ON MINIMAL COLLECTIONS OF INDEXES

Egor A. Timoshenko

We denote $s = [\frac{n}{2}]$, $l = [\frac{n+1}{2}]$, $M_n = C_n^s = C_n^l$; indexes built for the case of n columns (i.e., ordered subsets of the set $\{1, 2, \dots, n\}$) will be called *n-indexes*.

The length of an index a is designated by $\|a\|$. If $0 \leqslant i \leqslant \|a\|$, then $S_i(a)$ will denote the set consisting of the first i elements of the index a. Any set that has exactly i elements is said to be an *i-element set*. The cardinality of a set A will be designated by $|A|$.

Our aim is to build a collection of n-indexes U such that for any subset $A \subseteq \{1, 2, \dots, n\}$ there exists an index $a \in U$ satisfying the condition $S_{|A|}(a) = A$. For any $i \in \{0, 1, 2, \dots, n\}$ every i-element subset of $\{1, 2, \dots, n\}$ must agree with the beginning of at least one index, so the number of the indexes cannot be less than C_n^i, which is the number of different i-element subsets. In particular, if we consider the s-element subsets of $\{1, 2, \dots, n\}$, we obtain that the number of the indexes is at least $M_n = C_n^s$.

At least C_n^{s+1} of these M_n (or more) indexes of length s should be extended by one element. After that at least C_n^{s+2} of the extended indexes should be extended by one more element, etc. Finally, one ($C_n^n = 1$) of the indexes already having length $n-1$ should be extended by one more element. We obtain that the total length of the indexes cannot be less than $L_n = sC_n^s + C_n^{s+1} + C_n^{s+2} + \dots + C_n^n$.

Recall that for any $i \in \{1, 2, \dots, n\}$ the formula $C_{n+1}^i = C_n^{i-1} + C_n^i$ is valid. Next, for $i \in \{0, 1, 2, \dots, n\}$ we have $C_n^i = C_n^{n-i}$. Note also that $\sum\limits_{i=0}^{n} C_n^i = 2^n$. Then we obtain:

a) if $n = 2s + 1$, then $L_n = sC_n^s + \frac{1}{2} \cdot 2^n = sM_n + 2^{n-1}$;

b) if $n = 2s$, then $L_n = sC_n^s + \frac{1}{2}(2^n - C_n^s) = \left(s - \frac{1}{2}\right) C_n^s + 2^{n-1} = \left(s - \frac{1}{2}\right) M_n + 2^{n-1}$.

In both cases we have the equality $L_n = \frac{(n-1)M_n}{2} + 2^{n-1}$.

A collection U of n-indexes is said to be *minimal* if:

1) the collection satisfies the conditions of the original problem, i.e., for any subset $A \subseteq \{1, 2, \dots, n\}$ there exists an index $a \in U$ such that $S_{|A|}(a) = A$;

2) the collection has exactly M_n indexes;

3) the total length of the indexes of the collection is L_n.

The above argument shows that a minimal collection of n-indexes (if such a collection exists) satisfies the following:

I. Every index of the collection has length at least l (for an odd number n this follows from the equality $C_n^s = C_n^l$).

II. The collection has exactly C_n^{l+1} indexes whose length is at least $l+1$; exactly C_n^{l+2} indexes whose length is at least $l+2$; ... ; C_n^{n-1} indexes whose length is at least $n-1$, and one ($C_n^n = 1$) index whose length is n.

An n-index a of a minimal collection is called *long* if $\|a\| \geqslant l+1$ (the total number of such indexes is $D_n = C_n^{l+1}$). All other indexes of the collection will be called *short*, their total number being $K_n = M_n - D_n$.

A minimal collection U of n-indexes is said to be *good* if it has the property

$$\begin{array}{l}\text{for any natural number } i \leqslant n \text{ and any } i\text{-element set } A \subseteq \{1, 2, \dots, n\} \\ \text{there exists an index } a \in U \text{ such that } S_i(a) = A \text{ and } \|a\| \geqslant n - i.\end{array} \qquad (*)$$

Note that it suffices to check this condition for all $i < \frac{n}{2}$ (if $i \geqslant \frac{n}{2}$, then it is immediate that $\|a\| \geqslant i$ and so $\|a\| + i \geqslant n$). It is also clear that the condition $(*)$ holds for $i = 0$.

Let a be some n-index (we assume $\|a\| \geqslant \frac{n}{2}$). Denote by $F_{n+1}(a)$ the $(n+1)$-index which is obtained by inserting the element $n+1$ into the index a (the element must be inserted so that it becomes the m-th element of the index $F_{n+1}(a)$, where $m = n + 1 - \|a\| \leqslant \|a\| + 1$).

