On the Minimum Size of Systems of Building Blocks Expressing all Intervals, and Range-Restricted Queries

Hee-Kap Ahn * Peter Brass † Hyeon-Suk Na ‡ Chan-Su Shin §

March 8, 2007

Abstract

In this paper we show that the minimum total size of a system of intervals in $\{1, \ldots, n\}$ that allows to express any interval as a disjoint union of at most k intervals of the system is $\Theta(n^{1+\frac{2}{k}})$ for any fixed k. We also prove that the minimum number of intervals $k = k(n, c)$, for which there exists a system of intervals of total size cn with that property, satisfies $k(n, c) = \Theta(n^{\frac{1}{c}})$ for any fixed integer c.

This has applications to range-restricted queries for decomposable searching problems: if we can preprocess a set of size i in time preproc(i) to answer queries in time query(i), then we can preprocess all intervals of the system in time $O(n^{\frac{2}{k}}\,\text{preproc}(n))$ to answer any interval-restricted query in time $O(\log n + k\,\text{query}(n))$.

We also discuss the situation when $k = \Theta(\log n)$, as well as higher-dimensional orthogonal range searching problems posed by Bentley and Maurer [2].

*Department of Computer Science and Engineering, POSTECH, Pohang, Korea. Email: heekap@gmail.com

†Department of Computer Science, City College, New York, USA. Email: peter@cs.ccny.cuny.edu

‡School of Computing, Soongsil University, Seoul, Korea. Email: hsnaa@computing.ssu.ac.kr. Research supported by KRF funded by the Korean Government(R04-2004-000-10004-0).

§School of Electrical and Information Engineering, Hankuk University of Foreign Studies, Yongin, Korea. Email: cssin@hufs.ac.kr

1 Introduction

In this paper we study systems of intervals in $\{1,\ldots,n\}$ with the property that for any interval $\{i,\ldots,j\} \subseteq \{1,\ldots,n\}$ there is a way to express $\{i,\ldots,j\}$ as disjoint union of at most k intervals from the system. Among all such systems we want to find one that minimizes the total size, that is, the sum of the lengths of the intervals of the system. Our main result is:

Theorem 1 *The minimum total size of a system of intervals, that allows to express any interval in $\{1,\ldots,n\}$ as disjoint union of at most k intervals of the system, is $\Theta(n^{1+\frac{2}{k}})$, for any fixed integer $k \geq 1$.*

This bound on the total size, $\Theta(n^{1+\frac{2}{k}})$, holds only for constant k, where it is tight. If k is increasing as a function of n, the growth rate of the minimum total size will get smaller. But, no matter how many intervals we allow in the decomposition, the total size of the system will be at least n, since each number i must occur in some interval. So the lowest possible growth rate of the total size is linear, and we show that any system of linear size requires $\Omega(n^{\varepsilon})$ intervals of the system to express some interval. This bound is again tight.

Theorem 2 *The minimum number of intervals $k = k(n,c)$, for which there exists a system of intervals of total size cn, that allows to express any interval in $\{1,\ldots,n\}$ as disjoint union of at most k intervals from that system, satisfies $k(n,c) = \Theta(n^{\frac{1}{c}})$ for fixed integer $c \geq 1$.*

Between these two extreme situations, the constant number of pieces and the linear total size of the system, there is another important special case, that of logarithmic number of pieces, $k = \Theta(\log n)$. Here we have only a construction: if we allow $O(\log n)$ pieces, then a total size of $O(n \log n)$ is sufficient. Lemma 1 in Section 4 implies that there is an interval system which requires total size of $\Omega(n \log n / \log\log n)$); closing the gap between $\Omega(n \log n / \log\log n)$ and $O(n \log n)$ remains open.

Theorem 3 *For each $c > 0$, there is a system of intervals of total size $f(c) n \log n$ that allows to express any interval in $\{1,\ldots,n\}$ as disjoint union of at most $\lceil c \log n \rceil$ intervals of the system.*

Here $f(c) = O(c \cdot 2^{\frac{4}{c}})$ for $c < 2$, and $f(c) = O(1)$ for $c \geq 2$.

