On Asymptotic Notation with Multiple Variables

Rodney R. Howell Dept. of Computing and Information Sciences Kansas State University Manhattan, KS 66506 USA rhowell@ksu.edu

Technical Report 2007-4

January 18, 2008

Abstract

We show that it is impossible to define big-O notation for functions on more than one variable in a way that implies the properties commonly used in algorithm analysis. We also demonstrate that common definitions do not imply these properties even if the functions within the big-O notation are restricted to being strictly nondecreasing. We then propose an alternative definition that does imply these properties whenever the function within the big-O notation is strictly nondecreasing.

1 Introduction

Big-O notation for functions on one variable was introduced by Bachman in 1894 [1]. In the following decades, its properties became well-understood, and it has been widely used in the analysis of algorithms. Little-o notation for 1-variable functions was introduced by Landau in 1909 [7], and big- Ω , big- Θ , and little- ω were defined for single-variable functions by Knuth in 1976 [6]. As with big-Onotation for one variable, these asymptotic notations are all well-understood and widely used in algorithm analysis.

Many algorithms have more than one natural parameter influencing their performance. For example, the running times of most graph algorithms depend on both the number of vertices and the number of edges when the graph is represented by an adjacency list. However, most textbooks on algorithm analysis (e.g., [3, 4, 5, 8]) do not explicitly extend the definition of asymptotic notation to multi-variable functions. Instead, they include analyses using asymptotic notation with multiple variables, applying the properties shown for single-variable asymptotic notation. The implicit assumption is that these properties extend

```
\begin{array}{c} {\rm F}(m,n) \ \{ \\ {\rm \ for \ } i \leftarrow 0 \ {\rm to \ } m-1 \ {\rm do \ } \{ \\ {\rm \ G}(i,n) \\ \} \\ \} \end{array}
```

to more than one variable. In this paper, we demonstrate that this assumption, which seems to pervade the computing literature, is invalid.

Consider, for example, the algorithm shown in Figure 1.1, and suppose we are told that G(i, n) runs in O(in) time. Knowing this fact, most computer scientists would not hesitate to conclude that F(m, n) runs in $O(m^2n)$ time. However, we show in this paper that this conclusion does not necessarily follow when any of the common definitions of big-O notation are used. Specifically, we can find an algorithm G(i, n) such that, using a common definition of big-O, G(i, n) runs in O(in) time, but F(m, n) does not run in $O(m^2n)$ time. In fact, in order to obtain any upper bound on the running time of F(m, n), it is necessary to know more about the behavior of G(i, n) than standard asymptotic notation provides.

Ideally, we would like to define big-O so that the commonly-assumed properties all hold, even when multi-variable functions are used. However, we prove in this paper that it is impossible to define big-O notation for multi-variable functions in a way that implies all of these properties. Specifically, we first identify five properties of big-O notation for unary functions. We argue that these five properties are commonly assumed to hold for big-O notation on multiple variables, particularly in the analysis of algorithms. We then show that it is impossible to satisfy all five properties simultaneously.

The proof of this impossibility relies on a particular set O(g(m, n)) such that g(m, n) decreases at infinitely many points; specifically, there are infinitely many pairs (m, n) such that g(m, n) > g(m + 1, n). However, the functions used in algorithm analysis are nearly always strictly nondecreasing. Unfortunately, we demonstrate that the common definitions of big-O notation do not satisfy all five properties even when the functions within the notation are strictly nondecreasing. We then propose an alternative definition that implies all five properties, provided the functions within the asymptotic notation are strictly nondecreasing. Furthermore, given this restriction, we show that our definition implies all of the properties that are commonly assumed to hold. Finally, we extend this definition to big- Ω , big- Θ , little- ω , and prove similar results for these definitions.

The remainder of the paper is organized as follows. In Section 2, we identify five properties of big-O notation on one variable, and show that they are inconsistent for multi-variable functions. In Section 3, we present our proposed definition of big-O for multiple variables, and show that it satisfies the desired properties when the functions within the notation are strictly nondecreasing. In Section 4, we extend these results to big- Ω , and in Section 5, we extend our results to big- Θ , little-o, and little- ω . Finally, we give some concluding remarks in Section 6.

2 Inconsistent Properties

Let $f: \mathbb{N} \to \mathbb{R}^{\geq 0}$, where \mathbb{N} denotes the natural numbers and $\mathbb{R}^{\geq 0}$ denotes the nonnegative real numbers, and let $\mathbb{R}^{>0}$ denote the strictly positive real numbers. O(f(n)) is then defined to be the set of all functions $g: \mathbb{N} \to \mathbb{R}^{\geq 0}$ such that there exist $c \in \mathbb{R}^{>0}$ and $N \in \mathbb{N}$ such that for every $n \geq N$

$$g(n) \le cf(n).$$

If we wish to extend this definition to multi-variable functions $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$, we would expect, at a minimum, that the set $O(f(n_1, \ldots, n_k))$ should satisfy the following property:

Property 1 For every function $g(n_1, \ldots, n_k) \in O(f(n_1, \ldots, n_k))$, there exist $c \in \mathbb{R}^{>0}$ and $N \in \mathbb{N}$ such that for every $n_1 \geq N, n_2 \geq N, \ldots, n_k \geq N$,

$$g(n_1,\ldots,n_k) \le cf(n_1,\ldots,n_k)$$

It is this property that gives $O(f(n_1, \ldots, n_k))$ its "asymptotic" nature. Note that a single threshold N is sufficient, because if the inequality holds whenever each variable is sufficiently large, where "sufficiently large" is defined as a separate threshold for each variable, then we can pick N to be the maximum of these individual thresholds.

A common definition of $O(f(n_1, \ldots, n_k))$ is exactly the set of all functions satisfying Property 1 (see, e.g., [2, 9]). In order to distinguish this definition from others, let us denote it as $O_{\forall}(f(n_1, \ldots, n_k))$. We will now demonstrate that this definition is too weak.

