Discrete Mathematics 34 (1981) 311–313
North-Holland Publishing Company

NOTE

A SHORT PROOF FOR A THEOREM OF HARPER ABOUT HAMMING-SPHERES

P. FRANKL

CMS, 54 bd. Raspail, Paris 75007, France

Z. FÜREDI

Mathematical Institute of the Hungarian Academy of Sciences, Realtanoda u. 13–15, Budapest, Hungary

Received 27 June 1980

The *Hamming-distance* of two 0–1 sequences $\alpha=(\alpha_i)_{i=1,\ldots,n}$ and $\beta=(\beta_i)_{i=1,\ldots,n}$ is the number of different coordinates. In other terminology, the distance of two sets A and B is the cardinality of their symmetric difference, $d(A,B)=|A\,\Delta\,B|$. (With this distance the set-system $P(X)$ consisting of all subsets of the finite set X is a metric space).

A *Hamming-sphere* with *center* C is a set-system $\mathcal{S}\subset P(X)$ such that for some k:

$$\{S\subset X:\ d(S,C)\leqslant k\}\subset\mathcal{S}\subset\{S\subset X:\ d(S,C)\leqslant k+1\}.$$

The *d-neighbourhood* of a set-system $\mathcal{A}\subset P(X)$ is

$$\Gamma_d\mathcal{A}=\{Y\subset X:\ d(Y,\mathcal{A})=\min_{A\in\mathcal{A}} d(Y,A)\leqslant d\}.$$

It was Harper who first proved that the cardinality of $\Gamma_d\mathcal{A}$ is at least as large as the d-neighbourhood of some appropriate Hamming-sphere with the same cardinality $|\mathcal{A}|$. This theorem has important applications in information theory. Katona [3] gives a different proof. For a generalization see Margulis [5] (Blowing-up lemma). Here we give a new proof for Harper's theorem in an equivalent form.

Theorem. *Let $\mathcal{A}$ and $\mathcal{B}$ be set-systems on X and*

$$d(\mathcal{A},\mathcal{B})=\min\{d(A,B):\ A\in\mathcal{A},B\in\mathcal{B}\}=d.$$

Then there are two Hamming-spheres $\mathcal{A}_0$ with center X and $\mathcal{B}_0$ with center $\emptyset$ such that $|\mathcal{A}_0|=|\mathcal{A}|$, $|\mathcal{B}_0|=|\mathcal{B}|$ and $d(\mathcal{A}_0,\mathcal{B}_0)\geqslant d(\mathcal{A},\mathcal{B})$.

Proof. Consider the set of pairs

$$\{(A,A^*):\ A\in\mathcal{A},A^*\notin\mathcal{A},|A|<|A^*|\}$$

and

$$\{(B, B^*):\ B\in\mathscr{B}, B^*\notin\mathscr{B}, |B|>|B^*|\}.$$

If there are no such pairs, then $\mathscr{A}$ is an X-centered and $\mathscr{B}$ is an $\emptyset$-centered Hamming-sphere, and then there is nothing to prove.

Otherwise let us choose a pair (A, A^*) or (B, B^*) with minimal symmetric difference $|A\,\Delta\, A^*|$ or $|B\,\Delta\, B^*|$ resp. Assume this minimal pair is (A_0, A_0^*).

Set

$$A_0-A_0^*=U, \qquad A_0^*-A_0=V, \qquad |U|<|V|.$$

For these sets U and V we define the following two operations (Up and Down).

$$\mathscr{U}(A)=\begin{cases} A-U+V & \text{if } U\subset A,\ V\cap A=\emptyset,\ A-U+V\notin\mathscr{A},\\ A & \text{otherwise.}\end{cases}$$

$$\mathscr{D}(B)=\begin{cases} B-V+U & \text{if } V\subset B,\ U\cap B=\emptyset,\ B-V+U\notin\mathscr{B},\\ & \text{otherwise.}\end{cases}$$

It is clear that the mapping $\mathscr{U}$ and $\mathscr{D}$ are one-to-one and thus $|\mathscr{U}(\mathscr{A})|=|\mathscr{A}|$, $|\mathscr{D}(\mathscr{B})|=|\mathscr{B}|$, further $|\mathscr{U}(A)|\geqslant|A|$, $|\mathscr{D}(B)|\leqslant|B|$. Since $\mathscr{U}(A_0)=A_0^*$, the joint application $\mathscr{U}$ and $\mathscr{D}$ strictly increases the quantity $(\sum|A|-\sum|B|)$. We show $d(\mathscr{U}(\mathscr{A}), \mathscr{D}(\mathscr{B}))\geqslant d(\mathscr{A}, \mathscr{B})$, and thus the repeated applications of $\mathscr{U}$ and $\mathscr{D}$ finally lead to two Hamming-spheres.

