SEARCH USING SETS WITH SMALL INTERSECTION

György E. Katona

We consider the following type of intersection. Let \( S \) be a set of cardinality \( n \), and \( A_1, \ldots, A_n \) be subsets of \( S \) of the same cardinality \( k \). We assume that \( S = \bigcup_{i=1}^{n} A_i \) and the intersection \( \bigcap_{i=1}^{n} A_i \) is nonempty. Let \( S' \) be a set of cardinality \( m \), and \( B_1, \ldots, B_m \) be subsets of \( S' \) of the same cardinality \( k \). We assume that \( S' = \bigcup_{i=1}^{m} B_i \) and the intersection \( \bigcap_{i=1}^{m} B_i \) is nonempty. In this paper, we consider the case \( k = n = m \).

INTRODUCTION

Let \( S \) be a finite set of a universal and \( A \) be an element of element of \( S \). We can have the following type of information on \( A \). For some elements \( A \) \( C \) we may ask whether \( A \) is in \( S \) \( \subseteq S \) or \( A \) is not in \( S \) \( \not\subseteq S \). We thus trying to determine whether \( A \) or \( \not\subseteq S \). We are able to do this in a number of \( 2^n \) ways. For \( n \) \( \leq 5 \) the number is \( 32 \) and for \( n \) \( \leq 6 \) the number is \( 64 \). For larger values of \( n \), the number becomes very large. We may consider the number of \( 2^n \) ways to determine if \( A \) or \( \not\subseteq S \).
the following constraint is satisfied:

\[ b_i \geq b_{i+1} \leq b_{i+2} \leq \cdots \leq b_{k-1} \leq b_k \leq b_{k+1} \leq b_{k+2} \leq \cdots \leq b_n \leq b_{n+1} \]

The case \( k = 1 \) is completely solved. If \( k = 1 \) we have fully good lower and upper estimation. The other cases are partially verified.

It is worth while to mention that under the constraint \( b_k \leq \frac{1}{2} \), the lower inequality (1.1) becomes twice for all possible evaluations of \( \sum_{i = k}^{n} a_i \). Hence, it is in upper estimates under the present conditions \( a_i \leq b_i \leq 0 \)

**LONGER EDITIONS**

Let \( X \) be a finite set and \( \mathcal{A} \) be a subset of \( X \). We say that \( \mathcal{A} \) separates two subsets \( S \) and \( T \) of \( X \) if \( S \) contains exactly one of \( S \) and \( T \). The family \( \mathcal{F} = \{ F \} \) of subsets of \( X \) is called a separator system of \( \mathcal{A} \) if every subset of \( X \) which is not an element of \( \mathcal{A} \) contains a set of \( \mathcal{F} \). A separator system is a sequence of \( \mathcal{F} \) of \( \mathcal{A} \). A \( \mathcal{F} \)-separator system \( \mathcal{F} \) of \( \mathcal{A} \) is a sequence of \( \mathcal{F} \) of \( \mathcal{A} \). A separator system is a sequence of \( \mathcal{F} \) of \( \mathcal{A} \). A separator system is a sequence of \( \mathcal{F} \) of \( \mathcal{A} \). A separator system is a sequence of \( \mathcal{F} \) of \( \mathcal{A} \). A separator system is a sequence of \( \mathcal{F} \) of \( \mathcal{A} \). A separator system is a sequence of \( \mathcal{F} \) of \( \mathcal{A} \). A separator system is a sequence of \( \mathcal{F} \) of \( \mathcal{A} \). A separator system is a sequence of \( \mathcal{F} \) of \( \mathcal{A} \).

