Julia Frances

Professor Mark Boady

CS 502

13 July 2022

The Sieve of Eratosthenes

INTRODUCTION

A computer can not perform any task that a human has not already figured out. Although a computer can perform tasks many times faster than any human possibly could, a human had to tell the computer precisely what to do before it could do anything. Therefore, while some computer algorithms may seem esoteric and complex and impossible for a regular human mind, many of these algorithms are, in fact, digital translations of intuitive human processes and physical behavior. Someone took something that people already did or understood and coded it in such a way that a computer can copy our patterns.

The Sieve of Eratosthenes is one such algorithm. A sieve, or sifter, is a tool used to separate matter. You likely have one in your kitchen: a pasta strainer is a sieve that separates noodles from boiling water. This algorithm takes the idea of a sieve, which is simple and intuitive enough to have been in use for thousands of years, and copies the patterns of separation to achieve a significant mathematical goal. It is highly efficient on a small scale because it does not require even very straightforward mathematical calculations to achieve this goal. The original algorithm itself, also thousands of years old, is no longer considered practical for most contemporary purposes. Its importance lies in its intuitiveness, lack of computation, and role as an inspiration and benchmark for more advanced algorithms today.

Eratosthenes

Euclid, a Greek mathematician recognized today as the father of geometry, initialized the systematic study of primes with his proof of their infinitude in the 300s BC. His magnum opus, *Elements*, is referred to as the most influential textbook ever written with over a thousand editions reprinted since the invention of the printing press (Boyer 119). Euclid's scholarship sparked thousands of years of extensive use of prime numbers in multitudes of ologies. However numerous his contributions, he was unable to provide a primality test as efficient as the sieve.

The traditional, ancient version of this classic algorithm is commonly credited to Eratosthenes, a Greek polymath postdating Euclid by about a century. After studying mainly philosophy and mathematics in Athens (Roller 8-9), Eratosthenes lived and worked primarily in Alexandria, Egypt, where he was not only a librarian at the famous library of Alexandria but also a tutor to the future King Ptolemaious IV (Roller IX).

Eratosthenes lived during "the maximum political extent of the Ptolemaic empire and greatest flowering of Alexandria as an intellectual center" (Roller 9) and capitalized on this wave of intellectualism by devoting himself to multiple scholarly pursuits. He concentrated his earlier years mainly on mathematics, and connected with Archimedes, who sought his advice on various occasions (Roller 12). He largely dedicated his later years to the study of geography, and his efforts to legitimize geography as a "legitimate scholarly endeavor" led to the creation of the word 'geographia,' the root of which is still in use today (Roller IX).

Due to the nature of ancient history, little can be confirmed about the circumstances surrounding Eratosthenes' initial conception of the sieve other than giving credit where it is due.

Eratosthenes is remembered today not only for this algorithm—his sieve—but also for his work determining the circumference of the earth, his general geographic scholarship, his work on chronology, and—to a lesser extent—his poetry (Roller 14-15).

THE SIEVE

In essence, the Sieve of Eratosthenes is an algorithm used to obtain prime numbers. A primality test itself stands in contrast with factorization or a simple divisibility check, which are thought to be computationally difficult (in the case of factorization) and/or outright lazy (in the case of divisibility checks) (O'Neill 97). Primality tests determine whether or not a number is prime using a variety of methods, one of which being a sieve.

The physical sieve, a tool used for various purposes such as gold panning, separates particles based on size. The algorithmic sieve separates numbers based on a numeric quality. The method as conceived by Eratosthenes works like this:

1. List the natural numbers up to some number *n.*
2. Cross out 1.
3. Beginning at 2, determine multiples by counting up the list by 2 and crossing those numbers out (so 4, 6, 8, etc.).
4. Move to 3, and determine multiples by counting up the list by 3 and crossing out those numbers (so 6, 9, 12, etc.)
5. In this way, for a number *p*, count up the list by *p* and continue crossing out its multiples.
6. Continue until $p^2 > n$.

