PARTITIONS WITH RESTRICTED BLOCK SIZES, MÖBIUS FUNCTIONS AND THE K-OF-EACH PROBLEM.

SVANTE LINUSSON

Abstract.
Given a list of n numbers in $\mathbb{R}$, one wants to decide wether every number in the list occurs at least k times. I will show that $(1-\epsilon)n\log_3(n/k)$ is a lower bound for the depth of a linear decision tree determining this problem. This is done by using the Björner-Lovász method, which turns the problem into one of estimating the Möbius function for a certain partition lattice. I will also calculate the exponential generating function for the Möbius function of a partition poset with restricted block sizes in general.

Abstrait.
Etant donnée une liste de n nombres appartiennent à $\mathbb{R}$, on veut décider si chaqu'un est répété au moins k fois. Nous montons que $(1-\epsilon)n\log_3(n/k)$ est une borne inférieure à la profondeur d'un arbre de décision linéaire pour ce problème. Nous employons la méthode de Björner-Lovász, qui réduit le problème à l'estimation de la fonction Möbius d'un certain treillis de partitions. Aussi, nous calculons en général la fonction exponentielle génératrice pour la fonction Möbius d'un poset de partitions ayant tailles de bloques restraintes.

Introduction. Identifying the list of numbers with a point $\mathbf{x} \in \mathbb{R}^n$, the problem can be viewed as deciding whether $\mathbf{x}$ belongs to a certain subset $V_{n,k}$ of $\mathbb{R}^n$. $V_{n,k}$ can be described as the union of a subspace arrangement. Given n and k let $\mathcal{A}_{n,k}$ denote the set of all linear subspaces of $\mathbb{R}^n$ defined by some equations of type $x_{i_1} = x_{i_2} = \cdots = x_{i_r}$ where $r \geq k$ such that every coordinate occurs in one of the equations. Then

$$V_{n,k} = \cup_{A\in\mathcal{A}_{n,k}} A.$$

Now the problem is to decide whether $\mathbf{x}$ is in $V_{n,k}$ or not. Partially order the elements of $\mathcal{A}_{n,k}$ by reversed inclusion, adding $\mathbb{R}^n$ as $\hat{0}$ we get the *intersection lattice* denoted $L_{n,k}$. (For a discussion of lattices and subspace arrangements see [B] or [O].) $L_{n,k}$ is isomorphic to the partition lattice $\Pi_{n,k}$ consisting of partitions of $\{1,2,\ldots,n\}$ where block sizes $1,2,\ldots,k-1$ are forbidden (Prop. 2.1). The model for deciding if a point $\mathbf{x}$ belongs to the subspace arrangement or not is a *linear decision tree*, a rooted ternary tree where at every interior node a linear function is evaluated at $\mathbf{x}$, and the three edges leaving the node are labeled "$<$", "$=$" and "$>$". I will use Theorem 3.7 in [BL], which gives a lower bound on the number $l^-(V_{n,k})$ of no-leaves in the decision tree namely:

$$l^-(V_{n,k}) \geq |\mu_{L_{n,k}}(\hat{0},\hat{1})|$$

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TeX

In Section 1 I will calculate the exponential generating function for the Möbius function of a partition poset with an arbitrary set of forbidden block sizes. In Section 2 I will then use it to get a lower bound on $|\mu_{L_{n,k}}(\hat{0},\hat{1})|$, leading to the complexity-theoretic lower bound.

1. The Möbius function of partition posets with restricted block sizes.

The intersection lattices we will be interested in have a combinatorial description in terms of set partitions. We will derive the exponential generating function for such partition posets. This is done also in [BL,Section 4], but only in the case when singleton blocks are allowed. Here we will need the case when singleton blocks are forbidden. I will treat both cases simultaneously with a method different from the one used in [BL]. The arguments are purely combinatorial and we will work in a setting more general than the application requires.

