A THEOREM OF FINITE SETS

by

G. KATONA

Mathematical Institute of the Hungarian Academy of Sciences
Budapest, Hungary

§ 1. Introduction

Let $A_1, \ldots, A_n$ be a system of different subsets of a finite set H, where $|H| = h$ and $|A_i| = l$ $(1 \leq i \leq n)$ ($|A|$ denotes the number of elements of A). We ask for a system $A_1, \ldots, A_n$ (for given h, l, n) for which the number of sets B satisfying $|B| = l - 1$ and $B \subset A_i$ for some i is minimum. The first lower estimation for this minimum is given by SPERNER ([1], Hilfssatz). His estimation is $\dfrac{n \cdot l}{h - l + 1}$. This depends on h. However, if $n = \binom{N}{l}$, it is expected that the minimizing system is the system of all l-tuples chosen from a subset of N elements of H. In this case the number of B's is $\binom{N}{l-1}$ which does not depend on h. A. HAJNAL proved this statement in the case of $l = 3$ (unpublished). In this paper I prove for all cases that this is, indeed, the minimum, and find the (more complicated) minimum also for arbitrary n. The theorem is probably useful in proofs by induction over the maximal number of elements of the subsets in a system, as was SPERNER's lemma in his paper [1].

KLEITMAN told me in Tihany (Hungary) that he thought I could solve the following problem of ERDŐS by the aid of the above theorem and the "marriage problem": Let $A_1, \ldots, A_n$ be subsets of H, where $|H| = 2h$ and $|A_i| = h$. For what n's is it always possible to construct a system $B_1, \ldots, B_n$ with the properties $B_i \subset A_i$, $|B_i| = h - 1$ $(1 \leq i \leq n)$. § 3 contains the solution of this problem in a more general form.

§ 2. The main result

Before the exact formulation of the theorem we need the following simple but interesting

LEMMA 1. *If n and l are natural numbers, we can write the number n uniquely in the form*

$$n = \binom{a_l(n, l)}{l} + \binom{a_{l-1}(n, l)}{l-1} + \cdots + \binom{a_{t(n,l)}(n, l)}{t(n, l)}, \tag{1}$$

where $t(n, l) \geq 1$, $a_l > a_{l-1} > \ldots > a_{t(n,l)}$ are natural numbers and $a_i(n, l) \geq i$ $(i = t(n, l), t(n, l) + 1, \ldots, l)$.

Proof. The existence of form (1) is proved by induction over l. For $l = 1$ the statement is trivial. Assume that for $l = k - 1$ it is true also and prove for $l = k$. Let a_k be the maximal integer satisfying the inequality $\binom{a_k}{k} \leq n$. If here equality holds, we are ready. If it does not, using the induction hypothesis we have for the number $n - \binom{a_k}{k}$ the following expression:

$$n - \binom{a_k}{k} = \binom{a_{k-1}}{k-1} + \ldots + \binom{a_t}{t}, \tag{2}$$

where $t \geq 1$, $a_{k-1} > \ldots > a_t$, $a_i \geq i$ $(i = t, t+1, \ldots, k-1)$. (2) gives an expression for n, we have to verify only $a_k > a_{k-1}$ and $a_k \geq k$. If $a_k \leq \leq a_{k-1}$ held, then

$$n \geq \binom{a_k}{k} + \binom{a_{k-1}}{k-1} \geq \binom{a_k}{k} + \binom{a_k}{k-1} = \binom{a_k+1}{k}$$

would hold also, which contradicts choosing of a_k. On the other hand, $a_k \geq k$ follows from $a_k > a_{k-1}$ and $a_{k-1} \geq k - 1$.

The unicity of Form (1) is proved also by induction over l. For $l = 1$ the statement is trivial. Assume that for $l = k - 1$ it is also true and prove for $l = k$. If, on the contrary, there exist two forms:

$$n = \binom{a_k}{k} + \binom{a_{k-1}}{k-1} + \ldots + \binom{a_t}{t} = \binom{a'_k}{k} + \binom{a'_{k-1}}{k-1} + \ldots + \binom{a'_r}{r}, \tag{3}$$

we may separate two different cases. If $a_k = a'_k$, we can obtain two different forms of $n - \binom{a_k}{k}$, which contradict our induction hypothesis. If $a_k < a'_k$, the contradiction follows from

$$n \leq \binom{a_k}{k} + \binom{a_k - 1}{k-1} + \ldots + \binom{a_k - k + 1}{1} = \binom{a_k+1}{k} - 1 < \binom{a_k+1}{k} \leq$$

$$\leq \binom{a'_k}{k} \leq \binom{a'_k}{k} + \ldots + \binom{a'_r}{r}.$$

Thus we proved the lemma.

In the future we will use the following two notations:

$$E_l(n) = \binom{a_l(n,l) - 1}{l-1} + \binom{a_{l-1}(n,l) - 1}{l-2} + \ldots + \binom{a_{t(n,l)}(n,l) - 1}{t(n,l) - 1}$$

and

$$F_l(n) = \binom{a_l(n,l)}{l-1} + \binom{a_{l-1}(n,l)}{l-2} + \ldots + \binom{a_{t(n,l)}(n,l)}{t(n,l) - 1}.$$

These numbers are uniquely determined by Lemma 1.

Let us consider now the problem. Let H be a finite set with h elements, and

$$\mathcal{A} = \{A_1, \dots, A_n\}$$

a system of different subsets of H, where the number of elements of A_i is

$$|A_i| = l \qquad (1 \leq i \leq n).$$

Obviously, l is a fixed integer between 1 and h. Let $c(\mathcal{A})$ denote the following system

$$c(\mathcal{A}) = \{B : |B| = l-1 \text{ and } B \subset A_j \text{ for at least one } j\}.$$

The problem is to determine the minimum of $|c(\mathcal{A})|$, if h, n and l are given. Theorem 1 gives the exact solution of this problem.

THEOREM 1. *Let h, n and l be given integers with the properties*

$$h \geq 1, \qquad 1 \leq l \leq h \quad \text{and} \quad 1 \leq n \leq \binom{h}{l}.$$

If H is a set of h elements, and

$$\mathcal{A} = \{A_1, \dots, A_n\}, \quad |A_i| = l \qquad (i = 1, \dots, n)$$

a system of different subsets of H, then

$$\min |c(\mathcal{A})| = F_l(n),$$

where the minimum runs over all such systems $\mathcal{A}$.