Theorem. *For any natural number n there exists a good collection of n-indexes.*

Proof. The basis (for $n = 1$) is obvious.

The inductive step will be divided into two cases.

a) Suppose that we already have a good collection of $(2k-1)$-indexes, say, U. Define the collection V of $2k$-indexes by the formula $V = U \cup V'$, where $V' = \{F_{2k}(a) \mid a \in U\}$. Now we check that the collection V is minimal.

1) Let $A \subseteq \{1, 2, \dots, 2k\}$ and $B = A \setminus \{2k\}$, where $|B| = i$. By the inductive assumption there exists a $(2k-1)$-index $a \in U \subset V$ with the properties $S_i(a) = B$ and $\|a\| \geqslant 2k - 1 - i$. If $2k \notin A$, then $S_i(a) = A$. Now suppose that $2k \in A$. Then the $2k$-index $b = F_{2k}(a) \in V$ has the element $2k$ in the m-th place, where $m = 2k - \|a\| \leqslant i + 1$. Hence $S_{i+1}(b) = A$. So the collection V satisfies the conditions of the original problem.

2) The number of the indexes in the collection V is

$$|V| = |U| + |V'| = 2M_{2k-1} = 2C_{2k-1}^{k-1} = C_{2k-1}^{k-1} + C_{2k-1}^{k} = C_{2k}^{k} = M_{2k}.$$

3) The total length of the indexes of the collection U is L_{2k-1}, hence the total length of the indexes of the collection V' equals $L_{2k-1}+|V'|=L_{2k-1}+M_{2k-1}$. Then for the total length of the indexes of the collection V we obtain the expression

$$\begin{aligned} 2L_{2k-1}+M_{2k-1} &= 2\left(\frac{(2k-2)M_{2k-1}}{2}+2^{2k-2}\right)+M_{2k-1}= \\ &= (2k-2)M_{2k-1}+2^{2k-1}+M_{2k-1}= \\ &= (2k-1)M_{2k-1}+2^{2k-1}=\frac{(2k-1)M_{2k}}{2}+2^{2k-1}=L_{2k}, \end{aligned}$$

as desired. Thus the collection V is really minimal.

We show that the collection V is good. Let us choose $i\in\{1,2,\dots,k-1\}$ and some i-element set $A\subseteq\{1,2,\dots,2k\}$. Consider two cases.

- Let $2k\in A$; denote $B=A\setminus\{2k\}$, then $|B|=i-1$. By the inductive assumption there exists a $(2k-1)$-index $a\in U$ such that $S_{i-1}(a)=B$ and $\|a\|\geqslant 2k-i$. Then the $2k$-index $b=F_{2k}(a)\in V$ is obtained from a by inserting the element $2k$ into the m-th place, where $m=2k-\|a\|\leqslant i$. It means that $\|b\|\geqslant 2k-i$ and $S_i(b)=A$, i.e., the index b guarantees the validity of the condition $(*)$ for the collection V.
- Let $2k\notin A$. Then by the inductive assumption there exists a $(2k-1)$-index $a\in U$ with $S_i(a)=A$ and $\|a\|\geqslant 2k-1-i$. If $\|a\|\geqslant 2k-i$, then the $2k$-index $a\in V$ guarantees the validity of the condition $(*)$ for the collection V. And if $\|a\|=2k-1-i$, then the $2k$-index $b=F_{2k}(a)\in V$ differs from a only by the element $2k$ inserted into the $(i+1)$-th place. Hence $\|b\|=2k-i$ and $S_i(b)=A$. Thus the collection V satisfies the condition $(*)$.

b) Suppose that we already have a good collection of $2k$-indexes, say, U. Define the collection V of $(2k+1)$-indexes by the formula $V=V'\cup V''$, where V' is the set of all long indexes of the collection U and $V''=\{F_{2k+1}(a)\mid a\in U\}$. Let us check that V is minimal.

1) Let $A\subseteq\{1,2,\dots,2k+1\}$ and $B=A\setminus\{2k+1\}$, where $|B|=i$. By the inductive assumption there exists a $2k$-index $a\in U$ with $S_i(a)=B$ and $\|a\|\geqslant 2k-i$. If $2k+1\notin A$ and $\|a\|\geqslant k+1$, then $a\in V'\subset V$ and $S_i(a)=A$. If $2k+1\notin A$ and $\|a\|=k$, then the $(2k+1)$-index $b=F_{2k+1}(a)\in V$ differs from a only by the element $2k+1$ added to the end, so $S_i(b)=A$.

