. It is known (see e.g. [13]) that, for any fixed $k\geq 2$, $\lim\limits_{r\to\infty}\frac{T(r+k-1,k-1,k)}{\binom{r+k-1}{k-1}}$ is a positive constant c_k, and $\lim\limits_{k\to\infty}c_k=1$ (cf. [17]).

2. Parts (ii) and (iii) of the theorem assert $f(r,3)=T(r+2,2,3)$ and $f(2,k)=T(k+1,k-1,k)$, resp. That is, in these cases the lower bound given by (i) is sharp. One might be tempted to think that this is also true in general, i.e. $f(r,k)=T(r+k-1,k-1,k)$ for all $r\geq 1$, $k\geq 2$.

3. A set-system $\mathscr{H}$ is said to be of *rank* r, if $|E|\leq r$ for every $E\in\mathscr{H}$. Let $f(\leq r,k)=\max|\mathscr{H}|$, where the maximum is taken over all set-systems $\mathscr{H}$ of rank r, containing no k disjointly representable members. Then we obviously have $f(r,k)<f(\leq r,k)$. E.g., in case $k=3$ one can prove $f(\leq r,3)=\binom{r+2}{2}$, and an optimal construction is the following: $\mathscr{H}:=\{E_{ij}|1\leq i<j\leq r+2\}$, where $E_{ij}:=\{k|i<k<j$ or $j<k\leq r+2\}$.

2. Disjointly t-representable sets

Hajnal [11] has suggested the following modification of the above problem: What happens if, instead of k disjointly representable sets, we look for k sets such that each can be represented by more than one point?

Definition 2. *The sets $E_1, E_2, \dots, E_k \subseteq X$ are said to be disjointly t-representable if one can choose disjoint t-element subsets $X_i \subseteq E_i$ $(1 \leq i \leq k)$ such that $X_i \cap E_j = \emptyset$ whenever $i \neq j$.*

An apparent difficulty is that, given any $r \geq t > 1$, one can construct arbitrarily large r-uniform set-systems without 2 disjointly t-representable members. (E.g. a large Δ-system with a kernel of size at least $r-t+1$ will do.) However, the following analogue of Theorem 1 is valid.

Theorem 2. *Given any four natural numbers t, r, k, l $(r \geq t, k \geq 3, l \geq 2)$, let $f_t^l(r, k) = \max |\mathcal{H}|$, where the maximum is taken over all r-uniform set-systems $\mathcal{H}$ having the following two properties:*

(i) *$\mathcal{H}$ does not contain an l member Δ-system whose kernel is of size $> r-t$;*

(ii) *$\mathcal{H}$ does not contain k disjointly t-representable members.*

Then we have

$$c(k, l, t) r^{(k-1)t} \leq f_t^l(r, k) \leq c'(k, l, t) r^{kt-1},$$

where c and c' are positive constants independent of r.

Proof. The lower bound can be established by the following construction. Given an $(r+kt-1)$-element set X, we can select a system $\mathcal{F}$ of $(kt-1)$-tuples of X so that

(a) $|F_1 \cap F_2| < (k-1)t$ for every pair $F_1, F_2 \in \mathcal{F}$;

(b) $|\mathcal{F}| \geq \frac{1}{2} \binom{r+kt-1}{(k-1)t} \Big/ \binom{kt-1}{(k-1)t}$, (see Rödl [15]).

Choose a partition $X_1 \cup X_2 \cup \dots \cup X_{kt-1}$ of X so as to maximize the cardinality of the set-system

$$\mathcal{F}' := \{F \in \mathcal{F} \,|\, |F \cap X_i| = 1,\ i = 1, 2, \dots, kt-1\}.$$

We may assume (cf. [3], [4]) that

$$|\mathcal{F}'| \geq \frac{(kt-1)!}{(kt-1)^{kt-1}} |\mathcal{F}| \geq c(k, t) r^{(k-1)t}.$$

Observe, finally, that $\mathcal{H} = \{X - F \,|\, F \in \mathcal{F}'\}$ is an r-uniform set-system meeting both requirements (i) and (ii), (with $l=2$).

Next we prove the upper bound. Let $\mathcal{H}$ be a set-system satisfying (i) and (ii). Take $k-1$ distinct members $E_1, E_2, \dots, E_{k-1} \in \mathcal{H}$ such that $E_1 \cup E_2 \cup \dots \cup E_{k-1} = E^*$ is as large as possible. Then $|E \cap E^*| > r-t$ for every $E \in \mathcal{H}$. For if not, then either $E, E_1, E_2, \dots, E_{k-1}$ would be disjointly t-representable or replacing some E_i by E we would get a contradiction with the maximum property of E^*.