Given any set $T \subset \mathbb{Z}_+$, we consider the set $\Pi_{n,T}$ of partitions of $[n] = \{1,2,\ldots,n\}$ into blocks whose sizes are in T. Ordering the elements by refinement we get a poset, which is not a lattice in general. If $1 \notin T$ then we have to add the discrete partition $(1)(2)\ldots(n)$ to $\Pi_{n,T}$ as $\hat{0}$. We denote by $\mu_{n,T}$ the Möbius function of the poset $\Pi_{n,T}$, where the subscript n often will be suppressed. Let also $\mu_T(n) = \mu_{n,T}(\hat{0},\hat{1})$, if $n \in T$. It will be convenient to extend the definition of $\mu_{n,T}(\pi,\sigma)$ by setting it to 0 if either π or σ is not in $\Pi_{n,T}$. In particular $\mu_T(n) = 0$ if $n \notin T$.

First we need a basic recurrence formula.

Proposition 1.1. *If $n \in T\backslash\{1\}$ we have:*

$$(1) \qquad \mu_T(n) = -\sum_{\sum_{i \in T\backslash\{n\}} ic_i = n} \mu_T{}^{c_1}(1)\ldots\mu_T{}^{c_{n-1}}(n-1)\frac{n!}{\Pi(j!)^{c_j}c_j!}, \qquad if\ 1 \in T$$

$$(2) \quad \mu_T(n) = \sum_{\sum_{i \in T\backslash\{n\}} ic_i = n} (-1)^{\sum c_i}\mu_T{}^{c_2}(2)\ldots\mu_T{}^{c_{n-2}}(n-2)\frac{n!}{\Pi(j!)^{c_j}c_j!} - 1, \qquad if\ 1 \notin T$$

Proof: When $1 \in T$ we have

$$[\hat{0},(1,2,3,\ldots,l)(l+1,\ldots,n)] = [\hat{0},(1,2,\ldots,l)] \times [\hat{0},(l+1,\ldots,n)]$$

and hence $\mu_T(\hat{0},(1,2,\ldots,l)(l+1,\ldots,n)) = \mu_T(l)\mu_T(n-l)$.

We also know that there are $\frac{n!}{\prod_{j=1}^{n}(j!)^{c_j}c_j!}$ partitions of $[n]$ of type $c_1,\ldots,c_n$. By definition of the Möbius function we get the first formula.

If $1 \notin T$ we have instead

$$[\hat{0},(1,2,\ldots,l)(l+1,\ldots,n)]\backslash\{\hat{0}\} = ([\hat{0},(1,2,\ldots,l)]\backslash\{\hat{0}\}) \times ([\hat{0},(l+1,\ldots,n)]\backslash\{\hat{0}\}),$$

and hence

$$\begin{aligned}
&\mu_T(\hat{0},(1,2,\ldots,l)(l+1,\ldots,n)) = \\
&= - \sum_{\hat{0}<\pi\leq(1\ldots l)(l+1\ldots n)} \mu_T(\pi,(1,2,\ldots,l)(l+1,\ldots,n)) = \\
&= - \sum_{\substack{\hat{0}<\pi_1\leq(1\ldots l)\\ \hat{0}<\pi_2\leq(l+1\ldots n)}} \mu_T(\pi_1\times\pi_2,(1,2,\ldots,l)(l+1,\ldots,n)) = \\
&= - \sum_{\substack{\hat{0}<\pi_1\leq(1\ldots l)\\ \hat{0}<\pi_2\leq(l+1\ldots n)}} \mu_T(\pi_1,(1,2,\ldots,l))\mu_T(\pi_2,(l+1,\ldots,n)) = \\
&= -(-\mu_T(l))(-\mu_T(n-l)) = -\mu_T(l)\mu_T(n-l)
\end{aligned}$$

This gives the second equation, where the -1 term is for $\hat{0}=(1)(2)\ldots(n)$ which is not included in the sum. $\square$

The next step is to calculate the exponential generating function for each specific T. Define

$$F_T(x)=\sum_{n=1}^{\infty}\mu_T(n)\frac{x^n}{n!}$$

remembering that $\mu_T(n)=0$ if $n\notin T$. Define also for every $\pi\in\Pi_{n,\mathbb{Z}_+}$

$$s_T(\pi)=\sum_{\sigma\leq\pi}\mu_T(\hat{0},\sigma)$$

Also let $s_T(n)=s_T(\hat{1})$. By the definition of the Möbius function, we have $s_T(\pi)=0$ for all $\pi\in\Pi_{n,T},\pi\neq\hat{0}$. In particular $s_T(n)=0$ if $n\in T\backslash\{1\}$.

