THE SIZE OF A MINIMAL SEPARATING SYSTEM OF SUBSETS

BY

LOUIS GORDON

TECHNICAL REPORT NO. 11
JUNE 30, 1970

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT GP-15909

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
The Size of a Minimal Separating System of Subsets

by

Louis Gordon

TECHNICAL REPORT NO. 11
June 30, 1970

Prepared for the National Science Foundation Grant GP-15909

Reproduction in Whole or in Part is Permitted for any Purpose of the United States Government
This document has been approved for public release and sale; its distribution is unlimited

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
Abstract

Let $N = \{1, \ldots, n\}$. A complete separating system on $N$ is a collection of subsets of $N$, $\{S_1, \ldots, S_k\}$, for which each element of $N$ appears in some set and such that $\bigcap_{i \in S_j} S_j = \{i\}$ for each $i \in N$. The minimum number of subsets needed to completely separate a set of $n$ elements is denoted $f(n)$. It is shown that if $\left(\frac{k-1}{2}\right) < n \leq \left(\frac{k}{2}\right)$, then $f(n) = k$. 
The Size of a Minimal Separating System of Subsets

Let $S = \{S_1, \ldots, S_k\}$ be a set of subsets of the set $N = \{1, \ldots, n\}$. T. J. Dickson has shown in a paper in the Journal of Combinatorics (7, November, 1969), that $\exists f(n) \sim \log 2n$ where $f(n)$ is the minimum number of subsets necessary to form a complete separating system on $n$ elements. A complete separating system is a set of subsets such that $i, j \in N$ and $i \neq j$ implies that $\exists S_1, S_2 \in S$ such that $i \in S_1$, $j \notin S_1$, $j \in S_2$, and $i \notin S_2$. In this discussion we show that if

$$\left(\frac{k-1}{k-2}\right) < n < \left(\frac{k}{k-2}\right)$$

then $f(n) = k$.

Definition. Let $S = \{S_1, \ldots, S_k\}$ be a collection of subsets on $N = \{1, \ldots, n\}$. Define, for each $i \in N$, $A_i = \{j|i \in S_j \in S\}$. We shall see that $S$ is completely determined by the set of $A_i$.

Proposition. Let $\mathcal{A} = \{A_1, \ldots, A_n\}$ be a collection of non-empty subsets of $\{1, \ldots, K\}$ = $K$. Let $N = \{1, \ldots, n\}$.

Define $S_j = \{i \in N|j \in A_i\}$ for each $j \in K$. Then $S = \{S_j|j = 1, \ldots, k\}$ is a complete and separating system iff $p \nRightarrow q$ implies

$$A_p \notin A_q \text{ and } A_q \notin A_p.$$ 

Proof: Lest the meaning of the theorem be obscured by the notation, we first give a translation to English. The sets $A_i$ are addresses for the elements 1. That is, $A_i$ indexes those sets $S_j$ in which $i$ appears. Then for the set $S_\mathcal{A}$ to be complete and separating, each
element $i, 1 \leq i \leq n$ has some non-empty address $A_i$, and no element has its address contained in that of some other element.

Symbolically, we have the argument:

a) $S$ is complete on $\{1, \ldots, n\} = N$.

\[
i \in N \quad \Rightarrow \quad A_i \not\subseteq \emptyset \quad \text{by assumption}
\]

\[
\Rightarrow \quad \exists j \in A_i
\]

\[
\Rightarrow \quad i \in S_j
\]

b) $S$ is separating.

Let $p \not\models q$.

Then $A_p \not\subseteq A_q$ and $A_q \not\subseteq A_p$

\[
\iff \{j_p, j_q\} \subseteq K \exists
\]

\[
j_p \in A_p \setminus A_q \quad \text{and} \quad j_q \in A_q \setminus A_p
\]

\[
\iff \begin{cases}
p \in S_{j_p}, \quad q \in S_{j_q} \\
p \not\models S_{j_p}, \quad q \not\models S_{j_q}
\end{cases}
\]

\[
\iff \quad S \quad \text{is separating}
\]

Corollary. If $S = \{S_1, \ldots, S_k\}$ is complete and separating and $A_i = \{j \mid i \in S_j\}$ then

\[
A_p \not\subseteq \emptyset \quad \text{and} \quad A_p \not\models A_q \quad \text{for all} \quad p \models q
\]

\[
\{p, q\} \subseteq \{1, \ldots, n\} = N
\]

2
Proof: $A_p \not\subseteq \emptyset$ because $S$ is complete so that $p \in N \Rightarrow (\exists S_j \in S \exists p \in S_j) \Rightarrow j \in A_p$. The second assertion follows by Prop. 1.

