On a problem of Rényi and Katona

Éva Hosszu*

Department of Telecommunications and Media Informatics
Budapest University of Technology and Economics
hosszu@tmit.bme.hu

János Tapolcai*

Department of Telecommunications and Media Informatics
Budapest University of Technology and Economics
tapolcai@tmit.bme.hu

Gábor Wiener§

Department of Computer Science and Information Theory
Budapest University of Technology and Economics
wiener@cs.bme.hu

Abstract: We are dealing with the classical problem of determining the minimum size of a separating system consisting of sets of size k. The problem was raised by Rényi, the first and most important results are due to Katona; Wegener, Luzgin and Ahlswede also proved important bounds. We give a simple, short proof of a strengthening of Katona's main theorem determining the minimum size of a separating system of k-sets.

Keywords: Separating system, k-set, search

1 Introduction and results

A set system is said to be a separating system if any two elements of the underlying set can be separated by some set of the system. More formally:

Definition 1 *Let H be a finite set. The system $\mathcal{A} \subseteq 2^H$ is a separating system if for any $x, y \in H$, $x \neq y$: $\exists\, S \in \mathcal{A}$, such that $x \in S$, $y \notin S$ or $x \notin S$, $y \in S$.*

Separating systems were introduced by Alfréd Rényi [6] in 1961 concerning information-theoretic problems. The problem of finding the minimum size of a separating system containing sets of size k was also raised by Rényi.

Definition 2 *Let m and k be positive integers, such that $k < \frac{m}{2}$. Let us denote the smallest size of a separating system $\mathcal{A} \subseteq 2^{[m]}$ of sets of size exactly k, size at most k, and average size at most k, by $n(m,k)$, $n'(m,k)$, and $n^*(m,k)$, respectively.*

It is obvious that

Claim 3 $n^*(m,k) \leq n'(m,k) \leq n(m,k)$.

*Research is supported by the MTA-BME Future Internet Research Group.

§The work reported in the paper has been developed in the framework of the project "Talent care and cultivation in the scientific workshops of BME". This project is supported by the grant TÁMOP - 4.2.2.B-10/1–2010-0009.

Rényi's problem was to determine the number $\mathrm{n}(m,k)$. In 1966 Katona, using the main theorem in [2] showed

Theorem 4 (Katona) *For $k < \frac{m}{2}$, $n'(m,k) = n(m,k)$.*

In 2008 Ahlswede showed [1, Appendix]

Theorem 5 (Ahlswede) *For $k < \frac{m}{2}$, $n^*(m,k) = n(m,k)$.*

We give a short, simple proof of both theorems in Section 2. Katona's main theorem in [2] is the following.

Theorem 6 (Katona) *For $k < \frac{m}{2}$, $n(m,k)$ is equal to the least number n, for which there exists a system of non-negative integers $s_0, s_1, \ldots, s_n$ satisfying the following three conditions:*

$$\sum_{i=0}^{n} i \cdot s_i = kn, \tag{1}$$

$$\sum_{i=0}^{n} s_i = m, \tag{2}$$

$$s_i \leq \binom{n}{i} \quad i = 0, 1, \ldots, n. \tag{3}$$

We prove the following strengthening of this theorem.

Theorem 7 *For $k < \frac{m}{2}$, $n(m,k)$ is equal to the least number n, for which there exist natural numbers $j \leq n-1$ and $a < \binom{n}{j+1}$, such that*

$$\sum_{i=0}^{j} i \cdot \binom{n}{i} + a(j+1) \leq kn, \tag{4}$$

$$\sum_{i=0}^{j} \binom{n}{i} + a = m. \tag{5}$$

Katona mentions [2] that though Theorem 6 determines $\mathrm{n}(m,k)$ implicitly, it cannot be used to compute the value of $\mathrm{n}(m,k)$. On the other hand, using Theorem 7 it is easy to compute $\mathrm{n}(m,k)$: first fix n and k and find the maximum m satisfying (4) and (5). Let this maximum be $M(n,k)$. Condition (4) is equivalent to

$$\sum_{i=0}^{j-1} \binom{n-1}{i} + \frac{a(j+1)}{n} \leq k,$$

where $\frac{a(j+1)}{n} < \binom{n-1}{j}$, thus the maximum possible values for j and a are easy to find. Therefore, by (5) we have $M(n,k)$. Now $\mathrm{n}(m,k)$ is the smallest n, for which $m \leq M(n,k)$. In Section 2 we will also see that not just the size of a minimum separating system of k-sets is easy to determine but it is also easy to give such a system.

It is worth mentioning that a closed formula for $\mathrm{n}(m,k)$ is not known. The best known lower bound (based on a nice entropy approach) is due to Katona [2], while the best known upper bound is due to Wegener [7] and Luzgin [4]. In 2002 Katona showed [3] that Theorem 6 can be used to obtain really good approximate solutions, while in 2008 Ahlswede proved [1] that the entropy type bound of Katona is asymptotically tight.

2 Proofs

Let $\mathcal{H} \subseteq 2^{[m]}$ be a set system of size n and consider any linear order of its sets. The *incidence matrix* of $\mathcal{H}$ is the 0–1 matrix $M(\mathcal{H}) = (m_{ij})_{n,m}$, where m_{ij} is 1 if the i^{th} set of $\mathcal{H}$ contains the element j and 0 otherwise. Henceforth, all matrices in this paper are binary. A matrix will be called *simple*, if it does not contain identical columns. The *weight* of a row or a column A is defined as the number of 1's in A and is denoted by $w(A)$. We use the following two notions of Katona [2]: a matrix is called *admissible* if the weights of any two rows are the same, and a matrix is called *quasi-admissible* if the weights of any two rows differ by at most one.