Remark: Note that if $n\notin T$, $s_T(n)$ gets the value that $-\mu_T(n)$ would have had if n had belonged to T. Hence if $n\notin T$ one can replace $\mu_T(n)$ by $-s_T(n)$ in Proposition 1.1.

Proposition 1.2. *The exponential generating function for $\Pi_{n,T}$ is given by:*

(3) $$F_T(x)=\ln\left(1+\sum_{n\in\mathbb{Z}_+\backslash(T\backslash\{1\})} s_T(n)\frac{x^n}{n!}\right) \qquad ,if\ 1\in T$$

(4) $$F_T(x)=-\ln\left(e^x-\sum_{n\in\mathbb{Z}_+\backslash T} s_T(n)\frac{x^n}{n!}\right) \qquad ,if\ 1\notin T$$

Proof: Case 1: $1\in T$

Using the recurrence formula in Proposition 1.1 we get,

$$0 = \sum_{n \in T \setminus \{1\}} 0 \frac{x^n}{n!} = \sum_{n \in T \setminus \{1\}} \left(\sum_{\sum_{i \in T} i c_i = n} \mu_T{}^{c_1}(1) \ldots \mu_T{}^{c_n}(n) \frac{n!}{\Pi (j!)^{c_j} c_j!} \right) \frac{x^n}{n!} =$$

$$= \prod_{j \in T} \left(1 + \frac{\mu_T(j) x^j}{j!} + \frac{\mu_T{}^2(j) x^{2j}}{(j!)^2 2!} + \ldots \right) - 1 -$$

$$- \sum_{n \in \mathbb{Z}_+ \setminus (T \setminus \{1\})} \left(\sum_{\sum_{i \in T} i c_i = n} \mu_T{}^{c_1}(1) \ldots \mu_T{}^{c_n}(n) \frac{n!}{\Pi (j!)^{c_j} c_j!} \right) \frac{x^n}{n!} =$$

$$\overset{*}{=} \prod_{j \in T} e^{\mu_T(j) \frac{x^j}{j!}} - 1 - \sum_{n \in \mathbb{Z}_+ \setminus (T \setminus \{1\})} s_T(n) \frac{x^n}{n!} =$$

$$= e^{F_T(x)} - 1 - \sum_{n \in \mathbb{Z}_+ \setminus (T \setminus \{1\})} s_T(n) \frac{x^n}{n!}$$

and the equation follows. The $*$ equality (above and below) follows from the above Remark.

Case 2: $1 \notin T$

$$0 = \sum_{n \in T} \left(\sum_{\sum_{i \in T} i c_i = n} (-1)^{\sum c_i} \mu_T{}^{c_2}(2) \ldots \mu_T{}^{c_n}(n) \frac{n!}{\Pi (j!)^{c_j} c_j!} - 1 \right) \frac{x^n}{n!} =$$

$$= \sum_{n \in T} \left(\sum_{\sum_{i \in T} i c_i = n} \frac{(-\mu_T(2) x^2)^{c_2}}{(2!)^{c_2} c_2!} \frac{(-\mu_T(3) x^3)^{c_3}}{(3!)^{c_3} c_3!} \cdots \frac{(-\mu_T(n) x^n)^{c_n}}{(n!)^{c_n} c_n!} \right) - \sum_{n \in T} \frac{x^n}{n!} =$$

$$= \prod_{j \in T} \left(1 - \frac{\mu_T(j) x^j}{j!} + \frac{\mu_T{}^2(j) x^{2j}}{(j!)^2 2!} - \ldots \right) - 1 - \sum_{n \in T} \frac{x^n}{n!} -$$