We now define the function $g(k)$

$$g(k) = \max \left\{ n \mid \text{a complete separating system of} \right. \begin{array}{l}
k \text{ elements on } \{1, \ldots, n\}
\end{array} \right\}$$

Note the following relation between Dickson's $f(n)$ and our $g(k)$:

$$f(n) = k \iff g(k-1) < n \leq g(k) .$$

Lemma 2. Let $\mathcal{A} = \{A_1, \ldots, A_n\}$ be a collection of $n$ non-empty subsets of $\{1, \ldots, k\}$ such that $q \not\subseteq p \Rightarrow A_p \not\subseteq A_q$ and $A_q \not\subseteq A_p$. Then $\mathcal{A} = \{A_1^c, A_2^c, \ldots, A_n^c\}$ also has this property.

Proof: $A_p \not\subseteq A_q$ and $A_q \not\subseteq A_p \Rightarrow A_p^c \not\subseteq A_q^c$ and $A_q^c \not\subseteq A_p^c$. Also $A_p^c = \emptyset \Rightarrow A_p = \emptyset \Rightarrow A_p \supset A_q$ for all $q$.

Notation: Let $\nu(A)$ denote the size (cardinality) of the set $A$.

Lemma 3. Let $\mathcal{A} = \{A_1, \ldots, A_n\}$ be a collection of non-empty subsets of $\{1, \ldots, k\}$ with the property $p \not\subseteq q \Rightarrow A_p \not\subseteq A_q$ and $A_q \not\subseteq A_p$ then $\exists \hat{\mathcal{A}} = \{\hat{A}_1, \ldots, \hat{A}_n\} \exists \nu(\hat{A}_j) \geq \lceil \frac{k+1}{2} \rceil$ and $\max(\nu(\hat{A}_j)|j \leq n) = (\max(\nu(A_j)|A_j \in \mathcal{A})) \lor \lceil \frac{k+1}{2} \rceil$.

Proof by induction.

Denote by $||\mathcal{A}|| = \min(\nu(A_j)|1 \leq j \leq n)$

If $||\mathcal{A}|| \geq \lceil \frac{k+1}{2} \rceil$ there is nothing to prove.

If $||\mathcal{A}|| < \lceil \frac{k+1}{2} \rceil$ then denote $||\mathcal{A}|| = a$

Denote by $\mathcal{A}_a = \{A \in \mathcal{A} | \nu(A) = a\}$
Denote by $\mathcal{C}_{a+1} = \{ CC \mathcal{K} \big| \nu(C) = a + 1 \text{ and } \big[ \exists A \in \mathcal{A}_a \ni C \subseteq A \big] \}$. 

Note first that

$C \in \mathcal{C}_{a+1} \Rightarrow \text{ if } A \in \mathcal{A}, \text{ then } A \not\subseteq C \text{ because if } A \in \mathcal{A}_a \text{ then } \nu(C) = a + 1 > a = \nu(A) \text{ and if } A \notin \mathcal{A}_a \text{ then } A \supseteq C \subseteq A^* \in \mathcal{A}_a$. Contradicting the assumed property

$A \upharpoonright A_q \Rightarrow A \upharpoonright A_q \cdot$

We now show that $\nu(\mathcal{C}_{a+1}) \geq \nu(\mathcal{A}_a)$. Note that each element of $\mathcal{A}_a$ is contained in exactly $k-a$ $(a+1)$-subsets of $\{1, \ldots, k\}$. Hence the total number of containments of an element of $\mathcal{C}_{a+1}$ is $(k-a) \nu(\mathcal{A}_a)$. But, each element of $\mathcal{C}_{a+1}$ contains at most $\binom{a+1}{a}$ $a$-subsets so that the number of containments of elements of $\mathcal{A}_a$ by elements of $\mathcal{C}_{a+1}$ is at most $(a+1) \nu(\mathcal{C}_{a+1})$. They may be pictured by the following graph where the elements of $\mathcal{A}_a$ are listed on the left and those of $\mathcal{C}_{a+1}$ on the right. We connect an element $C$ with an element $A$ iff $A \subseteq C$.

![Graph showing containment relationships between sets $\mathcal{A}_a$ and $\mathcal{C}_{a+1}$](attachment:graph.png)
Hence

\[(k-a) \nu(\mathcal{A}_a) = \nu(\text{containments}) \leq (a+1) \nu(\mathcal{C}_{a+1})\]

but \(a \leq \left[\frac{k+1}{2}\right] - 1\) by assumption so that

\[k-a \geq k - \left[\frac{k+1}{2}\right] + 1 = \left[\frac{k}{2}\right] + 1 \geq \left[\frac{k+1}{2}\right] = \left[\frac{k+1}{2}\right] \geq a + 1\]

so \(\frac{a+1}{k-a} \leq 1\) and

\[\nu(\mathcal{A}_a) \leq \frac{a+1}{k-a} \nu(\mathcal{C}_{a+1}) \leq \nu(\mathcal{C}_{a+1})\]

Hence we may define a 1-1 mapping of

\[\mathcal{A}_a \longrightarrow \mathcal{C}_{a+1}\]

and denote \(A_j \longrightarrow \hat{A}_j \in \mathcal{C}_{a+1}\) for all \(A_j \in \mathcal{A}_a\).