This study was motivated by an application to the design of structures answering range-restricted queries. Suppose that we have a decomposable searching problem of evaluating a function $f(X,q)$ for some set X and query q, where the function has the property that for any partition $X = X_1 \cup X_2$ we can construct $f(X,q)$ from $f(X_1,q)$ and $f(X_2,q)$ in $O(1)$ time. Assume that we can build a structure for any set of i elements in time $\mathrm{preproc}(i)$, to evaluate the function in time $\mathrm{query}(i)$, with both $\frac{1}{i}\,\mathrm{preproc}(i)$ and $\mathrm{query}(i)$ monotone increasing. We want to preprocess X in such a way that we can construct the answer for *any* subset $Y \subseteq X$ from the answers for at most k preprocessed subsets of X whose disjoint union is Y. Then we can build the whole structure in time $O(n^{\frac{2}{k}}\,\mathrm{preproc}(n))$ to answer any query in time $O(\log n + k\,\mathrm{query}(n))$.

Data structure problems of this type occur in a number of situations. In [1, 5], a convex polygon is preprocessed such that for any halfplane, which cuts off a (cyclic) interval, one can find for any query point the farthest neighbor among the points of this interval. To compute farthest neighbor, for each interval I they construct an entire farthest Voronoi diagram, with point-location query structure; answering this point-location query takes not constant time, but time proportional to $\log|I|$, where $|I|$ is the size of the interval I. This is a decomposable searching problem stated above.

But the most classical application is the d-dimensional orthogonal range searching. For it, one considers the sequence of values of the first coordinate of the points, and builds some system of intervals such that each interval has a preprocessed $(d-1)$-dimensional orthogonal range searching structure for the points with the first coordinate in the interval. Then, for a query, one decomposes the interval of the first coordinate of the query range into a union of the preprocessed intervals, and executes the $(d-1)$-dimensional queries for those intervals, recursively.

The standard structure of this type is the orthogonal range trees [3], where any query interval can be represented as a disjoint union of $O(\log n)$ intervals in the system. This tree has the total size of $O(n \log^d n)$ and can answer the query in $O(\log^d n)$ time. But Bentley and Maurer [2] developed a fast range query structure, with the total size $O(n^{1+\varepsilon})$ and the query time $O(\log n)$. Their $(d-1)$-dimensional structure requires larger preprocessing time, but each query interval is decomposed in a constant number of intervals and thus the query time goes down to $O(\log n)$. This corresponds to our Theorem 1. At the other extreme, they gave a data structure in which the $(d-1)$-dimensional structure has the total size of $O(n)$ only, but each query must then be decomposed into many intervals, so the query time increases to $O(n^\varepsilon)$. This corresponds to our Theorem 2.

A formally related problem concerns the minimum *number* of intervals in an interval system allowing such decompositions, instead of the minimum total length. That problem was discussed in the context of the range counting problem in the semigroup model. In range counting problems, what we need from each preprocessed block is just one information, such as the number of items in there, or in the semigroup model the semigroup sum of the associated values, which is not the case in the applications we mentioned above.

In Section 2, we construct systems of intervals in $\{1, \ldots, n\}$, to prove the upper bounds of Theorem 1, 2 and 3, and then prove the corresponding lower bounds in Section 3 and 4. In Section 5, we discuss some related problems and applications.

2 The Upper Bounds: The Constructions

In what follows, we use $S(n, k)$ to denote any system of intervals in $\{1, \ldots, n\}$ with the property that any interval $\{i, \ldots, j\} \subseteq \{1, \ldots, n\}$ can be written as a disjoint union of at most k intervals from the system. The length of an interval I is the number of points in $\{1, \ldots, n\}$ contained in I, denoted by Length(I). The total size of $S(n, k)$ is the sum of the lengths of the intervals in $S(n, k)$, that is, $\text{Length}(S(n, k)) = \sum_{I \in S(n,k)} \text{Length}(I)$.

