Multidimensional divide-and-conquer maximin recurrences

Laurent Alonso, Edward M. Reingold, René Schott

INRIA

UNITÉ DE RECHERCHE
INRIA-LORRAINE

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P. 105
78153 Le Chesnay Cedex
France
Tél.:(1) 39 63 55 11

Rapports de Recherche

1992

25ème anniversaire

N° 1701

Programme 1
Architectures parallèles, Bases de données,
Réseaux et Systèmes distribués

MULTIDIMENSIONAL DIVIDE-AND-CONQUER MAXIMIN RECURRENCES

Laurent ALONSO
Edward M. REINGOLD
René SCHOTT

Mai 1992

RR.1701

Multidimensional Divide-and-Conquer Maximin Recurrences

Relations de Récurrence Multidimensionnelles*

Laurent Alonso[†]
Edward M. Reingold[‡]
René Schott[§]

Abstract

Bounds are obtained for the solution to the divide-and-conquer recurrence

$$M(n) = \max_{k_1+\cdots+k_p=n} (M(k_1) + M(k_2) + \cdots + M(k_p) + \min(f(k_1), \cdots, f(k_p))),$$

for nondecreasing functions f. Similar bounds are found for the recurrence with "min" replaced by "sum-of-all-but-the-max." Such recurrences appear in the analysis of various algorithms.

Résumé

Des bornes inférieures et supérieures sont données pour les solutions de la relation de récurrence du type "diviser pour régner" :

$$M(n) = \max_{k_1+\cdots+k_p=n} (M(k_1) + M(k_2) + \cdots + M(k_p) + \min(f(k_1), \cdots, f(k_p))),$$

De façon similaire, nous donnons des bornes supérieures et inférieures pour les relations obtenues en remplaçant le terme "min" par "somme des tous les termes moins le maximum". De telles récurrences apparaissent dans les analyses de divers algorithmes.

*This research was supported in part by INRIA and the NSF, through grant number NSF INT 90-16958.

[†]CRIN, INRIA-Lorraine, Université de Nancy I, 54506 Vandoeuvre-lès-Nancy, France and ENS, 45 rue d'Ulm, 75005 Paris, France

[‡]Department of Computer Science, University of Illinois at Urbana-Champaign, 1304 W. Springfield Avenue, Urbana, Illinois 61801, USA

[§]CRIN, INRIA-Lorraine, Université de Nancy I, 54506 Vandoeuvre-lès-Nancy, France

Multidimensional Divide-and-Conquer Maximin Recurrences*

Laurent Alonso[†]
Edward M. Reingold[‡]
René Schott[§]

Abstract. Bounds are obtained for the solution to the divide-and-conquer recurrence

$$M(n) = \max_{k_1+\cdots+k_p=n} (M(k_1) + M(k_2) + \cdots + M(k_p) + \min(f(k_1), \cdots, f(k_p))),$$

for nondecreasing functions f. Similar bounds are found for the recurrence with "min" replaced by "sum-of-all-but-the-max." Such recurrences appear in the analysis of various algorithms.

Key words. Recurrence relations, divide and conquer, algorithmic analysis

AMS(MOS) subject classifications. 68Q25, 68R05, 05A20, 26A12

1 Introduction

We consider the following four similar recurrences:

$$M(n) = \max_{k_1+\cdots+k_p=n,\ p\geq 2} \left(\sum_{i=1}^{p} M(k_i) + \min_{1\leq i\leq p} f(k_i) \right) \tag{1}$$

$$M(n) = \max_{k_1+\cdots+k_p=n,\ p\geq 2} \left(\sum_{i=1}^{p} M(k_i) + \operatorname*{sam}_{1\leq i\leq p} f(k_i) \right), \tag{2}$$

with $M(1)$ given, and

$$M(n) = \max_{k_1+\cdots+k_p=n} \left(\sum_{i=1}^{p} M(k_i) + \min_{1\leq i\leq p} f(k_i) \right) \tag{3}$$

$$M(n) = \max_{k_1+\cdots+k_p=n} \left(\sum_{i=1}^{p} M(k_i) + \operatorname*{sam}_{1\leq i\leq p} f(k_i) \right), \tag{4}$$

*This research was supported in part by INRIA and the NSF, through grant number NSF INT 90-16958.
[†]CRIN, INRIA-Lorraine, Université de Nancy I, 54506 Vandoeuvre-lès-Nancy, France and ENS, 45 rue d'Ulm, 75005 Paris, France
[‡]Department of Computer Science, University of Illinois at Urbana-Champaign, 1304 W. Springfield Avenue, Urbana, Illinois 61801, USA
[§]CRIN, INRIA-Lorraine, Université de Nancy I, 54506 Vandoeuvre-lès-Nancy, France

with $M(1)$, $M(2)$, ..., $M(p-1)$ given, where "sam" is the sum of all but the maximum, defined formally as

$$\operatorname*{sam}_{x \in S} f(x) = \sum_{x \in S} f(x) - \max_{x \in S} f(x).$$

Note that in recurrences (1) and (2) the maximum is over all partitions of n into *at least* two parts, while in recurrences (3) and (4) the maximum is over all partitions of n into *exactly* p parts. Divide-and-conquer recurrence relations of these types, for various functions f, occur in a variety of problems in the analysis of algorithms (all-nearest-neighbors [6], tree drawing algorithms [5], and so on). When $p = 2$, recurrence formulas (1)–(4) are identical; this case has been thoroughly investigated by Li and Reingold [4]. Our purpose is to obtain bounds for these recurrence formulas for general p, for nondecreasing f; in so doing, we sharpen one of the bounds in [4] and provide a solution to a problem left open there.[1]

In studying $M(n)$ as defined by the recurrence formulas (1)–(4), we will use trees to represent the recursive evaluation. Let $\mathcal{T}(n)$ be the set of *partition trees*: ordered trees with $n-1$ internal nodes and n external nodes (leaves) such that each internal node has at least two subtrees and such that the subtrees are in nondecreasing order, from left to right, by the number of leaves in the subtree. For a node N of a partition tree T, we denote by $\#N$ the number of leaves in the subtree rooted at N. We define the functions $\hat{F}(T)$ and $F(T)$ by

$$F(T) = \sum_{\substack{\text{leftmost} \\ \text{nodes } N \text{ of } T}} f(\#N),$$

$$\hat{F}(T) = \sum_{\substack{\text{nonrightmost} \\ \text{nodes } N \text{ of } T}} f(\#N).$$

If f is nondecreasing, the formation rule for partition trees makes the relationship between the recurrence formulas (1) and (2) and partition trees

$$M(n) = nM(1) + \max_{T \in \mathcal{T}(n)} F(T)$$

for recurrence (1) and

$$M(n) = nM(1) + \max_{T \in \mathcal{T}(n)} \hat{F}(T)$$

for recurrence (2). We will, therefore, be able to bound $M(n)$ by bounding $F(T)$ and $\hat{F}(T)$. Similar relationships hold for recurrences (3) and (4), respectively, but with the maxima taken over p-ary trees.

2 Recurrences (1) and (2)

Let $\mathcal{B}(n)$ be the set of binary partition trees; that is, partition trees in which every internal node has exactly two subtrees. Notice that for $T \in \mathcal{B}(N)$, $F(T) = \hat{F}(T)$. The following results tell us that we need only consider binary partition trees to bound $M(n)$ for recurrences (1) and (2).

[1] We note here that Li and Reingold [4] considered only the case when f is nondecreasing, claiming that the (less interesting) nonincreasing case is easily handled by induction once a simple observation has been made. This claim is wrong, as discussed in Alonso [1].