Consider, for example, the algorithm shown in Figure 1.1, and suppose we know that G(i, n) runs in $O_{\forall}(in)$ time. Applying standard algorithm analysis techniques, we would then conclude that F(m, n) runs in $O_{\forall}(m^2n)$ time. However, this conclusion is invalid. To see why, consider the function $g: \mathbb{N}^2 \to \mathbb{R}^{\geq 0}$ such that

$$g(m,n) = \begin{cases} 2^n & \text{if } m = 0\\ mn & \text{if } m \ge 1. \end{cases}$$
(2.1)

Because $g(m, n) \leq mn$ whenever both m and n are at least $1, g(m, n) \in O_{\forall}(mn)$, but

$$\sum_{i=0}^{m-1} g(i,n) = 2^n + \sum_{i=1}^{m-1} mn$$
$$= 2^n + \frac{mn(m-1)}{2}$$
$$\notin O_{\forall}(m^2n).$$

This example illustrates a property that is commonly assumed for asymptotic notation, but which does not hold for $O_{\forall}(f(n_1,\ldots,n_k))$:

Property 2 Let $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ be an eventually positive function; i.e., for some $N \in \mathbb{N}, f(n_1, \ldots, n_k) > 0$ whenever $n_1 \geq N, \ldots, n_k \geq N$. Then for every $g(n_1, \ldots, n_k) \in O(f(n_1, \ldots, n_k))$ and every positive integer $j \leq k$,

$$\sum_{i_{j}=0}^{n_{j}} g(n_{1}, \dots, n_{j-1}, i_{j}, n_{j+1}, \dots, n_{k})$$

$$\in O\left(\sum_{i_{j}=0}^{n_{j}} f(n_{1}, \dots, n_{j-1}, i_{j}, n_{j+1}, \dots, n_{k})\right).$$

We have required that f be eventually positive because otherwise, this property fails even for k = 1, as $g(n) \in O(0)$ may have finitely many positive values. This property is used whenever we analyze a **for** loop in which the index variable runs from 0 to n_j . It is easily shown to hold for big-O with k = 1, but as we have seen above, it does not necessarily hold for O_{\forall} with k > 1. This is actually a less general property than we really need, as it deals only with indices ranging from 0 to n_j ; however, it is sufficient to demonstrate the inconsistencies in the commonly-assumed properties.

Note from the above discussion that we have shown that Property 2 does not hold for O_{\forall} when f is chosen to be f(m, n) = mn — a strictly nondecreasing function. We therefore have the following theorem:

Theorem 2.1 Property 2 does not necessarily hold for $O_{\forall}(f(n_1, \ldots, n_k))$ even if $f(n_1, \ldots, n_k)$ is strictly nondecreasing.

A stronger definition of $O(f(n_1, \ldots, n_k))$ is the set of all functions $g : \mathbb{N} \to \mathbb{R}^{\geq 0}$ such that there exist $c \in \mathbb{R}^{>0}$ and $N \in \mathbb{N}$ such that

$$g(n_1,\ldots,n_k) \le cf(n_1,\ldots,n_k)$$

whenever $n_j \ge N$ for some positive integer $j \le k$. Let us denote this set as $O_{\exists}(f(n_1, \ldots, n_k))$. Functions in this set may violate the above inequality for at most finitely many points. As a result, it is not hard to show that Properties 1 and 2 both hold for this definition. On the other hand, because mn = 0

at infinitely many points and mn + 1 > 0 for all m and n, it follows that $mn + 1 \notin O_{\exists}(mn)$.

One property that is used to show that $n + 1 \in O(n)$ is the summation rule, namely, that $O(f_1(n)) + O(f_2(n)) = O(\max(f_1(n), f_2(n)))$. (In the expression, " $O(f_1(n)) + O(f_2(n))$ ", we are using the standard technique of adding two sets of functions; i.e., if A and B are sets of functions mapping k-tuples of natural numbers into the nonnegative reals, then A + B is the set of all functions $g_1(n_1, \ldots, n_k) + g_2(n_1, \ldots, n_k)$ such that $g_1 \in A$ and $g_2 \in B$.) It is not hard to show that when generalized to multiple variables, the summation rule holds for O_{\exists} . However, there is one more rule involved that allows us to conclude that $O(\max(n, 1)) \subseteq O(n)$ — namely, that if $g(n) \leq f(n)$ when n is sufficiently large, then $O(g(n)) \subseteq O(f(n))$. In order to generalize this rule to multiple variables, we need to consider how it applies to algorithm analysis.

We typically deal with the point-wise maximum of two functions either as a result of applying the summation rule when analyzing two blocks of code executed in sequence or as a result of analyzing an **if** statement. For example, consider again Figure 1.1. We can separate the initialization code consisting of the assignment of 0 to *i* from the remainder of the code. Suppose, for the sake of argument, that we have analyzed the loop (excluding the initialization) and discovered that it runs in O(mn) time. We then have code running in O(1)time followed by code running in O(mn) time. In order to avoid a complicated expression like O(mn + 1) or $O(\max(mn, 1))$, we would like to conclude that the algorithm runs in O(mn) time. We therefore need a rule that allows us to use the maximum of mn and 1 as both m and n become sufficiently large. We therefore need the following property:

Property 3 Let $f_1 : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ and $f_2 : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ be two functions such that for some $N \in \mathbb{N}$, and every $n_1 \geq N, \ldots, n_k \geq N$,

$$f_1(n_1,\ldots,n_k) \le f_2(n_1,\ldots,n_k).$$

Then

$$O(f_1(n_1,\ldots,n_k)) \subseteq O(f_2(n_1,\ldots,n_k)).$$

By taking $f_1(m,n) = \max(mn,1)$, and $f_2(m,n) = mn$, we can see that Property 3 does not hold for O_{\exists} . Furthermore, these functions are strictly nondecreasing. We therefore have the following theorem:

Theorem 2.2 Property 3 does not necessarily hold for $O_{\exists}(f_1(n_1,\ldots,n_k))$ and $O_{\exists}(f_2(n_1,\ldots,n_k))$ even if $f_1(n_1,\ldots,n_k)$ and $f_2(n_1,\ldots,n_k)$ are strictly nondecreasing.