2.1 Interval Systems for Fixed k

We show that there is a system $S(n, k)$ of total size $O(n^{1+\frac{2}{k}})$, which proves the upper bound of Theorem 1.

For $k = 1$, this is trivial: the set of all intervals in $\{1, \ldots, n\}$ has size $\sum_{i=1}^{n} i(n + 1 - i) = \Theta(n^3)$.

For $k = 2$, define $S(n, 2)$ as the set of all intervals of form

$$\begin{aligned} &\{a, \ldots, \left\lceil \frac{a}{2^b} \right\rceil 2^b\} && \text{for } 1 \le a \le n,\ 0 \le b \le \lfloor \log n \rfloor, \text{ and} \\ &\{\left\lfloor \frac{a}{2^b} \right\rfloor 2^b + 1, \ldots, a\} && \text{for } 1 \le a \le n,\ 0 \le b \le \lfloor \log n \rfloor. \end{aligned}$$

That is, we connect each point a upward to the nearest multiple of 2^b and downward to the nearest multiple of 2^b plus one, for each b. Any interval $\{i, \ldots, j\} \subseteq \{1, \ldots, n\}$ contains a unique point r which is divisible by the largest power of two, possibly 2^0, and $S(n, 2)$ contains the intervals $\{i, \ldots, r\}$ and $\{r + 1, \ldots, j\}$. Thus $S(n, 2)$ represents each interval as a union of two intervals of the system. Every point a has upward and downward intervals of length at most 2^b for each $b \le \log n$, so the sum of the length of intervals starting or ending at the point a is at most $2n$. Thus the total size of $S(n,2)$ is $O(n^2)$.

For $k \geq 3$, we construct $S(n,k)$ by three steps: first we divide the interval $\{1,\ldots,n\}$ into $n^{1-\frac{2}{k}}$ pieces of equal length $n^{\frac{2}{k}}$. Within each piece, construct the system of intervals by the previous construction for $S(n^{\frac{2}{k}},2)$. The total size of these intervals is $n^{1-\frac{2}{k}}O\left((n^{\frac{2}{k}})^2\right) = O(n^{1+\frac{2}{k}})$. For the second, we connect each a to the upper and lower endpoints of its containing piece, which contributes intervals again of size $O(n \cdot n^{\frac{2}{k}}) = O(n^{1+\frac{2}{k}})$. Finally we connect all the $n^{1-\frac{2}{k}}$ endpoints of the pieces by the construction for $S(n^{1-\frac{2}{k}},k-2)$, scaled up by a factor $n^{\frac{2}{k}}$, the length of the pieces. We now have that

$$\mathrm{Length}(S(n,k)) = n^{\frac{2}{k}} \cdot \mathrm{Length}(S(n^{1-\frac{2}{k}},k-2)) + O(n^{1+\frac{2}{k}}).$$

By the inductive assumption for $\mathrm{Length}(S(n^{1-\frac{2}{k}},k-2)) = O((n^{1-\frac{2}{k}})^{1+\frac{2}{k-2}}) = O(n)$, the total size of $S(n,k)$ is $O(n^{1+\frac{2}{k}})$.

Now we show that this system has the required property. Any interval $\{i,\ldots,j\}$ within a piece can be represented as a disjoint union of two intervals constructed in the first step. Otherwise, i and j are connected to the endpoints of pieces containing them, and the interval between the endpoints can represented by $k-2$ intervals of the third step, so $\{i,\ldots,j\}$ is represented as a disjoint union of at most k intervals of $S(n,k)$.