On the other hand, the above cited result of Erdős and Rado [7] yields that any $(r-t+1)$-element subset of E^* can be contained only in at most $(t-1)!\,l^{t-1}$

members of $\mathscr{H}$, otherwise (i) would be violated. Thus $|\mathscr{H}| \leqq (t-1)!\, t^{t-1} |\mathscr{H}^*|$, where $\mathscr{H}^*$ is the trace of $\mathscr{H}$ on E^*, i.e.

$$\mathscr{H}^* := \mathscr{H}|_{E^*} = \{E \cap E^* | E \in \mathscr{H}\}.$$

We need the following well-known result of Sauer [16] (see also [8], [9]):

Lemma 3. *Let $\mathscr{F}$ be a system of subsets of an n element set X. Then either $|\mathscr{F}| \leqq \sum_{i=0}^{K-1} \binom{n}{i}$, or there exists a K element subset $Y \subseteq X$ such that for every $Y' \subseteq Y$ one can find $F \in \mathscr{F}$ with $F \cap Y = Y'$.* ∎

Applying this to $\mathscr{H}^*$ (with $n = |E^*| \leqq (k-1)r$, $K = kt$) and noticing that, by condition (ii) in the theorem, the latter possibility is excluded, we obtain the desired upper bound on $|\mathscr{H}^*|$ (and hence on $|\mathscr{H}|$).

3. Traces of set-systems

Throughout this section let $\binom{k}{\geqq i}$, $\binom{k}{\leqq i}$ and $\binom{k}{= i}$ stand for the set-systems consisting of all *at least i-*, *at most i-* and *exactly i*-element subsets of a k-element set, resp.

Given two set-systems $\mathscr{H}$ and $\mathscr{F}$, $\mathscr{F}$ is called a *trace* of $\mathscr{H}$ (in short, $\mathscr{H} \to \mathscr{F}$), if $\mathscr{F}$ is isomorphic to a subsystem of

$$\mathscr{H}|_Y := \{E \cap Y | E \in \mathscr{H}\},$$

the restriction of $\mathscr{H}$ to a subset Y.

The problem investigated in the first section of this paper belongs to the following general pattern: Determine

$$m(n, r, \mathscr{F}) := \max |\mathscr{H}|,$$

where the maximum is taken over all r-uniform set-systems $\mathscr{H}$ on an n element ground-set, with the property $\mathscr{H} \not\to \mathscr{F}$.

Using these terms (and the notation in Theorem 1), we have $m\left(n, r, \binom{k}{=1}\right) = f(r, k)$, provided n is large enough.

Further, Lemma 3 (Sauer's theorem) immediately implies $m\left(n, r, \binom{k}{\geqq 0}\right) \leqq \sum_{i=0}^{k-1} \binom{n}{i}$. Next we show that the first $k-1$ terms of this sum can be omitted. The proof uses a simple linear algebraic approach developed in [9].

Theorem 3. *Let n, r, k be natural numbers ($n \geqq r \geqq k$), and let $\mathscr{H}$ be an r-uniform set-system on an n element ground-set X. Then either $|\mathscr{H}| \leqq \binom{n}{k-1}$, or $\mathscr{H} \to \binom{k}{\geqq 0}$.*

Proof. Let $Y_1, Y_2, \dots, Y_{\binom{n}{k-1}}$ be an enumeration of all $(k-1)$-tuples of X, and let

$E_1, E_2, \ldots, E_{|\mathscr{H}|}$ denote the members of $\mathscr{H}$. Define an $|\mathscr{H}|\times\binom{n}{k-1}$ matrix $A=(a_{ij})$, as follows.

$$a_{ij}:=\begin{cases}1 & \text{if } E_i \supseteq Y_j \\ 0 & \text{if } E_i \not\supseteq Y_j\end{cases}; \quad \left(1 \leqq i \leqq |\mathscr{H}|,\ 1 \leqq j \leqq \binom{n}{k-1}\right).$$

Assume $|\mathscr{H}|>\binom{n}{k-1}$. Then the rows of A cannot be linearly independent. That is, one can choose suitable reals $\alpha_1, \alpha_2, \ldots, \alpha_{|\mathscr{H}|}$ (not all of them being zero) such that

$$\sum_{i=1}^{|\mathscr{H}|} \alpha_i a_{ij} = 0 \quad \left(1 \leqq j \leqq \binom{n}{k-1}\right).$$