It is easy to see [5] that a set system $\mathcal{H}$ is separating if and only if $M(\mathcal{H})$ is a simple matrix and therefore $\mathrm{n}(m,k)$ ($\mathrm{n}'(m,k)$) is the smallest number n, such that an $n \times m$ simple matrix with row weights exactly (at most) k exists.

First we give a short proof of Theorem 4.

Proof of Theorem 4: By Claim 3 we only have to prove $\mathrm{n}(m,k) \leq \mathrm{n}'(m,k)$. For this, it suffices to show that if there exists an $n \times m$ simple matrix M with row weights at most k, then there exists an $n \times m$ simple matrix M', where every row has weight k. Let M be an $n \times m$ simple matrix M with row weights at most k, such that the number of 1's in M is maximum. We show that every row of M has weight k. Assume to the contrary that a row A of M exists, such that $w(A) < k$. For the sake of convenience let us assume that A is the first row of M. Since $w(A) < k < \frac{m}{2}$, the number of 0's is greater than the number of 1's in A. Therefore, there exists a column C of M, such that the first entry of C is 0 and M does not contain the column which differs from C only in the first entry. Thus if we change the first entry of column C to 1, we obtain a simple matrix M' with row weigths at most k, such that M' contains more 1's than M, a contradiction. □

Let $\mathrm{r}(m,k)$ be the least number n, for which there exist numbers j and a, such that $j \leq n-1$, $0 \leq a < \binom{n}{j+1}$ and equations (4) and (5) hold.

Lemma 8 $n^*(m,k) = r(m,k)$.

Proof: First we show that $\mathrm{n}^*(m,k) \leq \mathrm{r}(m,k)$. Let $n = \mathrm{r}(m,k)$, and j, a the numbers for which (4) and (5) hold. Let us consider a matrix M consisting of every column of length n and weight at most j and a different columns of length n and weight $j+1$. M is obviously simple and contains n rows, furthermore by (4) and (5) M contains m columns and at most kn 1's. The existence of such a matrix proves the inequality. In order to prove $\mathrm{r}(m,k) \leq \mathrm{n}^*(m,k)$ let $n = \mathrm{n}^*(m,k)$ and let M be a simple $n \times m$ matrix containing at most kn 1's, such that the number of 1's is minimum. We show that for some $j < n$ every column of M has weight at most $j+1$ and every column of weight at most j appears in M. Now if we let a be the number of columns of weight $j+1$ then it is easy to check that for j and a the equations (4) and (5) hold, from which the inequality follows. For this, we have to show that if a column A of length n appears in M, then every column B of length n and weight less than $w(A)$ also appears in M. This is easy to see: if $w(B) < w(A)$ and B is not a column of M, then by deleting A from M and adding B to M we would obtain an $n \times m$ simple matrix containing less 1's than M, a contradiction. □

To prove Theorems 5 and 7 we need a lemma of Katona, which appears as Step C in the proof of Theorem 6 in [2].

Lemma 9 (Katona) *Let n and b be positive integers, $b \leq n$. Let furthermore c be a positive integer satisfying $c \leq \binom{n}{b}$. Then there exists an $n \times c$ quasi-admissible matrix $M(n,b,c)$, where every column has weight exactly b.*

Now we prove Theorem 5, from which Theorem 7 (by Lemma 8) immediately follows.
Proof of Theorem 5: By Theorem 4 and Claim 3 it suffices to show that $\mathrm{n}'(m,k) \leq \mathrm{n}^*(m,k)$. For this, it is enough to show that if there exists an $n \times m$ simple matrix M containing at most kn 1's, then there exists an $n \times m$ simple matrix M', where every row has weight at most k. Let M be a simple $n \times m$

matrix containing at most kn 1's, such that the number of 1's is minimum. We have seen in the previous proof that for some $j < n$ every column of M has weight at most $j + 1$ and every column of weight at most j appears in M. Now let us delete the columns of weight $j + 1$ from M and add the columns of $M(n, j + 1, m - \sum_{i=0}^{j} \binom{n}{i})$ to M. The matrix M' obtained in this way is obviously an $n \times m$ simple, quasi-admissible matrix containing the same number of 1's as M, which is at most kn. Therefore (since M' is quasi-admissible), every row of M' has weight at most k, which finishes the proof. $\square$

References

[1] R. Ahlswede: Rate-wise optimal non-sequential search strategies under a cardinality constraint on the tests, *Discrete Applied Mathematics* **156** (2008), 1431-1443. Online version with the appendices: http://www.math.uni-bielefeld.de/ahlswede/homepage/public/223.pdf

[2] G. O. H. Katona: On Separating Systems of a Finite Set, *Journal of Combinatorial Theory* **1** (1966), 174-194.

[3] G. O. H. Katona: Search with small sets in presence of a liar, *Journal of Statictical Planning and Inference* **100** (2002), 319-336.

[4] V. N. Luzgin: Separating systems of partitions of a finite set, *Combinatorial Analysis* **5** (1980), 46-49. (in Russian)

[5] J. Spencer: Minimal completely separating systems, *Journal of Combinatorial Theory* **8**, (1970) 446-447.

[6] A. Rényi: On Random Generating Elements of a Finite Boolean Algebra, *Acta Sci. Math. (Szeged)* **52** (1961), 75-81.

[7] I. Wegener: On separating systems whose elements are sets of at most k elements, *Discrete Mathematics* **28** (1979), 219-222.