It is not surprising that Property 3 does not always hold for O_{\exists} , as Property 3 is very nearly the converse of Property 1; hence, when we combine Properties 1 and 3, we will have very nearly restricted our possible definitions to just O_{\forall} . Furthermore, because Property 2 fails for O_{\forall} , our properties are already very nearly inconsistent. However, we still need one additional property to ensure that we cannot trivially satisfy all the properties by making $O(f(n_1, \ldots, n_k)) = \emptyset$ for all $f(n_1, \ldots, n_k)$.

Property 4 For all functions $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$, $f(n_1, \ldots, n_k) \in O(f(n_1, \ldots, n_k))$.

Theorem 2.3 For k > 1, Properties 1–4 are inconsistent.

Proof: We will prove the theorem for k = 2, as it will be easy to see how it can be generalized to larger k. The proof is by contradiction.

Suppose O is defined so that all four properties hold. Let g(m, n) be as defined in (2.1). By Property 3, $O(g(m, n)) \subseteq O(mn)$; hence, by Property 4, $g(m, n) \in O(mn)$. By Property 2, we have

$$\sum_{i=0}^{m} g(i,n) \in O\left(\sum_{i=0}^{m} in\right)$$
$$2^{n} + \frac{mn(m+1)}{2} \in O\left(\frac{mn(m+1)}{2}\right),$$

which clearly contradicts Property 1.

Observe that in the above proof, the functions used within the big-O notation are all eventually nondecreasing; i.e., there is some natural number N such that whenever $m \ge N$ and $n \ge N$, $f(m, n) \le f(m+1, n)$ and $f(m, n) \le f(m, n+1)$. However, the function g(m, n) is not strictly nondecreasing. In particular, we used Properties 3 and 4 with O(g(m, n)). Because functions used within big-O notation in algorithm analysis typically are strictly nondecreasing, it might make sense to consider whether these properties are still inconsistent when the functions are restricted in this way.

Rather than proceeding to such an examination, however, we instead need to consider one additional property that also causes problems. This property is used when we analyze a loop in which the running time of the body of the loop is independent from the iteration number. In this case, we would typically compute the total running time by multiplying the running time of the body by the number of iterations. We therefore need the following property:

Property 5 For all functions $f_1 : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ and $f_2 : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$,

$$O(f_1(n_1,...,n_k))O(f_2(n_1,...,n_k)) \subseteq O(f_1(n_1,...,n_k)f_2(n_1,...,n_k)).$$

Theorem 2.4 For k > 1, Properties 1–5 are inconsistent even if Properties 1–3 are restricted so that the functions within the asymptotic notation are strictly nondecreasing.

Proof: Again, we will proceed by contradiction with k = 2. Suppose O is defined so that all five properties hold. Let g(m, n) be as defined in (2.1).

Observe that for all natural numbers m and n, $mn g(m, n) = m^2 n^2$; hence, by Property 5,

$$O(mn)O(g(m,n)) \subseteq O(m^2n^2).$$

Because both $\max(mn, 1)$ and mn are strictly nondecreasing, by Property 3, $O(\max(mn, 1)) \subseteq O(mn)$; hence,

$$O(\max(mn, 1))O(g(m, n)) \subseteq O(m^2n^2).$$

From Property 4, we therefore have

$$\max(mn, 1)g(m, n) \in O(m^2n^2).$$

Now by Property 2, we have

$$\begin{split} \sum_{i=0}^{m} (\max(in,1)g(i,n)) &\in O\left(\sum_{i=0}^{m} i^2 n^2\right) \\ g(0,n) + \sum_{i=1}^{m} in \, g(i,n)) &\in O\left(\frac{mn^2(m+1)(2m+1)}{6}\right) \\ 2^n + \frac{mn^2(m+1)(2m+1)}{6} &\in O\left(\frac{mn^2(m+1)(2m+1)}{6}\right), \end{split}$$

which clearly contradicts Property 1.

We conclude from the above theorem that in addition to restricting at least one of Properties 1–4, we must also restrict one of Properties 4 and 5. In the next section, we propose a definition big-O notation for multiple variables that implies Properties 1 and 4, as well as restricted versions of Properties 2, 3 and 5.

3 A Proposed Solution

In view of Theorems 2.1 and 2.2, we need a new definition of big-O. In view of Theorems 2.3 and 2.4, we cannot expect any alternative definition to satisfy Properties 1–5 without restriction. Furthermore, we would like for our definition to be a generalization of the accepted definition for a single variable.

It would appear that in order to meet these constraints, our definition must strengthen Property 1 by utilizing, in some way, the function values at points for which one or more of the parameters is small. To formalize this idea, we define, for any function $f: \mathbb{N}^k \to \mathbb{R}^{\geq 0}$,

$$f(n_1, \dots, n_k) = \max\{f(i_1, \dots, i_k) \mid 0 \le i_j \le n_j \text{ for } 1 \le j \le k\}.$$

We then define, for $f: \mathbb{N}^k \to \mathbb{R}^{\geq 0}$, $\widehat{O}(f(n_1, \ldots, n_k))$ to be the set of all functions $g: \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ such that there exist $c \in \mathbb{R}^{>0}$ and $N \in \mathbb{N}$ such that for all $n_1 \geq N, \ldots, n_k \geq N$, the following conditions hold:

$$g(n_1,\ldots,n_k) \le cf(n_1,\ldots,n_k); \tag{3.1}$$

$$\widehat{g}(n_1,\ldots,n_k) \le c\widehat{f}(n_1,\ldots,n_k). \tag{3.2}$$

Condition (3.1) is simply Property 1; hence,

$$O(f(n_1,\ldots,n_k)) \subseteq O_{\forall}(f(n_1,\ldots,n_k))$$

for every $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$. Condition (3.2) restricts the set so that function values for small parameters are considered. Thus, for example, it is not hard to see that $g(m,n) \notin \widehat{O}(mn)$, where g(m,n) is as defined in (2.1). On the other hand, even infinitely many large function values will not cause the function to be excluded if these values are not too large in comparison to growth of the function as all variables increase. Thus, for example, by taking N = 1 and c = 2, we can see that $mn + 1 \in \widehat{O}(mn)$.