We observe that this construction is slightly smaller than the one used by Bentley and Maurer [2] (also Falconer and Nickerson [6]), who only considered the case k odd, and gave a construction of total size $\Theta(n^{1+\frac{4}{k+1}})$. They used a construction with $\frac{k+1}{2}$ levels, where each piece of level i is divided in $n^{\frac{2}{k+1}}$ pieces of of level $i-1$, and the intervals of the system are formed by all unions of consecutive pieces of the same level that are contained in one piece of the next higher level.

2.2 Interval Systems for Increasing k

For $k = 2cn^{\frac{1}{c}}$, a system $S(n,2cn^{\frac{1}{c}})$ of total size cn consists of the intervals with base $n^{\frac{1}{c}}$ of form

$$S(n,2cn^{\frac{1}{c}}) = \left\{\{an^{\frac{b}{c}}+1,\ldots,(a+1)n^{\frac{b}{c}}\} \cap \{1,\ldots,n\} \mid a \geq 0,\, 0 \leq b \leq c-1\right\}.$$

This system is a union of collections I_b of intervals for $0 \leq b \leq c-1$, where I_b consists of $n^{1-\frac{b}{c}}$ disjoint intervals of length at most $n^{\frac{b}{c}}$ whose union is

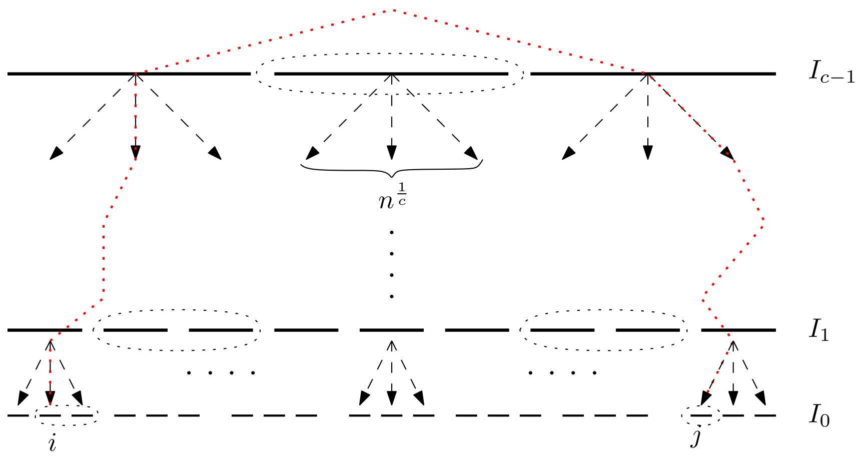

Figure 1: The structure of intervals in $S(n,k)$ when $k = 2cn^{\frac{1}{c}}$.

$\{1, \ldots, n\}$. The structure of this system is a $n^{\frac{1}{c}}$-ary tree (without the root) having intervals of I_b at the b-th level from the bottom as shown in Figure 1. Thus a query interval $\{i, \ldots, j\}$ for $i < j$ can be represented as a union of at most $2cn^{\frac{1}{c}}$ disjoint intervals, at most $2n^{\frac{1}{c}}$ taken from each level. We have a union of c collections, each of which has total size exactly n, thus the total size of this system is cn. This proves the upper bound in Theorem 2.

For $k = c \log n$, let $C = \max(2, 2^{\frac{2}{c}})$. Then the following interval system allows to express any interval as a union of at most $\lceil 2 \log_C n \rceil$ pieces (thus at most $\lceil c \log n \rceil$ pieces):

$$S(n, 2\log_C n) = \Big\{ \{a_1 C^{j-1} + bC^j + 1, \ldots, a_2 C^{j-1} + bC^j\} \cap \{1, \ldots, n\} \mid \\ 0 \le a_1 < a_2 \le C,\ b \ge 0,\ 1 \le j \le \lceil \log_C n \rceil \Big\}.$$

This system is a $\lceil \log_C n \rceil$-level structure as follows: for $j = \lceil \log_C n \rceil$ the system contains at most $\binom{C}{2}$ intervals, each of length at most n, for $j = \lceil \log_C n \rceil - 1$ the system contains C groups of at most $\binom{C}{2}$ intervals, each interval of length at most $\frac{n}{C}$, and in general for $j = \lceil \log_C n \rceil - i$, the system contains C^i groups of $\binom{C}{2}$ intervals, each interval of length at most n/C^i.