Now suppose $2k+1\in A$. Then the $(2k+1)$-index $b=F_{2k+1}(a)\in V$ has the element $2k+1$ in the m-th place, where $m=2k+1-\|a\|\leqslant i+1$. Hence $S_{i+1}(b)=A$. Thus V satisfies the conditions of the original problem.

2) The number of the indexes in the collection V is

$$|V| = |V'| + |V''| = D_{2k} + M_{2k} = C_{2k}^{k+1} + C_{2k}^{k} = C_{2k+1}^{k+1} = M_{2k+1}.$$

3) The total length of the indexes of the collection U is L_{2k}, so the total length of the indexes of the collection V'' is $L_{2k} + |V''| = L_{2k} + M_{2k}$. The total length of all short indexes of the collection U equals $k \cdot K_{2k}$, and the total length of all long indexes equals $L_{2k} - k \cdot K_{2k}$. Then for the total length of the indexes of the collection V we obtain the expression

$$2L_{2k} + M_{2k} - k \cdot K_{2k} = 2\left(\frac{(2k-1)M_{2k}}{2} + 2^{2k-1}\right) + M_{2k} - k(M_{2k} - D_{2k}) =$$
$$= (2k-1)M_{2k} + 2^{2k} + M_{2k} - kM_{2k} + kD_{2k} = kM_{2k} + 2^{2k} + kD_{2k} =$$
$$= kC_{2k}^{k} + 2^{2k} + kC_{2k}^{k+1} = kC_{2k+1}^{k+1} + 2^{2k} = kM_{2k+1} + 2^{2k} = L_{2k+1}.$$

Hence the collection V is minimal.

We show that the collection V is good. Choose $i \in \{1, 2, \dots, k\}$ and some i-element set $A \subseteq \{1, 2, \dots, 2k+1\}$. Consider two cases.

- Let $2k+1 \in A$; we designate $B = A \setminus \{2k+1\}$, then $|B| = i - 1$. By the inductive assumption there exists a $2k$-index $a \in U$ such that $S_{i-1}(a) = B$ and $\|a\| \geqslant 2k + 1 - i$. Then the $(2k+1)$-index $b = F_{2k+1}(a) \in V$ is obtained from a by inserting the element $2k+1$ into the m-th place, where $m = 2k + 1 - \|a\| \leqslant i$. It means that $\|b\| \geqslant 2k + 1 - i$ and $S_i(b) = A$, i.e., the index b guarantees the validity of the condition $(*)$ for the collection V.

- Let $2k+1 \notin A$. Then by the inductive assumption there exists a $2k$-index $a \in U$ with $S_i(a) = A$ and $\|a\| \geqslant 2k - i$. If $\|a\| \geqslant 2k + 1 - i$, then the $2k$-index a is long, i.e., the $(2k+1)$-index $a \in V' \subset V$ guarantees the validity of the condition $(*)$ for the collection V. And if $\|a\| = 2k - i$, then the $(2k+1)$-index $b = F_{2k+1}(a) \in V$ differs from a only by the element $2k+1$ inserted into the $(i+1)$-th place. Hence $\|b\| = 2k + 1 - i$ and $S_i(b) = A$. Thus the collection V satisfies $(*)$. $\square$

Good collections of n-indexes for small n

For $n = 1$:

(1)

For $n = 2$:

$(2, 1)$

(1)

For $n = 3$:

$(3, 2, 1)$

$(1, 3)$

$(2, 1)$

For $n = 4$:

$(4, 3, 2, 1)$

$(1, 4, 3)$

$(2, 4, 1)$

$(3, 2, 1)$

$(1, 3)$

$(2, 1)$

For $n = 5$:

$(5, 4, 3, 2, 1)$

$(1, 5, 4, 3)$

$(2, 5, 4, 1)$

$(3, 5, 2, 1)$

$(4, 3, 2, 1)$

$(1, 3, 5)$

$(2, 1, 5)$

$(1, 4, 3)$

$(2, 4, 1)$

$(3, 2, 1)$

For $n = 6$:

$(6, 5, 4, 3, 2, 1)$

$(1, 6, 5, 4, 3)$

$(2, 6, 5, 4, 1)$

$(3, 6, 5, 2, 1)$

$(4, 6, 3, 2, 1)$

$(5, 4, 3, 2, 1)$

$(1, 3, 6, 5)$

$(2, 1, 6, 5)$

$(1, 4, 6, 3)$

$(2, 4, 6, 1)$

$(3, 2, 6, 1)$

$(1, 5, 4, 3)$

$(2, 5, 4, 1)$

$(3, 5, 2, 1)$

$(4, 3, 2, 1)$

$(1, 3, 5)$

$(2, 1, 5)$

$(1, 4, 3)$

$(2, 4, 1)$

$(3, 2, 1)$