REMARK. It is interesting, that $\min|c(\mathcal{A})|$ does not depend on h. For example, SPERNER's estimation [1]:

$$c(\mathcal{A}) \geq \frac{n \cdot l}{h - l + 1}$$

depends on h.

Before the proof we shall give another theorem. We will prove them together.

THEOREM 2. *Let h, n and l be given integers with the properties*

$$h \geq 1, \qquad 1 \leq l \leq h \quad \text{and} \quad \binom{h}{l} \leq n \leq 2\binom{h}{l}.$$

Further G and H are disjoint sets of h elements. If

$$\mathcal{A} = \{A_1, \dots, A_n\}$$

is a system of A_i 's, where

$$A_i \subset G \quad \text{or} \quad A_i \subset H \qquad (1 \leq i \leq n)$$

and

$$|A_i| = l \qquad (1 \leq i \leq n),$$

then

$$\min |c(\mathcal{A})| = \binom{h}{l-1} + F_l\left(n - \binom{h}{l}\right).$$

PROOF. 1. First we construct the minimizing system of Theorem 1. Denote this system by $\mathscr{M}(h, n, l)$. Obviously, it is sufficient to construct the system $\mathscr{M}(a_l^*(n), n, l)$, where $a_l^*(n)$ is the least integer satisfying

$$\binom{a_l^*(n)}{l} \geq n .$$

The construction will be carried out by induction over l. If $l = 1$, $a_1^*(n) = n$ and $\mathscr{M}(a_1^*(n), n, 1)$ consists of all the sets of one element. Assume we constructed already the system $\mathscr{M}(a_{l-1}^*(n), n, l-1)$ for all n. Construct now $\mathscr{M}(a_l^*(n), n, l)$. If $n = \binom{a_l(n, l)}{l}$, then the minimizing system consists of all the subsets having l elements. If $n > \binom{a_l(n, l)}{l}$, let H be a set of $a_l^*(n) = a_l(n, l)+1$ elements, and e an element of H. Since $a_l > a_{l-1}$, we can construct the system $\mathscr{M}\left(a_l(n, l), n - \binom{a_l(n, l)}{l}, l-1\right)$ on $H - \{e\}$ by the induction hypothesis. Define the system $\mathscr{N}$ in the following manner:

$$\mathscr{N} = \left\{N \cup \{e\} : N \in \mathscr{M}\left(a_l(n, l), n - \binom{a_l(n, l)}{l}, l-1\right)\right\}.$$

If $\mathscr{P}$ denotes the system of all subsets of $H - \{e\}$, having l elements, then $\mathscr{P}$ and $\mathscr{N}$ form together the system $\mathscr{M}(a_l^*(n), n, l)$. Indeed, the number of sets is $\binom{a_l(n, l)}{l} + n - \binom{a_l(n, l)}{l} = n$ and we have only to verify

$$(4) \qquad |c(\mathscr{M}(a_l^*(n), n, l))| = \binom{a_l(n, l)}{l-1} + \ldots + \binom{a_{t(n,l)}(n, l)}{t(n, l)-1} = F_l(n) .$$

However, it is easy to see, that

$$|c(\mathscr{M}(a_l^*(n), n, l))| = \binom{a_l(n, l)}{l-1} + \left|c\left(\mathscr{M}\left(a_l(n, l), n - \binom{a_l(n, l)}{l}, l-1\right)\right)\right|$$

and by the induction hypothesis

$$\left|c\left(\mathscr{M}\left(a_l(n, l), n - \binom{a_l(n, l)}{l}, l-1\right)\right)\right| = \binom{a_{l-1}(n, l)}{l-2} + \ldots + \binom{a_{t(n,l)}(n, l)}{t(n, l)-1},$$

which proves (4).

2. The minimizing system of Theorem 2 consists of a complete system in G, and $\mathscr{M}\left(h, n - \binom{h}{l}, l\right)$ in H.

3. In the previous two points we showed that in the case of Theorem 1

$$\min |c(\mathscr{A})| \leq F_l(n) ,$$

and in the case of Theorem 2

$$\min |c(\mathscr{A})| \leq \binom{h}{l-1} + F_l\left(n - \binom{h}{l}\right).$$

Thus, it is sufficient to verify

$$|c(\mathscr{A})| \geq F_l(n) \tag{5}$$

and

$$|c(\mathscr{A})| \geq \binom{h}{l-1} + F_l\left(n - \binom{h}{l}\right), \tag{6}$$

respectively. These statements will be proved by induction over l. If $l = 1$, both statements are trivial. Assume we have proved for all numbers $< l$ and prove for l.

4. First we prove the inequality

$$F_l(n) \leq F_l(n_1) + F_{l-1}(n_2), \tag{7}$$

if

$$n = n_1 + n_2, \qquad n_1 \geq 0, \quad n_2 \geq 0 \tag{8}$$

are integers, and

$$n_2 \leq E_l(n). \tag{9}$$

The statement will be proved for fixed l and for every n, n_1, n_2 using the induction hypothesis for $l - 1$. For the sake of simplicity we use the following notations:

$$\begin{array}{llll} t = t(n, l) & a_i = a_i(n, l) & (t \leq i \leq l) & a_l^* = a_l^*(n), \\ r = t(n_1, l) & b_i = a_i(n_1, l) & (r \leq i \leq l) & b_l^* = a_l^*(n_1), \\ s = t(n_2, l-1) & c_i = a_i(n_2, l-1) & (s \leq i \leq l-1) & c_{l-1}^* = a_{l-1}^*(n_2). \end{array}$$

It follows from (8) and (9) that

$$n_1 \geq n - E_l(n) = \binom{a_l - 1}{l} + \ldots + \binom{a_t - 1}{t}. \tag{10}$$

Because of (10)

$$b_l \geq a_l - 1 \tag{11}$$

must hold, since in the contrary case it would be

$$n_1 \leq \binom{a_l - 2}{l} + \binom{a_l - 3}{l-1} + \ldots + \binom{a_l - (l+1)}{1} = \binom{a_l - 1}{l} - 1,$$

what contradicts (10). On the other hand

$$a_l \geq b_l \tag{12}$$

because of (8). Applying (11) and (12) we can distinguish two different cases: (a) $b_l = a_l$ and (b) $b_l = a_l - 1$.