So the total size of this system is at most $\binom{C}{2} n \lceil \log_C n \rceil$, which is $O(c \cdot 2^{\frac{4}{c}} \cdot n \log n)$ for $c < 2$ and $O(n \log n)$ for $c \ge 2$. This is the construction claimed in Theorem 3.

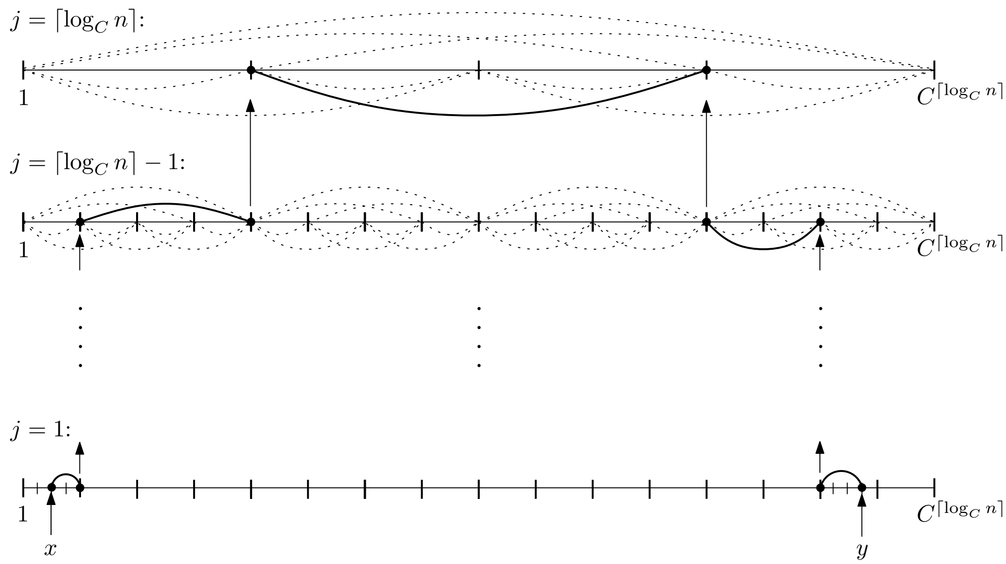

Figure 2: The structure of $S(n, 2\log_C n)$ for $C = \max(2, 2^{\frac{2}{c}})$. Note that $2\log_C n \le c \log n$.

3 The Lower Bound of Theorem 1

Here we prove that for fixed $k \ge 1$, any $S(n,k)$ must have at least $\Omega(n^{1+\frac{2}{k}})$ total size.

For $k = 1$, the lower bound $\Omega(n^3)$ is trivial; there is no choice, we have to select all intervals.

For $k = 2$ we use the following argument: the total size of $S(n,2)$ is

$$\sum_{I \in S(n,2)} \text{Length}(I) = \sum_{1 \le j \le n} |\{I \in S(n,2) \mid j \in I\}|.$$

Now consider j; if $j \le \frac{1}{2}n$, then j is contained in each of the j intervals $\{1,..,1+j\}$, ..., $\{j,\ldots,j+j\}$; note that all the left endpoints are distinct and all the right endpoints are distinct. Since each of these intervals is union of two intervals from $S(n,2)$ and any interval from $S(n,2)$ can occur here at most twice, there are at least $\frac{1}{2}j$ distinct intervals in $S(n,2)$ that contain j. In the same way, for $j \ge \frac{1}{2}n$, each of the $(n-j)$ intervals $\{j,\ldots,n\}$, $\{j-1,\ldots,n-1\}$, ..., $\{2j-n+1,..,j+1\}$ contains j and their left and right endpoints are all distinct, so there are at least $\frac{1}{2}(n-j)$ distinct intervals in

$S(n, 2)$ that contain j. Thus,

$$\sum_{1 \leq j \leq n} |\{I \in S(n,2) \mid j \in I\}| \geq \sum_{1 \leq j \leq n} \frac{1}{2} \min(j, n-j) = \Omega(n^2),$$

which is the lower bound for $k = 2$.