Now define the collection \(\mathcal{A} = \{\hat{A}_1, \ldots, \hat{A}_n\}\) by:

\[
\hat{A}_j = \begin{cases} 
A_j & \text{if } \nu(A_j) \geq a + 1 \\
\hat{A}_j & \text{if } A_j \in \mathcal{A}_a 
\end{cases}
\]

Note that \(\nu(\hat{A}_j) \geq a + 1\) for all \(j = 1, \ldots, n\), and so \(\hat{A}_j \neq \emptyset\) for any \(j\).
We now show that \( p \vdash q \Rightarrow \hat{A}_p \not\models \hat{A}_q \) and \( \hat{A}_q \not\models \hat{A}_p \). There are three cases:

a) \( \hat{A}_p = A_p \quad \hat{A}_q = A_q \).

\( \hat{A}_q \not\models \hat{A}_p \) and \( \hat{A}_p \not\models \hat{A}_q \) by assumption.

b) \( \hat{A}_p = A_p \quad A_q \subseteq A_a \).

Then \( \hat{A}_p \not\models \hat{A}_q \) because we already showed that \( A \not\models C \) for all \( A \in \mathcal{A}, C \in \mathcal{C}_{a+1} \). Now \( \hat{A}_q \not\models \hat{A}_p \) because \( p \vdash q \Rightarrow \hat{A}_p \not\models \hat{A}_q \) and \( \nu(\hat{A}_p) \geq a + 1 = \nu(\hat{A}_q) \) so that \( \hat{A}_q \not\models \hat{A}_p \Rightarrow \hat{A}_q = A_p \).

c) \( A_p \in \mathcal{A}, A_q \in \mathcal{A} \). Then \( \hat{A}_p \not\models \hat{A}_q \) since the map \( \mathcal{A} \rightarrow \mathcal{C}_{a+1} \) was injective. Since \( \nu(\hat{A}_p) = \nu(\hat{A}_q) \),

\( \hat{A}_p \not\models \hat{A}_q \) and \( \hat{A}_q \not\models \hat{A}_p \).

Hence, given \( \mathcal{A} = \{A_1 \ldots A_n\}, A_1 \not\vdash \emptyset \) and

\[ p \vdash q \Rightarrow A_p \not\models A_q, A_q \not\models A_p \quad \text{and} \quad ||A|| < \left[ \frac{n+1}{2} \right] \]

then \( \exists \hat{\mathcal{A}} = \{\hat{A}_1 \ldots \hat{A}_n\} \) with \( \hat{A}_1 \not\vdash \emptyset \)

and \[ ||\hat{\mathcal{A}}|| = ||\mathcal{A}|| + 1 \quad . \]

Applying this result \[ \left[ k+1 \right] - ||\mathcal{A}|| \] times yields the result.

Note that at no stage in the process was the set of \( A^* \) \[ \max(v(A_j)) = v(A^*) \] touched unless \( v(A^*) < \left[ \frac{k+1}{2} \right] \) and then all the sets in the final \( \hat{\mathcal{A}} \) had size \[ \left[ \frac{k+1}{2} \right] \]. Hence

\[ \max(v(\hat{A}_j)|A_j \in \hat{\mathcal{A}}) = \left[ \frac{k+1}{2} \right] \lor \max\{v(A_j)|A_j \in \mathcal{A}\} \quad . \]

Hence the lemma. \( \square \).

Proposition 4. If \( \mathcal{A} = \{A_1 \ldots A_n\} \) is a collection of subsets of \( K = \{1, \ldots, k\} \) such that
1) $\bar{A}_i \not\in \emptyset$ \hspace{1em} $i = 1, \ldots, n$

ii) $p \nmid q \Rightarrow A_p \not\in A_q$ and $A_q \not\in A_p$.

Then $\exists$ a set $\mathcal{A} = \{\bar{A}_1, \ldots, \bar{A}_n\}$ of subsets of $K$ such that

i) $\bar{A}_p \not\in \emptyset$ \hspace{1em} $\forall p = 1, \ldots, n$

ii) $p \nmid q \Rightarrow \bar{A}_p \not\in \bar{A}_q$ and $\bar{A}_q \not\in \bar{A}_p$

iii) $v(\bar{A}_p) = \left[\frac{k+1}{2}\right]$ \hspace{1em} $p = 1, \ldots, n$.