(a) In this case (7) has the form

$$\binom{a_l}{l-1} + \binom{a_{l-1}}{l-2} + \ldots + \binom{a_t}{t-1} \leq$$
$$\leq \binom{a_l}{l-1} + \binom{b_{l-1}}{l-2} + \ldots + \binom{b_r}{r-1} + F_{l-1}(n_2)\,.$$

Decreasing both sides by $\binom{a_l}{l-1}$ we have

$$F_{l-1}\left(n - \binom{a_l}{l}\right) \leq F_{l-1}\left(n_1 - \binom{a_l}{l}\right) + F_{l-1}(n_2)\,. \tag{13}$$

Let H_1 and H_2 be disjoint sets. Construct the system $\mathscr{M}\left(b_{l-1}+1,\, n_1 - \binom{a_l}{l},\right.$ $\left. l-1\right)$ on H_1 and the system $\mathscr{M}(c^*_{l-1}, n_2, l-1)$ on H_2. In this manner we obtain a system $\mathscr{N}$ on $H_1 \cup H_2$. Applying the induction hypothesis (Point **3**. (5)) for $\mathscr{N}$ and $l-1$ we have

$$F_{l-1}\left(n - \binom{a_l}{l}\right) \leq |c(\mathscr{N})| =$$
(14)
$$= \left| c\left(\mathscr{M}\left(b_{l-1}+1,\, n_1 - \binom{a_l}{l},\, l-1\right)\right)\right| + |c(\mathscr{M}(c^*_{l-1}, n_2, l-1))|\,.$$

However, we know (Point **1**. (4)) that

$$\left| c\left(\mathscr{M}\left(b_{l-1}+1,\, n_1 - \binom{a_l}{l},\, l-1\right)\right)\right| = F_{l-1}\left(n_1 - \binom{a_l}{l}\right) \tag{15}$$

and

$$|c(\mathscr{M}(c^*_{l-1}, n_2, l-1))| = F_{l-1}(n_2)\,. \tag{16}$$

Finally, (13) follows from (14), (15), and (16).

(b) $b_l = a_l - 1$. We separate this case into two subcases:

$$\text{(ba) } n_2 \geq \binom{a_l - 1}{l-1}, \qquad \text{(bb) } n_2 < \binom{a_l-1}{l-1}.$$

(ba) In this case (7) has the form

$$\binom{a_l}{l-1}+\binom{a_{l-1}}{l-2}+\ldots+\binom{a_t}{t-1}\leq\binom{a_l-1}{l-1}+\binom{b_{l-1}}{l-2}+\ldots+\binom{b_r}{r-1}+$$
$$+\binom{a_l-1}{l-2}+\binom{c_{l-2}}{l-3}+\ldots+\binom{c_s}{s-1},$$

since $c_{l-1}=a_l-1$, because of (9) and the supposition (ba). Decreasing both sides by $\binom{a_l}{l-1}=\binom{a_l-1}{l-1}+\binom{a_l-1}{l-2}$ we have

$$F_{l-1}\left(n-\binom{a_l}{l}\right)\leq F_{l-1}\left(n_1-\binom{a_l-1}{l}\right)+F_{l-2}\left(n_2-\binom{a_l-1}{l-1}\right). \tag{17}$$

We can prove (17) by using of the induction hypothesis if

$$n_2-\binom{a_l-1}{l-1}\leq E_{l-1}\left(n-\binom{a_l}{l}\right) \tag{18}$$

holds. However (9) gives

$$\binom{a_l-1}{l-1}+\binom{c_{l-2}}{l-2}+\ldots+\binom{c_s}{s}\leq\binom{a_l-1}{l-1}+ \tag{19}$$
$$+\binom{a_{l-1}-1}{l-2}+\ldots+\binom{a_t-1}{t-1}.$$

Decreasing both sides by $\binom{a_l-1}{l-1}$ we obtain

$$\binom{c_{l-2}}{l-2}+\ldots+\binom{c_s}{s}\leq\binom{a_{l-1}-1}{l-2}+\ldots+\binom{a_t-1}{t-1} \tag{20}$$

and (20) is equivalent to (18).

(bb) In this case (7) has the form

$$\binom{a_l}{l-1}+\binom{a_{l-1}}{l-2}+\ldots+\binom{a_t}{t-1}\leq\binom{a_l-1}{l-1}+$$
$$+\binom{b_{l-1}}{l-2}+\ldots+\binom{b_r}{r-1}+F_{l-1}(n_2).$$

Decreasing both sides by $\binom{a_l-1}{l-1}$ we have

$$\binom{a_l-1}{l-2}+F_{l-1}\left(n-\binom{a_l}{l}\right)\leq F_{l-1}\left(n_1-\binom{a_l-1}{l}\right)+F_{l-1}(n_2). \tag{21}$$

Let G and H be two disjoint sets of $a_l - 1$ elements. Construct the system $\mathcal{M}\left(a_l - 1, n_1 - \binom{a_l - 1}{l}, l - 1\right)$. We can it construct if $n_1 - \binom{a_l - 1}{l} \leq$ $\leq \binom{a_l - 1}{l - 1}$. But this follows from $a_l - 1 = b_l > b_{l-1}$, since

$$n_1 - \binom{a_l - 1}{l} = \binom{b_{l-1}}{l - 1} + \ldots + \binom{b_r}{r}.$$

Construct further the system $\mathcal{M}(c^*_{l-1}, n_2, l - 1)$ on H. The possibility of this construction follows from the assumption (bb). In this manner we obtain a system $\mathcal{N}$ on $G \cup H$. Applying the induction hypothesis (Point **3.** (6)) for $\mathcal{N}$ and $l - 1$ we have

$$\binom{a_l - 1}{l - 2} + F_{l-1}\left(n - \binom{a_l}{l}\right) \leq \left| c\left(\mathcal{M}\left(a_l - 1, n_1 - \binom{a_l - 1}{l}, l - 1\right)\right)\right| +$$

$$+ \left| c\left(\mathcal{M}(c^*_{l-1}, n_2, l - 1)\right)\right| . \tag{22}$$

However, we know (Point **1.** (4)) that

$$\left| c\left(\mathcal{M}\left(a_l - 1, n_1 - \binom{a_l - 1}{l}, l - 1\right)\right)\right| = F_{l-1}\left(n_1 - \binom{a_l - 1}{l}\right) \tag{23}$$

and

$$\left| c\left(\mathcal{M}(c^*_{l-1}, n_2, l - 1)\right)\right| = F_{l-1}(n_2) , \tag{24}$$

further, (21) follows from (22), (23) and (24). Thus we proved the inequality for l.