This procedure can be demonstrated visually with an *n* of 20 in the table below.

p	~~1~~	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	~~1~~	2	3	~~4~~	5	~~6~~	7	~~8~~	9	~~10~~	11	~~12~~	13	~~14~~	15	~~16~~	17	~~18~~	19	~~20~~
3	~~1~~	2	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~	11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~
5	~~1~~	2	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~	11	~~12~~	13	~~14~~	~~15~~	~~16~~	17	~~18~~	19	~~20~~
	~~1~~	**2**	**3**	~~4~~	**5**	~~6~~	**7**	~~8~~	~~9~~	~~10~~	**11**	~~12~~	**13**	~~14~~	~~15~~	~~16~~	**17**	~~18~~	**19**	~~20~~

Critically, little to no computation is taking place, as we are counting by *p* rather than performing any mathematical operation. The sieve works by marking the multiples of our current *p and* sifting them out of the running. Once the test concludes, the remaining numbers will all be prime. The importance of this method, according to a computing article from over fifty years ago, comes from the fact that it is easily adapted to mechanical procedures, is suited for the computation of large tables, and does not depend on the testing of individual numbers for divisibility (Dubisch 236-7).

Evolution of the Sieve

Obviously, this traditional sieve method is rather inefficient for larger sets of numbers and would not come close to verifying some of the primes known today (Mollin 19). We can make several straightforward improvements, one of which was even known and used by Eratosthenes himself, according to Nicomachus, a 1st-century mathematician and the first authority on Eratosthenes (Quesada).

The first simple improvement would be to begin counting at p^2 instead of *2p*. We can change our starting point in this way because we would have already crossed off lower multiples of *p* once we reach *p* for use as a counting interval (O'Neill 96). For example, if *p* is five, and we have already progressed through counting by twos and threes, then we know that 2x5, 3x5, and

4x5 will already be eliminated. Therefore, the ideal starting point for optimizing the algorithm will be 5x5, or 5^2.

The second simple improvement, the one known by Eratosthenes himself, would be to remove all even numbers before the procedure begins, thus narrowing the range of numbers we must sieve overall. Although this initial modification of the pool of numbers eliminates multiples of two just like the original algorithm, it does not technically follow the exact algorithmic pattern as we are not eliminating using the counting method. Therefore, it is considered a modification/improvement to the original algorithm (Quesada). Even though we are removing half the initial pool of numbers in an algorithm that works by counting progressively, the algorithm will still work because the counting pattern remains consistent and identifiable, as the table demonstrates below.

p	~~1~~	2	3	5	7	9	11	13	15	17	19
3	~~1~~	2	3	5	7	~~9~~	11	13	~~15~~	17	19
5	~~1~~	2	3	5	7	~~9~~	11	13	~~15~~	17	19
	~~1~~	**2**	**3**	**5**	**7**	~~9~~	**11**	**13**	~~15~~	**17**	**19**

Looking at the differences between the first and second tables, we can see that the number of steps necessary to sieve out multiples dramatically decreases once we remove the even numbers. It might seem natural to attempt to improve the algorithm again by further reducing the size of the initial set of numbers. However, suppose we try to eliminate the multiples of three in the same way. In this case, we run into a problem where the multiples of a number *p* are no longer spaced *p* units apart, and the original counting method falls apart.

Luckily, hope is not lost, and it is still possible to sieve using Eratosthenes' counting method even when initially removing multiples of three and greater. Instead of counting by *p* for multiples of *p*, cyclical counting patterns have been identified which are used to identify multiples of *p* even in a significantly reduced set of numbers. For example, after removing multiples of two and three, if we counted the intervals in the remaining subset between the remaining multiples of five, we would find ourselves counting seven spaces, three spaces, seven, three, seven, three, etc. This identifiable pattern has been used to improve the algorithm further. In fact, along with any *p* in this subset, an associated pattern of two intervals (d_1, d_2) between multiples can be used to cyclically count and thus identify the remaining multiples using Eratosthenes' original method. These first two simple improvements are known together as the first extension of the sieve (Quesada).

One glaring inefficiency of Eratosthenes' original algorithm is that a person (or machine) would waste a great deal of time crossing out numbers that have already been eliminated. For example, using the initial complete set of natural numbers in the first table, we can see that we cross out twelve when eliminating multiples of two, and again when eliminating multiples of three. This redundant step will be repeated numerous times throughout the algorithm, resulting in many wasted decisions and operations. The true, original sieve has a complexity of O(n log log n) (O'Neill 97).

Euler, an 18th-century Swiss mathematician, came up with his own version of the sieve in which only numbers which are a multiple of the current *p* and haven't already been eliminated are crossed off. Therefore, each number destined to be eliminated will be crossed off only once, thus saving a great deal of time. This version of the sieve has a complexity of O(n) (Salvo 1).

Euler's Sieve and several others created over the centuries since Eratosthenes are commonly known and used as substantially more efficient versions of the original sieve.