Proof. Apply Lemma 3 to the given sets to obtain $\hat{A}$. Then $\hat{A}_j \in \hat{A} \Rightarrow v(\hat{A}) \geq \left[\frac{k+1}{2}\right]$. Now apply Lemma 2 to $\hat{A}$ to get $\tilde{\mathcal{A}} = \{\hat{A}_j^c | \hat{A}_j \in \hat{A}\}$. Note that the $\{\hat{A}_j^c\}$ have the same properties as the original $\{A_j\}$, but note that $v(\hat{A}_j^c) = k - v(\hat{A}_j)$

$\leq k - \left[\frac{k+1}{2}\right]$

$= \left[\frac{k}{2}\right]$.

Hence $|\tilde{\mathcal{A}}| \leq \left[\frac{k}{2}\right]$ and $\max\{v(\hat{A}_j^c) | \hat{A}_j \in \hat{A}\} \leq \left[\frac{k}{2}\right]$. By applying Lemma 3 again we obtain a collection $\tilde{\mathcal{A}} = \{\tilde{A}_j | j = 1, \ldots, n\}$ such that

i) $\tilde{A}_j \not\in \emptyset$ \hspace{1em} $j = 1, \ldots, n$

ii) $p \nmid q \Rightarrow \tilde{A}_p \not\in \tilde{A}_q$ and $\tilde{A}_q \not\in \tilde{A}_p$

iii) $v(\tilde{A}_p) = \left[\frac{k+1}{2}\right]$ \hspace{1em} $p = 1, \ldots, n$ \hspace{1em} $\square$.

Theorem 5. Let $\mathcal{A} = \{A_1, \ldots, A_n\}$ be a collection of subsets of $K = \{1, \ldots, k\}$ such that $p \nmid q \Rightarrow A_p \not\in A_q$ and $A_q \not\in A_p$, then

$n \leq \left(\left[\frac{k+1}{2}\right]\right) = \binom{k}{\left[\frac{k}{2}\right]}$.

Proof: By Prop. 4, $\exists \tilde{\mathcal{A}} = \{\tilde{A}_j | j = 1, \ldots, n\}$ satisfying $p \nmid q \Rightarrow \tilde{A}_p \not\in \tilde{A}_q$ and $\tilde{A}_q \not\in \tilde{A}_p$ and $v(\tilde{A}_p) = \left[\frac{k+1}{2}\right]$. Since $p \nmid q \Rightarrow \tilde{A}_p \not\in \tilde{A}_q$, $\tilde{\mathcal{A}}$ is a subset containing distinct elements from the set of all $\left[\frac{k+1}{2}\right]$-combinations of $K$, indexed by $\{1, \ldots, n\}$. Hence $v(\tilde{\mathcal{A}}) = n \leq \left(\left[\frac{k+1}{2}\right]\right) = \binom{k}{\left[\frac{k}{2}\right]} \hspace{1em} \square$. 

7
Corollary. \( g(k) = n \) iff \( \left( \frac{k-1}{2} \right) < n \leq \left( \frac{k+1}{2} \right) \).

Proof: If \( S \) is a complete and separating system on \( \{1, \ldots, n\} \) apply Prop. 1 and its corollary to obtain \( A_S = \{A_1 \ldots A_n\} \). Let \( \nu(S) = k \).

note that \( p \perp q \Rightarrow A_p \not\perp A_q \) and \( A_q \not\perp A_p \)

\[ \Rightarrow A_p \not\perp \emptyset \quad \text{for all } p = 1, \ldots, n \]

and

\[ A_p \subseteq K = \{1, \ldots, k\} \]

Now apply theorem 5 to get \( n \leq \left( \frac{k}{2} \right) \). Hence \( g(k) \leq \left( \frac{k}{2} \right) \).

Now letting \( \mathcal{A} = \{A \subseteq K|\nu(A) = \frac{k+1}{2}\} \) yields a collection of \( \left( \frac{k}{2} \right) = \left( \frac{k+1}{2} \right) \) elements \( \exists A \perp B \Rightarrow A \not\perp B \) and \( \not\perp A \). Applying Prop. 1 yields \( S \), a complete and separating system on \( \left( \frac{k}{2} \right) \) elements consisting of \( k \) subsets. Hence \( g(k) \geq \left( \frac{k}{2} \right) \). So \( g(k) = \left( \frac{k}{2} \right) \).

Corollary. \( f(n) = k \) where \( \left( \frac{k-1}{2} \right) < n \leq \left( \frac{k}{2} \right) \).

Proof: follows from the definition of \( g \) and \( f \).
REFERENCE