Note that even though the definition uses only one real constant c and one natural number constant N, this is only for simplicity of the definition. Certainly, if a proof were to show the two conditions with different constants, we could use the larger of the two real values and of the two natural number values to conclude that the two conditions hold simultaneously.

Condition (3.2) of the definition of \hat{O} makes it a bit more complicated than either the definition of O_{\forall} or the definition of O_{\exists} . However, as we have already pointed out, the definition of big-O is rarely used in practice. Instead, properties of big-O are used to analyze algorithms. If we can show that \hat{O} satisfies the properties we expect, then the definition itself will be rarely needed. It will, however, give us a formal justification for asymptotic analyses of algorithms with more than one natural parameter.

Before we begin proving properties of \hat{O} , let us first consider how it compares, when k = 1, to the standard definition of big-O for single-variable functions. Unfortunately, it is not hard to see that $O(0) \not\subseteq \hat{O}(0)$, as $\hat{O}(0)$ contains only the constant function 0. However, we will show that, with this one exception, the definitions are equivalent.

Theorem 3.1 Let $f : \mathbb{N} \to \mathbb{R}^{\geq 0}$ such that for some $n \in \mathbb{N}$, f(n) > 0. Then $\widehat{O}(f(n)) = O(f(n))$.

Proof:

 \subseteq : Let $g(n) \in \widehat{O}(f(n))$. Then by Condition (3.1), $g(n) \in O(f(n))$.

 \supseteq : Let $g(n) \in O(f(n))$. Then there exist $c \in \mathbb{R}^{>0}$ and $N \in \mathbb{N}$ such that for every $n \ge N$,

$$g(n) \le cf(n);$$

hence, Condition (3.1) holds.

Because f(i) > 0 for some $i \in \mathbb{N}$, we can assume without loss of generality that for some i < N, f(i) > 0. We now let

$$c' = \max\left(c, \frac{\widehat{g}(N-1)}{\widehat{f}(N-1)}\right)$$

and

Then for $n \geq N$,

$$\widehat{g}(n) = \max(\widehat{g}(N-1), \max\{g(i) \mid N \le i \le n\})$$

$$\leq \max(c'\widehat{f}(N-1), \max\{cf(i) \mid N \le i \le n\})$$

$$\leq c'\widehat{f}(n).$$

Therefore, Condition (3.2) holds; hence, $g(n) \in \widehat{O}(f(n))$.

We will now begin examining the properties of \widehat{O} . First, the following theorem follows immediately from the definition of \widehat{O} :

Theorem 3.2 Properties 1 and 4 hold for \widehat{O} .

Rather than simply addressing the remaining properties from Section 2, we will instead consider the most basic properties of asymptotic notation first, then move on to more involved properties. Hence, we will now observe that the following theorem also follows immediately from the definition:

Theorem 3.3 For any positive integer k, positive real number c, and $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$, $\widehat{O}(f(n_1, \ldots, n_k)) = \widehat{O}(cf(n_1, \ldots, n_k))$.

In addition, the summation rule is easily shown to hold.

Theorem 3.4 For any positive integer $k, f_1 : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$, and $f_2 : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$,

$$\widehat{O}(f_1(n_1,\ldots,n_k)) + \widehat{O}(f_2(n_1,\ldots,n_k))$$

= $\widehat{O}(\max(f_1(n_1,\ldots,n_k),f_2(n_1,\ldots,n_k))).$

As we pointed out in Section 2, when we use the summation rule, we need a way of dealing with the point-wise maximum of two functions. Specifically, we would like for Property 3 to hold; however, Theorem 2.3 suggests that this may be impossible. Indeed, if we take g(m, n) as defined in (2.1), then whenever m and n are at least 1, $g(m, n) \leq mn$; however, as we have already pointed out, $g(m, n) \notin \hat{O}(mn)$. We therefore need to weaken the property.

Specifically, we will show that Property 3 holds for \widehat{O} when the function f_1 is strictly nondecreasing. First, however, because several of the properties will need to be weakened in this way, let us consider how a strictly nondecreasing function $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ impacts the definition of $\widehat{O}(f(n_1, \ldots, n_k))$. In particular, we have the following theorem.

Theorem 3.5 If $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ is strictly nondecreasing, then $\widehat{O}(f(n_1, \ldots, n_k))$ is exactly the set of functions satisfying Condition (3.2).

Proof: We will show that for strictly nondecreasing f, Condition (3.2) implies Condition (3.1). The theorem will then follow.

Because f is strictly nondecreasing,

~

$$f(n_1,\ldots,n_k)=f(n_1,\ldots,n_k)$$

for all natural numbers n_1, \ldots, n_k . Furthermore, for any function $g: \mathbb{N}^k \to \mathbb{R}^{\geq 0}$,

$$\widehat{g}(n_1,\ldots,n_k) \ge g(n_1,\ldots,n_k)$$

for all natural numbers n_1, \ldots, n_k . Therefore, whenever

$$\widehat{g}(n_1,\ldots,n_k) \le c\widehat{f}(n_1,\ldots,n_k),$$

it follows that

$$g(n_1,\ldots,n_k) \le cf(n_1,\ldots,n_k)$$

Using Theorem 3.5, we can now show the following theorem.

Theorem 3.6 Property 3 holds for \widehat{O} when the function f_1 is strictly nondecreasing.