5. However, we need (7) under the condition

$$n_2 \leq \frac{n \cdot l}{a^*_l} \tag{25}$$

instead of (9). Thus we are going now to prove the inequality

$$\frac{n \cdot l}{a^*_l} \leq E_l(n) . \tag{26}$$

We prove (26) by induction over l, but we should like to mention that the proof of (26) is independent from the whole proof of the theorems. For $l = 1$ the statement is trivial. Assume we proved it for the integers $< l$, and prove for l. If $n = \binom{a_l}{l}$ then $a^*_l = a_l$ and $E_l(n) = \binom{a_l - 1}{l - 1}$, thus (26) holds with equality. We may assume $a^*_l = a_l + 1$. Obviously

$$\frac{l}{a_l + 1} \binom{a_l}{l} \leq \binom{a_l - 1}{l - 1} \tag{27}$$

and by the induction hypothesis

(28) $$\frac{l-1}{a_{l-1}+1}\left[\binom{a_{l-1}}{l-1}+\ldots+\binom{a_r}{r}\right]\le\binom{a_{l-1}-1}{l-2}+\ldots+\binom{a_r-1}{r-1}.$$

If $\frac{l}{a_l+1}\le\frac{l-1}{a_{l-1}+1}$, summarizing (27) and (28) we obtain (26). In the contrary case

(29) $$\frac{l}{a_l+1}>\frac{l-1}{a_l}$$

holds because of $a_l\ge a_{l-1}+1$. Let us set out from the identity

$$\binom{a_l}{l-1}\left(\frac{l}{a_l+1}-\frac{l-1}{a_l}\right)=\binom{a_l-1}{l-1}-\frac{l}{a_l+1}\binom{a_l}{l}.$$

The expression in the bracket is positive because of (29), thus we can write

$$\left[\binom{a_{l-1}}{l-1}+\binom{a_{l-2}}{l-2}+\ldots+\binom{a_r}{r}\right]\left(\frac{l}{a_l+1}-\frac{l-1}{a_l}\right)<\binom{a_l-1}{l-1}-\frac{l}{a_l+1}\binom{a_l}{l},$$

since $a_l>a_{l-1}$. Write $\frac{l-1}{a_{l-1}+1}$ instead of $\frac{l-1}{a_l}$, and reorder the inequality

$$\frac{l}{a_l+1}\left[\binom{a_l}{l}+\binom{a_{l-1}}{l-1}+\ldots+\binom{a_r}{r}\right]<\binom{a_l-1}{l-1}+$$
$$+\frac{l-1}{a_{l-1}+1}\left[\binom{a_{l-1}}{l-1}+\ldots+\binom{a_r}{r}\right].$$

Finally, from the above inequality (26) follows by (28).

6. Now let us prove statement (5) for l by induction over h if $h=l$ is trivial. Assume we have proved (5) for all sets $|H|<h$, and prove for h. There exists an element e of H, contained by at most $\frac{n\cdot l}{h}$ sets A_i. We define the following systems:

$$\mathscr{B}=\{A:A\in\mathscr{A},\ e\notin A\}$$

and

$$\mathscr{C}=\{A-\{e\}:A\in\mathscr{A},\ e\in\mathscr{A}\}$$

where

(30) $$n_2=|\mathscr{C}|\le\frac{n\cdot l}{h}\le\frac{n\cdot l}{a_l^*(n)}.$$

Naturally,

$$c(\mathscr{B})\subset c(\mathscr{A})$$

and

$$c(\mathscr{C})\,(\cup)\,e\subset c(\mathscr{A}),$$

where $\mathcal{D}(\cup)a$ denotes in general the system $\{D \cup \{a\} : D \in \mathcal{D}\}$. Thus the inequality

(31) $$|c(\mathcal{A})| \geq |c(\mathcal{B})| + |c(\mathcal{C})|$$

holds. However, $\mathcal{B}$ is a system in $H - \{e\}$, we may apply the induction hypothesis for $h - 1$

(32) $$|c(\mathcal{B})| \geq F_l(n - n_2)\,.$$

Further, applying the induction hypothesis for $l - 1$ we obtain

(33) $$|c(\mathcal{C})| \geq F_{l-1}(n_2)\,.$$

It follows from (31), (32) and (33) that

(34) $$F_l(n - n_2) + F_{l-1}(n_2) \leq |c(\mathcal{A})|\,.$$

Using the result of Point **5**, inequality (5) follows from (34) and (7) by (30), since (7) is proved already for l.

7. Now prove statement (6) for l by induction over h. If $h = l$, it is trivial. Assume we have proved (6) for all sets $|G| = |H| < h$, and prove for h. The proof will be similar to the proof of the previous point.

Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be given by

$$\mathcal{A}_1 = \{A : A \in \mathcal{A},\, A \subset G\}\,,$$

and

$$\mathcal{A}_2 = \{A : A \in \mathcal{A},\, A \subset H\}\,.$$

If $|\mathcal{A}_1| = r$ and $|\mathcal{A}_2| = s$, there are two elements $e \in G$ and $f \in H$, such that e is contained by at most $\frac{r \cdot l}{h}$, and f is contained by at most $\frac{s \cdot l}{h}$ sets A_i. Define the following systems:

$$\mathcal{B} = \{A : A \in \mathcal{A},\, e \notin A,\, f \notin A\}\,,$$

$$\mathcal{C}_1 = \{A - \{e\} : A \in \mathcal{A}_1,\, e \in A\}$$

and

$$\mathcal{C}_2 = \{A - \{f\} : A \in \mathcal{A}_2,\, f \in A\}\,,$$

where

(35) $$r_2 = |\mathcal{C}_1| \leq \frac{r \cdot l}{h}$$

and

(36) $$s_2 = |\mathcal{C}_2| \leq \frac{s \cdot l}{h}\,.$$

Naturally,

$$c(\mathcal{B}) \subset c(\mathcal{A})\,,$$

$$c(\mathcal{C}_1)\,(\cup)\,e \subset c(\mathcal{A})$$

and

$$c(\mathcal{C}_2)\,(\cup)\,f \subset c(\mathcal{A})\,.$$

Thus the inequality

$$|c(\mathscr{A})| \geq |c(\mathscr{B})| + |c(\mathscr{C}_1)| + |c(\mathscr{C}_2)| = |c(\mathscr{B})| + |c(\mathscr{C}_1 \cup \mathscr{C}_2)| \tag{37}$$

holds. However $\mathscr{B}$ is a system in $G \cup H - \{e\} - \{f\}$, we may apply our induction hypothesis for $h - 1$:

$$|c(\mathscr{B})| \geq \binom{h-1}{l-1} + F_l\left(n - r_2 - s_2 - \binom{h-1}{l}\right). \tag{38}$$