$$- \sum_{n \notin T} \left(\sum_{\sum_{i \in T} i c_i = n} \frac{(-\mu_T(2) x^2)^{c_2}}{(2!)^{c_2} c_2!} \frac{(-\mu_T(3) x^3)^{c_3}}{(3!)^{c_3} c_3!} \cdots \frac{(-\mu_T(n-2) x^{n-2})^{c_{n-2}}}{((n-2)!)^{c_{n-2}} c_{n-2}!} \right) =$$

$$\overset{*}{=} \prod_{j \in T} e^{-\mu_T(j) \frac{x^j}{j!}} - 1 - \sum_{n \in T} \frac{x^n}{n!} + \sum_{n \notin T} (s_T(n) - 1) \frac{x^n}{n!} =$$

$$= e^{-F_T(x)} - e^x + \sum_{n \notin T} s_T(n) \frac{x^n}{n!}$$

and the proposition follows. $\square$

2. The k-of-each problem.

The problem posed is the following: given a list of n numbers in $\mathbb{R}$ one wants to decide whether every number in the list occurs at least k times.

Algorithm: The following algorithm shows that the k-of-each problem can be solved using a tree with depth $8n\log_3(n/k)$. Assume (for simplicity) that $n = 2^m k$. We start with determining the $(2^{m-1}k)$-th largest element; this takes $3n$ comparisons (see [R]). Then we go on with finding the $(2^{m-2}k)$-th largest elements among those smaller and also among those larger than this element. In the j-th phase, those elements found so far split all elements into blocks of size $2^{m-j}k$, and we find the element of each block which splits it into two equal parts (where each of these special elements is counted in the block before it).

After m phases, we have found the k-th, $2k$-th, ... ,$2^m k$-th largest elements. Now we check for each element if it is equal to the special elements before or after it. If it is not equal to either one of them, then there cannot be $k-1$ other elements equal to it and the answer is No. While checking each element the tree will remember how many of them that are equal to each special element (that might be equal). If there are k of each kind then the answer is Yes, otherwise No. The total number of linear tests performed is:

$$3nm + 2n = 3n\log_2 \frac{n}{k} + 2n \leq 8n\log_3 \frac{n}{k}. \quad \square$$

We will be interested in the partition poset $\Pi_{n,k}$ where block sizes $\{1, 2, \ldots, k-1\}$ are forbidden, with the discrete partition $(1)(2)\ldots(n)$ added as zero. Observe that $\Pi_{n,k}$ is a lattice with the same join-operator as Π_n. The meet-operation is that of Π_n (coarsest common refinement) unless one gets some block of size less than k, then the meet will be $\hat{0}$. Our interest in $\Pi_{n,k}$ comes from the following proposition.

Proposition 2.1. *$L_{n,k}$ is isomorphic to $\Pi_{n,k}$.*

Proof: If $\sigma \in \Pi_{n,k}$, let

$$B_\sigma = \{x \in \mathbb{R}^n | x_i = x_j \text{ if i and j are in the same block in } \sigma\}$$

We get that $\dim B_\sigma$ =Number of blocks in σ.

It is immediate that $B_\sigma \vee B_\pi = B_\sigma \cap B_\pi = B_{\sigma\vee\pi}$ and from this follows that $L_{n,k} \cong \Pi_{n,k}$ $\square$

When $T = \mathbb{Z}_+ \backslash \{1, 2, \ldots, k-1\}$, let $\mu_{n,k}$ denote $\mu_{n,T}$, $\mu_k(n)$ denote $\mu_T(n)$, and so on. In $\Pi_{n,k}$ we have that $s_k(n) = 1$ for all $n \leq k$, so we get

$$F_k(x) = -ln(e^x - p_k(x)),$$

where $p_k(x) = \sum_{n=1}^{k-1} \frac{x^n}{n!}$. Now we have come to the main theorem of this section. It says that the above algorithm is (up to a constant) the fastest possible in the worst case.

Theorem. *The number of no-leaves in a tree deciding the k-of-each problem will be at least*

$$\frac{n!}{4n(9k)^n}$$

for at least one n in every interval of length N, for some constant N. The depth will be bounded asymptotically below for every $\epsilon > 0$ by

$$(1-\epsilon)n\log_3\frac{n}{k}. \tag{5}$$

Proof: We use the result in [B-L] Theorem 3.7, which says that the number of no-leaves in a linear decision tree is bounded below by the absolute value of the Möbius function for the corresponding intersection lattice. The theorem will be shown in the following pages using three lemmas.