Proof: Let $g(n_1, \ldots, n_k) \in \widehat{O}(f_1(n_1, \ldots, n_k))$. Then for some $c \in \mathbb{R}^{>0}$ and $N \in \mathbb{N}$, whenever $n_1 \ge N, \ldots, n_k \ge N$,

$$g(n_1,\ldots,n_k) \le cf_1(n_1,\ldots,n_k),$$

and

$$\widehat{g}(n_1,\ldots,n_k) \le c\widehat{f}_1(n_1,\ldots,n_k)$$

Without loss of generality, we can assume that $f_1(n_1, \ldots, n_k) \leq f_2(n_1, \ldots, n_k)$ whenever $n_1 \geq N, \ldots, n_k \geq N$. Because f_1 is strictly nondecreasing, we also have

$$\widehat{g}(n_1, \dots, n_k) \le c\widehat{f}_1(n_1, \dots, n_k)$$

$$= cf_1(n_1, \dots, n_k)$$

$$\le cf_2(n_1, \dots, n_k)$$

$$\le c\widehat{f}_2(n_1, \dots, n_k).$$

From Theorem 3.5,

$$g(n_1,\ldots,n_k)\in O(f_2(n_1,\ldots,n_k))$$

Theorem 3.6 allows us to analyze sequential blocks of code and **if**-statements. For example, suppose we have an **if**-statement in which the guard requires $\hat{O}(1)$ time, one branch requires $\hat{O}(1)$ time, and the other branch requires $\hat{O}(mn)$ time. We can first use Theorem 3.6 to conclude that each branch runs in $\widehat{O}(mn)$ time and that execution of the guard also runs in $\widehat{O}(mn)$ time. Then using the summation rule, we can conclude that the entire statement runs in $\widehat{O}(mn)$ time.

We are now ready to examine Property 5. By Theorem 2.4, it would appear that we need to restrict this property, as we did Property 3. Indeed, it is not hard to show, following the first part of the proof of Theorem 2.4, that

$$O(mn)O(g(m,n)) \not\subseteq O(mn g(m,n)),$$

where g(m, n) is as defined in (2.1). We therefore will restrict both f_1 and f_2 in Property 5 to be nondecreasing.

Theorem 3.7 Property 5 holds for \widehat{O} when the functions f_1 and f_2 are strictly nondecreasing.

Proof: Let

$$g_1(n_1,\ldots,n_k) \in O(f_1(n_1,\ldots,n_k))$$

and

 $g_2(n_1,\ldots,n_k)\in \widehat{O}(f_2(n_1,\ldots,n_k)).$

Then there exist positive real numbers c_1 and c_2 and a natural number N such that whenever $n_1 \ge N, \ldots, n_k \ge N$, each of the following hold:

$$g_1(n_1, \dots, n_k) \le c_1 f_1(n_1, \dots, n_k);$$

$$g_2(n_1, \dots, n_k) \le c_2 f_2(n_1, \dots, n_k);$$

$$\widehat{g_1}(n_1, \dots, n_k) \le c_1 \widehat{f_1}(n_1, \dots, n_k);$$
 and

$$\widehat{g_2}(n_1, \dots, n_k) \le c_2 \widehat{f_2}(n_1, \dots, n_k).$$

Let $c = c_1 c_2$. Then for $n_1 \ge N, \ldots, n_k \ge N$, because f_1 and f_2 are strictly nondecreasing,

$$\begin{split} \widehat{g_1g_2}(n_1, \dots, n_k) &\leq \widehat{g_1}(n_1, \dots, n_k) \widehat{g_2}(n_1, \dots, n_k) \\ &\leq c_1 c_2 \widehat{f_1}(n_1, \dots, n_k) \widehat{f_2}(n_1, \dots, n_k) \\ &= c f_1(n_1, \dots, n_k) f_2(n_1, \dots, n_k) \\ &= c \widehat{f_1f_2}(n_1, \dots, n_k). \end{split}$$

Because $f_1 f_2$ must be strictly nondecreasing, by Theorem 3.5,

$$g_1(n_1, \dots, n_k)g_2(n_1, \dots, n_k) \in O(f_1(n_1, \dots, n_k)f_2(n_1, \dots, n_k)).$$

The only remaining property from Section 2 is Property 2. Unfortunately, this property does not hold for arbitrary $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ and $g : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$. For example, suppose we define

$$f(m,n) = \begin{cases} 2^n & \text{if } m = 0, n \text{ even} \\ mn & \text{otherwise.} \end{cases}$$

Then taking g as defined in (2.1), it is easily seen that $g(m,n) \in \widehat{O}(f(m,n))$. However, for every odd n,

$$\sum_{i=0}^{m} g(i,n) = 2^{n} + \frac{mn(m+1)}{2},$$

whereas

$$\sum_{i=0}^{m} f(i,n) = \frac{mn(m+1)}{2}.$$

Therefore,

$$\sum_{i=0}^{m} g(i,n) \notin \widehat{O}\left(\sum_{i=0}^{m} f(i,n)\right).$$

We therefore need to assume that f is strictly nondecreasing.

Furthermore, as we mentioned in Section 2, we often need a more general property in which the bounds of the summation index are arbitrary. Therefore, in what follows, we will show that this more general property holds when f is strictly nondecreasing. In order to simplify the notation, we will state the theorem so that the last variable is the summation index; however, it should be clear that the proof works when any variable is used as the summation index.

Theorem 3.8 Let $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ be a strictly nondecreasing function. Then for any $g(n_1, \ldots, n_k) \in \widehat{O}(f(n_1, \ldots, n_k))$, any natural number m, and any strictly nondecreasing unbounded function $h : \mathbb{N} \to \mathbb{N}$,

$$\sum_{i=m}^{h(n_k)} g(n_1,\ldots,n_{k-1},i) \in \widehat{O}\left(\sum_{i=m}^{h(n_k)} f(n_1,\ldots,n_{k-1},i)\right).$$

Proof: Let $c \in \mathbb{R}^{>0}$ and $N \in \mathbb{N}$ be such that whenever $n_1 \ge N, \ldots, n_k \ge N$,

$$g(n_1,\ldots,n_k) \leq cf(n_1,\ldots,n_k)$$

and

$$\widehat{g}(n_1,\ldots,n_k) \le c\widehat{f}(n_1,\ldots,n_k).$$

Without loss of generality, we can assume that $N \ge m$. In order to simplify the notation, we first define

$$f'(n_1, \dots, n_k) = \sum_{i=m}^{h(n_k)} f(n_1, \dots, n_{k-1}, i)$$

and

$$g'(n_1,\ldots,n_k) = \sum_{i=m}^{h(n_k)} g(n_1,\ldots,n_{k-1},i).$$

We therefore need to show that

$$g'(n_1,\ldots,n_k) \in \widehat{O}(f'(n_1,\ldots,n_k)).$$

Because f and h are strictly nondecreasing, it follows that f' is strictly nondecreasing; hence, by Theorem 3.5, we only need to show Condition (3.2). Note that for $0 \le i_1 \le n_1, \ldots, 0 \le i_k \le n_k$,

$$g'(i_1, \dots, i_k) = \sum_{i=m}^{h(i_k)} g(i_1, \dots, i_{k-1}, i)$$

$$\leq \sum_{i=m}^{h(n_k)} \widehat{g}(n_1, \dots, n_{k-1}, i);$$

hence,

$$\widehat{g'}(n_1,\ldots,n_k) \le \sum_{i=m}^{h(n_k)} \widehat{g}(n_1,\ldots,n_{k-1},i).$$