Lemma 1 *For recurrence* (1) *and* f *nonnegative and nondecreasing,*

$$M(n) = nM(1) + \max_{T \in \mathcal{B}(n)} F(T).$$

Proof. Since f is nondecreasing we know that

$$M(n) = nM(1) + \max_{T \in \mathcal{T}(n)} F(T).$$

We will prove, using a slight modification of Knuth's natural correspondence [3, page 333], that to each tree $T \in \mathcal{T}(n)$ there corresponds a binary tree $B \in \mathcal{B}(n)$ such that

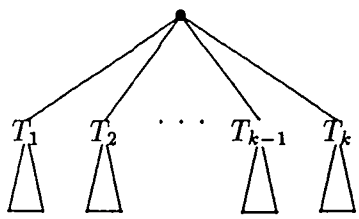
$$F(B) \geq F(T);$$

since $\mathcal{B}(n) \subset \mathcal{T}(n)$,

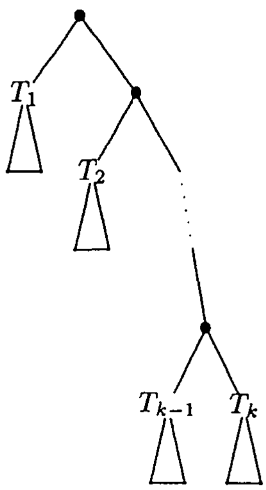
$$\max_{T \in \mathcal{B}(n)} F(T) \leq \max_{T \in \mathcal{T}(n)} F(T)$$

and we will be done.

We construct B from T inductively. If T has only binary nodes, then $B = T$. Otherwise, T has at least one internal node with three or more subtrees:



Replace this subtree with



Since f is nonnegative, this transformation does not decrease the value of F. □

Corollary 1 *For recurrence* (1) *and* f *nonnegative and nondecreasing,* $M(n)$ *satisfies recurrence* (3) *with* $p = 2$.

Lemma 2 *For recurrence* (2) *and* f *nondecreasing,*

$$M(n) = nM(1) + \max_{T \in \mathcal{B}(n)} F(T).$$

Proof. Since f is nondecreasing we know that

$$M(n) = nM(1) + \max_{T \in \mathcal{T}(n)} \hat{F}(T).$$

For any binary tree B, $\hat{F}(B) = F(B)$ since for binary trees the minimum is identical to the sum of all but the maximum. The same construction as in the previous lemma shows that for any $T \in \mathcal{T}(n)$ there corresponds a binary tree $B \in \mathcal{B}(n)$ such that $\hat{F}(T) = \hat{F}(B)$. Thus we have

$$F(B) = \hat{F}(B) = \hat{F}(T).$$

Since $\mathcal{B}(n) \subset \mathcal{T}(n)$,

$$\max_{B \in \mathcal{B}(n)} F(B) = \max_{T \in \mathcal{T}(n)} \hat{F}(T),$$

and the result follows. □

Corollary 2 *For f nondecreasing, recurrence* (2) *has the same solution as recurrence* (3) *with* $p = 2$.

Theorem 1 *For f nonnegative and nondecreasing in recurrence* (1), *and for f nondecreasing in recurrence* (2), *the solution $M(n)$ satisfies*

$$nM(1) + \sum_{j=1}^{\lceil \lg n \rceil} \lfloor n/2^j \rfloor f(2^{j-1}) + \sum_{j=1}^{l-1} f\left(\sum_{i=1}^{j} 2^{k_i}\right) \le M(n) \le nM(1) + \sum_{j=1}^{\lfloor \lg n \rfloor} \lfloor n/2^j \rfloor f(2^j) + \sum_{j=1}^{l-1} f(2^{k_j+1}),$$

where $n = 2^{k_1} + 2^{k_2} + \cdots + 2^{k_l}$, $0 \le k_1 < k_2 < \cdots < k_l$, *and* $l \le 1$.

Proof. This follows directly from the last two corollaries, together with Corollary 9 in [4]. □

Examples of these bounds applied to various functions f can be found in Table 1 in [4].

3 Recurrence (3)

In bounding the growth of $M(n)$ as defined by recurrences (3) and (4), it is necessary to make some assumptions about the initial values $M(1)$, $M(2)$, ..., $M(p-1)$. For example, we could assume that M is concave (or convex) on these values; without some such an assumption the asymptotic behavior of M would be obscured by idiosyncracies arising from these initial values. To avoid such difficulties, we will assume that M is defined *only* for n such that $(p-1)|(n-1)$, that is, n must be of the form $n = (p-1)l + 1$. This assumption is natural in the context of divide-and-conquer algorithms in which $O(p) = O(1)$ dummy elements are introduced to make the size of the input conform to the assumption. The assumption is also natural in the context of algorithms based on p-ary trees.

The tree transformation technique of the previous section does not work for recurrences (3) and (4). Instead, for recurrence (3), we will use counting arguments to bound the number of leftmost nodes with a certain range of descendant leaves. Let $\mathcal{P}(n)$ be the set of p-ary trees with $n = (p-1)l + 1$ leaves and let

$$R_i(n) = \max_{T \in \mathcal{P}(n)} \sum_{\substack{\text{leftmost nodes } N \text{ in } T \\ \text{with } p^{i-1} < \#N \le p^i}} 1.$$

Thus $R_0(n)$ is the largest possible number of leftmost leaves in a p-ary tree with n leaves, and $R_1(n)$ is the largest possible number of internal nodes that are leftmost children of their parents and have at least 2 but no more than p descendant leaves. $R_i(n) = 0$ for $i > \lfloor \log_p n \rfloor$ since a leftmost node with $p^{\lfloor \log_p n \rfloor} + 1$ or more descendant leaves has $p - 1$ siblings to its right, each of which has at least as many descendant leaves, for a total of

$$p(p^{\lfloor \log_p n \rfloor} + 1) = p^{\lfloor \log_p n \rfloor + 1} + p > p^{\log_p n} = n$$

descendant leaves, which is impossible.

We need the following generalization of the well-known observation that a binary tree with n external nodes contains $n - 1$ internal nodes.

Lemma 3 *A set of k p-ary trees with a total of I internal nodes contains a total of $I(p-1) + k$ external nodes.*

Proof. The I internal nodes have a total of Ip children, including the $I - k$ that are not the roots of the k trees. The remaining $Ip - (I - k) = I(p-1) + k$ children must be external nodes. □

We can now express an upper bound for $F(T)$ (and hence $M(n)$) for nondecreasing functions f in terms of R, since $f(p^i)$ is no smaller than the contribution of a node counted in $R_i(n)$. Thus

$$\begin{aligned} F(T) &= \sum_{\substack{\text{leftmost} \\ \text{nodes } N \text{ of } T}} f(\#N) \\ &= \frac{n-1}{p-1} f(1) + \sum_{\substack{\text{leftmost} \\ \text{nodes } N \text{ of } T}} [f(\#N) - f(1)], \end{aligned}$$

since by Lemma 3 there are $(n-1)/(p-1)$ internal nodes, each of which has a leftmost child. For f nondecreasing, $f(x) - f(1)$ is a nonnegative, nondecreasing function of x and hence

$$F(T) \leq \frac{n-1}{p-1} f(1) + \sum_{i=0}^{\lfloor \log_p n \rfloor - 1} R_i(n)[f(p^i) - f(1)] + R_{\lfloor \log_p n \rfloor}(n)[f(\lfloor n/p \rfloor) - f(1)],$$

because $f(\#N) - f(1) \leq f(p^i) - f(1)$ when $p^{i-1} < \#N \leq p^i$ and no leftmost node N can have $\#N > \lfloor n/p \rfloor$. However, for $i = 0$, $f(p^i) = f(1)$, so we have

$$F(T) \leq \frac{n-1}{p-1} f(1) + \sum_{i=1}^{\lfloor \log_p n \rfloor - 1} R_i(n)[f(p^i) - f(1)] + R_{\lfloor \log_p n \rfloor}(n)[f(\lfloor n/p \rfloor) - f(1)],$$

and hence

$$\begin{aligned} M(n) &= nM(1) + \max_{T \in \mathcal{P}(n)} F(T) \\ &\leq nM(1) + \frac{n-1}{p-1} f(1) + \sum_{i=1}^{\lfloor \log_p n \rfloor - 1} R_i(n)[f(p^i) - f(1)] \\ &\quad + R_{\lfloor \log_p n \rfloor}(n)[f(\lfloor n/p \rfloor) - f(1)] \end{aligned} \tag{5}$$

so that an upper bound on $R_i(n)$ will give us an upper bound on $M(n)$.