Let R_k denote the radius of convergence for $F_k(x)$ considered as a function on $\mathbb{C}$. Let $z_1, \bar{z}_1, \ldots, z_t, \bar{z}_t$ denote the nonreal zeroes of $e^z - p_k(z)$ with modulus R_k. The real number $-R_k$ might also be a zero, let $\delta = 1$ if this is the case, otherwise let $\delta = 0$. Write $z_j = R_k e^{i\theta_j}$ with $0 < \theta_j < \pi$ for $j = 1, \ldots, t$. Observe that there cannot be other zeroes with modulus arbitrarily close to R_k, since an entire function with an accumulationpoint of zeroes has to be identically zero. So we can speak of the next zero which will have strictly larger modulus than R_k. Let R' denote this value (it might be infinity), which will be the radius of convergence of

$$\ln\left(\frac{e^z - p_k(z)}{(z-(-R_k))^\delta \Pi_{j=1}^t (z-z_j)(z-\bar{z}_j)}\right) = \sum_n b_n z^n.$$

As long as we are only dealing with real powerseries with a nonzero constant there is no problem using the laws of logarithm. But when it comes to separating $(z - z_j)$ from $(z - \bar{z}_j)$ we have to take care. But with the usual branchcut along the negative real axis the following caculations are valid when z is a positive real number.

$$\begin{aligned}
&\ln(z^2 - 2\mathrm{Re}(z_j)z + R_k^2) = \ln((z + R_k e^{i(\theta_j-\pi)})(z + R_k e^{-i(\theta_j-\pi)})) = \\
&= \ln(z + R_k e^{i(\theta_j-\pi)}) + \ln(z + R_k e^{-i(\theta_j-\pi)}) = \\
&= \ln R_k e^{i(\theta_j-\pi)} + \ln\left(\frac{z}{R_k e^{i(\theta_j-\pi)}} + 1\right) + \\
&+ \ln R_k e^{-i(\theta_j-\pi)} + \ln\left(\frac{z}{R_k e^{-i(\theta_j-\pi)}} + 1\right) = \\
&= \ln R_k^2 + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}\left(-\frac{z}{z_j}\right)^n + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}\left(-\frac{z}{\bar{z}_j}\right)^n = \\
&= \ln R_k^2 - \sum_{n=1}^{\infty} 2\mathrm{Re}\left(\frac{1}{z_j}\right)^n \frac{z^n}{n}
\end{aligned}$$

The calculations for a real zero are easier and all together we get:

$$\ln(e^z - p_k(z)) = \sum_{j=1}^{t} \ln(z^2 - 2\mathrm{Re}(z_j)z + R_k^2) + \delta \ln(z - (-R_k)) +$$

$$+\ln\left(\frac{e^z - p_k(z)}{(z-(-R_k))^\delta \Pi_{j=1}^t (z-z_j)(z-\bar{z}_j)}\right) =$$

$$= \sum_{j=1}^{t}\left(\ln R_k^2 - \sum_{n=1}^{\infty} 2\mathrm{Re}\left(\frac{1}{z_j}\right)^n \frac{z^n}{n}\right) + \delta\left(\ln R_k - \sum_{n=1}^{\infty}\left(-\frac{1}{R_k}\right)^n \frac{z^n}{n}\right) + \sum_n b_n z^n .$$

Comparing coefficients we get the following bound for sufficiently large n, where $1 < c < \frac{R'}{R_k}$:

$$\frac{|\mu_k(n)|}{n!} \geq \frac{|2\sum_{j=1}^t \cos(-n\theta_j) + \delta\cos(-n\pi)|}{nR_k^n} - |b_n| \geq$$

$$\geq \frac{1}{R_k^n}\left(\frac{|2\sum_{j=1}^t \cos(-n\theta_j) + \delta\cos(-n\pi)|}{n} - \frac{1}{c^n}\right)$$