**LEMMA 1**. Let \( \mathcal{F} = \{ F_1, \ldots, F_m \} \) be a family of subsets of \( X \) and let \( \mathcal{A} = \{ a_1, \ldots, a_n \} \) be a subset of \( X \). Let \( \delta = \delta(\mathcal{F}, \mathcal{A}) \) be the number of subsets of \( X \) which are not in \( \mathcal{A} \). Then the number of subsets of \( X \) which are not in \( \mathcal{A} \) is equal to

\[ \delta = \sum_{i=1}^{m} |F_i| \]

and

\[ \sum_{i=1}^{m} |F_i| = \delta \]

for some \( 1 \leq k \leq \delta \), for \( 1 \leq n \leq m \) should be included in \( \sum_{i=1}^{m} |F_i| \).

**Proof.** Let \( \mathcal{F} = (F_1, \ldots, F_m) \) be the incidence matrix of \( \mathcal{F} \). Then, \( f_{ij} = 1 \) if \( x_i \in F_j \) and \( f_{ij} = 0 \) otherwise. Define by \( a_i \) the number of columns of \( \mathcal{F} \) having exactly \( r \) columns equal to \( \{ a_1, \ldots, a_n \} \). Then the sum equality can be verified

\[ \sum_{i=1}^{m} f_{ij} = \delta \]
\begin{equation}
\sum_{\mathbf{r}} \binom{n}{r} x^r = 0
\end{equation}

(2) follows from (1). Since any two elements of \( X \) are separated by at least \( r \), we have an \( r \)-column. If \( r \) is the smallest \( r \) such that \( x^r = 0 \), then the columns are different. The maximum number of different columns with \( r \) columns.

(3) follows from (2). The number of pairs \( x^r \) are \( r \), and if we can not reach \( x^r \), then we can not reach \( x^{r+1} \).

(4) follows from (3). If \( x^r \) satisfies (4) and \( s \) is a divisor of \( r \), then \( x^s \) satisfies (4). Consequently, \( x^r \) satisfies (4), and \( x^{r+s} \) satisfies (4).

(5) follows from (3) and (4). The left-hand side of (5) is bounded by

\begin{equation}
\binom{n}{r} x^r + \binom{n}{s} x^s < 0
\end{equation}

that is, \( x \) satisfies (5).

(6) The repeated application of the previous result will prove our statement. If

\begin{equation}
\alpha = \binom{n}{r} x^r + \binom{n}{s} x^s,
\end{equation}

and \( s \leq r \), then \( \alpha \) is the minimal value of \( x \) satisfying (6).

Lemma 1: The minimal \( x \) of the solution of (1) and (2) is obtained for the maximum \( x \) for which there is a solution in \( \mathbb{R}^n \).

Proof: \( x \) is a solution if the set of solutions of (1) \& (2) is contained in the set of solutions of (2) \& (3). Moreover, it is an interior of points and moreover these linear equations are defined for different \( x \). Finally, if \( x > 0 \), then the corresponding inequalities are satisfied in the opposite way. The assertion is then trivially satisfied.

Lemma 2: If \( x < 0 \), then the minimal value of (1) \& (2) is obtained in satisfying

\begin{equation}
\alpha = \binom{n}{r} x^r + \binom{n}{s} x^s
\end{equation}

(7)
Let us verify that

\[ \binom{r}{k} \leq \frac{\binom{n}{k}}{\binom{n-r}{n-k}}. \tag{7} \]

Theorem (7.3) follows from (2) and (2'), if \( n \geq 3 \) and \( r \geq 1 \). Indeed, the left-hand side of (2) implies

\[ \binom{r}{k} \leq \frac{\binom{n}{k}}{\binom{n-r}{n-k}} \leq 1, \tag{8} \]

(assuming \( n > m \geq r > 0 \)). Hence we obtain

\[ \binom{r}{k} \leq \frac{\binom{n}{k}}{\binom{n-r}{n-k}} \leq \frac{\binom{n}{k}}{\binom{n-r}{n-k}} \leq \binom{n}{k}, \tag{9} \]

and (7) follows from (6). (The case \( n = 3, m = 1 \) can be checked by easy computation.)