Let $i > 0$. Given a tree $T \in \mathcal{P}(n)$, contract it by deleting all external nodes and all internal nodes whose leftmost child is *not* counted in $R_i(n)$, preserving any parent-child relationships among internal nodes that are not deleted. Then, add an external node for every missing child among the remaining nodes so that each node is properly p-ary. The result of this contraction is a set of p-ary trees that contain among them exactly $R_i(n)$ internal nodes and, by Lemma 3, at least $R_i(n)(p-1)+1$ external nodes. By construction, each of these external nodes corresponds to a subtree of T with at least $p^{i-1}+1$ leaves (because *each* subtree of a node counted in $R_i(n)$ has at least as many descendant leaves as the leftmost subtree, which has at least $p^{i-1}+1$). Thus there must have been at least $[R_i(n)(p-1)+1](p^{i-1}+1)$ leaves in T, so that

$$n \geq [R_i(n)(p-1)+1](p^{i-1}+1), \tag{6}$$

or

$$R_i(n) \leq \left\lfloor \frac{n}{(p-1)(p^{i-1}+1)} - \frac{1}{p-1} \right\rfloor. \tag{7}$$

This bound can be strengthened when $i = 1$ since no subtree in a p-ary tree (except a leaf) can have fewer than p leaves so (6) becomes

$$n \geq [R_1(n)(p-1)+1]p,$$

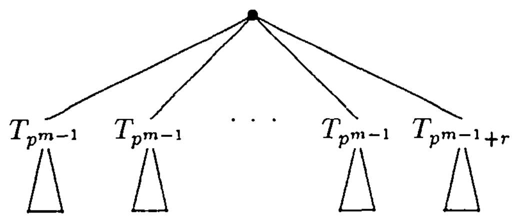
or

$$R_1(n) \leq \left\lfloor \frac{n}{p^2-p} - \frac{1}{p-1} \right\rfloor. \tag{8}$$

Thus (5), (7), and (8) combine to give us

$$\begin{aligned} M(n) \;\leq\; & nM(1) + \left(\frac{n-1}{p-1} - \left\lfloor \frac{n}{p^2-p} - \frac{1}{p-1} \right\rfloor - \sum_{i=2}^{\lfloor \log_p n \rfloor} \left\lfloor \frac{n}{(p^{i-1}+1)(p-1)} - \frac{1}{p-1} \right\rfloor \right) f(1) \\ & + \left\lfloor \frac{n}{p^2-p} - \frac{1}{p-1} \right\rfloor f(p) + \sum_{i=2}^{\lfloor \log_p n \rfloor - 1} \left\lfloor \frac{n}{(p^{i-1}+1)(p-1)} - \frac{1}{p-1} \right\rfloor f(p^i) \\ & + \left\lfloor \frac{n}{(p^{\lfloor \log_p n \rfloor - 1}+1)(p-1)} - \frac{1}{p-1} \right\rfloor f(\lfloor n/p \rfloor). \end{aligned} \tag{9}$$

To obtain a lower bound for $M(n)$ we will construct a tree $T_n \in \mathcal{P}(n)$ for which $F(T_n)$ is large. Let n be of the form $(p-1)l+1$; the tree T_n is defined recursively as follows. T_1 is the empty p-ary tree consisting of a single leaf. Given $n > 1$, let $m = \lfloor \log_p n \rfloor$ and $r = n - p^m$; T_n is formed by combining $p-1$ copies of $T_{p^{m-1}}$ on the left with a copy of $T_{p^{m-1}+r}$ on the right:



We have

$$M(n) \geq nM(1) + F(T_n), \tag{10}$$

so we need to compute $F(T_n)$.

Let $S_i(n)$ be the number of nodes N in T_n for which $\#N = p^i$; clearly,

$$S_i(p^m) = \lfloor p^{m-i} \rfloor.$$

Also, let $\hat{S}_i(n)$ be the number of nodes N in T_n for which $\#N = p^i$, but that are not on the rightmost boundary of T_n. We prove by induction on n that for $n \geq p^i$,

$$\hat{S}_i(n) = (p-1) \left\lfloor \frac{n - p^i}{(p-1)p^i} \right\rfloor . \tag{11}$$

As the basis, observe that for $n = p^i$ the formula correctly gives $\hat{S}_i(n) = 0$. Now suppose $n > p^i$ and let $m = \lfloor \log_p n \rfloor$ and $r = n - p^m$. Since $n > p^i$, we know $m \geq i$. For $m = i$, (11) correctly gives $\hat{S}_i(n) = 0$, so assume $m > i$ and we have

$$\begin{aligned}
\hat{S}_i(n) &= \hat{S}_i(p^m + r) \\
&= (p-1)S_i(p^{m-1}) + \hat{S}_i(p^{m-1} + r) \\
&= (p-1)\lfloor p^{m-1-i} \rfloor + (p-1) \left\lfloor \frac{p^{m-1} + r - p^i}{(p-1)p^i} \right\rfloor \\
&= (p-1)p^{m-1-i} + (p-1) \left\lfloor \frac{p^{m-1} + r - p^i}{(p-1)p^i} \right\rfloor \\
&= (p-1) \left\lfloor p^{m-1-i} + \frac{p^{m-1} + r - p^i}{(p-1)p^i} \right\rfloor \\
&= (p-1) \left\lfloor \frac{n - p^i}{(p-1)p^i} \right\rfloor ,
\end{aligned}$$

as desired.

Let v_i be the highest node along the right boundary of T_n such that $p^i \leq \#v_i < p^{i+1}$. We have

$$n = \#v_i + \hat{S}_i(n)p^i$$

because the leaves of T_n appear in subtrees of v_i or in one of the $\hat{S}_i(n)$ subtrees of size p^i. So (11) gives us

$$\begin{aligned}
\#v_i &= n - \hat{S}_i(n)p^i \\
&= n - (p-1) \left\lfloor \frac{n - p^i}{(p-1)p^i} \right\rfloor p^i \\
&= n - (p-1) \left(\frac{n - p^i}{(p-1)p^i} - \left\{ \frac{n - p^i}{(p-1)p^i} \right\} \right) p^i \\
&= p^i + \left\{ \frac{n - p^i}{(p-1)p^i} \right\} (p-1)p^i
\end{aligned}$$

where $\{x\} = x - \lfloor x \rfloor$ is the fractional part of x.