Hence we need to estimate R_k from above and the sum of cosines from below. R_k is easy to estimate when k is odd and at least 5, since then there is a real root to $e^x - p_k(x)$ in the interval $[-k, 0]$. But for the general case there is a need for some heavy artillery from complex analysis. For the following Lemma I'm in debt to Daniel Bertilsson:

Modulus Lemma. *There is a zero of $e^z - p_k(z)$ with modulus less than $9k$.*

Proof: The main ingredient in the argument is the following version of Landau's theorem (see [J] and [H]): *Suppose $f : D_1 = \{z \in \mathbb{C} \mid |z| < 1\} \to \mathbb{C}\backslash\{0,1\}$ is an analytic function. Then $|f'(0)| \leq 2|f(0)|(|\ln|f(0)|| + A)$ where $A = \frac{\Gamma(1/4)^4}{4\pi^2} \approx 4.45$.*

Suppose now $e^z - p_k(z) \neq 0$, for all $z, |z| < R$. Define an analytic function $g : D_R \to \mathbb{C}$ by

$$g(z)^k = 1 - e^{-z}p_k(z)$$

There is a $\omega, \omega^k = 1$ such that $g(z) \neq \omega$ for all $z, |z| < R$. (Otherwise g would assume all k-roots of unity as values, and hence $p_k(z) = 0$ for k different $z \in \mathbb{C}$.) Define $f(z) = \frac{g(Rz)}{\omega}$, for all $z \in D_1$. The function f does not take the values 0 and 1. Landau's theorem says,

$$\left|\frac{g'(0)}{\omega}R\right| \leq 2\left|\frac{g(0)}{\omega}\right|\left(\left|\ln\left|\frac{g(0)}{\omega}\right|\right| + A\right)$$

But $kg(0)g'(0) = \frac{d}{dz}(1 - e^{-z}p_k(z)) \mid_{z=0} = -p_k'(0) = -1$ and $g(0) = 1$, so $\frac{R}{k} \leq 2A$,i.e. $R \leq 2Ak \approx 8.9k$ We can now conclude that in the disc $|z| < 9k$ there is a zero to $e^z - p_k(z)$. $\square$

Now we need to estimate the sum of cosines. We cannot hope for a bound that is valid for all n, but we can show it for enough n to draw the conclusions of the the Theorem. The following lemma is what we need with $m = 2t$ or $2t + 1$.

Lemma of Cosines. *Let $\theta_1, \theta_2, \ldots, \theta_m$ be real numbers. Then there is an integer N such that in any set of N consecutive integers, there is an integer n such that*

$$\left|\sum_{i=1}^{m} \cos n\theta_i\right| > \frac{\sqrt{m}}{2}.$$

This should be a known lemma but we have not been able to locate it in the literature so we include a proof due to Mats Boij.

Proof: For all integers n we define $f(n) = \sum_{i=1}^{m} \cos n\theta_i$, and for all integers n and N we define $g_N(n) = \sum_{k=n}^{n+N-1} f(k)^2/N$. We can compute $f(n)^2$ as

$$f(n)^2 = \sum_{i=1}^{m} \cos^2 n\theta_i + \sum_{i\neq j} 2\cos n\theta_i \cos n\theta_j = \frac{m}{2} + \frac{1}{2}\sum_{i=1}^{m} \cos 2n\theta_i$$
$$+ \sum_{i\neq j} \cos n(\theta_i + \theta_j) + \cos n(\theta_i - \theta_j).$$

We now use the following well-known formula for cosines.

$$\sum_{k=n}^{n+N-1} \cos k\varphi = \frac{\sin \frac{N\varphi}{2} \cos \frac{N+2n-1}{2}\varphi}{\sin \frac{\varphi}{2}}$$

This shows that either $\sin \varphi/2 = 0$ and $\sum_{k=n}^{n+N-1} \cos k\varphi = N$ or

$$\left|\sum_{k=n}^{n+N-1} \cos k\varphi\right| \leq \frac{1}{\sin \frac{\varphi}{2}}. \tag{6}$$

