Composite Factorizations

Vignesh Nydhruva

March 2024

Contents

1 Abstract

This paper introduces an algorithm that can be used to find the distinct factorizations of a composite number and introduces a conjecture on the distribution of distinct factorizations among the composites.

2 Algorithm

2.1 Overview

The purpose of the algorithm is to find the number of distinct factorizations of a composite integer n.

Definition 1. *A factorization of an integer* n *is a set of positive integers that multiply to* n*. (Note: this set consists of prime and/or composite numbers)*

Definition 2. *A prime factorization of n is a set of prime numbers that multiply to n.*

The first step of the algorithm is to begin with the prime factorization of n. This is denoted by the function $\kappa(n)$, which is defined as such

$$\kappa : \mathfrak{C} \to \{\mathbb{N}\}$$

where $\mathbb{N}$ denotes the set of nonnegative integers, $\mathbb{P}$ denotes the set of primes, and $\mathfrak{C} = (\mathbb{N} - \mathbb{P}) \setminus \{0, 1\}$ denotes the set of composite numbers.

Definition 3. *F_k denotes the set containing factorization(s) of the $k^{th}-$level and let f_k denote a set containing factors of a certain factorization of the $k^{th}-$level, where $f_k \subsetneq F_k$ (generally) and $f_k \neq \emptyset$.*

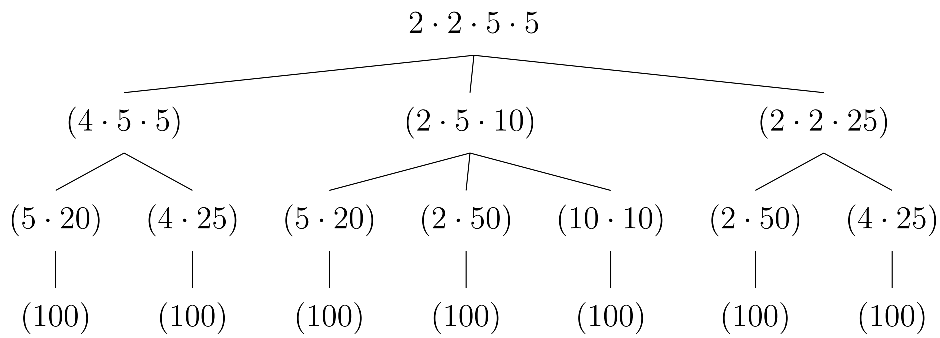
Definition 4. *A $k^{th}-$level factorization of n denotes a factorization of the $k^{th}-$level such that k represents the number of levels progressed from F_0.*

Based on Definition 3, let F_0 denote the set outputted by $\kappa(n)$ and let f_0 denote the subset of F_0 containing the prime factorization of n.

The next step of the algorithm is to find distinct combinations of factors within f_0 and subsequently create new factorizations of n. To proceed in a methodical manner, the cardinalities of each factorization in the next level will be $|f_0| - 1$. This is because exactly two factors are multiplied at a given iteration, thus leading to a factorization of the next level containing $|f_0| - 1$ factors. To illustrate more clearly, let the $0^{th}-$level factorization be $\kappa(n) = F_0 = \{f_0\} = \{\{p_1, p_2, p_3, p_4\}\}$, where p is a prime and $|f_0| = 4$. Let the $1^{st}-$level factorizations be the following set: $F_1 = \{\{p_1p_2, p_3, p_4\}$, $\{p_1p_3, p_2, p_4\}$, $\{p_1p_4, p_2, p_3\}$, $\{p_1, p_2p_3, p_4\}$, $\{p_1, p_2p_4, p_3\}$, $\{p_1, p_2, p_3p_4\}\}$. Notice that any f_1 is equal to $|f_0| - 1$. For this case (where $|f_0| = 4$), the cardinality of the set that contains all f_1 will be $|F_1| = \binom{4}{2} = \frac{4!}{2!(4-2)!} = 6$. In the general case, let F_k denote the set that contains all f_k, where $|F_k| = \frac{|F_{k-1}|!}{2!(|F_{k-1}|-2)!}$ (as long as any $|f_k| > 1$). In order to give numeric example of this step of the algorithm, take the following case. Let $n = 210$, where $F_0 = \{f_0\} = \{\{2, 3, 5, 7\}\}$. f_0 represents the $0^{th}-$level factorization of 210, it is 0 levels past $\kappa(210)$ (it is itself). The $1^{st}-$level factorizations of 210 would be the following set $F_1 = \{\{2 \cdot 3, 5, 7\}, \{2 \cdot 5, 3, 7\}, \{2 \cdot 7, 3, 5\}, \{2, 3 \cdot 5, 7\}, \{2, 3 \cdot 7, 5\}, \{2, 3, 5 \cdot 7\}\}$. Then, for each $1^{st}-$level factorization, the recursive chain continues, therefore creating factorizations of the 2^{nd}, 3^{rd}, and of the $k^{th}-$level. The algorithm must terminate when $|f_k| = 1$, implying that $f_k = \{n\}$.