Now any node u of T_n *not* on the right boundary of T_n, and hence any leftmost child of a node, has $\#u = p^i$ for some i. Hence we can write

$$F(T_n) = \sum_{i=0}^{\lfloor \log_p n \rfloor} \left(\begin{array}{c} \text{number of leftmost nodes} \\ N \text{ in } T_n \text{ with } \#N = p^i \end{array} \right) f(p^i), \tag{12}$$

and the coefficient of $f(p^i)$ splits into two parts—the leftmost children whose parent is not on the right boundary of T_n and the remaining leftmost children (that are in the subtree rooted at v_{i+1}). A node not on the right boundary of T_n has p equal-size children, so the parent of such a node with p^i descendant leaves must have p^{i+1} descendant leaves. Thus there are $\hat{S}_{i+1}(n)$ such parent nodes and hence that same number of leftmost nodes N in T_n with $\#N = p^i$ and the parent of N not on the right boundary of T_n. The remaining leftmost nodes N in T_n with $\#N = p^i$ have their parents on the right boundary of T_n, so that these parent nodes each have p children, $p-1$ of which are not on the right boundary. Each of those $p-1$ non-right-boundary children is node with p^i descendant leaves—we know there are $\hat{S}_i(\#v_{i+1})$ such nodes, by definition of $\hat{S}_i$, so we have a total of $\hat{S}_i(\#v_{i+1})/(p-1)$ such parent nodes on the right boundary, each of which has a leftmost child N with $\#N = p^i$. Therefore

$$\begin{aligned}
&\begin{pmatrix} \text{number of leftmost nodes} \\ N \text{ in } T_n \text{ with } \#N = p^i \end{pmatrix} \\
&= \hat{S}_{i+1}(n) + \frac{\hat{S}_i(\#v_{i+1})}{p-1}, \\
&= (p-1)\left\lfloor \frac{n-p^{i+1}}{(p-1)p^{i+1}} \right\rfloor + \left\lfloor \frac{p^{i+1} + \left\{\frac{n-p^{i+1}}{(p-1)p^{i+1}}\right\}(p-1)p^{i+1} - p^i}{(p-1)p^{i+1}} \right\rfloor \\
&= (p-1)\left\lfloor \frac{n-p^{i+1}}{(p-1)p^{i+1}} \right\rfloor + \left\lfloor \left\{\frac{n-p^{i+1}}{(p-1)p^{i+1}}\right\} p + 1 \right\rfloor \\
&= \left\lfloor \frac{n}{p^{i+1}} + \left\{\frac{n-p^{i+1}}{(p-1)p^{i+1}}\right\} \right\rfloor,
\end{aligned} \tag{13}$$

because

$$(p-1)\lfloor x \rfloor + \lfloor \{x\}p + 1 \rfloor = \lfloor (p-1)\lfloor x \rfloor + \{x\}p + 1 \rfloor = \lfloor (p-1)x + \{x\} + 1 \rfloor.$$

Combining (10), (12), and (13) gives

$$M(n) \geq nM(1) + \sum_{i=1}^{\lfloor \log_p n \rfloor} \left\lfloor \frac{n}{p^i} + \left\{\frac{n-p^i}{(p-1)p^i}\right\} \right\rfloor f(p^{i-1}), \tag{14}$$

which we combine with (9) to give

Theorem 2 *For f nondecreasing, the function defined by recurrence* (3) *for n of the form $(p-1)l+1$ with $M(1)$ given satisfies*

$$\begin{aligned}
nM(1) &+ \sum_{i=1}^{\lfloor \log_p n \rfloor} \left\lfloor \frac{n}{p^i} + \left\{\frac{n-p^i}{(p-1)p^i}\right\} \right\rfloor f(p^{i-1}) \leq M(n) \\
&\leq nM(1) + \left(\frac{n-1}{p-1} - \left\lfloor \frac{n}{p^2-p} - \frac{1}{p-1} \right\rfloor - \sum_{i=2}^{\lfloor \log_p n \rfloor} \left\lfloor \frac{n}{(p^{i-1}+1)(p-1)} - \frac{1}{p-1} \right\rfloor \right) f(1) \\
&\quad + \left\lfloor \frac{n}{p^2-p} - \frac{1}{p-1} \right\rfloor f(p) + \sum_{i=2}^{\lfloor \log_p n \rfloor - 1} \left\lfloor \frac{n}{(p^{i-1}+1)(p-1)} - \frac{1}{p-1} \right\rfloor f(p^i) \\
&\quad + \left\lfloor \frac{n}{(p^{\lfloor \log_p n \rfloor - 1}+1)(p-1)} - \frac{1}{p-1} \right\rfloor f(\lfloor n/p \rfloor).
\end{aligned}$$

For example, when $f(x) = x$, Theorem 2 tells us that

$$\frac{1}{p} n \log_p n - O(n) \le M(n) \le \frac{p}{p-1} n \log_p n + O(n).$$

Similarly, when $f(x) = \log_p x$ we get

$$\left(M(1) + \frac{1}{(p-1)^2}\right) n - O(1) \le M(n) \le \left(M(1) + \frac{3}{(p-1)^2} + \frac{1}{p(p-1)^3}\right) n + O(1).$$

We can compare the upper and lower bounds on $M(n)$ in general for f positive. Let $U(n)$ and $L(n)$ be, respectively, the upper and lower bounds in Theorem 2. For convenience, assume $M(1) = 0$ since the $M(1)$ term occurs with the identical coefficient in both $U(n)$ and $L(n)$. Then, from (9) and (14) we have

$$\begin{aligned} U(n) &\le \sum_{i=0}^{\lfloor \log_p n \rfloor - 1} \frac{n}{(p^{i-1}+1)(p-1)} f(p^i) + \frac{n}{(p^{\lfloor \log_p n \rfloor - 1}+1)(p-1)} f(\lfloor n/p \rfloor), \\ L(n) &\ge \sum_{i=0}^{\lfloor \log_p n \rfloor - 1} \left\lfloor \frac{n}{p^{i+1}} \right\rfloor f(p^i) \end{aligned}$$

Using

$$\left\lfloor \frac{n}{p^{i+1}} \right\rfloor \ge \frac{n}{2p^{i+1}},$$

we obtain

$$\begin{aligned} U(n) &\le \frac{2p^2}{p-1} L(n) + \frac{n f(\lfloor n/p \rfloor)}{(p^{\lfloor \log_p n \rfloor}+1)(p-1)} \\ &\le \frac{2p^2}{p-1} L(n) + \frac{p^2}{p-1} f(p^{\lfloor \log_p n \rfloor - 1}) \frac{f(\lfloor n/p \rfloor)}{f(p^{\lfloor \log_p n \rfloor - 1})} \\ &\le \frac{2p^2}{p-1} L(n) + \frac{p^2}{p-1} L(n) \frac{f(\lfloor n/p \rfloor)}{f(p^{\lfloor \log_p n \rfloor - 1})}, \end{aligned}$$

and so

$$\frac{U(n)}{L(n)} \le \frac{p^2}{p-1} \left(2 + \frac{f(\lfloor n/p \rfloor)}{f(p^{\lfloor \log_p n \rfloor - 1})}\right).$$

For $n = p^m$ this can be improved to

$$\frac{U(n)}{L(n)} \le \frac{p^2}{p-1}.$$

4 Recurrence (4)

As in the previous section, we will assume that M is defined only when $(p-1)|(n-1)$, that is, only for n of the form $n = (p-1)l + 1$.

A lower bound for $M(n)$ as given by recurrence (4) follows directly from our analysis of $F(T_n)$ in the previous section—in that analysis we counted the *leftmost* children of a node; here we need to count the *nonrightmost* children and hence $p-1$ times our value for $F(T_n)$ gives a bound for $\hat{F}(T_n)$ here. Thus (14) becomes

$$\begin{aligned} M(n) &= nM(1) + \max_{T\in\mathcal{P}(n)} \hat{F}(T) \\ &\geq nM(1) + \hat{F}(T_n) \\ &= nM(1) + (p-1)F(T_n) \\ &\geq nM(1) + (p-1)\sum_{i=1}^{\lfloor \log_p n \rfloor} \left\lfloor \frac{n}{p^i} + \left\{ \frac{n-p^i}{(p-1)p^i} \right\} \right\rfloor f(p^{i-1}), \end{aligned} \tag{15}$$

To obtain an upper bound on $M(n)$ as given by recurrence (4), we will follow the strategy used in the previous section and use counting arguments to bound the number of nonrightmost nodes with a certain range of descendant leaves. Let

$$L_i(T) = \sum_{\substack{\text{nonrightmost nodes} \\ N \text{ in } T \text{ with } \#N > i}} 1,$$