Further, applying the induction hypothesis for $l - 1$ we obtain

$$|c(\mathscr{C}_1 \cup \mathscr{C}_2)| \geq \binom{h-1}{l-2} + F_{l-1}\left(r_2 + s_2 - \binom{h-1}{l-1}\right). \tag{39}$$

It follows from (37), (38) and (39) that

$$\binom{h}{l-1} + F_l\left(n - r_2 - s_2 - \binom{h-1}{l}\right) + F_{l-1}\left(r_2 + s_2 - \binom{h-1}{l-1}\right) \leq |c(\mathscr{A})|. \tag{40}$$

Now we should like to use inequality (7) which is valid under condition (25) (Point **5**). For this reason we have to verify only

$$r_2 + s_2 - \binom{h-1}{l-1} \leq \frac{\left[n - r_2 - s_2 - \binom{h-1}{l} + r_2 + s_2 - \binom{h-1}{l-1}\right] l}{a_l^*\left(n - \binom{h}{l}\right)} = \frac{\left[n - \binom{h}{l}\right] \cdot l}{a_l^*\left(n - \binom{h}{l}\right)}. \tag{41}$$

However

$$r_2 + s_2 - \binom{h-1}{l-1} \leq \frac{\left[r + s - \binom{h}{l}\right] \cdot l}{h} = \frac{\left[n - \binom{h}{l}\right] \cdot l}{h} \tag{42}$$

is an immediate consequence of (35) and (36). Since $n \leq 2\binom{h}{l}$ is a condition of Theorem 2, $a_l^*\left(n - \binom{h}{l}\right) \leq h$ holds and (42) results (41). Thus we can use

(7) for this case:

$$(43)\quad F_l\left(n-\binom{h}{l}\right)\le F_l\left(n-r_2-s_2-\binom{h-1}{l}\right)+F_{l-1}\left(r_2+s_2-\binom{h-1}{l-1}\right).$$

Finally, (40) and (43) gives the desired inequality, and the whole proof is finished.

Now we consider a natural generalization of the problem of Theorem 1. The problem is to determine the minimum of $|c^k(\mathscr{A})|$, where $1\le k\le l$, $c^k(\mathscr{A})=c(c^{k-1}(\mathscr{A}))$ and $c^1(\mathscr{A})=c(\mathscr{A})$. It is not difficult to conjecture what is the result. To the theorem we need the following notation:

$$F_l^k(n)=\binom{a_l(n,l)}{l-k}+\binom{a_{l-1}(n,l)}{l-1-k}+\dots+\binom{a_{t(n,l)}(n,l)}{t(n,l)-k}\qquad(1\le k\le l),$$

where $\binom{a}{b}=0$ if $b<0$.

THEOREM 3. *Let h, n, l and k be given integers with the properties*

$$h\ge 1,\ 1\le k\le l\le h\quad\text{and}\quad 1\le n\le\binom{h}{l}.$$

If H is a set of h elements and

$$\mathscr{A}=\{A_1,\dots,A_n\},\qquad |A_i|=l\qquad(i=1,\dots,n)$$

a system of different subsets of H, then

$$\min|c^k(\mathscr{A})|=F_l^k(n),$$

where the minimum runs over all such systems $\mathscr{A}$.

PROOF. It is easy to see by induction over l, that $|c^k(\mathscr{M}(h,n,l))|=F_l^k(n)$. Thus, we have to prove only

$$(44)\qquad |c^k(\mathscr{A})|\ge F_l^k(n).$$

This will be proved by induction over k. For $k=1$ Theorem 3 gives Theorem 1. Assume now (44) is true for values smaller than k, and prove for k. Obviously,

$$c^k(\mathscr{A})=c(c^{k-1}(\mathscr{A}))$$

holds and using the induction hypothesis and Theorem 1 we obtain

$$(45)\qquad |c^k(\mathscr{A})|\ge F_{l-(k-1)}(F_l^{k-1}(n)).$$

(a) If $t(n,l)-(k-1)>0$, then

$$F_l^{k-1}(n)=\binom{a_l(n,l)}{l-(k-1)}+\dots+\binom{a_{t(n,l)}(n,l)}{t(n,l)-(k-1)}$$

is an expression of type (1). That is

$$(46)\qquad \begin{aligned} &t\left(F_l^{k-1}(n), l-k+1\right) = t(n, l) - k + 1 \\ &a_i\left(F_l^{k-1}(n), l-k+1\right) = a_{i+k-1}(n, l) \quad \left(t(n, l) - k + 1 \le i \le l - k + 1\right) \end{aligned}$$

and

$$(47)\qquad \begin{aligned} F_{l-k+1}\left(F_l^{k-1}(n)\right) &= \sum_{i=t(n,l)-k+1}^{l-k+1} \binom{a_i\left(F_l^{k-1}(n), l-k+1\right)}{i-1} = \\ &= \sum_{i=t(n,l)-k+1}^{l-k+1} \binom{a_{i+k-1}(n, l)}{i-1} = \sum_{j=t(n,l)}^{l} \binom{a_j(n, l)}{j-k} = F_l^k(n), \end{aligned}$$

which proves (44) and (45).