If $A\in\mathscr{U}(\mathscr{A})\cap\mathscr{A}$ and $B\in\mathscr{D}(\mathscr{B})\cap\mathscr{B}$, then clearly $d(A, B)\geqslant d$. Similarly, if $A'\in\mathscr{U}(\mathscr{A})-\mathscr{A}$, $B'\in\mathscr{D}(\mathscr{B})-\mathscr{B}$, then $A'=A-U+V$, $B'=B-V+U$ and thus $A'\,\Delta\, B'=A\,\Delta\, B$ where $A\in\mathscr{A}$, $B\in\mathscr{B}$. Therefore $|A'\,\Delta\, B'|=|A\,\Delta\, B|\geqslant d$. This settles the cases of two old or two new sets.

If one set is new and the other is unchanged, e.g.

$$A'\in\mathscr{U}(\mathscr{A})-\mathscr{A}, \qquad B\in\mathscr{D}(\mathscr{B})\cap\mathscr{B},$$

then $A'=A-U+V$ where $A\in\mathscr{A}$.

If $V\subset B$ and $U\cap B=\emptyset$, then B has not been changed to a smaller set by the operation $\mathscr{D}$ only because $\bar{B}=(B-V+U)\in\mathscr{B}$. Thus $A'\,\Delta\, B=A\,\Delta\,\bar{B}$ whence $d(A', B)=d(A, \bar{B})\geqslant d$.

If the condition $(V\subset B,\ U\cap B=\emptyset)$ is not satisfied and $U=\emptyset$, then $V\not\subset B$. Further $A_0\subset A_0^*$, thus the minimal choice of (A_0, A_0^*) implies $|V|=1$. We infer

$$A'\,\Delta\, B=(A+V)\,\Delta\, B=(A\,\Delta\, B)+V,$$

consequently $|A'\,\Delta\, B|\geqslant d+1$.

Finally, if $1\leqslant|U|<|V|$ and the condition $(V\subset B, U\cap B=\emptyset)$ is not satisfied then there are two elements $u\in U$, $v\in V$ such that at least one of the inclusions $v\in V-B$, $u\in U\cap B$ holds. Since

$$|\bar{A}|=:|A-(U-u)+(V-v)|=|A'|>|A| \quad\text{and}\quad |A\,\Delta\,\bar{A}|<|A_0\,\Delta\, A_0^*|,$$

the definition of A_0 implies that $\bar{A}\in\mathscr{A}$. Further $A'=(\bar{A}-u+v)$ and thus

$A' \Delta B = (\bar{A} - u + v) \Delta B$. If we delete the element u from $\bar{A}$, then $|\bar{A} \Delta B|$ increases or decreases by 1 according to whether $u \in B$ or not. Further if we adjoin the element v to $(\bar{A} - u)$ then $|(\bar{A} - u) \Delta B|$ increases or decreases by 1 according to whether $v \notin B$ or not. Thus in any case

$$|A' \Delta B| = |(\bar{A} - u + v) \Delta B| \geqslant |\bar{A} \Delta B| \geqslant d. \quad \square$$

If

$$\sum_{j=k+1}^{n} \binom{n}{j} < a \leqslant \sum_{j=k}^{n} \binom{n}{j},$$

then the exact computation of $\min\{|\Gamma_d \mathscr{A}| : |\mathscr{A}| = a\}$ has been reduced to the following problem: Given a set-system $\mathscr{F}$ of $(a - \sum_{j=k}^{n} \binom{n}{j})$ k-element sets, at least how many $(k-d)$-element subsets are contained in the sets of $\mathscr{F}$? This well-known problem is answered by the theorem of Kruskal and Katona [2, 4] which states that if

$$|\mathscr{F}| = \binom{a_k}{k} + \binom{a_{k-1}}{k-1} + \cdots + \binom{a_t}{t}$$

where $a_k > a_{k-1} > \cdots > a_t \geqslant t$ (the representation of $|\mathscr{F}|$ in this form is unique), then

$$|\{Y : |Y| = k - d, \exists F \in \mathscr{F} \quad Y \subset F\}| \geqslant \binom{a_k}{k-d} + \binom{a_{k-1}}{k-d-1} + \cdots + \binom{a_t}{t-d}.$$

References

[1] K.H. Harper, Optimal numberings and insperimetric problems on graphs, J. Combinatorial Theory 1 (1966) 385–393.

[2] G.O.H. Katona, A theorem of finite sets, in: Theory of Graphs, Proc. Colloq. Tihany, 1966 (Akademiai Kiadó, 1968) 187–207, MR 45 #76.

[3] G.O.H. Katona, The Hamming-sphere has minimum boundary, Studia Sci. Math. Hungar. 10 (1975) 131–140.

[4] J.B. Kruskal, The number of simplices in a complex, in: Math. Opt. Techniques (Univ. of Calif. Press, 1963) 251–278, MR 27 #4771.

[5] A.A. Margulis, Probabilistic properties of graphs with large connectivity, Probl. Peredachi Informacii 10 (2) (1974) 101–108.