(the root of the tree is considered a rightmost node) and let

$$\hat{L}_i(n) = \max_{T\in\mathcal{P}(n)} L_i(T).$$

Notice that $\mathcal{P}(n)$ is nonempty only for $n \equiv 1 \pmod{p-1}$, so that $\hat{L}_i(n)$ is defined only when $(p-1)|(n-1)$. We need the value of $\hat{L}_i(n)$ in what follows; to compute it we first observe,

Lemma 4 *The number of leaves in any tree (not necessarily of fixed arity) is one more than the number of nonrightmost nodes.*

Proof. Simple induction on the height of the tree. □

It follows from this lemma that $\hat{L}_0(n) = n-1$. Furthermore, using this lemma we can prove,

Theorem 3 *For $n > l(p-1)+1$, $n \equiv 1 \pmod{p-1}$,*

$$\hat{L}_{l(p-1)+1}(n) = \hat{L}_{l(p-1)+2}(n) = \cdots = \hat{L}_{l(p-1)+p-1}(n) = \left\lfloor \frac{n}{l(p-1)+p} \right\rfloor - 1.$$

Proof. Notice that the last of these, with $l=-1$, correctly gives $\hat{L}_0(n) = n-1$. We have

$$\hat{L}_{l(p-1)+1}(n) = \hat{L}_{l(p-1)+2}(n) = \cdots = \hat{L}_{l(p-1)+p-1}(n),$$

because $(p-1)|(\#N-1)$ for any node N in $T \in \mathcal{P}(n)$.

Take any tree $T \in \mathcal{P}(n)$, label each node N with $\#N$, and remove all nodes N of T labeled $l(p-1)+1$ or less; we thus obtain a tree T' with $L_{l(p-1)+1}(T)$ nonrightmost nodes. It follows from Lemma 4 that T' has $L_{l(p-1)+1}(T)+1$ leaves. Each of these leaves represents a subtree of T that has at least $(l+1)(p-1)+1 = l(p-1)+p$ leaves—otherwise that node would have been removed (and, the next larger possible number of leaves is $(l+1)(p-1)+1$). Thus

$$(l(p-1)+p)(L_{l(p-1)+1}(T)+1) \leq n,$$

or

$$L_{l(p-1)+1}(T) \leq \left\lfloor \frac{n}{l(p-1)+p} \right\rfloor - 1.$$

By the definition of $\hat{L}_{l(p-1)+1}(n)$, then,

$$\hat{L}_{l(p-1)+1}(n) \leq \left\lfloor \frac{n}{l(p-1)+p} \right\rfloor - 1.$$

To prove that this value is a lower bound on $\hat{L}_{l(p-1)+1}(n)$, we will construct a tree T with n leaves such that

$$L_{l(p-1)+1}(T) \geq \left\lfloor \frac{n}{l(p-1)+p} \right\rfloor - 1.$$

Let

$$u = \left\lfloor \frac{n}{l(p-1)+p} \right\rfloor$$

and

$$v = n \bmod (l(p-1)+p),$$

so that we have $u \geq 0$, $v \geq 0$,

$$n - 1 = u(l+1)(p-1) + (u+v-1),$$

and hence $u + v \equiv 1 \pmod{p-1}$ because $n \equiv 1 \pmod{p-1}$. Thus $\mathcal{P}(u+v)$ is not empty; let $T' \in \mathcal{P}(u+v)$. Replace each of the rightmost u leaves of T' by any tree from $\mathcal{P}(l(p-1)+p)$ to obtain a tree $T \in \mathcal{P}(n)$. When T is subjected to the pruning process described at the beginning of the previous paragraph, the result is a tree with at least u leaves, each representing a subtree of T that has at least $l(p-1)+p$ leaves. Hence by Lemma 4, T has at least $u-1$ nonrightmost nodes N with $\#N > l(p-1)+p$. □

Now for our analysis of (4). There are exactly $L_{i-1}(T) - L_i(T)$ nonrightmost nodes N in T with $\#N = i$ and so we have

$$\hat{F}(T) = \sum_{i\geq 1}[L_{i-1}(T) - L_i(T)]f(i). \tag{16}$$

This sum is actually finite because $L_i(T) = 0$ when $i > \lfloor (n-p)/2 \rfloor + 1$ and $T \in \mathcal{P}(n)$: If v is the nonrightmost node, aside from the root, with the largest number of descendant leaves, then v has a right sibling with at least as many descendant leaves and at least $p-2$ siblings with at least one descendant leaf each; thus

$$2\#v + p - 2 \leq n,$$

or

$$\#v \leq \left\lfloor \frac{n-p+2}{2} \right\rfloor,$$

so that

$$\#v \leq \left\lfloor \frac{n-p}{2} \right\rfloor + 1$$

for all nonrightmost nodes, as claimed.

We can use equation (16) and theorem 3 to obtain an upper bound on $\hat{F}(T)$.

Theorem 4 *Given an increasing sequence of integers* $\alpha_0 = 0 < \alpha_1 < \cdots < \alpha_k$, $\alpha_k > \lfloor (n-p)/2 \rfloor + 1$, *satisfying* $(p-1)|(\alpha_i - 1)$, $1 \le i \le k$, *and a corresponding sequence of functions* $L_i^*(n)$, $0 \le i \le k$, *satisfying, for all* $n \equiv 1 \pmod{p-1}$, $L_{i-1}^*(n) \ge L_i^*(n) \ge \hat{L}_{\alpha_i}(n)$, $L_0^*(n) = n-1$, *and* $L_k^*(n) = 0$, *then for nondecreasing* f,

$$\begin{aligned}\hat{F}(T) &\le \sum_{i=1}^{k}[L_{i-1}^*(n) - L_i^*(n)]f(\alpha_i) \\ &= \sum_{i=1}^{k-1}[L_{i-1}^*(n) - L_i^*(n)]f(\alpha_i) + L_{k-1}^*(n)f(\alpha_k),\end{aligned}$$

for all $T \in \mathcal{P}(n)$.

Proof. First we prove the result for integer-valued functions $L_i^*(n)$; afterward we will show how to remove this restriction. We have, by the definition of $\hat{F}$,

$$\hat{F}(T) = \sum_{\substack{\text{nonrightmost} \\ \text{nodes } N \text{ of } T}} f(\#N).$$

This sum of $n-1$ terms can be written as

$$\hat{F}(T) = \sum_{i=1}^{j}[L_{\beta_{i-1}}(T) - L_{\beta_i}(T)]f(\beta_i),$$

where $\beta_1 < \beta_2 < \cdots < \beta_j$ are the values assumed by $\#N$ as N ranges over the internal nodes of T; since $\#N$ assumes only values of the form $(p-1)l+1$, each β_i satisfies $(p-1)|(\beta_i - 1)$ and hence $\beta_{i-1} \le \beta_i - (p-1)$. Thus we can write $\hat{F}(T)$ as a sum of $n-1$ terms

$$\hat{F}(T) = f(\beta_1) + \cdots + f(\beta_j) \tag{17}$$

(a term $f(\beta_i)$ can occur more than one time, of course). The upper bound we want to prove has the same form, namely,

$$\sum_{i=1}^{k}[L_{i-1}^*(n) - L_i^*(n)]f(\alpha_i) = f(\alpha_1) + \cdots + f(\alpha_k), \tag{18}$$

also a sum of $\sum_{i=1}^{k}[L_{i-1}^*(n) - L_i^*(n)] = L_0^*(n) - L_k^*(n) = (n-1) - 0 = n-1$ terms. We will compare the sums in (17) and (18) term by term, showing that the tth term of (17) is less than or equal to the tth term of (18).