(b) If $t(n, l) - k + 1 \le 0$, then (46) does not hold. However in this case

$$F_l^{k-1}(n) - 1 = \binom{a_l(n, l)}{l-k+1} + \ldots + \binom{a_k(n, l)}{1}$$

and

$$\begin{aligned} &t\left(F_l^{k-1}(n) - 1, l-k+1)\right) = 1 \\ &a_i\left(F_l^{k-1}(n) - 1, l-k+1\right) = a_{i+k-1}(n, l) \quad (1 \le i \le l-k+1) \end{aligned}$$

hold. Further, the equation

$$(48)\qquad \begin{aligned} F_{l-k+1}\left(F_l^{k-1}(n) - 1\right) &= \sum_{i=1}^{l-k+1} \binom{a_i\left(F_l^{k-1}(n) - 1, l-k+1\right)}{i-1} = \\ &= \sum_{i=1}^{l-k+1} \binom{a_{i+k-1}(n, l)}{i-1} = \sum_{j=k}^{l} \binom{a_j(n, l)}{j-k} = \sum_{j=t(n,l)}^{l} \binom{a_j(n)}{j-k} = F_l^k(n) \end{aligned}$$

is true in this case instead of (47). If we prove

$$(49)\qquad F_{l-k+1}\left(F_l^{k-1}(n)\right) = F_{l-k+1}\left(F_l^{k-1}(n) - 1\right),$$

then (44) follows from (45), (49) and (48). (49) will be proved by the following simple lemma.

Lemma 2. *If* $t(m, r) = 1$, *then*

$$F_r(m+1) = F_r(m).$$

Proof. Let s be the least index such that $a_s(m, r) > a_{s-1}(m, r) + 1$ $(2 \le s \le r)$. If there is not such s, let s be equal to $r + 1$. Thus, we can write

$$\begin{aligned} m = \binom{a_r(m, r)}{r} + \ldots + \binom{a_{s-1}(m, r)}{s-1} + \binom{a_{s-1}(m, r) - 1}{s-2} + \ldots + \\ + \ldots + \binom{a_{s-1}(m, r) - (s-2)}{1} \end{aligned}$$

and

$$m+1=\binom{a_r(m,r)}{r}+\ldots+\binom{a_{s-1}(m,r)+1}{s-1}.$$

Now it is not difficult to see, that

$$F_r(m)=\binom{a_r(m,r)}{r-1}+\ldots+\binom{a_{s-1}(m,r)}{s-2}+\binom{a_{s-1}(m,r)-1}{s-3}+\ldots+$$

$$+\ldots+\binom{a_{s-1}(m,r)-(s-2)}{0}=\binom{a_r(m,r)}{r-1}+\ldots+\binom{a_{s-1}(m,r)+1}{s-2}=$$

$$=F_r(m+1),$$

which proves the lemma and Theorem 3.

§ 3. Solution of an Erdős-problem

Let H be a finite set of h elements, and $\mathscr{A}$ a system of subsets of H:

$$\mathscr{A}=\{A_1, A_2, \ldots, A_n\},\ A_i\subset H,\ |A_i|=l \qquad (1\leq i\leq n).$$

Erdős proposed the following problem. For which numbers n can we construct a system $\mathscr{B}$ with the properties

$$\mathscr{B}=\{B_1, B_2, \ldots, B_n\},\ B_i\subset A_i,\ |B_i|=l-k \qquad (1\leq i\leq n).$$

In the solution we use the well-known marriage problem. It is clear in this connection, that it is a very important question, in which cases does $F_l^k(n)<n$, $F_l^k(n)=n$ or $F_l^k(n)>n$ hold. The following sequence of lemmas deals with this problem.

Lemma 3. *If $1\leq k\leq l$ and x are positive integers, then*

$$f(x)=\binom{x}{l-k}-\binom{x}{l}$$

is a monotone increasing function between l and $2l-k-2$ but it is a monotone decreasing function from $2l-k-1$. The values $f(2l-k-2)$ and $f(2l-k-1)$ are equal.

Proof. Let $0\leq a<b\leq x$ be integers. It is easy to see that $\binom{x}{a}-\binom{x}{b}<0$, $\binom{x}{a}-\binom{x}{b}=0$ and $\binom{x}{a}-\binom{x}{b}>0$, respectively, if $a+b<x$, $a+b=x$ and $a+b>x$, respectively.

Consider the difference $f(x+1) - f(x) = \binom{x}{l-k-1} - \binom{x}{l-1}$. Using the above remark we obtain that

$$f(x+1) - f(x) < 0 \quad \text{if} \quad 2l - k - 2 < x,$$

$$f(x+1) - f(x) = 0 \quad \text{if} \quad 2l - k - 2 = x,$$

and finally,

$$f(x+1) - f(x) > 0 \quad \text{if} \quad 2l - k - 2 > x.$$

This completes the proof.

The following two lemmas are immediate consequences of Lemma 3.

LEMMA 3a. *If $1 \leq k \leq l$ and x are positive integers, then*

$$\binom{x}{l-k} - \binom{x}{l} \leq \binom{2l-k-1}{l-k} - \binom{2l-k-1}{l}.$$

LEMMA 3b. *If $1 \leq k \leq l$ and $x \geq 2l - k + 1$ are positive integers, then*

$$\binom{x}{l-k} - \binom{x}{l} \leq \binom{2l-k+1}{l-k} - \binom{2l-k+1}{l}.$$

LEMMA 4. *If $1 \leq k \leq m$, then*

$$\sum_{i=k}^{m-1} \left[\binom{2i-k-1}{i-k} - \binom{2i-k-1}{i} \right] \leq \binom{2m-k-1}{m-k} - \binom{2m-k-1}{m}.$$

PROOF. Let a and b be positive integers, where $\frac{a}{2} \leq b \leq a - 1$. Then

$$\binom{a}{b} - \binom{a}{b+1} = \binom{a}{b}\left[1 - \frac{a-b}{b+1}\right] = \binom{a}{b}\left[\frac{2b-a+1}{b+1}\right], \tag{50}$$

and similarly

$$\binom{a+2}{b+1} - \binom{a+2}{b+2} = \binom{a+2}{b+1}\left[1 - \frac{a-b+1}{b+2}\right] = \binom{a+2}{b+1}\left[\frac{2b-a+1}{b+2}\right]. \tag{51}$$

Further

$$\binom{a+2}{b+1}\left[\frac{2b-a+1}{b+2}\right] = \binom{a}{b}\left[\frac{2b-a+1}{b+1}\right] \cdot \left[\frac{(a+2)(a+1)}{(b+2)(a-b+1)}\right], \tag{52}$$

where

$$\frac{a+1}{b+2} \geq 1,$$

and

$$\frac{a+2}{a-b+1} \geq \frac{a+2}{\frac{a}{2}+1} = 2.$$

That is

$$\binom{a+2}{b+1} - \binom{a+2}{b+2} \geq 2\left[\binom{a}{b} - \binom{a}{b+1}\right] \tag{53}$$

follows from (50), (51) and (52).