2) From (7) we have \( \binom{n}{k} \leq n^{k} \), that is, by the inductive step, \( n^{k} \geq n^{k} \). \( \tag{10} \)

3) For \( n = 3 \) and \( k = 2 \), (7) follows from (2) and this completes the theorem. \( \tag{11} \)

Let us verify that

\[ \binom{r}{k} \leq \frac{\binom{n}{k}}{\binom{n-r}{n-k}} \leq \binom{n}{k}. \tag{12} \]

Let us verify that

\[ \binom{r}{k} \leq \frac{\binom{n}{k}}{\binom{n-r}{n-k}} \leq \binom{n}{k}. \tag{13} \]

For any solution of (7.5), suppose \( k \geq \frac{n}{2} \), then after all \( 1 \leq k \leq n \). Assume, in the contrary that \( k < \frac{n}{2} \). Hence and from (10) we have

\[ \binom{r}{k} \leq \frac{\binom{n}{k}}{\binom{n-r}{n-k}} \leq \binom{n}{k}. \tag{14} \]

The solution of the lemma is \( \binom{n}{k} \leq \binom{n}{k} \). Comparing it with the above inequality
we obtain
\[ \frac{n^2}{2} \leq \frac{a_k^2 + 2a_k}{2k + 1}, \]
and this is a contradiction for any \( k > 0 \).

Thus we have proved that we have to compute
\[ i = \binom{a}{2} \leq \binom{x - 1}{2} \]
rather than \( \binom{b}{2} \).

We prove now that (12) always have a (x integer) solution in x by comparing
\[ 1 \leq \binom{x}{2} \]
For a given m, \( x = \binom{m}{2} \) and \( x = \frac{m + 1}{2} \) determine an interval. We have to see that these intervals cover all the named numbers. Indeed, it is easy to see that
\[ \binom{m}{2} < \frac{m + 1}{2} < \frac{m + 2}{2} < \binom{m + 1}{2} \]
holds if \( k > 7 \) and \( m > 9 \). However, the beginning of the interval corresponding to \( m = 8 \) is \( 28 \). The interval over all the integers \( x > 25 \). By \( m > 10 \) there is no \( k \) satisfying \( 1 < k < \frac{25}{22} \).

We have proved that (12) has no solution for \( k = 1 \), and if \( x \) is a solution for (12), then the solution \( x = \binom{m}{2} \). This solution satisfies (12), and hence (12) has a solution for all \( m \). The converse is proved.

Case \( k = 1 \).

The left hand side of (11) \( \binom{m}{2} < \binom{x}{2} \). For \( k = 1 \) the controne
\[ 1 \leq \binom{m}{2} \]
It means that if \( x \) is a solution in the form \( 1 \leq \binom{m}{2} \) for some integer \( m \), then the \( m \) is the minimal solution of (11). By Lemma 2, otherwise there
\[ \binom{m}{2} \in \binom{a}{2} \leq \binom{x - 1}{2} \]
\[ 1 \leq \binom{m}{2} \leq \binom{x}{2} \]
\[ i = \binom{a}{2} < \binom{x}{2} \]
The minimal \( m \) in this case is the minimal \( m \) satisfying \( 1 \leq \binom{m}{2} \).
Symmetrizing the two cases, we obtain in satisfying (1) and (2) the obtained or satisfying $x < 1 + \frac{m}{3}$. By Lemma 2 this is a linear equation. The above equation is not. It only contains $E = \left(A_1, \ldots, A_m\right)$ by considering the corresponding $n \times n$ matrices $X$ of all different columns consisting of 1 and $m - 1$ entries of 2, where $0 < x < 1 - m$. One follows from the nonidentity of $x$ the $x \leq 1 + \frac{m}{3}$, the matrix obtained in the way obviously satisfies the condition that it has differences columns $x > 0$ and the column has at least two more 1's, $x < 1 - m$. This is then possible only if there exists an integer $k$ with $x < 1 - m$. Thus

$$\left| \begin{array}{c}
A_1
\vdots
\cdot
\cdot
\cdot
\vdots
n
\end{array} \right| < 0,$$

where (a) denotes the last integer $d$ and this is the best possible estimate.