Because h is strictly nondecreasing and unbounded, we can choose $N' \ge N$ so that for every $n \ge N'$, $h(n) \ge N$. Suppose $n_1 \ge N', \dots, n_k \ge N'$. Then because f is strictly nondecreasing,

$$\begin{aligned} \widehat{g'}(n_1, \dots, n_k) &\leq \sum_{i=m}^{h(n_k)} \widehat{g}(n_1, \dots, n_{k-1}, i) \\ &= \sum_{i=m}^N \widehat{g}(n_1, \dots, n_{k-1}, i) + \sum_{i=N+1}^{h(n_k)} \widehat{g}(n_1, \dots, n_{k-1}, i) \\ &\leq \sum_{i=m}^N \widehat{g}(n_1, \dots, n_{k-1}, N) + \sum_{i=N+1}^{h(n_k)} c\widehat{f}(n_1, \dots, n_{k-1}, i) \\ &\leq (N-m+1)c\widehat{f}(n_1, \dots, n_{k-1}, N) + c \sum_{i=N+1}^{h(n_k)} f(n_1, \dots, n_{k-1}, i) \\ &= c(N-m+1)f(n_1, \dots, n_{k-1}, N) + c \sum_{i=N+1}^{h(n_k)} f(n_1, \dots, n_{k-1}, i) \\ &\leq c(N-m+1)f'(n_1, \dots, n_k) \\ &= c(N-m+1)\widehat{f'}(n_1, \dots, n_k). \end{aligned}$$

Because c(N - m + 1) is a fixed positive constant, by Theorem 3.5,

$$g'(n_1,\ldots,n_k)\in \widehat{O}(f'(n_1,\ldots,n_k)).$$

Now taking m = 0 and h(n) = n, we have the following corollary.

Corollary 3.9 Property 2 holds for \widehat{O} when the function f is strictly nondecreasing.

4 **Big-** Ω

We can extend the definition of big- Ω to multiple variables by modifying the definition of \widehat{O} so that the directions of inequalities (3.1) and (3.2) are reversed. Let us refer to this extension as $\widehat{\Omega}$. Specifically, for $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$, $\widehat{\Omega}(f(n_1, \ldots, n_k))$ is defined to be the set of all functions $g : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ such that there exist $c \in \mathbb{R}^{>0}$ and $N \in \mathbb{N}$ such that for all $n_1 \geq N, \ldots, n_k \geq N$, the following conditions hold:

$$g(n_1,\ldots,n_k) \ge cf(n_1,\ldots,n_k); \tag{4.1}$$

and

$$\widehat{g}(n_1, \dots, n_k) \ge c\widehat{f}(n_1, \dots, n_k). \tag{4.2}$$

It is then easily seen that the standard duality between big-O and big- Ω holds.

Theorem 4.1 Let $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ and $g : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$. Then $f(n_1, \ldots, n_k) \in \widehat{O}(g(n_1, \ldots, n_k))$ iff $g(n_1, \ldots, n_k) \in \widehat{O}(f(n_1, \ldots, n_k))$.

As with big-O, we would like to know that $\widehat{\Omega}$ is the same as Ω when we restrict our attention to single-variable functions. However, as with big-O, there are cases in which this equality does not hold. Given Theorems 3.1 and 4.1, we might suspect that the exceptions would involve the constant function 0. Indeed, it is not hard to see that if f(n) is positive for at least one value but no more than finitely many values, then $0 \in \Omega(f(n))$, but $0 \notin \widehat{\Omega}(f(n))$. However, the following theorem shows that this is the only exception.

Theorem 4.2 Let $f : \mathbb{N} \to \mathbb{R}^{\geq 0}$. Then $\widehat{\Omega}(f(n)) - \{0\} = \Omega(f(n)) - \{0\}$; i.e., $\widehat{\Omega}(f(n))$ and $\Omega(f(n))$ are equal if we exclude the constant function 0 from both sets.

Proof:

 $\begin{array}{l} \subseteq: \mbox{ Let } g(n) \in \widehat{\Omega}(f(n)) - \{0\}. \mbox{ Then by Condition (4.1), } g(n) \in \Omega(f(n)) - \{0\}. \\ \supseteq: \mbox{ Let } g(n) \in \Omega(f(n)) - \{0\}. \mbox{ Then because } g(n) \in \Omega(f(n)), \mbox{ there exist} \\ c \in \mathbb{R}^{>0} \mbox{ and } N \in \mathbb{N} \mbox{ such that for every } n \geq N, \end{array}$

$$g(n) \ge cf(n);$$

hence, Condition (4.1) holds.

Because g(n) is not the constant function 0, g(n) > 0 for some $n \in \mathbb{N}$; hence, we can assume without loss of generality that for some i < N, g(i) > 0. We now let

$$c' = \begin{cases} c & \text{if } \widehat{f}(N-1) = 0\\ \min\left(c, \frac{\widehat{g}(N-1)}{\widehat{f}(N-1)}\right) & \text{otherwise.} \end{cases}$$

Note that $c \ge c' > 0$. Then for $n \ge N$,

$$\widehat{g}(n) = \max(\widehat{g}(N-1), \max\{g(i) \mid N \le i \le n\})$$

$$\geq \max(c'\widehat{f}(N-1), \max\{cf(i) \mid N \le i \le n\})$$

$$\geq c'\widehat{f}(n).$$

Therefore, Condition (4.2) holds; hence, $g(n) \in \widehat{\Omega}(f(n)) - \{0\}$.