We have that

$$g_N(n) = \sum_{k=n}^{n+N-1} f(k)^2/N = \frac{m}{2} + \frac{1}{N}\sum_{k=n}^{n+N-1} \frac{1}{2}\sum_{i=1}^{m} \cos 2k\theta_i$$
$$+ \frac{1}{N}\sum_{k=n}^{n+N-1} \sum_{i\neq j} (\cos k(\theta_i + \theta_j) + \cos k(\theta_i - \theta_j)).$$

Changing the order of summation gives together with (6) that there is an integer N such that $g_N(n) > m/4$ for all integers n. But then there is an integer n in every set of N consecutive integers, such that $f(n)^2 > m/4$, that is $|f(n)| > \sqrt{m}/2$, which proves the lemma. $\square$

Now we can prove the first part of the theorem. Let M be such that $\frac{1}{c^n} < \frac{1}{4n}$ whenever $n > M$. Using the lemmas above we get that for every integer $l > M$ there is an integer n such that $|n - l| < N$ and

$$|\mu_k(n)| \geq \frac{n!}{4n(9k)^n} \geq \left(\frac{n}{3}\right)^n \frac{1}{4n(9k)^n}$$
$$\Longrightarrow \log_3 |\mu_k(n)| > n \log_3 \frac{n}{k} - 4n.$$

The last tool we need is a monotonicity lemma to prove that the depth of the linear decision trees is almost monotone with respect to n.

Monotonicity Lemma. *The depth of an optimal linear decision tree for $A_{n,k}$ is at most n more than the depth of a tree for $A_{n+r,k}$ for all $r \geq k$.*

Proof: Given an $\mathbf{x} \in \mathbb{R}^n$, find a number $x_{n+1} = x_{n+2} = \cdots = x_{n+r}$ larger than all coordinates in $\mathbf{x}$. This can be done with n comparisons. Now one can run $(\mathbf{x}, x_{n+1}, \ldots, x_{n+r})$ in an optimal tree for $A_{n+r,k}$ and get a correct answer. $\square$

Now let $n > CN$, for some large constant C. Then the depth will be larger than

$$\left(1 - \frac{1}{C}\right) n \log_3 \frac{n}{k} - 5n.$$

This suffices to prove the theorem. $\square$

Remark 1: All the three lemmas and the calculations of Section 2 remains valid if one instead of forbidding the blocksizes $1, \ldots, k-1$ forbidds any finite set of sizes that includes 1. Hence the lower bound will be valid also here with k meaning the largest forbidden blocksize plus one.

Remark 2: The Modulus lemma is not needed for odd values of k since then there is a real root. This means that one could prove the theorem for odd k easier. And for even k one can by taking the input twice turn it into the $(2k-1)$-of-each problem. This way one would prove the theorem without the Modulus lemma but get a half in front of the lower bound. But with this proof Remark 1 would no longer be valid.

Acknowledgements: I wish to thank Anders Björner who suggested the problem and helped me all the way through. I am also very grateful to Michael Benedicks and Mattias Jonsson for invaluable discussions on the analysis part.

References

[B] Björner Anders, *Subspace Arrangements*, Proceedings of 1st European Congress of Mathematics (to appear).

[BL] Björner Anders and Lovász László, *Linear decision trees, subspace arrangements and Möbius functions*, Journal of the American Mathematical Society (to appear).

[H] Hempel Joachim A., *The Poincaré metric on the Twice Punctured Plane and the Theorems of Landau and Schottky*, Journal of the London Mathematical Society 20 (1979), 435–445.

[J] Jenkins James A., *On Explicit Bounds in Landau's Theorem 2*, Canadian Journal of Mathematics 33 (1981), 559–562.

[O] Orlik Peter, *Introduction to Arrangements*, Regional conference series in mathematics no. 72, AMS, 1989.

[R] Reingold E.M., *Computing the maxima and the median*, Proc. 12th IEEE Ann. Symp. on Switching and Automata Theory, IEEE, 1971, pp. 216–218.

Department of Mathematics
Royal Institute of Technology
S-100 44 Stockholm, Sweden
E-mail: linusson@math.kth.se