Applying (53) for $a = 2i = k - 1$, and $b = i - 1$, $(1 \leq k \leq i)$, we obtain

$$\binom{2(i+1)-k-1}{i} - \binom{2(i+1)-k-1}{i+1} \geq 2\left[\binom{2i-k-1}{i-1} - \binom{2i-k-1}{i}\right],$$

or

$$\binom{2(i+1)-k-1}{i+1-k} - \binom{2(i+1)-k-1}{i+1} \geq 2\left[\binom{2i-k-1}{i-k} - \binom{2i-k-1}{i}\right]. \tag{54}$$

Prove now the lemma by induction over m. If $m = k$, the statement is trivial. Let the lemma be true for m and prove it for $m + 1$.

$$\sum_{i=k}^{m}\left[\binom{2i-k-1}{i-k} - \binom{2i-k-1}{i}\right] = \sum_{i=k}^{m-1}\left[\binom{2i-k-1}{i-k} - \binom{2i-k-1}{i}\right] + \\ + \binom{2m-k-1}{m-k} - \binom{2m-k-1}{m}$$

and by induction hypothesis and (54)

$$\sum_{i=k}^{m}\left[\binom{2i-k-1}{i-k} - \binom{2i-k-1}{i}\right] \leq 2\left[\binom{2m-k-1}{m-k} - \binom{2m-k-1}{m}\right] \leq \\ \leq \binom{2(m+1)-k-1}{m+1-k} - \binom{2(m+1)-k-1}{m+1}$$

holds, which proves Lemma 4.

LEMMA 5. *If* $1 \leq k \leq l$ *and* $2l - k < a_l(n, l)$ *then*

$$F_l^k(n) < n. \tag{55}$$

PROOF. We may use Lemma 3b:

$$\binom{a_l(n,l)}{l-k} - \binom{a_l(n,l)}{l} \leq \binom{2l-k+1}{l-k} - \binom{2l-k+1}{l}. \tag{56}$$

On the other hand, by Lemma 3a

$$\binom{a_i(n,l)}{i-k} - \binom{a_i(n,l)}{i} \leq \binom{2i-k-1}{i-k} - \binom{2i-k-1}{i} \qquad (k \leq i \leq l-1)$$

holds and summarizing it we obtain

$$\sum_{i=k}^{l-1}\left[\binom{a_i(n,l)}{i-k}-\binom{a_i(n,l)}{i}\right]\leq\sum_{i=k}^{l-1}\left[\binom{2i-k-1}{i-k}-\binom{2i-k-1}{i}\right].$$

Applying now Lemma 4 and (54):

$$\sum_{i=k}^{l-1}\left[\binom{a_i(n,l)}{i-k}-\binom{a_i(n,l)}{i}\right]\leq\binom{2l-k-1}{l-k}-\binom{2l-k-1}{l}<$$

$$<\binom{2(l+1)-k-1}{l+1-k}-\binom{2(l+1)-k-1}{l+1}.$$

Obviously,

$$\sum_{i=1}^{l-1}\left[\binom{a_i(n,l)}{i-k}-\binom{a_i(n,l)}{i}\right]<\binom{2(l+1)-k-1}{l+1-k}-\binom{2(l+1)-k-1}{l+1} \tag{57}$$

also holds, since we added a nonpositive number to the left side. If we sum (56) and (57) the obtained inequality

$$F_l^k(n)-n<\binom{2l-k+1}{l-k}-\binom{2l-k+1}{l}+\binom{2(l+1)-k-1}{l+1-k}-$$

$$-\binom{2(l+1)-k-1}{l+1}=0$$

results (55).

Lemma 6. *If $1\leq k\leq l$ and $2l-k>a_l(n,l)$ then*

$$F_l^k(n)>n. \tag{58}$$

Proof. We know that

$$a_i(n,l)\leq a_l(n,l)-(l-i). \tag{59}$$

If $l>i\geq a_l(n,l)-(l-k)$, then $a_l(n,l)-(l-i)\leq 2i-k$ and by (59)

$$a_i(n,l)\leq 2i-k$$

holds. In this case, obviously

$$\binom{a_i(n,l)}{i-k}-\binom{a_i(n,l)}{i}\geq 0 \tag{60}$$

follows. If $k\leq i<a_l(n,l)-(l-k)$, then by (59) and Lemma 3

$$\binom{a_i(n,l)}{i-k}-\binom{a_i(n,l)}{i}\geq\binom{a_l(n,l)-(l-i)}{i-k}-\binom{a_l(n,l)-(l-i)}{i} \tag{61}$$

holds, but it is trivially true for $i<k$, too.

Sum (60) and (61)

$$\sum_{i=1}^{l-1}\left[\binom{a_i(n,l)}{i-k}-\binom{a_i(n,l)}{i}\right]\geq\sum_{i=1}^{a_l(n,l)-l+k-1}\left[\binom{a_l(n,l)-(l-i)}{i-k}-\right.$$

$$\left.-\binom{a_l(n,l)-(l-i)}{i}\right]=\binom{2a_l(n,l)-2l+k}{a_l(n,l)-l-1}-\binom{2a_l(n,l)-2l+k}{a_l(n,l)-l+k-1}+1.$$

That is

$$F_l^k(n)-n\geq\binom{2a_l(n,l)-2l+k}{a_l(n,l)-l-1}-\binom{2a_l(n,l)-2l+k}{a_l(n,l)-l+k-1}+1+ \tag{62}$$

$$+\binom{a_l(n,l)}{l-k}-\binom{a_l(n,l)}{l}$$

s true. Here

$$\binom{2a_l(n,l)-2l+k}{a_l(n,l)-l-1}-\binom{2a_l(n,l)-2l+k}{a_l(n,l)-l+k-1}\geq\binom{a_l(n,l)-1}{a_l(n,l)-l-1}- \tag{63}$$

$$-\binom{a_l(n,l)-1}{a_l(n,l)-l+k-1}$$

because of Lemma 3. However we can write the right hand side of (63) in the form

$$\binom{a_l(n,l)-1}{a_l(n,l)-l-1}-\binom{a_l(n,l)-1}{a_l(n,l)-l+k-1}= \tag{64}$$

$$=\binom{a_l(n,l)-1}{l}-\binom{a_l(n,l)-1}{l-k}.$$

Here

$$\binom{a_l(n,l)-1}{l}-\binom{a_l(n,l)-1}{l-k}=\left[\binom{a_l(n,l)}{l}-\binom{a_l(n,l)}{l-k}\right]-\left[\binom{a_l(n,l)-1}{l-1}-\right.$$

$$\left.-\binom{a_l(n,l)-1}{l-k-1}\right]$$

and since $2l-k-1>a_l(n,\ l)-1$ by supposition of the lemma, thus

$$\binom{a_l(n,l)-1}{l}-\binom{a_l(n,l)-1}{l-k}\geq\binom{a_l(n,l)}{l}-\binom{a_l(n,l)}{l-k}. \tag{65}$$

Finally, (65), (64) (63) and (62) give

$$F_l^k(n)-n\geq 1$$

which proves our lemma.