For $k \geq 3$, we first divide $\{1, \ldots, n\}$ into $\alpha n^{1-\frac{2}{k}}$ pieces of length $\frac{1}{\alpha} n^{\frac{2}{k}}$ for some α. We call a point *local*, if all intervals that start or end at this point do not extend beyond this piece and its immediate neighboring pieces. We call a piece *local* if it contains at least one local point; otherwise, when every point of the piece is start- or endpoint of at least one interval that extends beyond the immediate neighboring pieces, the piece is called *nonlocal*. A nonlocal piece contributes at least $\frac{1}{\alpha^2} n^{\frac{4}{k}}$ to the total size of the interval system, which, for a system of minimum total size, is less than $\beta n^{1+\frac{2}{k}}$. If we choose $\alpha = \frac{1}{2\beta}$, then the contribution of single nonlocal piece is at least $4\beta^2 n^{\frac{4}{k}}$, so of the $\frac{1}{2\beta} n^{1-\frac{2}{k}}$ pieces, at most half can be nonlocal.

Let m be the total number of local pieces. Note that $m \geq \frac{1}{4\beta} n^{1-\frac{2}{k}}$. Define a mapping ϕ that maps all the pieces to $\{0, 1, \ldots, m\}$ as follows: if the a-th piece is local, then $\phi(a)$ is the *rank* of the a-th piece counted only among the local pieces. Otherwise, i.e., if it is nonlocal, $\phi(a)$ is the rank of the last local piece preceding the a-th piece; if there is no local piece preceding the a-th piece, then let $\phi(a) = 0$. To handle the rank 0, we put a dummy interval $\{0, \ldots, 0\}$ into $S(n, k)$, which is local.

We now define a new interval system T on $\{0, \ldots, m\}$. For each interval in the original interval system $S(n, k)$ that starts in the a-th piece and ends in the b-th piece for $b \geq a \geq 0$, our new interval system T contains the intervals starting at a point in $\{\phi(a) - 1, \phi(a)\}$ and ending at a point in $\{\phi(b), \phi(b) + 1\}$. Then each interval in $S(n, k)$ contributes at most four intervals in T.

We first prove that T allows to express any interval in $\{1, \ldots, m\}$ as a union of at most $k-2$ intervals in T. To express an interval $\{\phi(a), \ldots, \phi(b)\} \subseteq \{1, \ldots, m\}$, consider the interval $\{i, \ldots, j\} \subseteq \{1, \ldots, n\}$ where i and j are the local points in the a-th and b-th pieces, respectively, which are both local. The interval $\{i, \ldots, j\}$ is expressed as a union of at most k intervals from $S(n, k)$. Then these at most k intervals of $S(n, k)$ representing $\{i, \ldots, j\}$ correspond to at most k intervals representing $\{\phi(a), \ldots, \phi(b)\}$ of T. However, we claim that the first two intervals among them in T can be replaced by one interval in T. Assume that the first interval of $S(n, k)$ starts at i and ends at i', i.e., $\{i, \ldots, i'\}$. Since the a-th piece is local, i' can be either in the same piece as i or in the next (the $(a+1)$-st piece). We now have two cases.

- Case(i): If i' is in the same (the a-th) piece as i or is in the next (the $(a+1)$-st) piece which is nonlocal, then the first interval $\{i, \ldots, i'\}$ in $S(n,k)$ corresponds to the first interval $\{\phi(a), \phi(a)\}$ in T because the a-th piece is local but $(a+1)$-st one is nonlocal. Thus the first interval in T can be discarded.