CURRENT APPLICATIONS

Primality is exceedingly and increasingly important in our information-based society because it is used in the secure transmission of data (Mollin 29). Before the advent of digital computing, cryptosystems encrypted and decrypted information using private keys, which had to be secretly decided ahead of time; this made it challenging to keep messages secure over a long time or distance. In the 1970s, as technology accelerated, a more secure system known as the RSA cryptosystem was invented. Although the fundamental idea behind the cryptosystem remains the same, the security increases significantly due to the use of the product of two large prime numbers in these keys (Rowland 1-2).

Modern RSA ciphers use prime numbers that are hundreds of digits long. Even today, no known algorithms can readily determine the prime factors of these huge numbers before the sun engulfs our planet (Rowland 5). This impossibility makes RSA ciphers incredibly secure and makes discovering increasingly larger prime numbers of exceptional interest to people in the cybersecurity field.

The Sieve of Eratosthenes, one of the first methods of determining primality still in use today, remains incredibly efficient at finding smaller prime numbers. However, its importance today mainly comes from the idea of using a sieve at all, as it has inspired and been surpassed by many more powerful algorithms (O'Neill 104). In 1971 the longest known prime number was 6,002 digits long. A decade later, after the advent of the RSA cryptosystem, the longest known prime number was 39,751 digits long (Rowland 11). Today, the longest known prime number is 24,862,048 digits long (Palca). Although the algorithms for determining these new primes may look and operate very differently from Eratosthenes' original sieve, he laid a vital piece of the groundwork for developing these algorithms in the first place.

CONCLUSION

While no longer the most efficient algorithm for finding useful prime numbers, the original Sieve of Eratosthenes inspired and guided many of the subsequent versions of primality test algorithms. In fact, up until as late as the 1980s, variations on the original sieve remained the only procedures for identifying prime numbers. This algorithm's implementations continue to be used today as convenient benchmarks to test some aspects of computer performance or efficiency (Bokhari 1). Eratosthenes had no way of predicting prime numbers' contemporary importance and consequence. Yet, thousands of years ago, he drafted a momentous algorithm relevant to this day. Due to the significance of prime numbers for things like blockchains, his algorithm continues to be referenced, implemented, thought about, and improved.

In my opinion, this algorithm is uniquely beautiful because it describes an exceedingly simple and intuitive procedure. It does not even require the operator to perform basic mathematical calculations. If you know how to count, you can sieve in the exact same way as Eratosthenes.

Works Cited

Bokhari, S H. *Multiprocessing the Sieve of Eratosthenes*. United States: N. p., 1987. Web. doi:10.1109/MC.1987.1663535.

Boyer, Carl B. “Euclid of Alexandria.” *A History of Mathematics*. 2nd ed., John Wiley & Sons, 1991.

Dubisch, Roy. “The Sieve of Eratosthenes.” *The Arithmetic Teacher*, vol. 18, no. 4, 1971, pp. 236–37. *JSTOR*, http://www.jstor.org/stable/41186370. Accessed 28 Jul. 2022.

Mollin, Richard A. “A Brief History of Factoring and Primality Testing B. C. (Before Computers).” *Mathematics Magazine*, vol. 75, no. 1, 2002, pp. 18–29. *JSTOR*, https://doi.org/10.2307/3219180. Accessed 28 Jul. 2022.

O'Neill, Melissa E. "The Genuine Sieve of Eratosthenes." *Journal of Functional Programming*, vol. 19, no. 1, 2009, pp. 95-106. *ProQuest*, http://ezproxy2.library.drexel.edu/login, doi:https://doi.org/10.1017/S0956796808007004.

Palca, Joe. “The World Has A New Largest-Known Prime Number.” *NPR*, 21 Dec. 2018, www.npr.org/2018/12/21/679207604/the-world-has-a-new-largest-known-prime-number.

Quesada, Antonio R. "Recent Improvements to the Sieve of Eratosthenes." *The Mathematics Teacher*, vol. 90, no. 4, 1997, pp. 304-307. *ProQuest*, http://ezproxy2.library.drexel.edu/login.

Roller, Duane W. *Eratosthenes' “Geography,”* Princeton University Press, 2010.

Salvo, Ivano, and Agnese Pacifico. *Three Euler's Sieves and a Fast Prime Generator (Functional Pearl).* Cornell University Library, arXiv.org, Ithaca, 2018. *ProQuest*, http://ezproxy2.library.drexel.edu/login?url=https://www.proquest.com/working-papers/threeeulers-sieves-fast-prime-generator/docview/2138927584/se-2.