LEMMA 7. *If*

$$n > \binom{2l-k}{l} + \binom{2(l-1)-k}{l-1} + \cdots + \binom{k}{k}, \tag{66}$$

then

$$F_l^k(n) < n.$$

On the other hand, if

$$n \leq \binom{2l-k}{l} + \binom{2(l-1)-k}{l-1} + \cdots + \binom{k}{k}, \tag{67}$$

then

$$F_l^k(n) \geq n$$

with equality only if

$$n = \binom{2l-k}{l} + \binom{2(l-1)-k}{l-1} + \cdots + \binom{2s-k}{s} \tag{68}$$

for some s $(k \leq s \leq l)$.

PROOF. Consider first the case of (66). If $a_i(n, l) = 2i - k$ $(k \leq i \leq l)$, then $t(n, l) < k$ and

$$n = \binom{2l-k}{l} + \cdots + \binom{k}{k} + \binom{k-1}{k-1} + \cdots + \binom{t(n, l)}{t(n, l)}.$$

Obviously,

$$F_l^k(n) = \binom{2l-k}{l} + \cdots + \binom{k}{k},$$

thus $F_l^k(n) < n$ holds.

In the contrary case

$$a_r(n, l) > 2r - k$$

$$a_i(n, l) = 2i - k \qquad (r < i \leq l)$$

hold for some r $(k \leq r < l)$. Since

$$n - F_l^k(n) = \left[\binom{a_r(n, l)}{r} + \cdots\right] - \left[\binom{a_r(n, l)}{r-k} + \cdots\right],$$

the statement follows by Lemma 5:

$$\binom{a_r(n, l)}{r} + \cdots > F_r^k\left[\binom{a_r(n, l)}{r} + \cdots\right].$$

The case (67) may occur in two different ways.

1. If (68) holds, then obviously $F_l^k(n) = n$
2. For some r $(k < r \leq l)$,

$$a_r(n, l) < 2r - k,$$

and

$$a_i(n, l) - 2i - k \qquad (r < i \leq l).$$

Since

$$n - F_l^k(n) = \left[\binom{a_r(n, l)}{r} + \dots\right] - \left[\binom{a_r(n, l)}{r-k} + \dots\right],$$

the statement follows by Lemma 6:

$$\binom{a_r(n, l)}{r} + \dots < F_r^k\left[\binom{a_r(n, l)}{r} + \dots\right].$$

THEOREM 4. *Let $1 \leq k \leq l \leq h$ be positive integers, H a set of h element and*

$$\mathcal{A} = \{A_1, \dots, A_n\}, \quad |A_i| = l \qquad (1 \leq i \leq n$$

a system of subsets of H. If

$$n \leq \binom{2l-k}{l} + \binom{2(l-1)-k}{l-1} + \dots + \binom{k}{k}, \tag{69}$$

there exists a system

$$\mathcal{B} = \{B_1, \dots, B_n\}, \quad |B_i| = l - k, \quad B_i \subset A_i \quad (1 \leq i \leq n) \tag{70}$$

but in the case of

$$n > \binom{2l-k}{l} + \binom{2(l-1)-k}{l-1} + \dots + \binom{k}{k} \tag{71}$$

not necessarily .

PROOF. First we prove the latter case. If (71) holds then by Lemma 7 $F_l^k(n) < n$. We know (Theorem 1) that there exists a system $\mathcal{A}$ such that $|c^k(\mathcal{A})| = F_l^k(n)$. Thus, a system $\mathcal{B}$ satisfying (70) does not exist.

In the proof of the existence of $\mathcal{B}$ in the case of (69) we use the well-known marriage problem [2]:

THEOREM OF ORE. *Let E and F be disjoint sets and G a graph on $E \cup F$ Assume G has the property that for arbitrary $D \subset E$ there is a set $H \subset F$ such that every element of H is connected with at least one element of D and $|H| \geq |D|$. Then there exists a one-to-one mapping between E and a subset K of F, such that the associating vertices are connected in G.*

In our case $E = \mathcal{A}$, $F = c^k(\mathcal{A})$ and $A \in \mathcal{A}$, $B \in c^k(\mathcal{A})$ are connected if and only if $A \supset B$. Thus, it is sufficient to verify that for every subsystem

$$\mathcal{C} = \{A_{i_1}, \dots, A_{i_m}\} \subset \mathcal{A}$$

there are at least m sets in $c^k(\mathcal{A})$, which are contained in one of $A_{i_j} (1 \leq j \leq \leq m)$. However, $m \leq n$, thus by (69)

$$m \leq \binom{2l-k}{l} + \dots + \binom{k}{k}$$

and Lemma 7 gives

$$F_l^k(m) \geq m. \tag{72}$$

Use now Theorem 1:

$$|c^k(\mathscr{C})| \geq F_l^k(m).$$

This and (72) results $|c^k(\mathscr{C})| \geq m$, which means that our graph has the property prescribed in the used theorem. Applying the theorem the obtained one-to-one mapping gives just the desired system $\mathscr{B}$.

Corollary. If $2l - k \geq h$, then (69) always holds and a system $\mathscr{B}$ satisfying (70) always exists.

This is an immediate consequence of the inequality

$$\binom{h}{l} \leq \binom{2l-k}{l} \leq \binom{2l-k}{l} + \dots + \binom{k}{k}$$

and the fact that $\mathscr{A}$ has at most $\binom{h}{l}$ elements.

REFERENCES

[1] Sperner, E.: Ein Satz über Untermengen einer endlichen Menge, *Math. Z.* **27** (1928) 544—548.
[2] Ore, O.: Graphs and matching theorems, *Duke Math. J.* **22** (1955) 625—639.