- Case(ii): If i' is in the next (the $(a+1)$-st) piece which is local, then the second interval in $S(n,k)$ starting at i' must end at i'' in the $(a+1)$-st piece or the $(a+2)$-nd piece. If i'' is in the $(a+1)$-st piece, the second interval in T is degenerated, so it can be discarded. If i'' is in the $(a+2)$-nd piece, then the first two intervals in T are $\{\phi(a), \phi(a)+1\}$ and $\{\phi(a)+1, \phi(a)+2\}$. We want to replace these two intervals by one interval $\{\phi(a), \phi(a)+1, \phi(a)+2\}$. Since $\{i, \ldots, i'\}$ is in $S(n,k)$, four intervals starting at $\phi(a)-1$, $\phi(a)$ and ending at $\phi(a)+1$, $\phi(a)+2$ are included in T. Thus the interval $\{\phi(a), \phi(a)+1, \phi(a)+2\}$ is in T, so we can replace.

Similarly, the last two intervals in T can be replaced by one interval in T. Since the new interval system T has now proven to have the required property for $k-2$, it follows by induction that its total size is at least $\Omega\left(m^{1+\frac{2}{k-2}}\right) = \Omega\left(\left(\frac{1}{4\beta}n^{1-\frac{2}{k}}\right)^{1+\frac{2}{k-2}}\right) = \Omega(n)$. But each interval of length at least two in the new system T corresponds to an interval in the original system $S(n,k)$ which was longer by a factor $n^{\frac{2}{k}}$, and each interval in the original system contributed at most 4 intervals in the new system. So the total size of the original system is at least $\Omega(n^{1+\frac{2}{k}})$. This completes the proof of Theorem 1.

4 The Lower Bound of Theorem 2

We prove the following stronger statement, from which Theorem 2 follows for $k = n^{\frac{1}{c}} - 1$.

Lemma 1 *For any $S(n,k)$, it holds*

$$\mathrm{Length}(S(n,k)) \geq n \log_{k+1} n.$$

Proof. We assign for each point $i \in \{1, \dots, n\}$ a word over the alphabet $\{1, \dots, k+1\}$. We first represent the interval $\{1, \dots, n\}$ as a union of at most k intervals from $S(n, k)$. We number the intervals used in this representation from 1 to at most k; if $i \in \{1, \dots, n\}$ belongs to the j-th interval, the word corresponding to i starts with j. Now for each interval that is used in the representation and has at least two points, we repeat the following step: we append the letter $k+1$ to the word corresponding to the last point of that interval and remove the last point from the interval. We represent the remaining part by at most k further intervals from $S(n, k)$, numbering these intervals from 1 to at most k, and appending the number to the word of each point of that interval.

By this we subdivide the intervals of $S(n, k)$, minus the last point, into further intervals of $S(n, k)$, of which we again delete the last point, and so on, until only intervals of length one are left. All the intervals we generate in this process are distinct, and the sum of their lengths is the sum of the lengths of the words we constructed. But these words form a prefix-code with n words over an alphabet of size $k+1$, so the average length of these words is at least $\log_{k+1} n$, which proves the lemma. □

5 Related Problems and Algorithmic Applications

The same type of question can be asked for any other set system: what is the minimum total size of a system of subsets (blocks) that allows the expression of all sets in the system as union of at most k blocks? The underlying model is that of range spaces: we have a universe U, a family of subsets $\mathcal{R}$, the ranges, and a set $P \subset U$ of n points. We want to express all sets $R \cap P$ for $R \in \mathcal{R}$ as a disjoint union of at most k building blocks, and minimize the total size of the building blocks we use. The measure of the size of a building block is the number of points of P it contains.

Even for the case of intervals, which we discussed in this paper, there are some open problems:

- Does the construction of Theorem 3 give the correct growth rate? Is it true that any system that allows expression of any interval with $c \log n$ building blocks has total size at least $f(c) n \log n$ for some $f(c) > 0$?