On the righthand side of (17) the $(n - L_{\beta_{i-1}}(T))$th to the $(n - L_{\beta_i}(T) - 1)$st terms are $f(\beta_i)$, while on the righthand side of (18), the $(n - L_{i-1}^*(n))$th to the $(n - L_i^*(n) - 1)$st terms are $f(\alpha_i)$ We will show that the tth term of (17) is no more than the tth term of (18). Suppose the tth term of (17) is $f(\beta_{\hat{i}})$. Then

$$n - L_{\beta_{\hat{i}-1}}(T) \le t < n - L_{\beta_{\hat{i}}}(T).$$

However, $L_{\beta_{\hat{i}-1}}(T) = L_{\beta_{\hat{i}}-(p-1)}(T)$ because, by definition of the β_i there are no nodes N satisfying $L_{\beta_{i-1}}(T) < \#N < L_{\beta_i}(T)$, but $\beta_{\hat{i}-1} \le \beta_{\hat{i}} - (p-1) < \beta_{\hat{i}}$. Thus,

$$n - L_{\beta_{\hat{i}}-(p-1)}(T) \le t.$$

Since $\hat{L}_i(n) \geq L_i(T)$ for any $T \in \mathcal{P}(n)$, we have

$$n - \hat{L}_{\beta_{\hat{\imath}}-(p-1)}(n) \leq t.$$

Let u be the least index for which $\alpha_u > \beta_{\hat{\imath}} - (p-1)$, and hence $\alpha_u \geq \beta_{\hat{\imath}}$; such an index u exists because $L_k^*(n) = 0$ and T has, by definition of the β_i, a node N with $\#N \geq \beta_{\hat{\imath}} - (p-1)$. But $L_{u-1}^*(n) \geq \hat{L}_{\alpha_{u-1}}(n)$ by hypothesis and $\alpha_{u-1} \leq \beta_{\hat{\imath}} - (p-1)$, so that $\hat{L}_{\alpha_{u-1}}(n) \geq \hat{L}_{\beta_{\hat{\imath}}-(p-1)}(n)$ and hence $L_{u-1}^*(n) \geq \hat{L}_{\beta_{\hat{\imath}}-(p-1)}(n)$. Thus

$$n - L_{u-1}^*(n) \leq t,$$

so that the tth term of (18) is $f(\alpha_t) \geq f(\alpha_u) \geq f(\beta_{\hat{\imath}})$ (because f is nondecreasing), which is what we wanted to prove.

We now show how to reduce the non-integer-valued case to the integer-valued case. Let the functions $L_i^*(n)$ be given and define

$$\Lambda_i^*(n) = \lfloor L_i^*(n) \rfloor.$$

We have

$$L_i^*(n) \geq \hat{L}_{\alpha_i}(n),$$

but the $\hat{L}_{\alpha_i}(n)$ are integer-valued so that

$$\Lambda_i^*(n) \geq \hat{L}_{\alpha_i}(n),$$

and we can apply our theorem, giving

$$\hat{F}(T) \leq \sum_{i=1}^{k} [\Lambda_{i-1}^*(n) - \Lambda_i^*(n)] f(\alpha_i),$$

for all $T \in \mathcal{P}(n)$. But

$$\sum_{i=1}^{k} [\Lambda_{i-1}^*(n) - \Lambda_i^*(n)] f(\alpha_i) \leq \sum_{i=1}^{k} [L_{i-1}^*(n) - L_i^*(n)] f(\alpha_i),$$

for consider the difference

$$\begin{aligned}
&\sum_{i=1}^{k} [L_{i-1}^*(n) - L_i^*(n)] f(\alpha_i) - \sum_{i=1}^{k} [\Lambda_{i-1}^*(n) - \Lambda_i^*(n)] f(\alpha_i) \\
&= \sum_{i=1}^{k} [L_{i-1}^*(n) - \Lambda_{i-1}^*(n) - L_i^*(n) + \Lambda_i^*(n)] f(\alpha_i) \\
&= \sum_{i=1}^{k} [\{L_{i-1}^*(n)\} - \{L_i^*(n)\}] f(\alpha_i) \\
&= \{L_0^*(n)\} f(\alpha_1) + \sum_{i=1}^{k-1} [f(\alpha_i) - f(\alpha_{i-1})] \{L_i^*(n)\} - \{L_k^*(n)\} f(\alpha_k) \\
&= \sum_{i=1}^{k-1} [f(\alpha_i) - f(\alpha_{i-1})] \{L_i^*(n)\} \\
&\geq 0,
\end{aligned}$$

since $L_0^*(n) = n - 1$, $L_k^*(n) = 0$, and f is nondecreasing. □

This last theorem has several interesting corollaries. First, there is a p-dimensional analogue of Corollary 9 in [4]:

Corollary 3 *For f nondecreasing, the function defined by recurrence (4) for n of the form $(p-1)l+1$ with $M(1)$ given satisfies*

$$M(n) \leq nM(1) + n(p-1) \sum_{i=1}^{\lfloor \log_p((n-p+2)/2) \rfloor} \frac{f(p^i)}{p^i} + \left(\frac{n}{p^{\lfloor \log_p((n-p+2)/2) \rfloor}} - 1 \right) f\left(\left\lfloor \frac{n-p+2}{2} \right\rfloor \right).$$

Proof. Apply Theorem 4 with $k = \lfloor \log_p((n-p+2)/2) \rfloor$, $\alpha_0 = 0$, $\alpha_i = p^i$, $1 \leq i < k$, $\alpha_k = (n-p+2)/2$, $L_i^*(n) = n/p^i - 1$, $1 \leq i < k$, $L_k^*(n) = 0$. □

Next, a much tighter upper bound:

Corollary 4 *For f nondecreasing, the function defined by recurrence (4) for n of the form $(p-1)l+1$ with $M(1)$ given satisfies*

$$\begin{aligned} M(n) \leq nM(1) \quad &+ \quad n \sum_{i=0}^{\lfloor (l-3)/2 \rfloor} \left(\frac{1}{(p-1)i+1} - \frac{1}{(p-1)i+p} \right) f((p-1)i+1) \\ &+ \quad \left(\frac{n}{(p-1)\lfloor (l-1)/2 \rfloor + 1} - 1 \right) f\left((p-1) \left\lfloor \frac{(l-1)}{2} \right\rfloor + 1 \right). \end{aligned}$$

Proof. Apply Theorem 4 with $k = \left\lfloor \frac{l+1}{2} \right\rfloor = \left\lfloor \frac{n+p-2}{2(p-1)} \right\rfloor$, $\alpha_i = (i-1)(p-1)+1$ and $L_i^*(n) = n/(\alpha_i + p - 1) - 1$, for $1 \leq i < k$, $\alpha_0 = 0$ and $L_0^*(n) = n - 1$, and $\alpha_k = \lfloor \frac{n-p+2}{2} \rfloor$ and $L_k^*(n) = 0$. □

For $p = 2$ the bound in Corollary 4 becomes

$$M(n) \leq nM(1) + n \sum_{i=1}^{\lfloor n/2 \rfloor - 1} \frac{f(i)}{i(i+1)} + O(f(\lfloor n/2 \rfloor)), \tag{19}$$

a big improvement over the upper bound given in Corollary 9 in [4]. For example, when $f(x) = x$ and $M(1) = 0$ we know (see [2, eq. 2.50], for example), that $M(n) = \frac{1}{2} n \log_2 n + O(n)$. The upper bound (19) gives $M(n) \leq n \ln n + O(n) \approx 0.693 \cdots n \log_2 n + O(n)$, while the result in [4] gives only $M(n) \leq n \log_2 n + O(n)$. When $f(x) = \log_2 x$ and $M(1) = 0$, (19) gives $M(n) \leq 1.137 \cdots n + O(\log n)$, while the result in [4] gives only $M(n) \leq 2n + O(\log n)$.