By observing that $0 \notin \Omega(f(n)) \cup \widehat{\Omega}(f(n))$ when f(n) is eventually positive, we obtain the following corollary.

Corollary 4.3 Let $f : \mathbb{N} \to \mathbb{R}^{\geq 0}$ be an eventually positive function. Then $\widehat{\Omega}(f(n)) = \Omega(f(n))$.

Let us now consider the properties of $\widehat{\Omega}$. We first observe that the following two theorems are easily shown.

Theorem 4.4 For any positive integer k, positive real number c, and $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$, $\widehat{\Omega}(f(n_1, \ldots, n_k)) = \widehat{\Omega}(cf(n_1, \ldots, n_k)).$

Theorem 4.5 For any positive integer $k, f_1 : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$, and $f_2 : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$,

$$\widehat{\Omega}(f_1(n_1,\ldots,n_k)) + \widehat{\Omega}(f_2(n_1,\ldots,n_k))$$

= $\widehat{\Omega}(\max(f_1(n_1,\ldots,n_k),f_2(n_1,\ldots,n_k)))$

For each of the five properties presented in Section 2, there is an analogous property for big- Ω . For example, the analog to Property 1 is simply (4.1). Furthermore, it is obvious that when we replace big-O with big- Ω in Property 4, the resulting property holds for $\hat{\Omega}$. In what follows, we will show that each of the remaining analogous properties holds for $\hat{\Omega}$ when the functions inside the asymptotic notation are strictly nondecreasing.

We begin by considering the impact of a strictly nondecreasing function f on the definition of $\hat{\Omega}$. By examining the proof of Theorem 3.5, we can see that reversing the direction of the inequalities in the definition allows us to prove that Condition (4.1) implies Condition (4.2). Let us define $\Omega_{\forall}(f(n_1, \ldots, n_k))$ to be the set of functions satisfying Condition (4.1). We then have the following theorem.

Theorem 4.6 If $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ is strictly nondecreasing, then

$$\Omega(f(n_1,\ldots,n_k)) = \Omega_\forall (f(n_1,\ldots,n_k)).$$

The following analog for Property 3 then follows immediately from Theorem 4.6.

Theorem 4.7 Let $f_1 : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ and $f_2 : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ be two functions such that for some $N \in \mathbb{N}$, and every $n_1 \geq N, \ldots, n_k \geq N$,

$$f_1(n_1,\ldots,n_k) \ge f_2(n_1,\ldots,n_k).$$

Further suppose that f_2 is strictly nondecreasing. Then

$$\Omega(f_1(n_1,\ldots,n_k)) \subseteq \Omega(f_2(n_1,\ldots,n_k))$$

Likewise, Theorem 4.6 implies that the analog for Property 5 holds for $\hat{\Omega}$ when the functions involved are strictly nondecreasing.

Theorem 4.8 Let $f_1 : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ and $f_2 : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ be two strictly nondecreasing functions. Then

$$\widehat{\Omega}(f_1(n_1,\ldots,n_k))\widehat{\Omega}(f_2(n_1,\ldots,n_k))\subseteq \widehat{\Omega}(f_1(n_1,\ldots,n_k)f_2(n_1,\ldots,n_k)).$$

Finally, we will show that the analog of Theorem 3.8 holds for $\widehat{\Omega}$. From this it will follow that the analog of Property 2 holds for $\widehat{\Omega}$ as well.

Theorem 4.9 Let $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ be a strictly nondecreasing function. Then for any $g(n_1, \ldots, n_k) \in \widehat{\Omega}(f(n_1, \ldots, n_k))$, any natural number m, and any strictly nondecreasing unbounded function $h : \mathbb{N} \to \mathbb{N}$,

$$\sum_{i=m}^{h(n_k)} g(n_1,\ldots,n_{k-1},i) \in \widehat{\Omega}\left(\sum_{i=m}^{h(n_k)} f(n_1,\ldots,n_{k-1},i)\right).$$

Proof: Let $c \in \mathbb{R}^{>0}$ and $N \in \mathbb{N}$ be such that whenever $n_1 \ge N, \ldots, n_k \ge N$,

$$g(n_1,\ldots,n_k) \ge cf(n_1,\ldots,n_k).$$

Furthermore, we can assume without loss of generality that $N \ge m$. Because h is strictly nondecreasing and unbounded, we can choose $N' \ge N$ such that $h(N') \ge 2N$.

Suppose $n_1 \ge N', \ldots, n_k \ge N'$. Then

$$\sum_{i=m}^{h(n_k)} g(n_1, \dots, n_{k-1}, i) \ge \sum_{i=N}^{h(n_k)} g(n_1, \dots, n_{k-1}, i)$$
$$\ge \sum_{i=N}^{h(n_k)} cf(n_1, \dots, n_{k-1}, i)$$
$$\ge \frac{c}{2} \sum_{i=m}^{h(n_k)} f(n_1, \dots, n_{k-1}, i)$$

because f is strictly nondecreasing and $h(n_k) \ge 2N$. From Theorem 4.6,

$$\sum_{i=m}^{h(n_k)} g(n_1,\ldots,n_{k-1},i) \in \widehat{\Omega}\left(\sum_{i=m}^{h(n_k)} f(n_1,\ldots,n_{k-1},i)\right).$$

5 Other Asymptotic Notation

For single-variable functions f(n), it is common to define $\Theta(f(n))$ to be

 $O(f(n)) \cap \Omega(f(n)).$

For functions $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$, we can therefore define $\widehat{\Theta}(f(n_1, \ldots, n_k))$ to be

$$\widehat{O}(f(n_1,\ldots,n_k))\cap \widehat{\Omega}(f(n_1,\ldots,n_k)).$$

It then follows from Theorems 3.1 and 4.2 that if $f : \mathbb{N} \to \mathbb{R}^{\geq 0}$ is eventually positive, then $\widehat{\Theta}(f(n)) = \Theta(f(n))$. Likewise, it is easily seen that the remaining theorems in Sections 3 and 4 extend to $\widehat{\Theta}$.