- Is it necessary to assume that the building blocks themselves are intervals? Or would we gain anything by allowing arbitrary sets? In general there might be a difference, but we believe not for intervals.

An example where it makes a big difference whether we might use sets different from those we want to express is the set system $\{1\}$, $\{1,2\}$,...,$\{1,2,\dots,n\}$. For the data structure application, we must be able to answer queries on the sets used as building blocks, which might give restrictions on the types of building blocks we can use.

- Is it necessary to assume that the building blocks are disjoint? Again we believe that in general range spaces it might make a difference, but not for intervals.

The most obvious range space one should consider are the higher-dimensional orthogonal ranges. For n points in d-dimensional space, the maximum number of distinct ranges is $\Theta(n^{2d})$, and the maximum total size of these distinct ranges is $\Theta(n^{2d+1})$ [7]. So if we want to express these sets using k building blocks, the minimum size of the building blocks for $k=1$ is $\Theta(n^{2d+1})$, and for $k=n$ is $\Theta(n)$. In between, the iteration of our construction, or that of [2] and [6], gives some upper bounds. If we just iterate a c-level subdivision, where each block on one level is divided in $n^{\frac{1}{c}}$ blocks on the level below, we obtain a structure with $k=(2c-1)^d$ and total size $O(n^{1+\frac{2d}{c}})$, so the minimum size of the block system that allows expression with k blocks is $O(n^{1+f(k,d)})$ where $f(k,d)=O(\frac{2d}{k^{1/d}})$, but we do not know whether the dependence of k is of the correct order.

Another interesting case are the intersections of the vertices of a convex polyhedron with halfspaces; one might, e.g., query for nearest or farthest vertices among these sets [4]. This looks like an immediate generalization of the convex polygon question [5], but it is not obvious that such an intersection can be expressed as a disjoint union of a small number of other intersections. Consider a near-spherical polyhedron with a large number n of vertices, all of which lie on the sphere. We project from a point on the sphere these vertices into a plane; then the subsets of polyhedron vertices cut off by a plane correspond to subsets of points in the plane cut out by some ellipse. But we cannot express a big ellipse as union of a finite number of small ellipses. So for this range space, there is no set of building blocks, for fixed k, that allows expression with at most k blocks and is significantly smaller than just taking all sets. But for $k=n^{\alpha}$, there might be such a set of blocks; we do not have to express exactly the ellipse, but only its intersection with the set of n points, which allows us some freedom of approximation.

References

[1] B. Aronov, P. Bose, E.D. Demaine, J. Gudmundsson, J. Iacono, S. Langerman, M. Smid, Data Structures for Halfplane Proximity Queries and Incremental Voronoi Diagrams. In: *LATIN 2006* (Proceedings of Latin American Theoretical Informatics), Springer LNCS **3887** (2006) 80–92.

[2] J.L. Bentley, H.A. Maurer, Efficient Worst-Case Query Structures for Range Searching, *Acta Informatica* **13** (1980) 155–168.

[3] M. de Berg, O. Cheong, M. van Kreveld, M. Overmars, *Computational Geometry: Algorithms and Applications*, 2nd Ed., Springer Verlag, (2000).

[4] O. Cheong, C.-S. Shin, A. Vigneron, Computing Farthest Neighbors on a Convex Polytope, *Theoretical Computer Science* **25** (2003) 161–170.

[5] O. Daescu, N. Mi, C.-S. Shin, A. Wolff, Furthest-Point Queries with Geometric and Combinatorial Constraints, *Computational Geometry: Theory and Applications* **33** (2006) 174–185.

[6] S.M. Falconer, B.G. Nickerson, On Multi-Level k-Ranges for Range Search, *International Journal of Computational Geometry & Applications* **15** (2005) 565–573.

[7] J.B. Saxe. On the Number of Range Queries in k-Space. *Discrete Applied Mathematics* **1** (1979) 217–225.