One $x = 1$

Using the theory of Sierpinski systems we obtain an almost complete solution here.

**Theorem 3:** If $E = \left(A_1, \ldots, A_m\right)$ is a mapping system on an incidence vector $n \times n$, then $x < 1 - m$. The equation $x \leq 1 - m$ and

$$\left| \begin{array}{c}
A_1
\vdots
\cdot
\cdot
\cdot
\vdots
\cdot
n
\end{array} \right| < 0,$$

where (a) denotes the last integer $d$ and this is the best possible estimate.

One $x = 1$

Using the theory of Sierpinski systems we obtain an almost complete solution here.

**Theorem 3:** If $E = \left(A_1, \ldots, A_m\right)$ is a mapping system on an incidence vector $n \times n$, then $x < 1 - m$. The equation $x \leq 1 - m$ and

$$\left| \begin{array}{c}
A_1
\vdots
\cdot
\cdot
\cdot
\vdots
\cdot
n
\end{array} \right| < 0,$$

where (a) denotes the last integer $d$ and this is the best possible estimate.
If there is a finite triple system for \( m = 1 \), let us arbitrarily delete \( \binom{m-1}{2}/3 = k \) right- and left-hand pairs. This construction proves the theorem in this case.

If there is a finite triple system for \( m = 2 \), delete a pair with the \( m+2 \) triples containing it. Then we have \( \binom{m-1}{2}/3 = n \) triples on \( m \) points with the desired property. However, if the number is \( \binom{m-1}{2}/3 \) we can delete some triples until we have only \( n \) triples on \( m \) points. As \( \binom{m-2}{2}/3 \), \( m = \binom{m-1}{2}/3 \) (this case is also admitted)

If \( m = \frac{1}{2} \sqrt{4m - 1} \) (there is exactly one such \( m \) for each \( n \)). We shall construct the incidence matrix \( A \) by its columns. It will have one column for each of the \( \binom{m}{2} \) columns with two \( 1 \)'s. It is easy to see that the number of remaining columns

\[ s = \binom{m}{2} \cdot \left( \binom{m-1}{2}/3 \right) \]

Choose now a different column with three \( 1 \)'s in each, according to the previous result. Hence, if \( 2d \) columns of \( A \) are mutually orthogonal or pairwise orthogonal, then \( 2d \) is divisible by \( 3 \) and the condition is proven.

The same proof works for the following more general case.

**Theorem:** For \( d = (a_1, \ldots, a_k) \) be a partition on \( n \) points and

\[ (a_i, \sum_{j=0}^{i-1} a_j) < a_i \quad (i = 1, \ldots, m, 1 \neq j \neq i) \]

then the minimum of \( m \) is

\[ \frac{d}{2} \]
\[
\begin{align*}
W &= \left( 1 + \sqrt{\frac{1}{2}\pi - 2\pi^2} - 2\pi \right)^
\frac{1}{4} + 2 \\
&= \left( \frac{1}{2} - \frac{\sqrt{1}}{2} \right)^
\frac{1}{4} + 2
\end{align*}
\]

where \( W \) is the first term. \( W \) and \( h \) are the numbers of triplets in a set \( \mathcal{W} \) and \( \mathcal{H} \), respectively.

The problem can be formulated in the following way:

\textbf{Problem 1.} Let \( Y \) be a set of \( n \) elements. What is the maximum number of different subsets of \( Y \), such that any pair of elements is contained in at most one subset?

Another, more helpful problem is to find the solution of the case \( \mathcal{H} \).

\textbf{Problem 2.} The number \( n \) of triplets in an \( n \)-element set is given. What is the number of the subset of elements in the set \( Y \)?

\textbf{Problem 3.} If it is true that for any \( w \) and \( n \), the \( n \) triplets can be chosen in the minimum of \( w \), at least by \( 2 \).

\textbf{REFERENCES:}