Let Y be a minimal subset of X with $\sum_{E_i \supseteq Y} \alpha_i \neq 0$. We clearly have $|Y| \geqq k$, and for every $Z \subseteqq Y$

$$\sum_{E_i \cap Y = Z} \alpha_i = (-1)^{|Y-Z|} \sum_{E_i \supseteq Y} \alpha_i \neq 0$$

holds, as can be shown by an easy backwards induction on $|Z|$. This yields, in particular, that for every $Z \subseteqq Y$ there exists $E_i \in \mathscr{H}$ with $E_i \cap Y = Z$. ∎

Corollary. *Given natural numbers n, k $(n \geqq k)$, we have*

$$(*) \qquad \binom{n-1}{k-1} \leqq m\left(n, k, \binom{k}{\geqq i}\right) \leqq \binom{n}{k-1} \quad \textit{for every} \quad i<k. \ ∎$$

Remarks 4. The upper bound in (*) cannot be improved without any further restriction on n and k. (Set $n=2k-1$.)

On the other hand, the lower bound can also be attained. If $i=k-1$, for instance, then the proof of Theorem 1 (i) applied to the system $\overline{\mathscr{H}}:=\{X-E \mid E \in \mathscr{H}\}$ shows that

$$m\left(n, k, \binom{k}{\geqq k-1}\right) = \binom{n-1}{k-1}.$$

5. One can easily prove that

$$(**) \qquad T(n-1, k-1, k) \leqq m\left(n, k, \binom{k}{= k-1}\right) \leqq T(n, k-1, k).$$

As a matter of fact, we suspect that in both (*) and (**) the lower bound is attained if n is sufficiently large. In particular, for $i=0$, this would yield an interesting generalization of the Erdős—Ko—Rado theorem [5].

Acknowledgement. The authors would like to express their gratitude to Paul Erdős, András Hajnal and Norbert Sauer for their valuable comments.

References

[1] C. Berge, *Graphs and Hypergraphs,* North-Holland, 1973.

[2] B. Bollobás, On generalized graphs, *Acta Math. Hung.* **16** (1965), 447—452.

[3] P. Erdős, On bipartite subgraphs of a graph (in Hungarian), *Matematikai Lapok* **18** (1967), 283—288.

[4] P. Erdős and D. J. Kleitman, On coloring graphs to maximize the proportion of multicolored k-edges, *J. Combinatorial Th.* **5** (1968), 164—169.

[5] P. Erdős, Chao Ko and R. Rado, Intersection theorems for systems of finite sets, *Quart. J. Math. Oxford (2),* **12** (1961), 313—320.

[6] P. Erdős and L. Moser, An extremal problem in graph theory, *J. Austral. Math. Soc.* **11** (1970), 42—47.

[7] P. Erdős and R. Rado, Intersection theorems for systems of sets, *J. London Math. Soc.* **35** (1960), 85—90.

[8] P. Frankl, On the trace of finite sets, *J. Combinatorial Th. Ser. A,* **34** (1983), 41—45.

[9] P. Frankl and J. Pach, On the number of sets in a null-t-design, *European J. Comb.* **4** (1983), 21—33.

[10] F. Jaeger and C. Payan, Determination du nombre d'aretes d'un hypergraphe τ-critique, *C. R. Acad. Sc. Paris* **273** (1971), 221—223.

[11] A. Hajnal, Personal communication.

[12] G. O. H. Katona, Solution of a problem of A. Ehrenfeucht and J. Mycielsky, *J. Comb. Th.* **17** (1974), 265—266.

[13] G. Katona, T. Nemetz and M. Simonovits, On a graph problem of Turán (in Hungarian), *Matematikai Lapok* **15** (1964), 228—238.

[14] L. Lovász, Topological and algebraic methods in graph theory, in: *Graph Theory and Related topics* (J. A. Bondy and U. S. R. Murty, eds.), Academic Press, New York 1979, 1—14.

[15] V. Rödl, Almost Steiner systems always exist, *to appear in European J. of Comb.*

[16] N. Sauer, On the density of families of sets, *J. Combinatorial Th. Ser. A,* **13** (1972), 145—147.

[17] A. F. Sidorenko, On the Turán number $T(n, 5, 4)$ and on the number of monochromatic 4-cliques in a two-coloured 3-graph (in Russian), *Voprosi Kibernetiki, Komb. Anal. i Teoria Grafov,* Nauka, Moscow 1980, 117—124.

[18] P. Turán, On the theory of graphs, *Coll. Math.* **3** (1954), 19—30.

P. Frankl

CNRS
15 Quai Anatole France,
75007 Paris, France

J. Pach

Mathematical Institute of the
Hungarian Academy of Sciences
Budapest, H-1395, Pf. 428, Hungary