When $f(x) = \lceil \log_2 x \rceil$ we know from [4] that $M(n) = nM(1) + n - \lceil \log_2 n \rceil - 1$ and we can use summation by parts (see, for example, [2, eq. 4.65] or [3, ex. 1.2.7-10]) with (15) and (19) to obtain

$$\begin{aligned} n\, M(1) + n + O(\log^2 n) \leq M(n) \quad &\leq \quad nM(1) + n \sum_{i=0}^{\infty} \frac{1}{2^i + 1} + O(\log n) \\ &\approx \quad nM(1) + 1.2645 \cdots n + O(\log n). \end{aligned}$$

Most interesting is the case $f(x) = \lfloor \log_2 x \rfloor$. We know from [4] that $M(n) = nM(1) + n - \lfloor \log_2 n \rfloor - \nu(n)$, where $\nu(n) = O(\log n)$ is the number of 1-bits in the binary representation of n. Using summation by parts with (15) and (19), we obtain the sharp result that $M(n) = nM(1) + n + O(\log^2 n)$.

Corollary 5 *For f nondecreasing, the function defined by recurrence (4) for n of the form $(p-1)l+1$ with $M(1)$ given satisfies*

$$M(n) \leq nM(1) + \sum_{i=0}^{\lfloor (l-1)/2 \rfloor} \left(\left\lfloor \frac{n}{(p-1)i+1} \right\rfloor - \left\lfloor \frac{n}{(p-1)i+p} \right\rfloor \right) f((p-1)i+1).$$

Proof. Apply Theorem 4 with $k = \left\lfloor \frac{l+1}{2} \right\rfloor = \left\lfloor \frac{n+p-2}{2(p-1)} \right\rfloor$, $\alpha_i = (i-1)(p-1)+1$ and $L_i^*(n) = \hat{L}_{\alpha_i}(n) = \left\lfloor \frac{n}{(i-1)(p-1)+p} \right\rfloor - 1$, for $1 \leq i \leq k$, and $\alpha_0 = 0$ and $L_0^*(n) = n-1$. □

The upper bound in Corollary 5 is sharper than that in Corollaries 3 and 4; in fact, it is sharper than *any* other upper bound of the same form. Let n be fixed and consider the partial order on the set $\mathcal{U}(n)$ of upper bounds for $M(n)$ of the form

$$V(f) = nM(1) + \sum_{i=1}^{k} v_i f(i) \tag{20}$$

that hold for all nondecreasing functions f. Upper bounds V and W are comparable, $V \prec W$, if $V(f) \leq W(f)$ for all nondecreasing functions f; V and W are incomparable if there exist nondecreasing functions f and g such that $V(f) < W(f)$ and $V(g) > W(g)$.

Lemma 5 *Let*

$$V(f) = nM(1) + \sum_{i=1}^{k} v_i f(i)$$

and

$$W(f) = nM(1) + \sum_{i=1}^{k} w_i f(i).$$

Then $V \prec W$ if and only if for all j, $1 \leq j \leq k$,

$$\sum_{i=j}^{k} v_i \leq \sum_{i=j}^{k} w_i. \tag{21}$$

Proof. If $V \prec W$, then (21) follows by considering the step function

$$f_j(x) = \begin{cases} 0 & x < j, \\ 1 & x \geq j. \end{cases}$$

On the other hand, suppose that (21) holds. To prove that $V \prec W$, we must show that $V(f) \leq W(f)$ for all nondecreasing functions f. Let $\hat{f}(x) = f(x) - f(1)$; $\hat{f}(x)$ is a non-negative, nondecreasing function of x and moreover,

$$W(f) - V(f) = W(\hat{f}) - V(\hat{f}),$$

so we need only prove that $V(f) \leq W(f)$ for non-negative, nondecreasing functions f. Such a function can be written as a linear combination of the step functions $f_j(x)$,

$$f(x) = \sum_{j=1}^{k} c_j f_j(x) + h(x),$$

where $h(x) = 0$ when x is an integer, $1 \leq x \leq k$, and the c_i are non-negative. Because V only uses f at integers, we have $V(f) = V(f - h)$, and we have

$$\begin{aligned} V(f) &= V(f-h) \\ &= nM(1) + \sum_{i=1}^{k} v_i[f(i) - h(i)] \\ &= nM(1) + \sum_{i=1}^{k} v_i \sum_{j=1}^{k} c_j f_j(i) \\ &= nM(1) + \sum_{i \geq j} v_i c_j, \end{aligned}$$

since $f_j(i)$ is 1 if $i \geq j$, and 0 otherwise. Thus,

$$V(f) = nM(1) + \sum_{j=1}^{k} c_j \sum_{i=j}^{k} v_i,$$

and so by (21),

$$\begin{aligned} V(f) &\leq nM(1) + \sum_{j=1}^{k} c_j \sum_{i=j}^{k} w_i \\ &= W(f), \end{aligned}$$

by a similar argument. □

Lemma 6 *Let*

$$V(f) = nM(1) + \sum_{i=1}^{k} v_i f(i).$$

Then, for $1 \leq j \leq k$, $\sum_{i=j}^{k} v_i \geq \hat{L}_{j-1}(n)$.

Proof. Suppose $\sum_{i=j}^{k} v_i < \hat{L}_{j-1}(n)$. Consider the step function

$$f_j(x) = \begin{cases} 0 & \text{if } x < j, \\ 1 & \text{otherwise.} \end{cases}$$

We have

$$\begin{aligned} M(n) &= nM(1) + \max_{T \in \mathcal{P}(n)} \hat{F}(T) \\ &= nM(1) + \max_{T \in \mathcal{P}(n)} \sum_{\substack{\text{nonrightmost} \\ \text{nodes } N \text{ of } T}} f_j(\#N) \\ &= nM(1) + \max_{T \in \mathcal{P}(n)} \sum_{\substack{\text{nonrightmost nodes} \\ N \text{ in } T \text{ with } \#N > j-1}} 1 \end{aligned}$$

$$
\begin{aligned}
&= nM(1) + \max_{T\in\mathcal{P}(n)} L_{j-1}(T) \\
&= nM(1) + \hat{L}_{j-1}(n) \\
&> nM(1) + \sum_{i=j}^{k} v_i \\
&= V(f_j),
\end{aligned}
$$

contradicting the fact that V is an upper bound. □

Theorem 5 *The upper bound of Corollary* 5,

$$V(f) = nM(1) + \sum_{i=0}^{\lfloor (l-1)/2 \rfloor} \left(\left\lfloor \frac{n}{(p-1)i+1} \right\rfloor - \left\lfloor \frac{n}{(p-1)i+p} \right\rfloor \right) f((p-1)i+1),$$

is the minimum element of the partial order $\mathcal{U}(n)$*; that is,* V *is in* $\mathcal{U}(n)$ *and is less than or equal to any other element in* $\mathcal{U}(n)$.

Proof. The upper bound V is of the form (20) with $k = (p-1)\lfloor (l-1)/2 \rfloor + 1$, $l = (n-1)/(p-1)$,

$$v_i = \begin{cases} \hat{L}_{i-1}(n) - \hat{L}_i(n) & \text{if } (p-1)|(i-1), \\ 0 & \text{otherwise.} \end{cases}$$

so that V is in $\mathcal{U}(n)$.

