Discrete Mathematics 34 (1981) 195–198
North-Holland Publishing Company

"DART CALCULUS" OF INDUCED SUBSETS

Pavel TOMASTA
Mathematical Institute SAV, Obrancovmieru 49, 88625 Bratislava, Czechoslovakia

Received 21 June 1978
Revised 29 May 1980

The paper contains some basic results from "dart calculus" of induced subsets. We obtain as their consequence a negative answer to Hajnal's hypothesis:

$$\left(m, 1+\sum_{i=0}^{n-1}\binom{m}{i}\right) \to \left(m-1, 1+\sum_{i=0}^{n-1}\binom{m-1}{i}\right).$$

Hajnal proposed the symbol $(m, k) \to (p, q)$ to denote the truth of the statement:

Given k distinct subsets $A_1, \ldots, A_k$ of an m-set S (that is, a set of m elements), there exists a p-subset P of S such that the family $\{A_i \cap P\}$, $(1 \le i \le k)$ contains at least q distinct members.

Using this dart notation Sauer's result [3] has the form:

$$\left(m, 1+\sum_{i=0}^{n-1}\binom{m}{i}\right) \to (n, 2^n). \tag{1}$$

By Erdös the density of a hypergraph $G=(S, \mathcal{A})$ is the maximal cardinality of the set $V \subset S$ for which

$$\operatorname{card}\{V \cap \mathcal{A}\} = 2^{\operatorname{card} V}.$$

That is, the density of a hypergraph with $1+\sum_{i=0}^{n-1}\binom{m}{i}$ edges is at least n.

Hajnal suggested a nice possible generalization of (1), namely

$$\text{(HH)} \quad \left(m, 1+\sum_{i=0}^{n-1}\binom{m}{i}\right) \to \left(m-1, 1+\sum_{i=0}^{n-1}\binom{m-1}{i}\right) \quad \text{if } m > n.$$

If true, this need simply be applied $m-n$ times to yield Sauer's result.

Unfortunately, one of the results (Theorem 1) of our paper gives the negative answer to Hajnal's hypothesis (HH).

At first we prove the Principle of Duality of this "dart calculus".

Proposition 1. *Let m, p, q be naturals. Then*

$$(m, p) \to (m-1, q) \quad \textit{iff} \quad (m, 2^m - p) \to (m-1, 2^{m-1} - p + q).$$

Proof. (I) Let $(m, p) \to (m-1, q)$ and $\mathcal{F}$ be an arbitrary family of $2^m - p$ subsets of the m-set S.

By the hypothesis there exists some $(m-1)$-set P such that at most $p-q$ sets of the family $\bar{\mathcal{F}}$ (complement of $\mathcal{F}$) have the property

$$B \in \bar{\mathcal{F}} \quad \text{and} \quad \{B \cup \{x\}\} \in \bar{\mathcal{F}}, \quad \text{where } x \notin P.$$

From this we have that the family $\{\mathcal{F} \cap P\}$ contains at least $2^m-(p-q)$ distinct sets, because either $B \in \mathcal{F}$ or $\{B \cup \{x\}\} \in \mathcal{F}$ besides at most $p-q$ sets. The proof of the first part is finished.

(II) Substitute in the previous 2^m-p for p and $2^{m-1}-(p-q)$ for q and calculate:

From $(m, 2^m-p) \to (m-1, 2^{m-1}-(p-q))$ we have

$$(m, 2^m-(2^m-p)) = (m, p) \to (m-1, 2^{m-1}-(2^m-p-(2^{m-1}-(p-q)))) = (m-1, q)$$

and the proof is completed.

From this Principle of Duality as a consequence one can easily obtain Bondy's result [1] (which he proved by a graph theoretical argument) namely,

$$(m, n) \to (m-1, m).$$

Corollary. $(m, m) \to (m-1, m)$ *iff* $(m, 2^m-m) \to (m-1, 2^{m-1})$.

Proof. On the right hand there is a Sauer's result (1) for $m-1=n$ and the Principle of Duality implies the equivalence.

Define the labeled graph $G_{\mathcal{A}}$ as follows: the vertices of $G_{\mathcal{A}}$ are the subsets $A_1, \ldots, A_n \in \mathcal{A}$ of the m-set S; and A_i is joined to A_j $(i \neq j)$ by an edge labeled x if either $A_i = A_j \cup \{x\}$ or $A_j = A_i \cup \{x\}$.

Let $\mathcal{B}$ be a family of all subsets of the m-set S. Then $G_{\mathcal{B}}$ is an m-dimensional labeled cube C_m and for every $\mathcal{A} \subset \mathcal{B}$, $G_{\mathcal{A}}$ is an induced subgraph of C_m.