In addition, it is common (see, e.g., [3]) to define o(f(n)) to be the set of all functions $g: \mathbb{N} \to \mathbb{R}^{\geq 0}$ such that for every positive real number c, there is a natural number N such that

$$g(n) < cf(n)$$

whenever $n \geq N$. For functions $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$, we can therefore define $\widehat{o}(f(n_1, \ldots, n_k))$ to be the set of all functions $g : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ such that for every positive real number c, there is a natural number N such that

$$g(n_1, \dots, n_k) < cf(n_1, \dots, n_k) \tag{5.1}$$

and

$$\widehat{g}(n_1, \dots, n_k) < cf(n_1, \dots, n_k) \tag{5.2}$$

whenever $n_1 \geq N, \ldots, n_k \geq N$.

Similarly, we can define $\widehat{\omega}(f(n_1, \ldots, n_k))$ to be the set of all functions $g : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ such that for every positive real number c, there is a natural number N such that

$$g(n_1,\ldots,n_k) > cf(n_1,\ldots,n_k) \tag{5.3}$$

and

$$\widehat{g}(n_1, \dots, n_k) > c\widehat{f}(n_1, \dots, n_k) \tag{5.4}$$

whenever $n_1 \ge N, \ldots, n_k \ge N$. Note that for single-variable functions f(n), $\omega(f(n))$ is defined as above, but without Condition (5.4).

The following theorem is now easily shown.

Theorem 5.1 Let $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$ and $g : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$. Then $f(n_1, \ldots, n_k) \in \widehat{o}(g(n_1, \ldots, n_k))$ iff $g(n_1, \ldots, n_k) \in \widehat{\omega}(f(n_1, \ldots, n_k))$.

It is not hard to modify the proofs of Theorems 3.1 and 4.2 to apply to \hat{o} and $\hat{\omega}$, respectively. However, it is interesting to note that we don't need to restrict f to be greater that 0 at some point, as in Theorem 3.1, because $\hat{o}(0) = o(0) = \emptyset$. Likewise, we do not need to remove the constant function 0 from $\omega(f(n))$ and $\hat{\omega}(f(n))$, as in Theorem 4.2, because these sets never contain 0. We therefore have the following theorems.

Theorem 5.2 For every function $f : \mathbb{N} \to \mathbb{R}^{\geq 0}$, $\hat{o}(f(n)) = o(f(n))$.

Theorem 5.3 For every function $f : \mathbb{N} \to \mathbb{R}^{\geq 0}$, $\widehat{\omega}(f(n)) = \widehat{\omega}(f(n))$.

The other results from Sections 3 and 4 extend to \hat{o} and $\hat{\omega}$, respectively, in a fairly straightforward way, provided we are careful in how we state the desired properties. For example, the analogs of Property 4 should be that for all functions $f : \mathbb{N}^k \to \mathbb{R}^{\geq 0}$, $f(n_1, \ldots, n_k) \notin o(f(n_1, \ldots, n_k))$, and $f(n_1, \ldots, n_k) \notin \omega(f(n_1, \ldots, n_k))$.

6 Conclusion

We have shown that the properties of big-O notation typically assumed in the analysis of algorithms are inconsistent when applied to functions on more than one variable. Furthermore, we have shown that even if the functions inside the big-O notation are restricted to be strictly nondecreasing, these properties do not all hold for the standard definition. In order to overcome this deficiency, we have proposed an alternative definition and showed that, when the functions inside the notation are nondecreasing, the desired properties hold for this definition. Furthermore, we have extended these definitions to other asymptotic notation commonly used in the analysis of algorithms, and have shown that these definitions yield similar results.

Because the worst-case (or expected-case) running times of most algorithms are strictly nondecreasing, it seems unlikely that the problems exposed by this paper have led to many (if any) incorrect conclusions in the analysis of agorithms in the literature. However, there do exists algorithms whose running times are not strictly nondecreasing. For example, Euclid's algorithm for finding the greatest common divisor of two positive integers m and n (see, e.g., [3]) takes more time for m = 5 and n = 3 than for m = 5 and n = 5. Thus, in terms of the values of m and n, the running time is not strictly nondecreasing. It therefore would be too restrictive to restrict the set of functions comprising O(f(m, n)) to be strictly nondecreasing. Note, however, that we can find a strictly nondecreasing upper bound (e.g., $\widehat{O}(\lg(\max(u, v))))$ on the running time of Euclid's algorithm.

The main contribution of our proposed definition is to provide a mathematical foundation for performing algorithm analysis with more than one parameter. Without this foundation, analyses cannot be formally justified, and therefore run a risk of leading to incorrect conclusions. Furthermore, because the behavior of our \hat{O} is closer to what is typically expected than is the behavior of O_{\forall} , most analyses in the literature should be valid for this new definition. The only aspect that needs to be checked is whether the functions inside the notation are nondecreasing when this property is required.

References

- Paul G. H. Bachmann. Analytische Zahlentheorie, Bd 2: Die Analysische Zahlentheorie. Teubner, Leipzig, Germany, 1894.
- [2] Gilles Brassard and Paul Bratley. Fundamentals of Algorithmics. Prentice Hall, 1996.
- [3] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*. McGraw-Hill, 2nd edition, 2001.
- [4] Michael T. Goodrich and Roberto Tamassia. Algorithm Design: Foundations, Analysis, and Internet Examples. Wiley, 2002.
- [5] Jon Kleinberg and Éva Tardos. Algorithm Design. Addison Wesley, 2006.
- [6] Donald E. Knuth. Big omicron and big omega and big theta. ACM SIGACT News, 8:18–24, 1976.
- [7] Edmund Landau. Handbuch der Lehre von der Verteilung der Primzahlen. Teubner, Leipzig, Germany, 1909.
- [8] Anany Levitin. Introduction to The Design & Analysis of Algorithms. Addison Wesley, 2nd edition, 2007.
- [9] Joachim von zur Gathen and Jürgen Gerhard. Modern Computer Algebra. Cambridge University Press, 2nd edition, 2003.