2.2 A Computer Program Implementation

In this example, let $n = 100$. For reference, n has the following factorization tree, with the root being the prime factorization of n. Duplicates in F_1 and any other subtrees have been removed, but not duplicates within a common level.



The first step is to obtain the prime factorization of n and store it inside list or other convenient data type, in accordance with the definition of F_0 made in Section 2.1.

The next step is to create F_1, which will be the input of the central recursive method of the algorithm.

Definition 5. *Let C_k denote the set that contains one or more sets, where each set in C_k represents a combination of two factors in a particular set f_k of level k. Let $c_k \subsetneq C_k$ (generally) where $c_k \neq \emptyset$.*

Definition 6. *Let P_k denote the set that contains the products of combinations of factors of a factorization in the $k^{th}-$level.*

Definition 7. *Let $S_F \subsetneq \kappa(n = 100)$, where $S_F \neq \emptyset$. Regarding the example, S_F is permanently defined as $\{2, 2, 5, 5\}$ and not $\{\{2, 2, 5, 5\}\}$.*

First, find the combinations of the factors present in the prime factorization (treating each factor as a unique element), thus creating the set $C_0 = \{\{2, 2\}, \{2, 5\}, \{2, 5\}, \{2, 5\}, \{2, 5\}, \{5, 5\}\}$. Remove duplicate subsets of C_0 if necessary. Now, $C_0 = \{\{2, 2\}, \{2, 5\}, \{5, 5\}\}$. Next, create a set P_0 that contains the products of the elements inside each set present in C_0, where $P_0 = \{4, 10, 25\}$. Then, create the set F_1 that contains $|P_0|$ sets where each set contains all the elements of S_F where a distinct element of P_0 is appended to each of the sets. After performing this operation, $F_1 = \{\{2,2,5,5,4\},\{2,2,5,5,10\},\{2,2,5,5,25\}\}$. Notice that $|C_0| = |P_0| = |F_1|$. Next, traverse F_1 and let c_{0_i} denote the $i^{th}-$combination present in C_0, where the indices of F_1 and C_0 are in synchronization. Note that $c_{0_i} \subsetneq S_F$. Now for each set in F_1, add the elements of $(S_F - c_{0_i})$ and remove the first $|S_F|$ consecutive elements that correspond to the prime factorization of n. Now, sort each set in F_1 in non-decreasing order (for clarity), thus making $F_1 = \{\{4,5,5\},\{2,5,10\},\{2,2,25\}\}$.

Variations of the logic presented in the previous paragraph can then be applied to find the rest of the nodes in the tree. In order to find distinct factorizations, sort the elements of each factorization and remove any duplicates. An implementation of this algorithm can be found here: (link).

3 Distinct Factorizations

3.1 Conjecture

Definition 8. *A distinct factorization d (represented by a set) of n (where $n \in \mathfrak{C}$) is a unique factorization of n consisting of $|d|$ factors, such that $2 \leq |d| \leq |\mathscr{P}|$ where $\mathscr{P} \subsetneq \kappa(n)$ and $\mathscr{P} \neq \emptyset$. $\mathscr{P}$ represents the set containing the numbers present in the prime factorization of n ($\mathscr{P}$ is a generalization of S_F, which was defined in the previous section). Let D denote the set of all distinct factorizations of n, where $d_i \subsetneq D$, $d \neq \emptyset$, and i is an indexing variable for d.*

It can be deduced that, as n increases, $|\mathscr{P}|$ increases. This is due to the fact that a large increase in n corresponds with a better chance of writing n as the product of more prime factors. More concretely, a doubling of n corresponds to an increase of $|\mathscr{P}|$ by a unit of one factor, namely, the factor 2. This implies that a small increase in $|\mathscr{P}|$ (on the order of 10^0) results in a relatively large increase in n (when n is on the order of 10^1 or higher). This difference in magnitudes between an increment in $|\mathscr{P}|$ and an increase in n is more obvious when larger primes are added to $\mathscr{P}$.