Now let $F(m, t)$ be the smallest number of the vertices of a subgraph of C_m, in which every x occurs at least t times. Then the following is true:

Proposition 2. *Let m, t be naturals, t be a power of two and* $\log_2 2t$ *divides m. Then*

$$F(m, t) \leq 1 + \frac{m(2t-1)}{\log_2 2t}.$$

Proof. Decompose the m-set S into $m/\log_2 2t$ disjoint subsets S_i. Each of them generates a subcube with 2t vertices. These subcubes have the empty set as a common vertex in a graph induced by a system of all subsets of the sets S_i, and so

$$F(m, t) \leq 1 + \frac{m(2t-1)}{\log_2 2t},$$

because every x occurs exactly t times in each graph constructed like that. The proof is completed.

If (HH) is true, then for every subgraph $G_{\mathcal{A}}$ of the cube C_m which has $1+\sum_{i=0}^{n-1}\binom{m}{i}$ vertices each x occurs in $G_{\mathcal{A}}$ at most

$$1+\sum_{i=0}^{n-1}\binom{m}{i}-\left[1+\sum_{i=0}^{n-1}\binom{m-1}{i}\right]=\sum_{i=1}^{n-1}\binom{m-1}{i-1} \quad \text{times.}$$

Calculate:

$$F\left(m, 1+\sum_{i=1}^{n-1}\binom{m-1}{i-1}\right)\leq F(m, 2^{r-1})\leq\frac{m(2^r-1)}{r}\sim C\frac{m^{n-1}}{\log_2 m},$$

where r is the smallest natural number for which

$$1+\sum_{i=1}^{n-1}\binom{m-1}{i-1}\leq 2^{r-1},$$

C is some constant and $n\geq 3$ is fixed.

On the other hand $1+\sum_{i=0}^{n-1}\binom{m}{i}\sim C_1 m^{n-1}$ and we prove

Theorem 1. *If $3\leq n\leq m-2$ then for every such n there exists a sufficiently large m_0 such that for every $m\geq m_0$*

$$\left(m, 1+\sum_{i=0}^{n-1}\binom{m}{i}\right)\not\rightarrow\left(m-1, 1+\sum_{i=0}^{n-1}\binom{m-1}{i}\right).$$

Example. Let $n=3$ and $m=15$. Then from (HH) we have:

$$\left(15, 1+1+15+\binom{15}{2}\right)=(15, 122)\rightarrow\left(14, 1+1+14+\binom{14}{2}\right)=(14, 107).$$

However, $F(15, 16)\leq 94$ (Proposition 2). Since $F(15, 16)\leq 122$ there exists a family $\mathcal{A}$ of 122 subsets of 15-set S such that every $x\in S$ occurs in $G_{\mathcal{A}}$ at least 16 times.

That is, there exist at most $122-16=106$ distinct sets in the family $\{\mathcal{A}\cap P\}$, where P is an arbitrary 14-subset of S.

So we obtain a counterexample to (HH) also for small m.

For the following results we will use Burtin's result [2]. He proved that $G_{\mathcal{A}}$ on p vertices has at most $\frac{1}{2}p\log_2 p$ edges. The equality is valid for subcubes only.

Theorem 2. *Let $m\geq 2$, p be naturals and $p\leq 2^m$. Then*

$$(m, p)\rightarrow\left(m-1, \left]p\left(1-\frac{\log_2 p}{2m}\right)\right[\right),$$

where $]x[$ is the post office function.

Proof. Since

$$\left]p\left(1-\frac{\log_2 p}{2m}\right)\right[= p-\left[\frac{p\log_2 p}{2m}\right],$$

take q such that

$$q=\left[\frac{p\log_2 p}{2m}\right]$$

where $[x]$ is the greatest integer less or equal to x. Then

$$q>\frac{p\log_2 p}{2m}-1,$$

which implies

$$\tfrac{1}{2}p\log_2 p<m(q+1).$$

From this inequality and Burtin's result we have: for every $G_{\mathscr{A}}$ on p vertices there exists at least one x which occurs in $G_{\mathscr{A}}$ at most q times.

That is, $(m,p)\rightarrow(m-1,p-q)$ and the proof is finished.

There exist examples for which these results are best possible. Still it is not clear how sharp the results actually are.

References

[1] J.A. Bondy, Induced subsets, J. Combinatorial Theory (B) 12 (1977) 201–202.

[2] J.D. Burtin, O verojatnosti svjaznosti slučajnogo podgrafa n-mernogo kuba, Problemy Peredači Inf. (13) (2) (1977) 90–95.

[3] N. Sauer, On the density of famlies of sets, J. Combinatorial Theory (A) 13 (1972) 145–147.