Given any element $W \in \mathcal{U}(n)$, $W(f) = nM(1) + \sum_{i=1}^{k} w_i f(i)$, pad the shorter of V and W with zeroes so the two are the same length. By Lemma 6, for $1 \le j \le k$,

$$
\begin{aligned}
\sum_{i=j}^{k} w_i &\ge \hat{L}_{j-1}(n) \\
&= \sum_{i=j}^{k} [\hat{L}_{i-1}(n) - \hat{L}_i(n)] \\
&= \sum_{i=j}^{k} v_i,
\end{aligned}
$$

so that $V \prec W$ by Lemma 5. □

Combining all of these results yields

Theorem 6 *For* f *nondecreasing, the function defined by recurrence* (4) *for* n *of the form* $(p-1)l+1$ *with* $M(1)$ *given satisfies*

$$
\begin{aligned}
& nM(1) + (p-1) \sum_{i=1}^{\lfloor \log_p n \rfloor} \left\lfloor \frac{n}{p^i} + \left\{ \frac{n-p^i}{(p-1)p^i} \right\} \right\rfloor f(p^{i-1}) \\
& \le M(n) \\
& \le nM(1) + \sum_{i=0}^{\lfloor (l-1)/2 \rfloor} \left(\left\lfloor \frac{n}{(p-1)i+1} \right\rfloor - \left\lfloor \frac{n}{(p-1)i+p} \right\rfloor \right) f((p-1)i+1)
\end{aligned}
$$

$$\leq \; nM(1) + n \sum_{i=0}^{\lfloor (l-3)/2 \rfloor} \left(\frac{1}{(p-1)i+1} - \frac{1}{(p-1)i+p} \right) f((p-1)i+1)$$

$$+ \left(\frac{n}{(p-1)\lfloor (l-1)/2 \rfloor + 1} - 1 \right) f\left((p-1) \left\lfloor \frac{(l-1)}{2} \right\rfloor + 1 \right)$$

$$\leq \; nM(1) + n(p-1) \sum_{i=1}^{\lfloor \log_p \frac{n-p+2}{2} \rfloor} \frac{f(p^i)}{p^i} + \left(\frac{n}{p^{\lfloor \log_p \frac{n-p+2}{2} \rfloor}} - 1 \right) f\left(\left\lfloor \frac{n-p+2}{2} \right\rfloor \right).$$

Proof. All of the inequalities except that between the two larger upper bounds follow immediately from our preceding discussion. Let $W(f)$ be the largest of the three upper bounds and let $V(f)$ be the second largest of the three upper bounds. We must show that $V \prec W$ in order to prove the theorem. Write $V(f)$ in the form

$$V(f) = \sum_{i=1}^{\lfloor \frac{n-p+2}{2} \rfloor} (v_i - v_{i+1}) f(i)$$

and write $W(f)$ in the form

$$W(f) = \sum_{i=1}^{\lfloor \frac{n-p+2}{2} \rfloor} (w_i - w_{i+1}) f(i),$$

with

$$v_i = \frac{n}{j(p-1)+1} - 1$$

when $(j-1)(p-1)+1 < i \leq j(p-1)+1$,

$$w_i = \frac{n}{p^{j-1}} - 1$$

when $p^{j-1} < i \leq p^j$, and

$$v_{\lfloor \frac{n-p+4}{2} \rfloor} = w_{\lfloor \frac{n-p+4}{2} \rfloor} = 0.$$

Since $v_i \leq w_i$ for each i, it follows that $V \prec W$ by Lemma 5. □

For example, when $f(x) = x$, using the lower bound and the middle of the three upper bounds in Theorem 6 tells us that

$$\frac{p-1}{p} n \log_p n + O(n) \leq M(n) \leq n \ln n + O(n).$$

Similarly, when $f(x) = \log_p x$, we get

$$\left(M(1) + \frac{1}{p-1} \right) n + O(\log^2 n) \leq M(n) \leq \left(M(1) + \frac{1}{p-1} + \frac{1}{(p-1)\ln p} \right) n + O(\log n),$$

when $f(x) = \lceil \log_p x \rceil$ we get

$$\left(M(1) + \frac{1}{p-1} \right) n + O(\log^2 n) \leq M(n) \leq \left(M(1) + \frac{3}{2p-2} \right) n + O(\log n),$$

and when $f(x) = \lfloor \log_p x \rfloor$, we get the sharper result that

$$M(n) = \left(M(1) + \frac{1}{p-1}\right) n + O(\log^2 n).$$

We can compare the upper and lower bounds in Theorem 6 on $M(n)$ for f positive. Let $U(n)$ be the largest of the three upper bounds in Theorem 6 and let $L(n)$ be the lower bound in Theorem 6. For convenience, assume $M(1) = 0$ since the $M(1)$ term occurs with the identical coefficient in both $U(n)$ and $L(n)$. Then, from Corollary 3 and (15), the same calculations that we did for recurrence (3) lead to

$$\frac{U(n)}{L(n)} \leq 2p \left(1 + \frac{f(\lfloor \frac{n-p+2}{2} \rfloor)}{f(p^{\lfloor \log_p (n-p) \rfloor - 1})}\right)$$

for recurrence (4) because

$$L(n) \geq \frac{n(p-1)}{2p} \sum_{i=0}^{\lfloor \log_p n \rfloor - 1} \frac{f(p^i)}{p^i}$$

and

$$U(n) \leq n(p-1) \sum_{i=1}^{\lfloor \log_p \frac{n-p+2}{2} \rfloor} \frac{f(p^i)}{p^i} + pf\left(\left\lfloor \frac{n-p+2}{2} \right\rfloor\right).$$

5 Conclusions

It is worth noting that, in contradistinction to the binary case explored in [4], even as strong a property as the concavity of f is insufficient to determine the exact location of the maximum in recurrences (3) and (4). For example, in recurrence (3) with $p = 3$ and f nondecreasing and concave, direct calculation gives unique values for $M(n)$, $3 \leq n \leq 53$, n odd, but gives

$$\begin{aligned} M(55) &= 55M(1) + \max\{19f(1) + 5f(3) + f(5) + 2f(9),\ 19f(1) + 6f(3) + f(9) + f(13)\} \\ &= 55M(1) + 19f(1) + 5f(3) + f(9) + \max\{f(5) + f(9),\ f(3) + f(13)\}, \end{aligned}$$

which is indeterminate given only that f is nondecreasing and concave: For $f(x) = x$, $f(3) + f(13)$ is larger while for $f(x) = \ln x$, $f(5) + f(9)$ is larger; both of these functions are nondecreasing and concave. In recurrence (4) with $p = 3$ and f is nondecreasing and concave, direct calculation gives unique values for $M(n)$, $3 \leq n \leq 13$, n odd, but gives

$$\begin{aligned} M(15) &= 15M(1) + \max\{11f(1) + 2f(3) + f(7),\ 10f(1) + 4f(3)\} \\ &= 15M(1) + 10f(1) + 2f(3) + \max\{f(1) + f(7),\ 2f(3)\}, \end{aligned}$$

which is similarly indeterminate.

Thus we leave it as an open problem to find general conditions on f under which the exact location of the maximum in recurrences (3) and (4) is determined. Such a condition would likely involve the signs of the differences Δf, $\Delta^{(2)} f, \ldots, \Delta^{(p)} f$, just as the case $p = 2$ involves conditions on the signs of Δf (that is, f nondecreasing) and $\Delta^{(2)} f$ (that is, f concave or convex).

References

[1] Alonso, L., Thèse de doctorat, Université de Paris 11, Orsay, 1992.

[2] Greene, D. H. and D. E. Knuth, *Mathematics for the Analysis of Algorithms*, 2nd ed., Birkhäuser, Boston, 1982.

[3] Knuth, D. E., *The Art of Computer Programming: Volume I, Fundamental Algorithms*, 2nd ed., Addison-Weslsey Publishing Co., Reading, MA, 1973.

[4] Li, Z. and E. M. Reingold, "Solution of a Divide-and-Conquer Maximin Recurrence," *SIAM J. Comp.* **18** (1989) 1188–1200.

[5] Tilford, J. S., "Tree Drawing Algorithms," Report UIUCDCS-R-81-1055, Department of Computer Science, University of Illinois, Urbana, IL, 1981. Much (but not all) of the contents of this report appears in "Tidier Drawings of Trees" by E. M. Reingold and J. S. Tilford, *IEEE Trans. Software Engineering* **7** (1981), 223–228.

[6] Vaidya, P., "An $O(n \log n)$ Algorithm for the All-Nearest-Neighbors Problem," *Disc. Comp. Geom.* **4** (1989), 101–115.

Imprimé en France
par
l'Institut National de Recherche en Informatique et en Automatique

ISSN 0249-6399