Definition 9. *Let ${}_{\mathscr{D}}\hat{F}_k$ denote the set of distinct factorizations in the non-trivial $k^{th}-$level of the factorization tree of n, where the tree has a total of $|\mathscr{P}| - 1$ levels. This is built on the definitions made in Section 2.1.*

The implication is that as n increases, the number of distinct factorizations of n also increases. As pointed out, a large increase in n corresponds to an increase in $|\mathscr{P}|$ (by a smaller degree). An increase in $|\mathscr{P}|$ implies an increase in $|D|$, due to the following equation

$$|D| = \left| \bigcup_{i=0}^{|\mathscr{P}|-1} {}_{\mathscr{D}}\hat{F}_i \right|$$

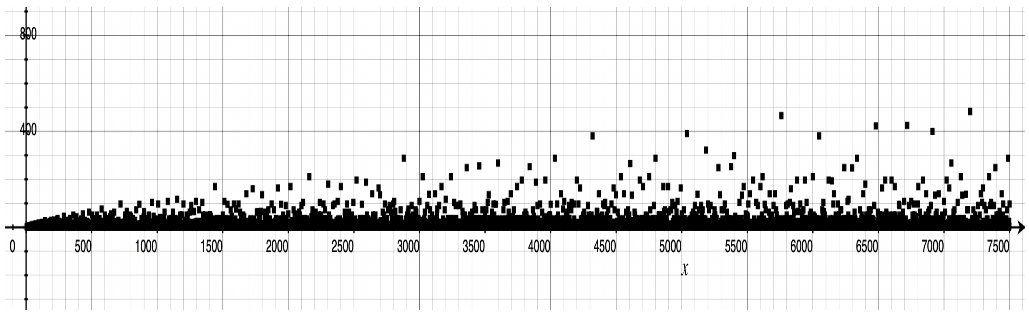
Incrementing $|\mathscr{P}|$ by one would result in an additional level in the factorization tree of n. This would result in $|D|$ increasing exponentially due to the following reason. For simplicity, assume in this case that every node in the tree is distinct, thus resulting in no duplicate nodes. The number of items added to D increases exponentially from level to level, since each particular level produces several distinct factorizations, which compound as k increases. Therefore, we have the following equation, where δ is an integer corresponding to the change in $|D|$, $\mathscr{P}$ is updated to contain an additional prime factor, and F_{k-1} reflects the new factorization tree generated from the incrementing of $|\mathscr{P}|$.

$$|D|_{new} = |D| + \delta = \sum_{k=2}^{|\mathscr{P}|-1} \frac{|F_{k-1}|!}{2!\,(|F_{k-1}| - 2)!}$$

To also consider the case where there exist duplicate nodes in the factorization tree after an additional prime is added to $\mathscr{P}$, it is obvious that initially, an exponential amount would be added to $|D|$. However, this means that the duplicate factorizations have not yet been added to D, so therefore, the removal of duplicate factorizations, subsequently, would have no effect on $|D|$.

It must be noted that as $\mathscr{P}$ gains a prime factor, the order on which n increases is greater than the order on which $|D|$ increases. This is because n grows with respect to any additional prime factors in $\mathscr{P}$, while $|D|$ grows with respect to purely a change in $|\mathscr{P}|$. Furthermore, virtually all primes are on a higher order than an increment of 1 to $|\mathscr{P}|$. So, the inclusion of a relatively large prime p to $\mathscr{P}$ increases n exponentially by base p, while the order of p has no effect on $|D|$ (since $|D|$ increases by 1 due to the inclusion of a single prime in $\mathscr{P}$, regardless of its order). In general, the inclusion of a single prime to $\mathscr{P}$ increases n by a higher exponential amount than it increases $|D|$ by.

The following graph displays for each composite integer from 0 to 7500 (on the horizontal axis), the number of distinct factorizations (on the vertical axis). The graph was generated using the application, *Grapher*, native to macOS.



Based on these conclusions, I present the following conjecture.

Distinct Factorizations Conjecture. *There exists a curve $g(x)$, with some subset $g_s : \mathbb{R}_0^+ \to \mathbb{R}_0^+$, where $\forall x \in \mathbb{R}_0^+ \left(\frac{dg_s}{dx} > 0 \wedge \frac{d^2 g_s}{dx^2} < 0\right)$ such that all points of the form $(n, |D_n|) \in \mathbb{N}^+ \times \mathbb{N}^+$ are contained by the curves g_s and $y = 0$.*